
Caracterização e Análise de Texturas Estáticas e Dinâmicas Utilizando Redes Complexas, Caminhadas Determinísticas e Campos Vetoriais

Leandro Nogueira Couto



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

Uberlândia
2018

Leandro Nogueira Couto

**Caracterização e Análise de Texturas Estáticas
e Dinâmicas Utilizando Redes Complexas,
Caminhadas Determinísticas e Campos
Vetoriais**

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: Profa. Dra. Celia Aparecida Zorzo Barcelos

Uberlândia
2018

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

C871c Couto, Leandro Nogueira, 1986-
2018 Caracterização e análise de texturas estáticas e dinâmicas utilizando
redes complexas, caminhadas determinísticas e campos vetoriais /
Leandro Nogueira Couto. - 2018.
115 f. : il.

Orientadora: Celia Aparecida Zorzo Barcelos.
Tese (Doutorado) - Universidade Federal de Uberlândia, Programa
de Pós-Graduação em Ciência da Computação.
Disponível em: <http://dx.doi.org/10.14393/ufu.te.2018.754>
Inclui bibliografia.

1. Computação - Teses. 2. Processamento de imagens - Teses. I.
Barcelos, Celia Aparecida Zorzo. II. Universidade Federal de
Uberlândia. Programa de Pós-Graduação em Ciência da Computação.
III. Título.

CDU: 681.3

Maria Salete de Freitas Pinheiro – CRB6/1262

*A Cynthia, minha amada esposa, amiga e companheira,
a meus amados pais, Leomam e Regina, e minha amada irmã Deborah,
e a Deus, com toda minha gratidão.*

Agradecimentos

O autor e orientadora deste trabalho gostariam de agradecer ao Prof. Dr. André Ricardo Backes da Universidade Federal de Uberlândia pela disponibilização do código de geração de Redes Complexas e colaborações com a pesquisa. Agradecemos também ao Prof. Dr. Eraldo Ribeiro do Florida Institute of Technology pela disponibilização do código para extração de padrões singulares de campos vetoriais.

O autor agradece à sua orientadora Profa. Dra. Celia Aparecida Zorzo Barcelos pela orientação, atenção e zelo; agradece também à Faculdade de Computação (FACOM) na pessoa de seu diretor Prof. Dr. Ilmério Reis da Silva; agradece ainda à Universidade Federal de Uberlândia por graciosamente permitir a conciliação de suas atividades como docente e como aluno de Doutorado; agradece ao Prof. MSc. Thiago Pirola Ribeiro pelas colaborações na pesquisa, execução e geração de resultados de diversos experimentos e a Cynthia Martins Villar Couto pela colaboração com a geração de diversas figuras.

Resumo

A classificação de texturas envolve a extração de características descritivas da imagem. Este trabalho propõe novas abordagens para modelagem e caracterização computacional de texturas estáticas e dinâmicas de maneira discriminativa, aplicadas à classificação de imagens e vídeos. Um novo descritor proposto é baseado em estatísticas extraídas de caminhadas determinísticas em uma transformação da textura inspirada em redes complexas, focada na representação dos padrões locais de vizinhança dos *pixels* para construir o vetor de características. Extensões do método para classificação de texturas dinâmicas são apresentadas juntamente um descritor discriminativo baseado em histogramas de padrões singulares de fluxo óptico são apresentadas. Experimentos aplicando os métodos à classificação de textura em diversas bases de dados difundidas na literatura mostram que os descritores propostos são eficazes, em diversos casos aprimorando taxas de classificação corretas quando comparadas a outros métodos do estado-da-arte enquanto usam um vetor de característica de tamanho menor.

Palavras-chave: Caracterização de Texturas. Caminhadas Determinísticas. Texturas Dinâmicas. Redes Complexas. Fluxo Óptico. Campos Vetoriais.

Abstract

Texture classification involves acquiring descriptive features from the image. This work proposes new approaches to mathematical characterization and computational modeling of static and dynamic textures in a discriminating way, applied to classification and recognition of images and videos. A novel proposed descriptor is based on statistics generated from deterministic walks on a complex network inspired transformation of the texture, focusing on the representation of the neighborhood pattern of the pixels to build the feature vector. An extension of the method for dynamic texture classification and a novel approach based on optical flow recurring pattern histograms are also presented. Experiments applying the methods for texture classification on several widespread data sets show that the proposed descriptors improves correct classification rates compared to other state-of-the-art methods while using a smaller feature vector.

Keywords: Deterministic Walks. Texture Characterization. Dynamic Textures. Complex Networks. Optical Flow. Vector Fields.

Lista de ilustrações

Figura 1 – Cada raio de vizinhança r oferece um conjunto de n vizinhos. Por exemplo, se $r = \sqrt{2}$, um pixel i possui os 8 vizinhos indicados em b . . .	28
Figura 2 – A partir do pixel i marcado em (a), estabelece-se os vizinhos candidatos em (b). Após a aplicação do limiar t , os vizinhos do pixel i são mostrados em cinza em (c) (Adaptado de Backes <i>et al.</i> (BACKES; CASANOVA; BRUNO, 2013)).	29
Figura 3 – Divisão de um vídeo em quadros. Respectivamente, da esquerda para a direita, divisão em quadros $[X, Y]$, quadros $[X, T]$ e quadros $[Y, T]$. (Adaptado de Gonçalves e Bruno (GONÇALVES; BRUNO, 2013b)).	30
Figura 4 – a) mostra a vizinhança de um pixel. b) resulta da aplicação da Equação 6 para gerar os valores binários da vizinhança do pixel, que é lida no sentido da seta (do bit menos significativo para o mais significativo). c) mostra os pesos usados na conversão do vetor binário para um valor decimal, e a conversão em si é mostrada em d).	31
Figura 5 – Exemplo de caminhada com $\mu = 3$ começando no pixel superior esquerdo, em que os pixels do transiente estão mostrados em verde (os 7 primeiros passos da caminhada) e os do período em azul (5 últimos passos), seguindo a dinâmica min	35
Figura 6 – Histograma de tamanhos de transiente e período da caminhada aplicado à base Brodatz (BRODATZ, 1966). Note como texturas de um mesmo objeto fornecem histogramas semelhantes. Extraído de (BACKES <i>et al.</i> , 2010).	36
Figura 7 – Dicionário de fluxos base para padrões singulares. A combinação linear desses padrões pode aproximar o comportamento de um campo vetorial em um ponto qualquer. Extraído de (LIU; RIBEIRO, 2012).	41

Figura 8 – Passos para geração do grafo D de graus médios. Uma rede G_t é criada baseado nas intensidades dos pixels da imagem I e em um valor de limiar t . O grau de cada vértice G_t gera o valor de intensidade do pixel correspondente na imagem transformada D_t . A média das imagens D_t gera D . (Adaptado de Backes <i>et al.</i> (BACKES; CASANOVA; BRUNO, 2013))	45
Figura 9 – Visão geral do método proposto. Da imagem I obtém-se a imagem D . Uma caminhada é gerada para cada valor de dyn , μ e cada pixel i . Cada passo w da caminhada W para cada pixel i é acrescentado a um histograma de direções de passos $H_{\mu,dyn}$, incrementando o <i>bin</i> de H correspondente à direção do passo. $H_{\mu,dyn}$ descreve a imagem inteira.	47
Figura 10 – Partindo da textura original em a) e usando a dinâmica <i>min</i> , temos atratores de caminhadas determinísticas gerados pelos valores de memória: b) $\mu = 1$, c) $\mu = 3$ e d) $\mu = 7$. Os padrões em f), g) e h) foram gerados pelos mesmos valores de memória que b), c) e d), respectivamente, mas usando dinâmica de caminhada <i>max</i> . Adaptado de (BACKES et al., 2010).	48
Figura 11 – Composição do vetor de características F_I a partir de suas partes. O vetor final é uma combinação dos histogramas obtidos para diferentes parâmetros, unido às estatísticas S_t dos histogramas.	50
Figura 12 – Comparação de diferentes vetores de características para dois conjuntos de três imagens cada, cada conjunto de um classe diferente (representadas por cores diferentes). O eixo horizontal mostra os índices do vetor, e o eixo vertical mostra seu valor.	51
Figura 13 – Exemplos de histogramas de direção de passos arranjados em um eixo de direções dir e um eixo de tamanhos de memória μ . As imagens a) e b) pertencem à mesma classe, enquanto c) e d) são um par de outra classe.	51
Figura 14 – Vizinhança tridimensional de pixel, considerando 26 vizinhos. Adaptado de https://clouard.users.greyc.fr/Pandore/c++doc/html . (Acessado em 03/04/2018).	55
Figura 15 – Sumário do método de caminhadas volumétricas. A partir da transformação baseada em redes complexas dos quadros do vídeo V em quadros D , as caminhadas volumétricas são aplicadas em cada pixel (cada passo de caminhada tem 3 coordenadas), gerando por fim o histograma de direções de passos de caminhada descritor de V	56

Figura 16 – Sumário do método de caminhadas TOP (três planos ortogonais). A partir da transformação baseada em redes complexas dos quadros do vídeo V em quadros D , caminhadas são aplicadas em cada plano, gerando um histograma de direções de passos de caminhada para cada plano, que concatenados compõem o descritor de V	57
Figura 17 – Exemplo de execução do método de detecção de padrões singulares (LIU; RIBEIRO, 2012) sobre vídeo de uma bandeira. a) ilustra o quadro original, b) apresenta linhas onde o fluxo óptico é acentuado entre os quadros, c) apresenta os vetores do fluxo óptico, d) ilustra os padrões detectados pelo método; o tamanho do círculo representa magnitude e a cor representa a magnitude em logaritmo da energia singular do padrão e e) apresenta o subconjunto de padrões mais fortes (de maior energia singular).	60
Figura 18 – Exemplo de execução do método de detecção de padrões singulares (LIU; RIBEIRO, 2012) sobre vídeo de uma piscina. a) ilustra o quadro original, b) apresenta linhas onde o fluxo óptico é acentuado entre os quadros, c) apresenta os vetores do fluxo óptico, d) ilustra os padrões detectados pelo método; o tamanho do círculo representa magnitude e a cor representa a magnitude em logaritmo da energia singular do padrão e e) apresenta o subconjunto de padrões mais fortes (de maior energia singular).	61
Figura 19 – Resumo da geração do vocabulário na metodologia <i>bag-of-features</i> aplicada aos padrões singulares de campos vetoriais.	63
Figura 20 – Agrupamento dos pontos de interesse, ou <i>features</i> , adquiridos pelo detector de padrões singulares no vocabulário gerado da <i>bag-of-features</i> , de onde é obtido o histograma de ocorrências.	64
Figura 21 – Sumário da geração do vetor de características por histogramas de coeficientes a_k . Neste método os pontos de interesse obtidos do fluxo óptico tem os coeficientes de seus vetores agrupados em histogramas, cada histograma agrupando as ocorrências de valores para um coeficiente a_k	65
Figura 22 – Exemplos de imagens da base de texturas Brodatz.	69
Figura 23 – Exemplos de quadros extraídos de vídeos da base de texturas dinâmicas Dyntex.	70
Figura 24 – Histograma linear e $\log \times \log$ de 8 amostras de classes da base Brodatz. A análise mostra que a distribuição não segue uma lei de potência, mas em alguns casos aproxima uma Gaussiana.	72
Figura 25 – Histogramas de ocorrências de valores de intensidade de pixels das amostras da Brodatz usadas na análise de distribuição de graus.	73

Figura 26 – Taxas de acerto para valores de memória μ isolados na base Brodatz, considerando as 3 combinações de dinâmicas.	76
Figura 27 – Exemplo de matriz de confusão para caminhadas na base Brodatz. Ao lado dela há amostras dos pares de classes que causaram maior confusão ao classificador, as classes 44 e 107 (4 classificações incorretas).	78
Figura 28 – Tamanhos médios de caminhadas para cada tamanho de memória μ , na base Brodatz. As barras verticais indicam o desvio padrão.	79
Figura 29 – Tamanhos médios de caminhadas volumétricas para diferentes tamanhos de memória μ , na base DynTex. As barras verticais indicam o desvio padrão.	84
Figura 30 – Exemplo de vetor de características para caminhadas volumétricas. As caminhadas com dinâmica <i>min</i> , mostradas na metade esquerda do vetor, valorizam de forma pronunciada as direções $V(x, y, t + 1)$ ou $V(x, y, t - 1)$ na maioria das texturas dinâmicas da base DynTex, quando comparadas às caminhadas com dinâmica <i>max</i> , à direita.	86
Figura 31 – a) Exemplo de matriz de confusão para caminhadas TOP na base DynTex. b) Amostras dos pares de classes que causaram maior confusão mútua do classificador, de cima para baixo: o par 15 e 72 (ambas texturas com movimentação circular, com 4 classificações incorretas), e o par 62 e 65 (momentos diferentes do mesmo cenário, com 3 classificações incorretas).	88
Figura 32 – O histograma de valores de a_k amostrados de múltiplos vídeos da DynTex mostra que a distribuição de valores de a_k segue aproximadamente uma distribuição normal (Gaussiana).	90
Figura 33 – Soma dos erros quadrados para cada <i>cluster</i> por número de grupos k . É possível ver o "cotovelo" da curva do gráfico, que indica um valor adequado de k	92
Figura 34 – Comparação de matrizes de confusão para testes com a) inclusão e b) exclusão das classes problemáticas (que possuem fluxo óptico de baixa magnitude). Em c) temos amostras destas classes (classes 17, 19, 33 e 35).	95

Lista de tabelas

Tabela 1 – Sumário das bases de dados usadas na validação dos métodos da proposta.	70
Tabela 2 – Comparação de taxas de classificação na base Brodatz com e sem ordenação do vetor de características. Colunas assinaladas com * indicam resultados obtidos para vetores ordenados.	75
Tabela 3 – Resultados de classificação na base Brodatz para dinâmicas isoladas e combinadas, com e sem as estatísticas <i>St</i> , e apenas com as estatísticas <i>St</i> .	77
Tabela 4 – Comparação das taxas de classificação para os três diferentes conjuntos de tamanhos de memória. Melhores resultados em negrito.	78
Tabela 5 – Comparação do método proposto com outros métodos da literatura. Métodos das 8 primeiras linhas foram executados com a mesma implementação do LDA. Melhores resultados em negrito.	80
Tabela 6 – Comparação do método proposto com métodos da literatura sobre as bases KTH _TIPS, UIUCTex e CURet (melhores resultados em negrito).	80
Tabela 7 – Resultados para caminhadas sobre a base de dados Brodatz com raio $r_D = 2\sqrt{2}$. Melhores resultados em negrito	81
Tabela 8 – Resultados de classificação, na base Brodatz com o raio $r_D = 2\sqrt{2}$ e $r_D = \sqrt{2}$.	81
Tabela 9 – Resultados para o descritor de histograma de direções de passos aplicados sobre imagens <i>I</i> sem transformação, para cada dinâmica (melhores resultados e negrito).	82
Tabela 10 – Resultados de classificação, na base Brodatz com intervalo de memória $[1, 8]$, das caminhadas sobre a imagem original <i>I</i> e sobre a transformada <i>D</i> .	82
Tabela 11 – Resultados de classificação com caminhadas volumétricas (sobre as 3 dimensões (X, Y, T)), na base Dyntex, com e sem adição dos valores de estatísticas <i>St</i> ao descritor.	84

Tabela 12 – Resultados de classificação com caminhadas volumétricas (sobre as 3 dimensões (X, Y, T)), na base Dyntex, com e sem adição do vetor de estatísticas St , sobre a transformação baseada em rede complexa da imagem.	85
Tabela 13 – Resultados de classificação com caminhadas TOP (sobre os 3 planos ortogonais (X, Y) , (X, T) e (Y, T)), na base Dyntex, com e sem adição do vetor de estatísticas St , sobre a transformação baseada em rede complexa da imagem.	87
Tabela 14 – Comparação do melhor resultado do nosso método com o melhor resultado do $LBP - TOP[8, 8, 8]$, para cada plano individualmente e para os 3 planos juntos. Melhor resultado em negrito.	88
Tabela 15 – Taxas de classificação correta para o descritor de padrões singulares em fluxos ópticos sobre a base Dyntex. Melhor resultado em negrito. . .	94
Tabela 16 – Taxas de classificação correta para o descritor de padrões singulares em fluxos ópticos sobre a base Dyntex. Melhor resultado em negrito. . .	96

Sumário

1	INTRODUÇÃO	19
1.1	Textura como característica da imagem	19
1.2	Visão Geral da Tese	22
1.3	Objetivos	23
1.3.1	Objetivo Geral	23
1.3.2	Objetivos Específicos	23
2	FUNDAMENTAÇÃO TEÓRICA	25
2.1	Representação Computacional de Imagens	25
2.2	Representação de Imagens e Vídeos como Grafos ou Redes Complexas	26
2.3	Métodos Afins e Caminhada Determinísticas	30
2.3.1	Padrões Binários Locais (<i>Local Binary Patterns</i>)	30
2.3.2	Caminhadas Determinísticas	33
2.3.3	Estatísticas de Histograma	37
2.4	Fluxo óptico e padrões singulares em campos vetoriais	39
2.4.1	Fluxo óptico	39
2.4.2	Padrões singulares em campos vetoriais	40
3	PROPOSTA E METODOLOGIA	43
3.1	Geração de rede para caminhadas	43
3.1.1	Parâmetros da transformação de rede complexa	45
3.2	Caminhadas Determinísticas Auto-Repelentes sobre a Rede	46
3.2.1	Complexidade computacional	50
3.2.2	Parâmetros de caminhadas determinísticas	52
3.3	Caminhadas determinísticas tridimensionais	53
3.3.1	Complexidade computacional	58

3.4	Padrões singulares de campos vetoriais aplicados a texturas dinâmicas	58
3.4.1	Vetor de características por <i>bag-of-features</i>	62
3.4.2	Vetor de características por histograma de coeficientes a_k	64
3.4.3	Complexidade computacional	66
4	EXPERIMENTOS E ANÁLISE DOS RESULTADOS	69
4.1	Bases de dados de textura	69
4.2	Análise da Rede Complexa	71
4.3	Experimentos do método das caminhadas estáticas	74
4.3.1	Parâmetros dos experimentos	74
4.3.2	Resultados e discussão	75
4.4	Experimentos com caminhadas em texturas dinâmicas	82
4.5	Experimentos de descrição de texturas dinâmicas por padrões singulares de campos vetoriais	89
4.5.1	Parâmetros dos experimentos	89
4.5.2	Resultados e discussão	92
5	CONCLUSÃO	97
5.1	Principais Contribuições	97
5.2	Trabalhos Futuros	98
5.2.1	Melhor invariância a rotação e escala	98
5.2.2	Melhor invariância a rotação e escala	98
5.2.3	Avaliação da evolução dos limiares	99
5.2.4	Integração com LBP	99
5.2.5	Análise fractal	99
5.2.6	Padrões singulares com <i>Three Orthogonal Planes</i>	100
5.3	Contribuições em Produção Bibliográfica	100
REFERÊNCIAS		103

CAPÍTULO 1

Introdução

Dados visuais, como imagens e vídeos, contêm grande riqueza de informações. Essas informações, porém, estão implícitas nos dados, e sua obtenção a partir da imagem é uma tarefa complexa. A aquisição de informação relevante e aplicável a partir de uma imagem ou vídeo digital é um grande desafio atual da computação, e requer muitas vezes técnicas de processamento de imagens elaboradas e computacionalmente exigentes. O paralelo com a visão humana evidencia a complexidade e significatividade da informação visual, já que a visão é um dos sentidos mais importantes e complexos do ser humano. De fato, muitas metodologias novas para realização de visão computacional e processamento de imagens baseiam-se em compreender e emular o funcionamento da visão humana (JULESZ, 1975) (LANDY; GRAHAM, 2004). Com os avanços contínuos da capacidade de processamento dos computadores, torna-se cada vez mais possível o estudo aprofundado de métodos para análise e obtenção de informação de imagens e vídeos digitais, e, recentemente, abordagens distintas têm oferecido bons resultados em diversas aplicações.

1.1 Textura como característica da imagem

A textura pode ser uma característica discriminativa poderosa de uma imagem ou de um vídeo. Enquanto há um entendimento natural do conceito, pode ser difícil formalizar uma definição para o termo textura. Hájek (HÁJEK, 2006) propõe que texturas são padrões visuais complexos compostos de elementos espacialmente organizados, com características particulares como brilho, cor, forma e tamanho, entre outras. Tais padrões locais podem ser descritos por propriedades taxonômicas como granularidade, contraste, direcionalidade, regularidade, aspereza, entre outros (TAMURA; MORI; YAMAWAKI, 1978)(LAWS, 1980) (HÁJEK, 2006). Um conceito importante da imagem é a autossimilaridade (FLORINDO; BRUNO, 2012).

Características baseadas em texturas vêm sendo aplicadas a diversas tarefas de visão computacional com sucesso há muito tempo. Entre as aplicações da análise de textura na computação, podemos citar as seguintes (sem pretensões de fazer uma listagem exaustiva):

- ❑ Recuperação de imagens e vídeos baseada em conteúdo (HIREMATH; PUJARI, 2007): chamada de CBIR ou CBVR (*Content Based Image Retrieval* ou *Content Based Video Retrieval*), a recuperação de imagens e vídeos baseada em conteúdo, ao invés de usar metadados, requer que os elementos da imagem sejam caracterizados. Características baseadas em textura são algumas das descrições possíveis.
- ❑ Segmentação de imagens e vídeos (BELONGIE et al., 1998) (GAETANO; SCARPA; POGGI, 2009) (GONÇALVES; BRUNO, 2013b): a textura é uma característica importante para definir bordas e regiões de uma imagem.
- ❑ Síntese e reconstrução de imagens e vídeo (EFROS; LEUNG, 1999) (GUO et al., 2013) (KATO; HARADA, 2014) a caracterização de uma textura permite que ela seja estendida, replicada e modificada, com aplicações em computação gráfica e edição de imagens.
- ❑ Agrupamento e classificação de imagens (BACKES et al., 2010) (ZHAO et al., 2013) (SERRA; GRANA; CUCCHIARA, 2014): o agrupamento e categorização de padrões como a textura é uma função importante do aprendizado de máquina.
- ❑ Reconstrução 3D (PAYET; TODOROVIC, 2011): Aspectos das texturas de uma imagem ou vídeo real são indicadores de atributos da cena original como distâncias relativas e perspectiva.
- ❑ Compressão de imagens (PATIL et al., 2013): Representações mais sucintas de texturas permitem que informação complexa e extensa seja expressa em um menor espaço.

Estas tarefas computacionais se fazem aplicáveis na prática em campos importantes como segurança e detecção de eventos (MA; CISAR, 2009), análise de imagens biológicas ou médicas (CASTELLANO et al., 2004) (KASSNER; THORNHILL, 2010)(NANNI; LUMINI; BRAHNAM, 2010) (PUJARI; YAKKUNDIMATH; BYADGI, 2015) (WAN et al., 2017), análise de comportamento social (GHIDONI; CIELNIAK; MENEGATTI, 2013) e robótica móvel (BLAS et al., 2008) (SHINZATO; WOLF, 2011)(SOUZA; OSORIO; WOLF, 2013), para citar alguns exemplos.

Quando se trata de processamento de vídeos, sequências de imagens consecutivas de modo que existe uma componente temporal, a percepção de auto-similaridade e periodicidade das texturas estáticas, bidimensionais, pode ser estendido para incluir padrões recorrentes na dimensão temporal, como vídeo que apresentam padrões de movimentação (CHETVERIKOV; PÉTERI, 2005). Tais texturas são chamadas de texturas dinâmicas (DORETTO et al., 2003). São texturas cuja análise, para que seja completa, deve incluir as variações no tempo, porque as características distintivas dos elementos da textura dinâmica não se limitam a padrões estacionários. Sequências de imagens que exemplificam

texturas dinâmicas incluem imagens com elementos cuja dinâmica pode ser analisada de forma agregada, como tráfego de pedestres e veículos, ou processos naturais que apresentam comportamento dinâmico estocástico, como fogo, água e vegetação ao vento (DER-PANIS; WILDES, 2012). A dimensão temporal traz como consequência complexidade adicional à análise, mas também maior quantidade de informação em potencial que pode ser usada para fazer distinção entre texturas diferentes.

Para que seja possível a análise computacional de texturas, é necessária em primeiro lugar uma modelagem computacional da textura, que represente as características e informação de textura a partir dos dados visuais. Esta é uma tarefa complexa pois os padrões da textura, embora reconhecíveis ao olho humano, não necessariamente se traduzem em periodicidade dos dados, havendo muitas texturas com arranjos ruidosos e com características estocásticas, e a percepção dos padrões pode até mesmo mudar em escalas distintas (SIQUEIRA; SCHWARTZ; PEDRINI, 2013). Há uma vasta variedade de métodos de descrição de texturas, que na literatura recente têm sido agrupados em quatro categorias principais (GONÇALVES; MACHADO; BRUNO, 2012) (ANDREARCZYK; WHELAN, 2017): métodos baseados em movimento (exclusivos de texturas dinâmicas) (FAZEKAS; CHETVERIKOV, 2007), métodos de análise espectral processamento de sinais (filtros e transformadas) (JAIN; FARROKHNA, 1991), métodos baseados em modelos (REED; DUBUF, 1993)(TUCERYAN; JAIN, 1998)(CHETVERIKOV; PÉTERI, 2005) e métodos estatísticos (HARALICK, 1979). Estes últimos estudam a distribuição e o comportamento estatístico de propriedades geométricas e características locais da imagem (OJALA; PIETIKAINEN; MAENPAA, 2002) (VARMA; ZISSERMAN, 2005) (ZHAO et al., 2013) (ZHANG et al., 2015) , enquanto outros métodos visam caracterizar texturas através de processos que lidam com a identificação e análise de elementos de mais alto-nível da textura (TODOROVIC; AHUJA, 2009). Métodos locais tem obtido resultados significativos, com alto poder discriminativo, mas sofrem com ruído na imagem. No caso do LBP (Padrões Binários Locais) (OJALA; PIETIKAINEN; MAENPAA, 2002) (ZHAO; PIETIKAINEN, 2007), o método apresenta custos computacionais proibitivos para descrever vizinhanças maiores da imagem (LIU et al., 2016b). Apesar disso, métodos como o LBP e variantes tem sido amplamente explorados na literatura, enquanto métodos como caminhadas determinísticas representam uma área menos estudada da análise de texturas, sendo um dos campos que este trabalho pretende explorar.

Metodologias como aquelas baseadas em descrição de movimento por fluxo óptico (FAZEKAS; CHETVERIKOV, 2007), são por natureza exclusivas a texturas dinâmicas, dada sua dependência da dimensão temporal, mas nas demais categorias há muitas abordagens de análise de texturas estáticas que apresentam análogos aplicados à análise de texturas dinâmicas. Métodos estatísticos, em especial, prestam-se particularmente à descrição de ambas texturas estáticas e dinâmicas e podem ser adaptados sem muita dificuldade. Por outro lado, o volume de dados e a demonstradamente alta dimensionalidade característica

da análise de texturas em relação a outras tarefas de visão computacional são obstáculos severos para o uso de diversas ferramentas (BASU et al., 2018). Métodos globais podem ser mais eficientes, por sua vez, mas nem sempre capturam as nuances necessárias para classificação de textura. Uma forma de contornar isso é combinar métodos locais em um descritor global. Este trabalho também pretende explorar o potencial de descrição global de metodologias para combinação de descritores locais de fluxo óptico aplicados a texturas dinâmicas.

1.2 Visão Geral da Tese

Esta tese consiste da pesquisa e desenvolvimento de métodos para caracterização de texturas estáticas e dinâmicas de maneira a gerar descritores precisos e abrangentes. Uma das abordagens propostas baseia-se na representação de imagens como grafos ou redes complexas. Existem diversas métricas e estatísticas que podem ser aplicadas na caracterização de grafos e redes complexas. Métodos estatísticos são propícios para caracterização de texturas devido à periodicidade que estas apresentam, e um método de análise de grafos e redes complexas que é de particular interesse para esse projeto é o método das caminhadas determinísticas auto-repelentes, também conhecidas como caminhadas do turista (BACKES et al., 2006). As caminhadas fundamentam-se no emprego de um agente que realiza um percurso sobre a topologia da rede e sobre os dados. A caminhada resultante pode colocar em evidência características da rede, e esta tese propõe novas informações descritivas sobre as características da caminhada. Outro método estatístico relevante a este trabalho é a análise de imagens baseada na transformação de redes complexas proposta por Backes *et al.* (BACKES; CASANOVA; BRUNO, 2013). Outra abordagem de relevância para este projeto é a análise baseada em fluxo óptico aplicada a texturas dinâmicas. O fluxo óptico uma categoria muito prolífica em termos de quantidade de métodos para caracterização de texturas dinâmicas (CHETVERIKOV; PÉTERI, 2005). O trabalho proposto apresenta um viés diferenciado pois propõe uma análise baseada em componentes geradores de campos vetoriais. Entre os trabalhos de caracterização de campos vetoriais, o trabalho de Liu e Ribeiro (LIU; RIBEIRO, 2012) se destaca pelo alto potencial de condensação e fidelidade de representação da informação do campo vetorial, que é modelado por padrões singulares formados por uma combinação de componentes extraídos de um dicionário de fluxos base. Dado que o fluxo óptico é um campo vetorial, este projeto propõe que, devido à periodicidade da textura, uma análise estatística da frequência de ocorrência e composição destes padrões singulares deverá fornecer descritores característicos.

A próxima seção elenca os objetivos gerais e específicos deste trabalho. O Capítulo 2 consiste da apresentação e análise de conceitos e trabalhos da literatura relacionados a este trabalho, e de uma revisão da bibliografia e trabalhos do estado-da-arte da área. Em

seguida, o Capítulo 3 baseia-se no arcabouço teórico derivado da literatura para descrever a metodologia e abordagens usadas nas propostas de descrição de texturas desta tese. No Capítulo 4 apresentamos os experimentos realizados e parâmetros experimentais utilizados para validação da metodologia desenvolvida, em tarefas de classificação de texturas estáticas e dinâmicas, sobre bases de dados de ampla difusão na literatura; apresentamos os resultados juntamente com nossa análise, e realizamos um estudo comparativo dos resultados obtidos com os de outros métodos do estado-da-arte. Por fim, o Capítulo 5 contém considerações finais, as contribuições e conclusões obtidas por esta tese, além de diversas sugestões para trabalhos futuros que desenvolvam sobre as contribuições deste trabalho.

1.3 Objetivos

1.3.1 Objetivo Geral

O objetivo geral desta tese é caracterizar texturas estáticas e dinâmicas através de métodos estatísticos inéditos. Os métodos propostos são fundamentados nos métodos da literatura de caminhadas determinísticas, análise de redes complexas e fluxo óptico. A partir destes métodos da literatura, a tese busca propor novas informações com alto poder discriminativo para serem extraídas de imagens ou vídeos.

1.3.2 Objetivos Específicos

Os objetivos específicos da tese são:

- ❑ Explorar uma transformação de imagem baseada em análise de redes complexas como potencializadora da capacidade discriminativa de outros métodos;
- ❑ Caracterizar texturas estáticas e dinâmicas através da extração de novas estatísticas discriminativas a partir do método de caminhadas determinísticas;
- ❑ Descrever texturas dinâmicas com um descritor estatístico global baseado na detecção de padrões singulares locais do fluxo óptico agrupados com diferentes estratégias de histograma;
- ❑ Verificar o poder discriminativo de estatísticas de histograma como potencializador dos vetores de características dos métodos propostos;

CAPÍTULO **2**

Fundamentação Teórica

Neste capítulo, realizaremos uma revisão bibliográfica da literatura seminal que inspirou as propostas deste trabalho, e do estado-da-arte de soluções de interesse para o problema de classificação de texturas. O detalhamento de certos aspectos de alguns dos métodos será realizado no Capítulo 3, onde os conceitos serão necessários.

2.1 Representação Computacional de Imagens

A representação digital mais comum para uma imagem em níveis de cinza é a de uma função bidimensional $I(x, y)$, onde x e y representam as coordenadas espaciais (no plano) dos pontos da imagem e a amplitude de I em um par de coordenadas (x, y) representa a intensidade da imagem naquele ponto. Esta intensidade é dada por um dentre um conjunto finito de valores em um intervalo $[0, L]$. A cada elemento discreto da imagem, representado por coordenada e valor, dá-se o nome de pixel. O processo de digitalização de uma imagem contínua em termos de suas coordenadas é chamado de amostragem, enquanto a digitalização dos valores de amplitude, ou intensidade luminosa, da imagem contínua é chamado de quantização (GONZALES; WOODS; EDDINS, 2004).

O processo de digitalização é por natureza uma aproximação, o que traz diversas implicações em relação ao processamento da imagem digital final. O número de valores de amplitude possíveis é chamado de profundidade de cores (ou *bit depth*), e afeta os valores de intensidade da imagem. Uma imagem em tons de cinza com profundidade de 8-bits apresenta 256 (ou 2^8) valores distintos possíveis, o que é suficiente na maioria dos casos para preservar os valores relativos entre pixels. Um problema mais significativo, no caso da análise de texturas, é a frequência de amostragem espacial, chamada de resolução. Uma vez realizada a amostragem, quaisquer padrões sub-pixel são perdidos. Para uma representação completa da imagem contínua, o teorema da amostragem de Nyquist-Shannon declara que a frequência de amostragem deve ser no mínimo duas vezes maior que a maior frequência da imagem original (HAYKIN; VEEN, 2007). Ainda com esses cuidados, variações de resolução tem efeito significativo na vizinhança dos pixels, o que

afeta a representação da textura, como um padrão fundamentalmente local. A partir de agora, sempre que nos referirmos a imagens estaremos nos referindo a imagens digitais, exceto onde notado.

Outra consideração a fazer é que a textura não é uma informação explícita na imagem. As únicas informações explícitas na imagem são a intensidade e coordenada de cada pixel. Assim, toda aplicação de visão computacional que pretende utilizar os atributos de textura da imagem requer que, em primeiro lugar, a textura seja modelada computacionalmente. Essa extração de características a partir dos dados é fundamental na análise de texturas. Após isso, em alguns métodos a informação extraída é codificada; um exemplo de codificação são processos estatísticos como a criação de um histograma agrupando e compilando os dados obtidos em um conjunto de metadados, como veremos neste capítulo.

Uma vez que os passos anteriores foram realizados, haverá um descritor da textura que representa suas características explicitamente. Em aplicações de classificação, por exemplo, descritores de texturas são usado para o treinamento de um classificador que seja capaz de separar as classes e estimar a classe de novos descritores. Se o treinamento é realizado a partir de um conjunto de classes pré-determinado e conhecido, a categorização é um processo de aprendizado supervisionado chamado de classificação; se a categorização ocorre por aglomeração de texturas similares sem determinação prévia de classes, o aprendizado é chamado de não-supervisionado, e é chamado de agrupamento, ou *clustering* (DUDA; HART; STORK, 2012). Métodos para aprendizado de máquina são diversos e profundamente estudados, já que sua aplicação independe da origem dos dados, o que os torna adequados para muitas aplicações. A extração de características de textura, por sua vez, não é assim. Características adequadas para descrever textura de forma discriminativa são intrinsecamente ligadas às características de texturas como padrões locais periódicos. Um bom descritor deve considerar essas particularidades.

2.2 Representação de Imagens e Vídeos como Grafos ou Redes Complexas

Recentemente, métodos envolvendo a aplicação de análise de grafos ou redes complexas à descrição de texturas (estáticas ou dinâmicas) têm obtido bons resultados. Redes complexas são grafos com topologias não-lineares, que não possuem uniformidade ou descrição trivial (BARABASI, 2003) (COSTA et al., 2007).

Para analisar as características da imagem como um grafo ou rede, é necessário que a imagem seja primeiramente modelada como tal. Considere a função bidimensional $I(x, y)$, representativa da imagem I . Para que essa função seja modelada como um grafo, uma abordagem encontrada frequentemente na literatura consiste em representar cada pixel como um vértice em um grafo. Considere dois pixels $i = (x, y)$ e $j = (x', y')$, ambos pertencentes à mesma imagem I . Considere ainda um grafo não-direcionado $G = (V, E)$,

construído pela associação de cada pixel $(x, y) \in I$ a um vértice $v_{x,y} \in V$, denotado por $v_{x,y}$ do grafo G . Ainda é preciso decidir a composição do conjunto de arestas E , isto é, se um par de vértices está conectado por uma aresta, ou seja, se $(v, v') \in E$. Uma forma de determinar a presença ou ausência de arestas é baseada na distância entre as coordenadas dos pixels (BACKES et al., 2010) (BACKES; CASANOVA; BRUNO, 2010) (GONÇALVES; BRUNO, 2013a). Pixels com coordenadas cuja distância Euclidiana é menor que um raio D geram nós vizinhos no grafo, como na Equação 1. Se $r_D = \sqrt{2}$, cada pixel (exceto nos casos particulares de pixels na borda da imagem) tem 8 vizinhos, como mostra a Figura 1.

Essa abordagem gera um grafo regular, ou seja, que apresenta simetria e periodicidade topológica, também conhecido como grafo *lattice* (da palavra do inglês que denota este tipo de padrão mosaico regular). Nesse caso a diferença entre o grafo gerado por uma imagem e outra não está na topologia, mas sim nos valores dos pixels.

$$\eta(x, y) = \{(x', y') | \sqrt{(x - x')^2 + (y - y')^2} \leq r_D\} \quad (1)$$

Uma vantagem de tratar a imagem como um grafo é que isso torna possível aplicar métricas de análise topológica de grafos ou de redes. Uma revisão dessas métricas pode ser encontrada em (COSTA et al., 2007). Tal informação pode ser utilizada para caracterização da estrutura da rede, que é representativa da textura a partir da qual ela foi gerada. Há trabalhos que usam uma abordagem baseada em redes para estabelecer a conectividade dos pixels no grafo (BACKES; CASANOVA; BRUNO, 2010) (BACKES; CASANOVA; BRUNO, 2013) (RIBEIRO et al., 2015). Nestes trabalhos, a representação de pixels de I como vértices de G permanece a mesma, mas a conectividade entre dois vértices $v_{x,y}$ e $v_{x',y'}$ do grafo G é determinada por uma medida de distância que denota a dissimilaridade entre os dois pixels que os vértices representam, $i = (x, y)$ e $j = (x', y')$. Tal métrica, nesse caso, não depende apenas das coordenadas do pixel mas também de seu valor $I(x, y)$, como mostra a Equação (3), e satisfaz os critérios de distância de simetria, não-negatividade, identidade de idênticos e desigualdade triangular. Gonçalves et al. (GONÇALVES et al., 2016) usam redes similares, porém direcionadas, para caracterizar textura com uma métrica de difusão.

Dois vértices $\omega(v_{x,y}, v_{x',y'})$ são representados por nós conectados em G se a distância Euclidiana entre eles é inferior a um raio r_G (ou seja, o par $(v_{x,y}, v_{x',y'}) \in E$, como mostra a Equação (3)), e se a medida de distância $\omega(v_{x,y}, v_{x',y'})$ entre eles, dada por (2), for menor que um limiar t (do inglês, *threshold*). O conjunto de arestas de G é denotado por E^* , um subconjunto de E descrito em (4). Desta forma, a topologia dos grafos de duas imagens distintas pode ser totalmente diferente. O uso de diferentes limiares caracteriza uma análise multi-escala do grafo (GONÇALVES et al., 2012). O trabalho de Ribeiro

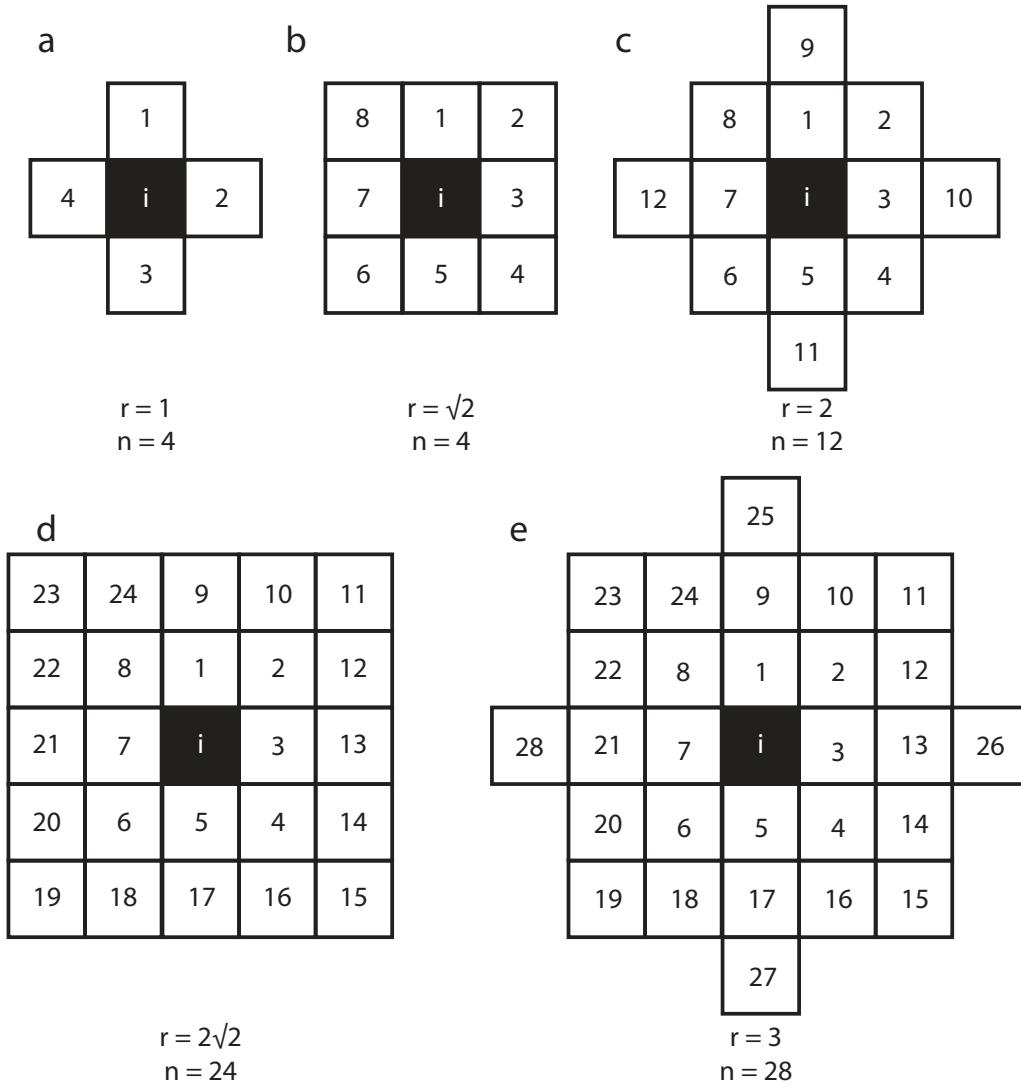


Figura 1 – Cada raio de vizinhança r oferece um conjunto de n vizinhos. Por exemplo, se $r = \sqrt{2}$, um pixel i possui os 8 vizinhos indicados em b.

et al. (RIBEIRO et al., 2015) realiza uma análise aprofundada dos valores de limiares t que oferecem grafos de maior capacidade discriminativa, propondo uma abordagem para seleção automática de limiares baseada em estatísticas da rede.

$$\omega(v_{x,y}, v_{x',y'}) = (x - x')^2 + (y - y')^2 + r_G^2 \frac{|I(x,y) - I(x',y')|}{L} \quad (2)$$

$$E = \left\{ (v_{x,y}, v_{x',y'}) \in I \times I \mid \sqrt{(x - x')^2 + (y - y')^2} \leq r_G \right\} \quad (3)$$

$$E^* = \{e \in E \mid w(e) \leq t\} \quad (4)$$

Um exemplo de vizinhança de um pixel dada pelas regras acima é dado na Figura 2, que mostra os valores dos pixels em (a), o conjunto E em (b) e o sub-conjunto E^* de E , de vértices que satisfazem a regra de (4), em (c).

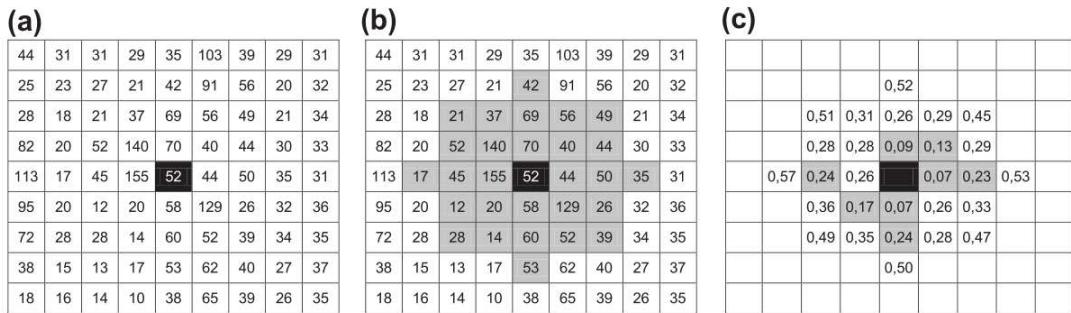


Figura 2 – A partir do pixel i marcado em (a), estabelece-se os vizinhos candidatos em (b). Após a aplicação do limiar t , os vizinhos do pixel i são mostrados em cinza em (c) (Adaptado de Backes *et al.* (BACKES; CASANOVA; BRUNO, 2013)).

Na teoria de redes complexas, podemos caracterizar essa rede como uma rede geográfica ou espacial, onde a posição dos vértices é relevante e pares de vértices próximos tem maior chance de apresentarem arestas (COSTA *et al.*, 2007). Além disso, a restrição de que a ligação entre vértices apenas possa ocorrer dentro de uma região de raio r aumenta a probabilidade de formação de *hubs*, que são vértices altamente conectados, e de comunidades bem distintas (COUTO; COMIN; COSTA, 2017). Regiões homogêneas da imagem serão altamente conectadas, enquanto pixels de valor distinto em relação à vizinhança (é o caso de certas texturas e de pixels de ruído, por exemplo) serão pouco conectados. Cabe notar que a escolha de métrica de dissimilaridade é dependente da aplicação; no caso da descrição de padrões locais, faz sentido utilizar o valor dos pixels e a distância Euclidiana entre eles.

Modelagens baseadas em grafos como a descrita anteriormente já foram estendida para vídeos de diversas maneiras. O trabalho de Gonçalves *et al.* (GONÇALVES; BRUNO, 2013b) propõe uma possível adaptação das caminhadas determinísticas em imagens para vídeos. A solução proposta considera que cada pixel possui coordenada (x, y, t) , já que o vídeo introduz uma componente temporal. O valor de t representa o quadro do vídeo em que a imagem se encontra. Múltiplos grafos são gerados a partir do vídeo; um grafo para cada imagem $I_{x,y}$ no plano $[X, Y]$ (num total de t imagens), um grafo para cada imagem $I_{x,t}$ no plano $[X, T]$ (num total de y imagens), e um grafo para cada imagem $I_{y,t}$ no plano $[Y, T]$ (num total de x imagens). Isso é ilustrado na Figura 3. Essa abordagem trata o tempo da mesma forma que as dimensões espaciais, gerando diversas imagens. Isso contrasta com trabalhos com diversas aplicações que tratam de redes variáveis no tempo, chamadas de redes temporais (HOLME; SARAMÄKI, 2012).

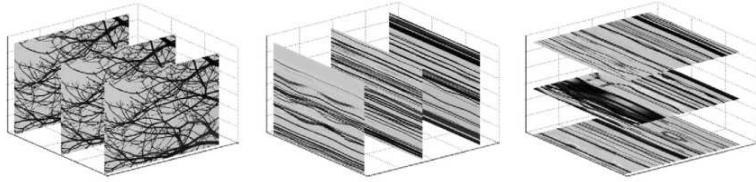


Figura 3 – Divisão de um vídeo em quadros. Respectivamente, da esquerda para a direita, divisão em quadros $[X, Y]$, quadros $[X, T]$ e quadros $[Y, T]$. (Adaptado de Gonçalves e Bruno (GONÇALVES; BRUNO, 2013b)).

A abordagem usada para criação da rede a partir da imagem é uma decisão importante e influí na informação que a rede contém a respeito da imagem. Os experimentos de Backes *et al.* (BACKES; CASANOVA; BRUNO, 2013) demonstram que a abordagem preserva informação discriminativa da textura original.

2.3 Métodos Afins e Caminhada Determinística

2.3.1 Padrões Binários Locais (*Local Binary Patterns*)

A estrutura periódica da textura faz com que métodos estatísticos sejam apropriados na extração de suas informações. Métodos estatísticos são por natureza eficientes em fazer emergir padrões. Observando-se a imagem completa, os padrões repetitivos da textura ficam evidentes. Um dos métodos mais populares e eficazes no sentido de expor características periódicas da imagem são os Padrões Binários Locais, ou LBP (Local Binary Patterns) (OJALA; PIETIKÄINEN; MÄENPÄÄ, 2002). Apresentaremos o LBP por ser um método importante no reconhecimento de texturas, e que juntamente com suas variantes tem obtido altas taxas de acerto quando validado em tarefas de classificação e segmentação de texturas. Além disso, é um método baseado nas derivadas direcionais do valor de intensidade do pixel relativo à sua vizinhança local, que é a mesma medida usada como critério para decisão do próximo passo nas caminhadas determinísticas.

O LBP avalia a vizinhança de cada pixel, gerando um padrão local baseado na diferença entre os valores dos pixels. Considere, por exemplo, os pixels $i = (x, y)$ da imagem I , onde $I(i)$ é a intensidade do pixel. Tomando-se uma vizinhança de raio $r = \sqrt{2}$, o pixel i possui 8 vizinhos, denotados por $N_i = \{n_{i,1}, n_{i,2}, \dots, n_{i,8}\}$, de modo semelhante ao mostrado na Figura 1 mostrada anteriormente. O método LBP propõe a comparação do pixel i com cada um de seus vizinhos em determinada ordem, criando um vetor de 8 dígitos binários T , mostrado na Equação 5. Em T , cada algarismo recebe valor 0 se o valor de intensidade do pixel vizinho de índice correspondente é maior que a $I(i)$, e recebe valor 1 caso contrário, de acordo com a Equação 6.

$$T = \{s([I(i) - I(n_{i,1})]), s([I(i) - I(n_{i,2})]), \dots, s([I(i) - I(n_{i,8})])\} \quad (5)$$

$$s(d) = \begin{cases} 0, & d < 0 \\ 1, & d \geq 0 \end{cases} \quad (6)$$

Considerando 8 vizinhos, e eliminando-se padrões repetidos que são versões rotacionadas uns dos outros, há 36 combinações de binários possíveis. Esse vetor de 8 dígitos binários, se analisado como um *byte*, pode ser convertido da base binária para um valor decimal correspondente. Este valor decimal descreve a configuração de intensidade relativa da vizinhança do pixel. Há um total de 256 valores distintos possíveis (pois são 2^8 combinações), no intervalo [0, 255], e chamamos esse valor de valor ou codificação LBP do pixel. A Figura 4 exemplifica o processo para geração do valor LBP para um pixel.

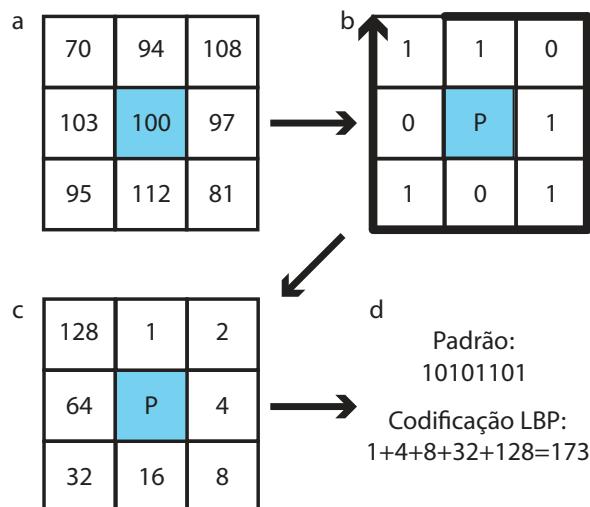


Figura 4 – a) mostra a vizinhança de um pixel. b) resulta da aplicação da Equação 6 para gerar os valores binários da vizinhança do pixel, que é lida no sentido da seta (do bit menos significativo para o mais significativo). c) mostra os pesos usados na conversão do vetor binário para um valor decimal, e a conversão em si é mostrada em d).

Entre as primeiras adaptações do método LBP está o LBPPROT, uma tentativa de alcançar invariância a rotação no LBP (PIETIKÄINEN; OJALA; XU, 2000). O método LBPPROT considera rotações de um mesmo padrão binário como padrões idênticos; por exemplo, os valores "00001010" e "00000101" são considerados o mesmo padrão, mas rotacionado (uma rotação de 45°, neste caso). Essa adaptação resulta em um vetor de

características resistentes a rotações e significativamente menor, com 36 padrões. Os resultados, porém, mostram que esse tipo de invariância a rotação não apresenta bom poder de discriminação (PIETIKÄINEN; OJALA; XU, 2000) (OJALA; PIETIKÄINEN; MÄENPÄÄ, 2002), e seu potencial de uso é limitado a rotações de 90° que não afetam o resultado da quantização da imagem. Os autores concluem que essa codificação é muito restritiva.

Uma adaptação de invariância a rotação mais bem sucedida foi o uso de padrões uniformes (OJALA; PIETIKÄINEN; MÄENPÄÄ, 2002). A ideia é baseada na observação de que um conjunto pequeno de padrões ocorre na maioria das texturas com muito mais frequência que todos os outros. São padrões que possuem alta uniformidade U , que é uma medida que corresponde ao número de transições entre 0 e 1 no padrão LBP, ou "byte". Por exemplo, o padrão "11111111" é perfeitamente uniforme pois não há transições ($U = 0$), e o padrão "00111000" tem duas transições, e portanto $U = 2$. O agrupamento de padrões de uniformidade $U \leq 2$ em um único *bin* do histograma, mantendo padrões não-uniformes em seus *bins* individuais como antes, promoveu melhora significativa nas taxas de classificações corretas, mesmo em texturas rotacionadas. O motivo é que rotações de quaisquer ângulos não afetam a medida U de um pixel.

Ao longo dos anos, diversas outras variantes do LBP foram criadas, para diversas aplicações (HUANG et al., 2011). Digno de nota, pelas altas taxas de classificação alcançadas, é o recente MRELBP (LIU et al., 2016b), uma variante que incorpora outros aprimoramentos do método e o aplica o LBP a valores de medianas locais da imagem ao invés de aos valores originais de intensidade. Esta variante busca lidar com a baixa tolerância a ruído e dificuldade computacional de análise de vizinhanças maiores do método LBP.

O sucesso do método LBP em descrever texturas estáticas elicitou adaptações para aplicação do conceito de padrões locais binários para dados espaço-temporais. O trabalho mais relevante que surgiu tratando de LBP para texturas dinâmicas foi o de Zhao et al. (ZHAO; PIETIKAINEN, 2007), que expôs duas maneiras de utilizar o LBP na classificação de texturas dinâmicas, o VLBP (*Volumetric LBP*) e o LBP-TOP (*LBP - Three Orthogonal Planes*).

Considerando um vídeo $V(x, y, t)$, o VLBP é baseado em uma análise volumétrica, que considera a vizinhança de cada pixel nas três dimensões $[X, Y, T]$ do vídeo para compor sua codificação LBP. Com mais vizinhos, é evidente que aumenta o número de codificações possíveis. Em paralelo ao VLBP, o LBP-TOP é similar ao LBP bidimensional, porém o histograma utilizado como descritor é calculado com base em todos os quadros do vídeo. Para capturar no vetor de características o comportamento da textura dinâmica em cada plano, este processo é repetido para cada par de dimensões $[X, Y]$, $[X, T]$ e $[Y, T]$, e os vetores obtidos concatenados. Um processo similar é aplicado em nossa proposta para texturas dinâmicas, de modo que detalharemos mais esta metodologia no próximo

capítulo.

2.3.2 Caminhadas Determinísticas

Embora o LBP e suas variantes (NANNI; LUMINI; BRAHNAM, 2012) gerem bons resultados para muitas texturas, aumentos no tamanho do raio de avaliação são via de regra prejudiciais às taxas de classificação, e o número de padrões unívocos possíveis cresce de modo explosivo com o aumento do raio, dada sua natureza combinatória. Recentemente, abordagens baseadas em caminhadas determinísticas tem tido sucesso em descrever vizinhanças (BACKES et al., 2010) (COUTO et al., 2015) e elementos de textura de maior nível (GONÇALVES; BRUNO, 2013a).

Considerando a imagem como um grafo, como exposto na seção anterior, é possível aplicar métricas de análise topológica de grafos ou de redes (COSTA et al., 2007). Tal informação pode ser utilizada para caracterização da estrutura da rede, que é representativa da textura a partir da qual ela foi gerada. Um método de descrição de redes que é muito usado é método das caminhadas. O uso de caminhadas aleatórias é um dos métodos usados para detecção de saliências ou segmentação de imagens e oferece bons resultados (GRADY, 2006) (GOPALAKRISHNAN; HU; RAJAN, 2009). Há ainda muitos trabalhos relativos a caminhadas determinísticas (TERÇARIOL; MARTINEZ, 2005), que recentemente têm sido aplicadas à caracterização de texturas com bons resultados. Caminhadas consistem em percorrer os dados com um agente que visita os dados seguindo uma regra de decisão, usualmente evitando repetir visitas a dados recentemente visitados (TERÇARIOL; MARTINEZ, 2005) (CAMPITELI et al., 2006). Diversas heurísticas podem determinar a decisão de passo de caminhada, mas são via de regra baseadas na intensidade dos pixels e em objetivos atribuídos aos agentes. Alguns trabalhos impõe limites de desempenho a tais agentes em seu percurso, e frequentemente os chamam de "*crawlers*" (MARTINEZ; KINOUCHI; RISAU-GUSMAN, 2004) (ZHANG; CHEN, 2005) (GONÇALVES; MACHADO; BRUNO, 2014).

Uma caminhada consiste de um conjunto de passos tomados consecutivamente. Note que o método das caminhadas pode ser aplicado a diversos tipos de dados. Aqui, descreveremos a aplicação sobre imagens representadas por grafos. Considere um vértice arbitrário $v(x_i, y_i)$ de um grafo G obtido a partir de uma imagem I . Escolhe-se o vértice como o primeiro passo da caminhada, w_0 . Em um dado passo w_s , a caminhada é representada por um conjunto que contém o passo atual e todos os passos anteriores, $W_s = \{w_0, w_1, \dots, w_s\}$, $w_k \in I$, onde cada passo é um vértice que foi visitado. De cada vértice em G obtém-se uma caminhada distinta; portanto, para alcançar uma caracterização completa da imagem, as caminhadas são realizadas começando em todos os vértices (o que significa que para cada pixel em uma dada imagem I é realizada uma caminhada começando no vértice de G que o representa).

Se um dado passo w_k atualizou a posição atual da caminhada para um vértice v , o próximo passo w_{k+1} levará a posição atual da caminhada a algum outro vértice. A decisão sobre que vértice escolher como próximo passo w_{k+1} é realizada elegendo-se o melhor vértice dentre os candidatos, seguindo algumas restrições. Em primeiro lugar, passos só podem ser realizados seguindo arestas de G , ou seja, o vértice do passo seguinte deve ser adjacente ao vértice atual $\eta(D(w_k))$. Em segundo lugar, para ser um candidato a próximo passo o vértice não pode estar na memória recente do agente da caminhada representada por M_s^μ . M_s^μ é o conjunto dos μ passos mais recentes em uma caminhada de s passos, como definido em (9). Diremos que μ é o tamanho da memória. Um vértice não pode ser visitado enquanto estiver na memória, para forçar que o agente visite vértices ainda não visitados. Isso previne que caminhadas fiquem presas em laços muito pequenos. Depois de estar armazenado na memória por $\mu + 1$ passos, um vértice é removido de M_s^μ e pode ser visitado novamente no futuro. Por esse motivo tais caminhadas são chamadas de parcialmente auto-repelentes; vértices recentemente percorridos são evitados. Uma caminhada sem memória entraria rapidamente em um laço infinito envolvendo poucos pixels de valor próximo.

Por fim, dentre os vértices candidatos restantes, a caminhada escolhe para o próximo passo o vértice mais adequado, baseado na dinâmica da caminhada. A dinâmica é a regra de decisão da caminhada. As duas dinâmicas de caminhada mais usadas na literatura são a diferença mínima e diferença máxima (abreviadas como dinâmicas *min* e *max*). Usar a dinâmica *min* significa que o melhor candidato para próximo passo é o vértice que representa o pixel de valor de intensidade de menor diferença, em módulo, em relação ao pixel do passo atual. A dinâmica *max*, por sua vez, elege o vértice cujo pixel tem maior diferença, em módulo, de valor em relação ao pixel do vértice atual. Como cada vértice é representado por um pixel, a diferença entre dois vértices pode ser representada pelo valor de seus pixels: $|I(x, y) - I(x', y')|$. Em caso de empates na diferença entre os valores, opta-se por uma das direções arbitrariamente. As diretivas citadas para decisão de que vértice visitar no passo $w_k + 1$ estão resumidas em (7) e (8), que mostram o caso usando a dinâmica *min* e a dinâmica *max*, respectivamente.

$$w_{s+1} = \arg \min_{j \in \eta(w_s), j \notin M_s^\mu} |I(w_s) - I(j)| \quad (7)$$

$$w_{s+1} = \arg \max_{j \in \eta(w_s), j \notin M_s^\mu} |I(w_s) - I(j)| \quad (8)$$

Onde M_s^μ é dado por:

$$M_s^\mu = \bigcup_{k=s-\mu}^s w_k \quad (9)$$

Se um vértice já foi percorrido antes na caminhada, mas ele já não está na memória M_s^μ , ele pode ser revisitado. Nesse caso, é possível que a caminhada fique presa em um laço infinito de passos repetidos. De fato, toda caminhada eventualmente atinge um estágio de repetição periódica. À parte da caminhada que não se repete dá-se o nome de transiente t , e à parte periódica dá-se o nome de período, ou atrator, p . Note que o transiente pode ter tamanho 0, enquanto o período tem pelo menos tamanho igual ao tamanho da memória $\mu + 1$. Assim, as caminhadas estão inherentemente ligadas ao valor de μ usado. A Figura 5 mostra um exemplo de caminhada sobre uma imagem I .

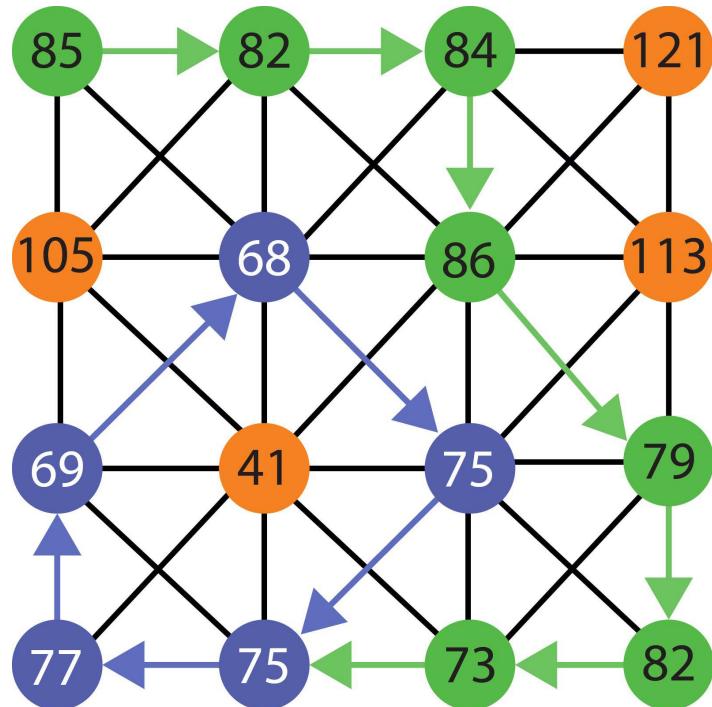


Figura 5 – Exemplo de caminhada com $\mu = 3$ começando no pixel superior esquerdo, em que os pixels do transiente estão mostrados em verde (os 7 primeiros passos da caminhada) e os do período em azul (5 últimos passos), seguindo a dinâmica min .

Devido à regularidade da textura, a realização de caminhadas a partir de todos os pixels da imagem faz emergir padrões descritivos que caracterizam a textura. Ribas *et al.* (RIBAS; BRUNO, 2016) apresentam um método para reduzir o custo computacional da execução das caminhadas, com perdas controladas em termos de taxas de acertos, realizando caminhadas em apenas um sub-conjunto dos pixels, mas via de regra, cada pixel é usado como início de pelo menos uma caminhada. Estatísticas das caminhadas

oferecem a informação para composição do descritor da textura (BACKES et al., 2010). Histogramas são frequentemente usados para representar tal informação pois são adequados para representação de tendências no conjunto de dados e periodicidade. Há trabalhos onde um histograma de frequências de tamanhos das caminhadas determinísticas é usado para caracterização de texturas estáticas (BACKES et al., 2010) (GONÇALVES et al., 2012) e dinâmicas (GONÇALVES; BRUNO, 2013b). A cardinalidade do conjunto de passos de uma caminhada W_s é dada por s . Parte dos passos da caminhada pertencem ao transiente e parte ao período da caminhada, de forma que $W_s = W_t \cup W_p$. O tamanho t do transiente e o tamanho p do período são descritivos da região da imagem percorrida. Considera-se as caminhadas sobre todos os vértices e agrupa-se seus tamanhos em um histograma bidimensional de ocorrências de pares (t, p) . Texturas de uma mesma classe geram histogramas similares, como pode ser visto na Figura 6.

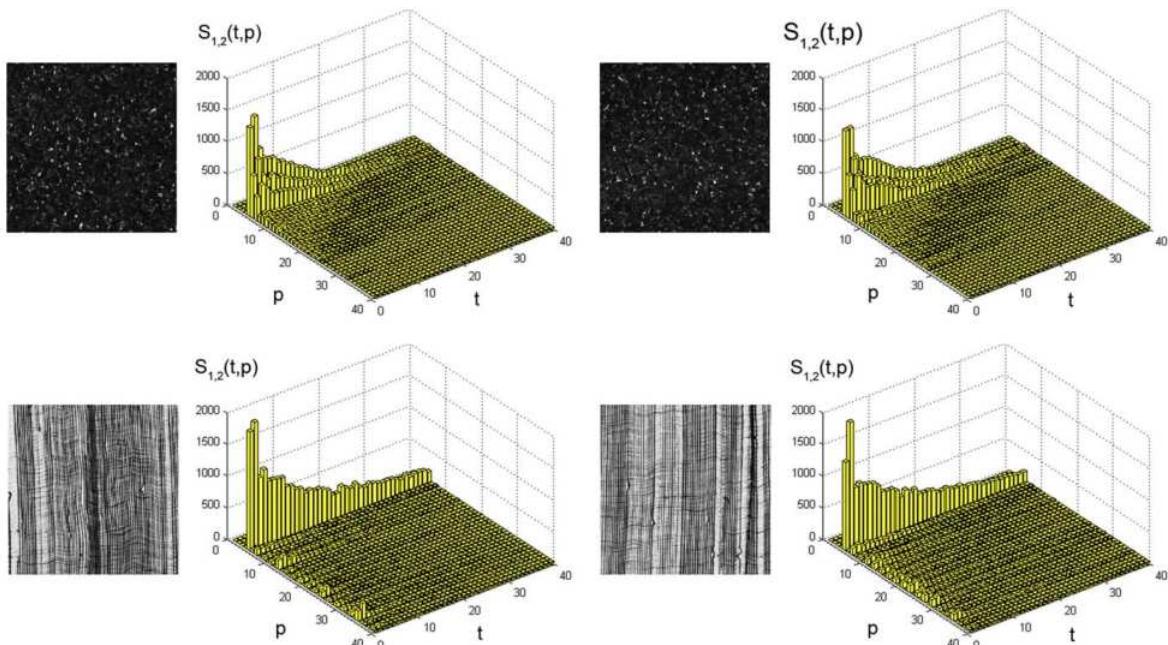


Figura 6 – Histograma de tamanhos de transiente e período da caminhada aplicado à base Brodatz (BRODATZ, 1966). Note como texturas de um mesmo objeto fornecem histogramas semelhantes. Extraído de (BACKES et al., 2010).

O histograma é uma representação de natureza estatística, que compila em uma única distribuição múltiplos dados. Se o número de observações registradas no histograma for variável, é indicado que o histograma seja normalizado, ou equalizado, para que esta diferença numérica não afete as comparações entre histogramas. Em todos os casos em que a normalização de um histograma H de n bins é pertinente, a Equação 10 é utilizada (GONZALES; WOODS; EDDINS, 2004), que garante que a área sob a curva do histograma seja 1, independente do número de amostras, e neste caso o histograma pode ser tratado como uma função de densidade de probabilidades.

$$H_{Norm}(i) = \frac{H(i)}{\sum_{j=1}^n H(j)}, \quad \forall i \in \{1, 2, \dots, n\} \quad (10)$$

Uma das caminhadas do método das caminhadas sobre o método LBP é que o aumento do raio das caminhadas não afeta o tamanho do descritor do método, enquanto no caso do LBP, devido à natureza combinacional do método, o tamanho do descritor aumenta exponencialmente com o tamanho da vizinhança, tornando impraticáveis na maioria das aplicações raios maiores, que são importantes pra descrever padrões de maior escala. Apesar disso, o método LBP tem maior proeminência na literatura de análise de texturas que as caminhadas determinísticas. Acreditamos que há campo para explorar as caminhadas mais a fundo. No capítulo seguinte apresentaremos propostas para uso de outras medidas de caminhadas como descritores de texturas.

2.3.3 Estatísticas de Histograma

Como visto, o resultado da caminhada sobre uma rede é uma estrutura da qual se obtém um histograma. Dos valores brutos dos *bins* do histograma, há estatísticas que podem ser geradas a partir de tais distribuições numéricas.

Normalmente tais medições estatísticas são aplicadas, quando se trata de processamento de imagens, sobre histogramas de intensidade extraídos diretamente da imagem. A forma do histograma oferece muitas pistas sobre a imagem da qual ele foi gerado, e é uma representação sucinta e particular da imagem. A presente tese propõe o emprego dessas estatísticas sobre o histograma gerado. As estatísticas denotam a topologia dos histogramas, e por consequência, podem ser usadas para caracterizar histogramas gerados pelo método das caminhadas proposto nesta tese (mais detalhes na seção 3.2).

Dentre as propriedades estatísticas mais populares na literatura por sua capacidade discriminativa comprovada, destacaremos 7 propriedades: energia, entropia, *skewness* (do inglês, algo como "viés" ou "inclinação"), contraste, média, variância e kurtosis (HARALICK; SHANMUGAM; DINSTEIN, 1973) (MATERKA; STRZELECKI et al., 1998) (AGGARWAL; AGRAWAL, 2012) (MALIK; BAHARUDIN, 2013). Note que propriedades de primeira ordem são quantidades unidimensionais úteis para representação global do histograma, mas não proveem informações a respeito das posições relativas dos valores na estrutura que originou o histograma. Isso é em contraste com o histograma de direções de caminhadas proposto por esta tese na seção 3.2, que traz justamente esse tipo de informação.

Considere o histograma $h(i)$, com $i \in [1, G]$ representando o índice de cada *bin*, onde G é o número de *bins*. Dividindo-se os valores do histograma pelo número de pixels, obtém-se a distribuição de probabilidade $p(i) = h(i)/N$, onde N é o número total de observações. Ao conjunto das 7 estatísticas daremos o nome de *St*. As Equações 11 a 17 mostram as fórmulas para o cálculo de cada uma das medidas.

Média:

$$\mu = \sum_{i=1}^G ip(i) \quad (11)$$

Variância:

$$\sigma^2 = \sum_{i=1}^G (i - \mu)^2 p(i) \quad (12)$$

Skewness:

$$\mu_3 = \sigma^{-3} \sum_{i=1}^G (i - \mu)^3 p(i) \quad (13)$$

Kurtosis:

$$\mu_4 = \sigma^{-4} \sum_{i=1}^G (i - \mu)^4 p(i) - 3 \quad (14)$$

Contraste:

$$C = \sum_{i=1}^G ip(i)^2 \quad (15)$$

Energia:

$$E = \sum_{i=1}^G p(i)^2 \quad (16)$$

Entropia:

$$H = - \sum_{i=1}^G p(i) \log_2(p(i)) \quad (17)$$

Faremos a seguir uma breve descrição de cada uma das métricas propostas, a fim de oferecer uma percepção em termos qualitativos do tipo de informação do histograma que cada medida expressa:

A Média é a medida do valor médio dos *bins* do histograma, e representa onde se concentram os valores na distribuição. No histograma de uma imagem a Média representa se a imagem é mais escura ou mais clara em termos de intensidade dos pixels. Variância, *Skewness* e Kurtosis são chamados de momentos de segunda, terceira e quarta ordem, respectivamente. A Variância mede o grau de desvio do histograma em torno do eixo da Média. Skewness mede a assimetria do histograma em torno da Média (um histograma simétrico no eixo da Média tem Skewness igual a zero). Kurtosis mede o quanto "suave" é a

curva do histograma é, ou seja, quanto bruscas são as transições entre os *bins*. O Contraste mede a amplitude do desvio do histograma ao longo do eixo dos *bins*, e portanto, quanto mais concentrados ou espalhados são os valores. A medida de Energia mostra a uniformidade do histograma (um histograma com variações grandes de valores nos *bins* terá energia menor, um histograma em que todos os *bins* tem o mesmo valor tem energia alta). A Entropia é o contrário da Energia e mede o quanto aleatória é a distribuição dos valores do histograma (MATERKA; STRZELECKI et al., 1998) (AGGARWAL; AGRAWAL, 2012) (MALIK; BAHARUDIN, 2013). Todas estas medidas se prestam a uma descrição mais detalhada do histograma.

2.4 Fluxo óptico e padrões singulares em campos vetoriais

2.4.1 Fluxo óptico

O fluxo óptico é uma informação visual relativa à percepção da movimentação de elementos visuais de uma cena por um observador. Tal percepção de movimento baseia-se na projeção do vetor de movimento de um objeto da cena sobre um plano de observação, e é uma ferramenta importante na detecção e estimativa de movimento e rastreio de objetos. Métodos de cálculo de fluxo óptico representam na forma de um campo vetorial a relação entre elementos de um quadro e outro. Por isso, tais métodos assumem que a variação entre os quadros seja contínua e pequena (um pressuposto razoável quando tratamos de texturas dinâmicas e vídeos de cenas reais).

A computação do fluxo óptico a partir de uma sequência de quadros é um problema complexo e com diversas soluções propostas, as mais populares sendo o método de Lucas-Kanade (LUCAS; KANADE et al., 1981) e o método de Horn-Schunck (HORN; SCHUNCK, 1981). Estes métodos resultam em um campo vetorial F . Um campo vetorial é definido por uma função que atribui um vetor com magnitude e direção própria a cada ponto do seu domínio. No caso do fluxo óptico, atribui-se cada vetor do campo a um pixel de coordenada (x, y) , e o vetor denota a estimativa do deslocamento daquele pixel de um quadro a outro.

O objetivo do método Lucas-Kanade, usado nesse trabalho, é alinhar uma imagem de entrada (ou *input*) $I(x, y)$ a uma imagem modelo (ou *template*) $T(x, y)$. Considere o conjunto $W(x, p)$ de transformadas possíveis de T para I , onde p é um vetor de parâmetros. O alinhamento de I com T é feito mapeando cada pixel (x, y) da imagem modelo T para um sub-pixel $W((x, y); p)$ de I (BAKER; MATTHEWS, 2004). Uma transformação de $W((x, y); p)$ pode ser, por exemplo, uma translação da coordenada (x, y) original para uma nova coordenada $(x + p_1, y + p_2)$, e nesse caso teríamos $p = p_1, p_2$. Outras transformadas possíveis para cada aplicação também podem ser previstas, como rotações

ou transformações afim (*affine*). Para encontrar o mapeamento dos sub-pixels busca-se os parâmetros de transformação que minimizem a soma do erro quadrado entre duas imagens.

Assumindo n como o conjunto de parâmetros de transformações e NxN , ou N^2 , como o número de pixels da imagem, a complexidade do algoritmo Lucas-Kanade é $O(n^2N^2 + n^3)$ (BAKER; MATTHEWS, 2004). O número de transformações consideradas pode ser arbitrariamente grande, mas, em aplicações criticamente sensíveis a tempo, poucas são usadas, devido à complexidade cúbica. Um aspecto significativo é que a complexidade computacional é linear em relação ao número de pixels.

2.4.2 Padrões singulares em campos vetoriais

Campos vetoriais criados por processos naturais e contínuos, como o fluxo óptico do vídeo de uma cena natural (em oposição a um campo vetorial gerado aleatoriamente, por exemplo), apresentam tendências e padrões locais característicos. Em uma textura dinâmica, há periodicidade temporal, o que significa que os campos vetoriais obtidos também serão aproximadamente periódicos (VIDAL; RAVICHANDRAN, 2005). Métodos de análise de fluxo óptico são usados na literatura para caracterização e classificação de texturas dinâmicas (CHETVERIKOV; PÉTERI, 2005). Se é possível representar e classificar campos vetoriais, então é possível caracterizar o fluxo óptico e usar essa informação como critério de similaridade entre texturas dinâmicas. Tal caracterização do campo vetorial tem aplicações expressivas em áreas tais como análise de fluidos (LI; YAU; WANG, 2006), meteorologia (KIHL; TREMBLAIS; AUGEREAU, 2008) e reconhecimento de impressões digitais (FAN et al., 2008).

O fluxo óptico tem sido extensivamente usado como um descritor de texturas dinâmicas (FAZEKAS; CHETVERIKOV, 2007) (CHAO; GU; NAPOLITANO, 2014). Ainda há diversos trabalhos que lidam com a caracterização de fluxos ópticos em diversas tarefas de reconhecimento de padrões (CHAUDHRY et al., 2009) (LIU et al., 2016). Uma das abordagens mais recentes para modelar regiões características de um campo vetorial é buscar nele elementos macroscópicos de interesse, como descrito por Liu e Ribeiro (LIU; RIBEIRO, 2010) (LIU; RIBEIRO, 2012). O método é particularmente eficaz na fidelidade da representação, evidenciada pelos experimentos realizados pelos autores de reconstrução de campos vetoriais a partir do modelo, enquanto é uma solução de baixa dimensionalidade. O método visa decompor o fluxo em padrões singulares através do uso de um conjunto de símbolos, um dicionário de padrões cuja combinação linear ponderada pode aproximar qualquer campo vetorial. A escolha de quais padrões comporão esse dicionário simbólico é importante, e não há definições claras para a maioria dos padrões visíveis em campos vetoriais, como vórtices, fontes ou sumidouros (JIANG; MACHIRAJU; THOMPSON, 2005). Rao e Jain (RAO; JAIN, 1992), em seu trabalho seminal, propuseram como dicionário 6 padrões distintos onde o campo se anula (ou seja, o vetor resultante tem

magnitude zero), já que esta é uma característica comum de diversos padrões de interesse em campos vetoriais. Liu e Ribeiro (LIU; RIBEIRO, 2012) propõe um conjunto mais abrangente de padrões que, incluindo os padrões clássicos onde o campo vetorial se anula, define uma função de valores complexos para levar em conta outras possíveis singularidades, introduzindo também robustez contra variações de rotação e escala.

Chamaremos estes pontos de interesse e sua vizinhança de padrões singulares. O aspecto mais significativo para a proposta desta tese é o fato de que uma combinação de padrões pode servir como assinatura para descrição de um determinado campo vetorial.

O trabalho de Liu e Ribeiro (LIU; RIBEIRO, 2012) estabelece um dicionário Φ de fluxos base. Um determinado campo vetorial F é decomposto linearmente em um somatório de k componentes de coeficientes a_k e monômios z^k , gerando um campo vetorial aproximado f , de acordo com a Equação 18. Essa aproximação é semelhante à aproximação de uma função por polinômios de Taylor. Isto é, os autores usam os monômios $\{z^k\}_{k=1}^N$ como os fluxos base $\Phi_k(z)$.

$$F(z) \approx f(z) = \sum_k a_k z^k \quad (18)$$

A Figura 7 mostra o dicionário de fluxos base para padrões singulares proposto na literatura.

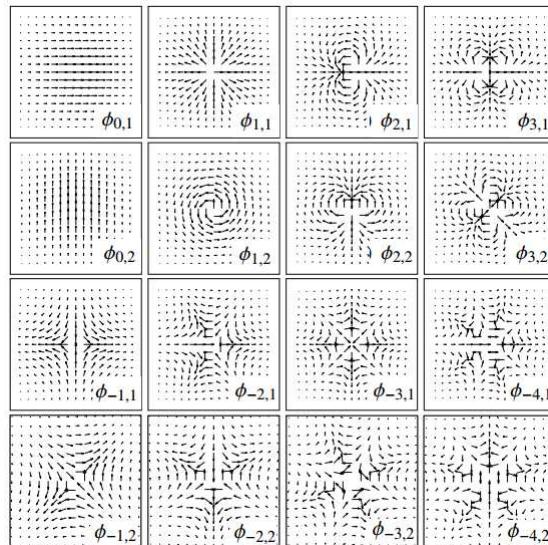


Figura 7 – Dicionário de fluxos base para padrões singulares. A combinação linear desses padrões pode aproximar o comportamento de um campo vetorial em um ponto qualquer. Extraído de (LIU; RIBEIRO, 2012).

Os coeficientes a_k são calculados por correlação cruzada, projetando-se o campo vetorial F sobre os fluxos base Φ_k . O valor dos coeficientes aumenta se há similaridade entre os padrões. Assim, pontos onde ocorrem máximos locais na soma dos coeficientes a_k são eleitos como padrões singulares, ou seja, regiões representativas do campo, que

apresentam alta similaridade em relação aos fluxos base. Os padrões gerados em (LIU; RIBEIRO, 2012) são pontos característicos invariantes a transformações como rotação e translação. A invariância é alcançada por método em extratores de pontos característicos como a transformada SIFT (LOWE, 2004) e SURF (BAY; TUYTELAARS; GOOL, 2006): a invariância a escala se dá porque os pontos com alta similaridade (no caso, os padrões singulares) são procurados em múltiplos espaços de escala, e a invariância a rotação se dá porque para cada ponto é calculada e atribuída uma orientação, o que permite o reconhecimento de padrões rotacionados.

Aplicações práticas de extratores de pontos de interesse como SIFT, SURF e o extrator baseado em padrões singulares baseiam-se via de regra no pareamento de pontos de interesse (LI et al., 2008), (LIU; RIBEIRO, 2012) (COUTO, 2012) (LIU et al., 2014). Quando o objetivo é a descrição de objetos ou cenas mais elaboradas, há estratégias para agrupar diversos pontos de interesse em representações mais ricas de complexidade e tamanho maiores (BOUREAU et al., 2010). A literatura apresenta um conjunto de estratégias de agrupamento que podem ser descrito pelos termos gerais *bag-of-keypoints* e *bag-of-features* (O'HARA; DRAPER, 2011) (SCHMID, 2011). A técnica também é conhecida, no caso de aplicações em visão computacional, pelo nome de *bag-of-visual-words*, devido à sua origem na recuperação de documentos de texto por conteúdo (SIVIC; ZISSERMAN, 2003) (CSURKA et al., 2004). Considerando o tamanho reduzido do descritor baseado em *bag-of-features* em relação aos dados originais, ele possui grande potencial descritivo, sendo usado na descrição de objetos e cenas complexas (LAZEBNIK; SCHMID; PONCE, 2006) (ZHANG et al., 2007), em tarefas de larga escala como CBIR (JÉGOU; DOUZE; SCHMID, 2010) e até mesmo na reconstrução sintética de imagens (KATO; HARADA, 2014).

O fluxo óptico é uma propriedade da textura dinâmica com potencial para contornar os efeitos que ruído causam em descritores de vizinhanças menores como o LBP. Os padrões singulares também são avaliados em diversas escalas e tem potencial para detectar padrões visuais em uma escala maior que outros descritores locais, porém para descrever uma textura dinâmica globalmente é necessário que estes padrões locais sejam combinados em um descritor global. Duas propostas para esta solução serão apresentadas no capítulo seguinte.

CAPÍTULO 3

Proposta e Metodologia

Para cumprir os objetivos elicitados no Capítulo 1, este capítulo apresenta propostas para geração de descritores para texturas ainda não exploradas pela literatura, fundamentadas na teoria já apresentada. As primeiras seções apresentam as propostas para texturas estáticas, seguidas das propostas para texturas dinâmicas ¹.

3.1 Geração de rede para caminhadas

Em alguns trabalhos da literatura (BACKES et al., 2010) (GONÇALVES; BRUNO, 2013b), as caminhadas determinísticas são realizadas sobre o grafo modelado a partir da imagem. Nestes trabalhos, cada pixel da imagem corresponde a um nó da rede complexa, e determina-se a existência de arestas entre os nós estabelecendo-se um raio máximo. Há trabalhos que mostram diferentes formas de geração de redes complexas a partir de imagens (BACKES; CASANOVA; BRUNO, 2013) (GONÇALVES; BRUNO, 2013a). Redes distintas evidenciam características distintas da imagem, e portanto podem gerar caminhadas determinísticas que podem oferecer um conjunto de características com maior potencial discriminativo. Caminhadas determinísticas sobre uma transformação em que o valor de cada pixel é dado pelo grau dos vértices gerados a partir das redes criadas com o método usado por Backes *et al.* (BACKES; CASANOVA; BRUNO, 2013) obtiveram resultados promissores. No trabalho citado, um descritor é criado a partir de medidas estatísticas realizadas sobre o grafo gerado, sem o uso de caminhadas. Considere a rede complexa G , que representa uma imagem. A partir da análise de G gera-se uma conjunto de múltiplos D_t , que são subgrafos de G que tem como parâmetro principal um limiar t particular, como mostrado na revisão da literatura. Cada vértice $v_{D_t}(x, y)$ de cada rede D_t , possui um conjunto de vértices vizinhos ∂_{v_t} (ou seja, vértices ligados por arestas a v_t), dado pela Equação 19, e a partir desse conjunto de vizinhos obtemos o grau $\deg(v_t(x, y))$ do

¹ A implementação de cada uma das propostas deste trabalho pode ser encontrada no endereço: <https://drive.google.com/open?id=1_TSPkRa-ODKDXlhDMHBmFLHkjkbKf4rU> , acessado em 03/04/2018

vértice, que expressa o número de arestas incidentes no vértice, e é dado pela Equação 20 (BACKES; CASANOVA; BRUNO, 2010), em que $\omega(v, v')$ é a métrica de dissimilaridade entre vértices definida na Equação 2.

$$\eta_{v_t} = \{v' \in V | (v, v') \in E \text{ tal que } \omega(v, v') \leq t\} \quad (19)$$

$$\deg(v_t) = |\partial_{v_t}| \quad (20)$$

Neste trabalho, nossa proposta consiste de utilizar uma transformação da imagem baseada na rede apresentada, e usar esta transformação para realização de caminhadas determinísticas. Note como, na rede apresentada, o grau de um vértice carrega informação sobre o quanto semelhante o pixel correspondente é a seus vizinhos, e depende essencialmente do valor de t escolhido. O grau é um valor poderoso que, no caso da imagem transformada, denota a relação de um pixel com sua vizinhança, uma informação crucial na descrição dos padrões locais de intensidade da imagem. Ainda assim, a execução das caminhadas sobre transformações derivadas de diversos limiares envolve alto custo computacional. Propomos, portanto, a criação de um grafo D uniformemente conectado com o mesmo número de vértices que D_t e raio de vizinhança fixo r . Isto é, D segue a regra de conectividade descrita na Equação (1). Como o valor do limiar t é determinante na topologia da rede gerada, diversas redes D_t são geradas, para diferentes valores de t em um conjunto T . Nossa proposta é que no grafo D , o valor de cada vértice $v_D(x, y)$ seja a média aritmética dos graus dos vértices correspondentes nos diversos D_t obtidos, como mostra a Equação (21). O uso da média é feito para que todos os valores de t sejam considerados em um único grafo D . Nossa proposta ainda empresta a ideia central de método como o LBP (OJALA; PIETIKAINEN; MAENPAA, 2002), em que a relação entre o valor de intensidade de um determinado pixel e todos os seus vizinhos é resumida em um único número, no caso, $v_D(x, y)$.

$$v_D(x, y) = \frac{1}{|T|} \sum_{t \in S} \deg(v_{D_t}(x, y)) \quad (21)$$

A Figura 8 mostra a geração de D que este trabalho propõe. Parte-se de uma imagem I e gera-se uma rede G_t . A análise de G_t sob os diversos valores de limiar t gera diversos D_t (BACKES; CASANOVA; BRUNO, 2013). D é a média de D_t para $t \in S$. Em outras palavras, D é uma imagem transformada cujos pixels possuem os valores médios dos graus de cada G_t , que representa a similaridade de cada pixel de I com sua vizinhança.

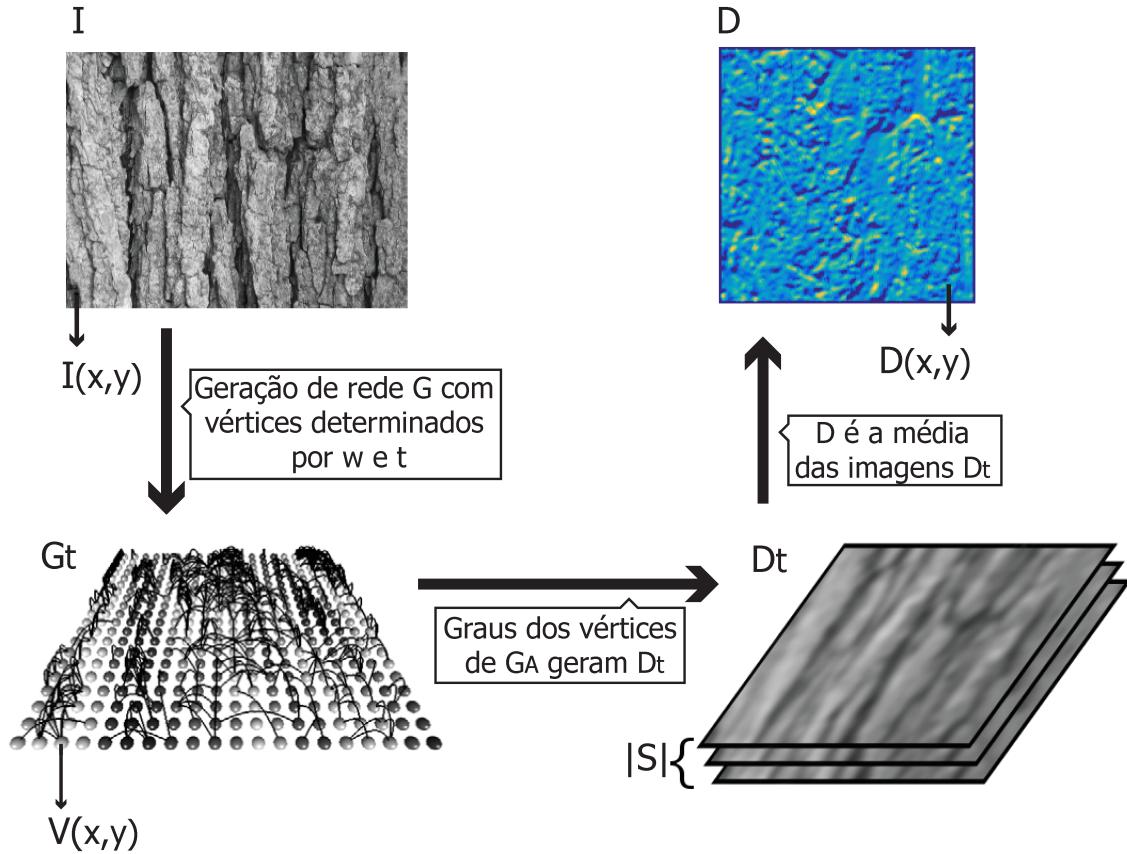


Figura 8 – Passos para geração do grafo D de graus médios. Uma rede G_t é criada baseado nas intensidades dos pixels da imagem I e em um valor de limiar t . O grau de cada vértice G_t gera o valor de intensidade do pixel correspondente na imagem transformada D_t . A média das imagens D_t gera D . (Adaptado de Backes *et al.* (BACKES; CASANOVA; BRUNO, 2013))

3.1.1 Parâmetros da transformação de rede complexa

Um parâmetro importante na geração da rede complexa é o raio máximo que permite conectividade. De acordo com resultados de publicações anteriores (BACKES; CASANOVA; BRUNO, 2013), o raio $r_G = 3$ oferece resultados significativamente mais discriminativos que raios menores, enquanto raios maiores não acrescentam arestas suficientes para contrabalancear a maior carga de processamento e dados gerados. Isso ocorre pois nós mais distantes tem menor chance de gerar arestas, e isso ocorre de forma quadrática em relação à distância, como visto na Equação (2), na seção 2.2.

É evidente que a adequação do valor de raio de geração da rede depende da escala da textura, mas segundo a análise de Backes et al. (BACKES; CASANOVA; BRUNO, 2013) envolvendo diversas bases de dados, $r_G = 3$ foi o valor que ofereceu descritores de qualidade mais consistente. Isso significa que a rede complexa obtida a partir da textura terá nós de grau $v_{x,y}$ de no máximo 28 arestas. A faixa S de valores usados para os limiares t usados foi 0.005, 0.020, ..., 0.530, com incrementos de 0.015, perfazendo um total de 36 grafos D_t gerados para cada imagem.

3.2 Caminhadas Determinísticas Auto-Repelentes sobre a Rede

Trabalhos que envolvem caminhadas determinísticas sobre a rede extraída da imagem (BACKES et al., 2006) (GONÇALVES; BRUNO, 2013b) extraem o tamanho da caminhada como informação para criação do vetor de características. Enquanto tal informação demonstrou ter boa capacidade discriminativa, ela parece ignorar grande quantidade de informação potencialmente significativa da caminhada. Pode-se utilizar da dimensão fractal com o propósito de descrever as formas complexas das caminhadas e aproveitar essa propriedade importante na classificação (GONÇALVES; BRUNO, 2013a), num processo mais custoso computacionalmente e que gera vetores de características de maior tamanho.

Este trabalho propõe as direções da caminhada como característica discriminante a ser extraída da caminhada. Em um grafo comum, a informação de direção e sentido da caminhada não é aplicável; a topologia do grafo não assume qualquer relação especial entre os vértices. Em um grafo gerado a partir de uma imagem, porém, onde cada vértice representa um pixel que possui coordenada específica no espaço (ou espaço-tempo, no caso de texturas dinâmicas), é possível usar essa informação direcional adicional para compor a assinatura de uma textura. Considerando uma representação de uma imagem como um grafo onde cada vértice possui n vizinhos, (de acordo com o raio de vizinhança r_G escolhido), cada passo da caminhada pode ter uma de n direções diferentes. A cada direção atribuiremos um índice.

Considera-se que um passo do pixel (x_a, y_a) para o pixel (x_b, y_b) tem a mesma direção que um passo de (x_c, y_c) para (x_d, y_d) se $(x_a - x_b) = (x_c - x_d)$ e $(y_a - y_b) = (y_c - y_d)$. Podemos representar a direção de um passo de caminhada como $dir(w_k, w_{k-1})$. Se considerarmos que cada uma das n direções possíveis pode ser representada por um índice no intervalo $[1, n]$, podemos construir um histograma de direções H . Cada passo em uma determinada direção será adicionado ao histograma H no *bin* correspondente ao índice daquela direção, de acordo com a Equação (22). Assim, cada passo na n -ésima direção significa um incremento ao n -ésimo *bin* de H . Em nossa proposta, em casos de empate a decisão de próximo passo, o vizinho de índice mais baixo é selecionado.

$$H_{\mu, dyn}(i) = \sum_{i=1}^n \delta(dir(w_k, w_{k-1}), i), \quad w_k \in W_{\mu, dyn}^{x, y} \quad (22)$$

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

$$\delta(j, i) = \begin{cases} 1, & j = i \\ 0, & j \neq i \end{cases} \quad (23)$$

Um resumo do método proposto, desde a imagem original até a geração de um único histograma de caminhadas com um conjunto de parâmetros específicos, pode ser encontrado na Figura 9. A imagem original I é transformada em D usando o método baseado em redes complexas descrito na Figura 8, e para cada pixel i , tamanho de memória μ e dinâmica dyn , obtém-se uma caminhada $W_{\mu, dyn, i} = \{w_1, w_2, \dots, w_n\}$. Note que, na prática, convém analisar uma mesma imagem sob diversas combinações de parâmetros para obter uma descrição completa da imagem.

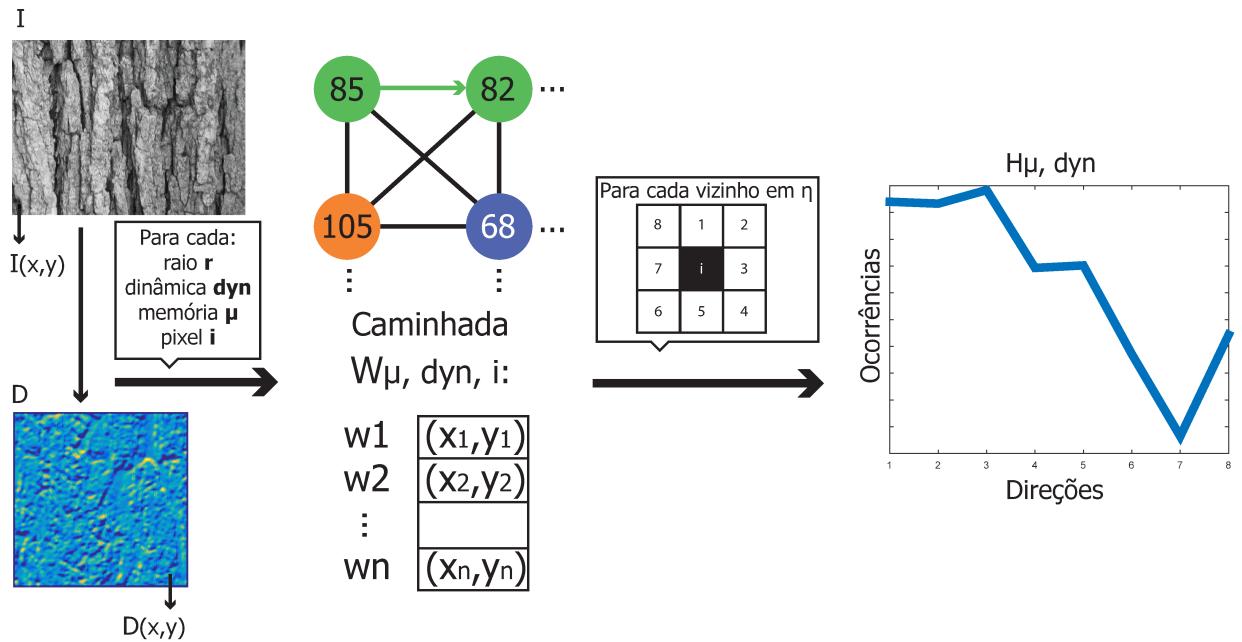


Figura 9 – Visão geral do método proposto. Da imagem I obtém-se a imagem D . Uma caminhada é gerada para cada valor de dyn , μ e cada pixel i . Cada passo w da caminhada W para cada pixel i é acrescentado a um histograma de direções de passos $H_{\mu, dyn}$, incrementando o *bin* de H correspondente à direção do passo. $H_{\mu, dyn}$ descreve a imagem inteira.

$H_{\mu, dyn}$ é, portanto, o histograma de direções de caminhadas calculado para cada valor de memória μ e para cada dinâmica de caminhada dyn , onde $\mu \in \{1, 2, \dots, \mu_\Theta\}$ e $dyn \in \{min, max\}$. Valores de memória menores favorecem padrões de menor escala em uma imagem, enquanto valores de memória maiores distinguem padrões de maior escala. Por sua vez, a dinâmica de caminhada *min* faz com que a caminhada tenda a percorrer regiões homogêneas da imagem, enquanto a dinâmica de caminhada *max* favorece regiões heterogêneas (BACKES et al., 2010). A influência desses dois parâmetros é ilustrada na Figura 10.

Para uma descrição integral das características da textura, é, portanto, de interesse que um conjunto de diferentes combinações de valores de memória e dinâmicas de caminhada seja usada. Para todo pixel na transformada D da imagem original I , temos um vetor ϕ_D

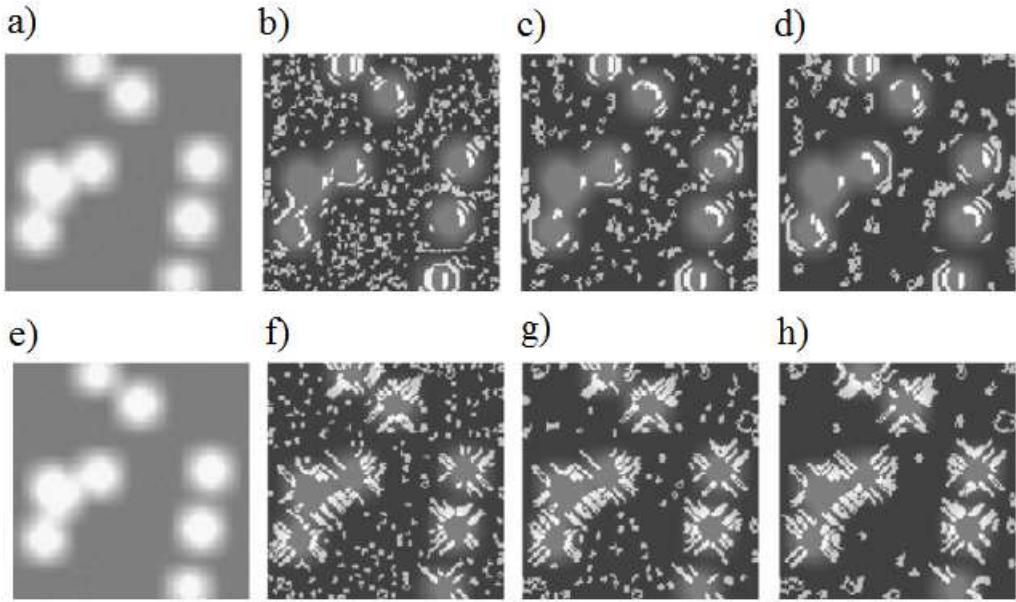


Figura 10 – Partindo da textura original em a) e usando a dinâmica min , temos atratores de caminhadas determinísticas gerados pelos valores de memória: b) $\mu = 1$, c) $\mu = 3$ e d) $\mu = 7$. Os padrões em f), g) e h) foram gerados pelos mesmos valores de memória que b), c) e d), respectivamente, mas usando dinâmica de caminhada max . Adaptado de (BACKES et al., 2010).

dado na Equação 24, calculado pela concatenação de histogramas de $|\mu_{theta}|$ tamanhos diferentes de memória, e usando uma das dinâmicas de caminhada $dyn \in \{min, max\}$.

$$\phi_D = \{H_{1,dyn}, H_{2,dyn}, \dots, H_{\mu_\Theta,dyn}\} \quad (24)$$

Note que as direções das caminhadas possuem dependência intrínseca à orientação das imagens. Isso pode ser ou não ser desejável, a depender da aplicação. Essa sensibilidade a rotação é comum à maioria dos descritores baseados em imagem, de forma que alternativas são buscadas para contornar esta característica. A solução encontrada por métodos como LBP, SIFT ou o extrator de padrões singulares de campos vetoriais para alcançar invariância a rotação envolvem uma descrição relativa da vizinhança do ponto em questão. O método LBP considera versões rotacionadas do padrão encontrado como idênticas (OJALA et al., 2002). Assim, no caso de LBP com 8 vizinhos, as 256 possíveis combinações de vizinhanças se tornam 32, já que cada padrão pode ser rotacionado de 8 maneiras diferentes em torno do pixel central, e os padrões resultantes não são considerados novos padrões, mas versões rotacionadas do primeiro. Isso é uma solução particularmente eficaz para rotações em ângulos múltiplos de 90° , enquanto outras rotações causam problemas ao método, alterando as vizinhanças locais devido à amostragem sub-pixel. O SIFT, por sua vez, realiza uma abordagem análoga atribuindo ao ponto característico extraído uma

orientação principal (LOWE, 2004). Essa atribuição de orientações principais é análoga à solução usada no método do extrator de padrões singulares de campos vetoriais (LIU; RIBEIRO, 2012). Ao comparar dois pontos de interesse, os dois métodos alinham as orientações principais dos pontos.

A maioria dos métodos de classificação de texturas enfrentam complicações ao tratar de texturas rotacionadas. Para que o método aqui proposto seja menos sensível a rotação, nossa proposta é que cada componente $H_{\mu,dyn}$ do histograma seja ordenado e também acrescentado ao vetor de características. Isso particularmente melhora a descrição de texturas pouco direcionais, com padrões que podem recorrer rotacionados. Além disso, assim como o LBP e métodos de caminhada da literatura, nossa proposta também é intrinsecamente invariante tanto a alterações monotônicas no nível de cinza da imagem (brilho) quanto a alterações de contraste.

O histograma de direções de caminhadas apresenta grande riqueza de informações de forma relativamente sucinta. Os valores dos *bins* do histograma compõem o descritor e são uma fonte concisa de informações estatísticas sobre as características das caminhadas. Outras informações sobre as caminhadas podem ser extraídas do histograma na forma de metadados. As métricas mencionadas na seção 2.3 oferecem estatísticas globais que descrevem propriedades topológicas diferentes. Chamaremos de St o conjunto das 7 propriedades apresentadas no capítulo anterior: energia, entropia, *skewness*, contraste, média, variância e kurtosis. O conjunto de estatísticas St é extraído para o histograma de cada dinâmica e anexado ao vetor de características. É necessário notar que o cálculo de St para um conjunto de histogramas ϕ_D (já mencionado na Equação 24), que denotaremos por $St(\phi_{D,\Theta,dyn})$, deve ser realizado antes da ordenação do vetor de características, para preservar a topologia original do histograma na representação. As estatísticas são acrescentadas ao vetor de característica final, que é representado pela Equação 25.

$$F_I = \{\phi_{D,\Theta,min}, St(\phi_{D,\Theta,min}), \phi_{D,\Theta,max}, St(\phi_{D,\Theta,max})\} \quad (25)$$

A composição do vetor de características final F_I , que descreve a imagem I , está representada na Figura 11. Esta figura retoma o produto final da Figura 9, o histograma $H_{\mu,dyn}$, obtido através das caminhadas por todos os pixels com um conjunto de parâmetros específico, e mostra como o histograma obtido para cada conjunto de parâmetros é utilizado, juntamente com suas estatísticas, na composição do vetor completo.

Um exemplo do potencial discriminativo do vetor de características proposto está exemplificado na Figura 12, que apresenta vetores de características gerados com os mesmos parâmetros (8 possíveis direções e $\mu \in [1, 8]$) para três imagens de textura de uma mesma classe, comparando-os com três vetores de textura de uma outra classe. Vetores

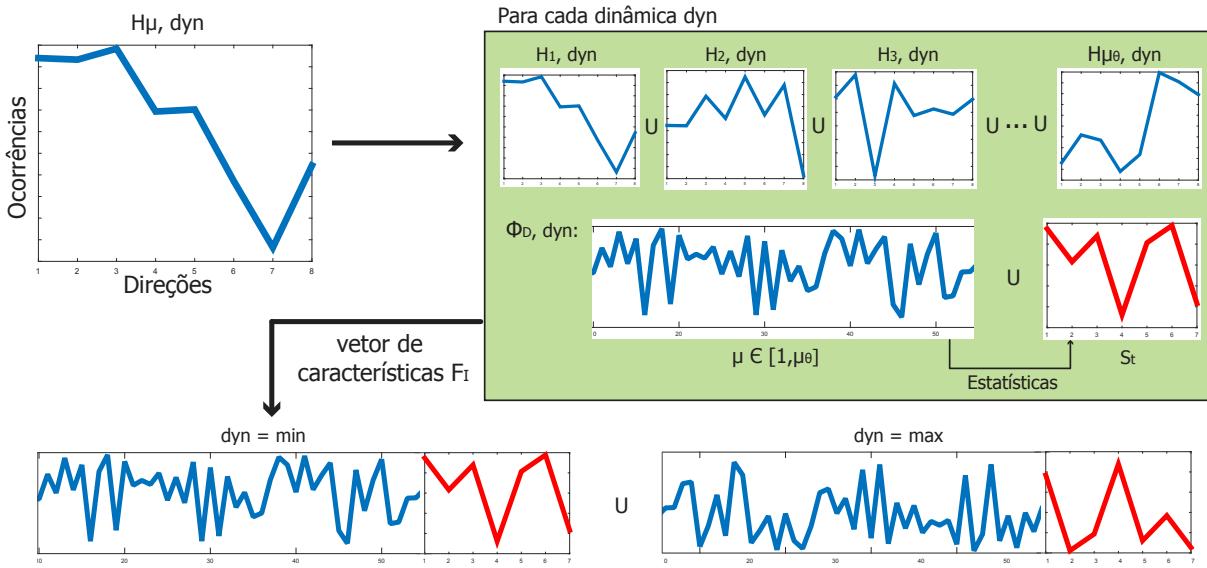


Figura 11 – Composição do vetor de características F_I a partir de suas partes. O vetor final é uma combinação dos histogramas obtidos para diferentes parâmetros, unido às estatísticas S_t dos histogramas.

de características da mesma classe estão representados nas mesmas cores, e diferenciados pelo tipo de traçado da linha. É possível visualizar a diferença entre vetores de diferentes classes, assim como a similaridade entre diferentes vetores de uma mesma classe, indicativos de um método com bom potencial de precisão e revocação (*recall*), se tomados os devidos cuidados com o dilema viés-variância na classificação. A Figura 13 também demonstra esse potencial, apresentando dois pares de histogramas arranjados em um eixo de direção de passos (considerando 8 possíveis direções) e um eixo de tamanhos de memória μ (considerando os oito possíveis tamanhos de memória no intervalo $[1, 8]$), e apresentando o resultado na forma de uma superfície tridimensional.

Resultados parciais referentes a esta proposta podem ser encontrados na literatura em dois trabalhos de Couto *et al.* (COUTO et al., 2015) (COUTO; BACKES; BARCELOS, 2017).

3.2.1 Complexidade computacional

Dado que as aplicações do algoritmo envolvem com frequência grandes volumes de dados, cabe aqui uma análise de sua complexidade computacional.

Consideremos uma imagem I de tamanho $N \times N$. A geração da imagem transformada D baseada em redes depende da comparação de cada um dos N^2 pixels com cada vizinho dentro do raio r_G . Isso é realizado uma vez para cada limiar $t \in S$, ou seja, $|S|$ vezes. Como os valores de r_G e de $|S|$ são muito menores que os valores de N^2 , podemos afirmar que

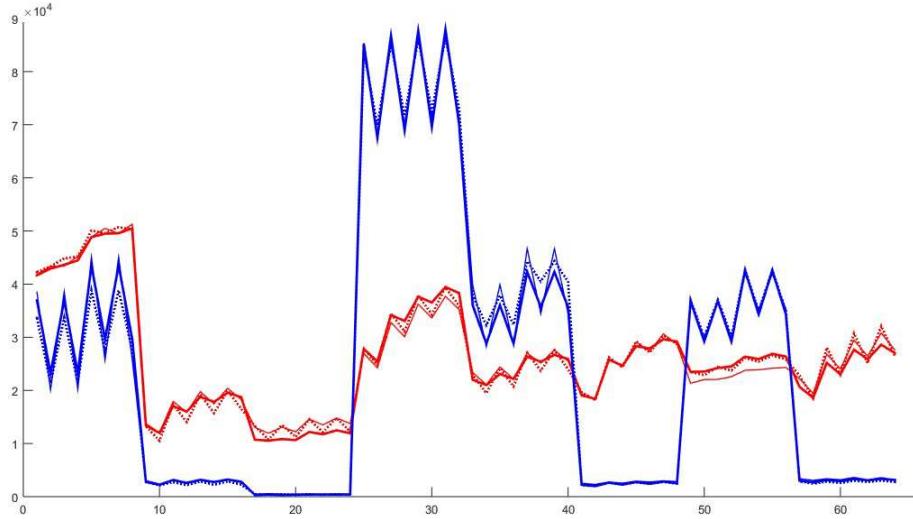


Figura 12 – Comparação de diferentes vetores de características para dois conjuntos de três imagens cada, cada conjunto de um classe diferente (representadas por cores diferentes). O eixo horizontal mostra os índices do vetor, e o eixo vertical mostra seu valor.

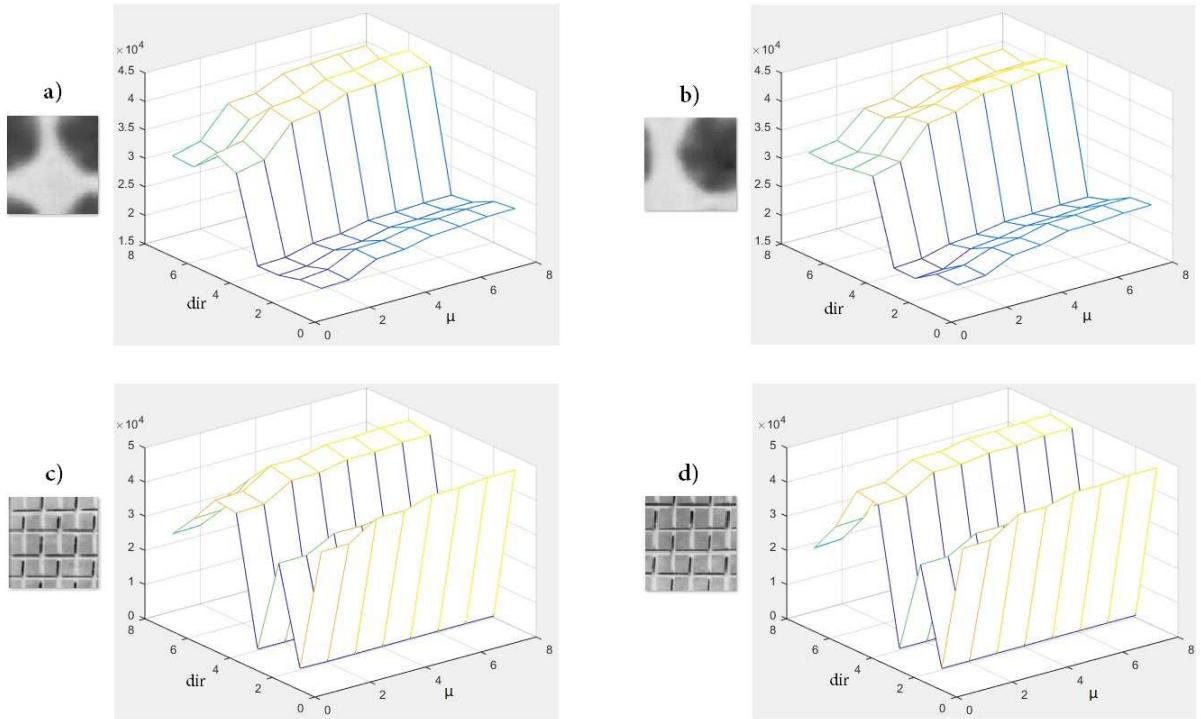


Figura 13 – Exemplos de histogramas de direção de passos arranjados em um eixo de direções dir e um eixo de tamanhos de memória μ . As imagens a) e b) pertencem à mesma classe, enquanto c) e d) são um par de outra classe.

a complexidade da geração de cada D_t a partir da rede tem complexidade computacional $O(N^2)$, como demonstrado em Backes *et al.* (BACKES; CASANOVA; BRUNO, 2013). Sendo D a média aritmética de cada D_t , D é gerada por $|S|$ operações de média para N^2

pixels. Novamente uma comparação entre os valores de $|S|$ e N^2 nos dá uma complexidade de $O(N^2)$ a esse passo, que não afeta a complexidade original.

No caso das caminhadas, o tamanho da imagem transformada D é o mesmo que o da imagem original I . O algoritmo realiza uma caminhada para cada vértice, e portanto para cada pixel de D , de modo que em uma imagem $N \times N$ haverá N^2 caminhadas. Cada conjunto de passos $|W|$ tem um número variável de passos $|W|$; no mínimo, uma caminhada terá o tamanho de memória μ passos, já que um atrator menor que o tamanho de memória não pode ser encontrado, e no máximo, a caminhada terá o tamanho da imagem, N^2 . Outro parâmetro que influencia o valor de $|W|$ é a dinâmica *dyn*, mas nesse caso o efeito causado pela dinâmica é menos previsível e altamente dependente das propriedades da textura percorrida. O valor de $|W|$ é um atributo particular de cada caminhada, mas se assumirmos um valor médio para $|W|$, temos a complexidade computacional para execução do método em uma imagem inteira: para cada uma das N^2 caminhadas W existem $|W|$ pixels visitados, o que resulta em uma complexidade de $O(N^2|W|)$.

A cardinalidade de W é diretamente afetada pelo tamanho de memória μ , já que valores de menores de μ permitem atratores de menor tamanho e maior quantidade, como demonstrado pelos experimentos de Backes et al. (BACKES et al., 2010). Portanto, o tempo de processamento pra valores de memória menores é significativamente inferior. De fato, no próximo capítulo demonstraremos que, em nossos experimentos na base Brodatz, o tamanho de caminhadas apresenta relação aproximadamente linear em relação ao tamanho de memória μ . A literatura mostra que valores menores de μ são mais computacionalmente viáveis e oferecem melhor potencial descritivo. Assim, a cardinalidade $|W|$ pouco contribui para a ordem de grandeza da complexidade do algoritmo, que pode ser aproximada por $O(N^2)$, já que em média o valor de N é significativamente maior que o valor de $|W|$.

Por fim, a complexidade para computação do vetor de características é $O(|F_I|)$, onde $|F_I|$ é o tamanho do vetor de características, ou, se pensarmos no vetor como um histograma, o número de *bins*.

3.2.2 Parâmetros de caminhadas determinísticas

Um parâmetro relevante para as caminhadas é o raio de vizinhança dos pixels em D . O valor usado para os raios foi $r_D = \sqrt{2}$, que significa 8 vizinhos para cada pixel, o que mantém viável a dimensão do vetor de características, já que o número de elementos do vetor é proporcional ao número de vizinhos. Cada vizinho adicional também significa mais uma comparação por passo da caminhada, o que é significativo em termos de tempo de processamento. Por isso, testes extensivos foram realizados com esse tamanho de raio de caminhada, para diversas bases de dados. Os resultados são apresentados no capítulo 4. Com o raio de $\sqrt{2}$, um pixel tem 8 vizinhos possíveis, o que, considerando

valores de memória μ no intervalo $[1, 8]$, resulta em um vetor de características F_I de 142 elementos (64 para cada dinâmica, mais 7 estatísticas para cada dinâmica). Apesar do maior número de elementos no vetor de características, testes com raio $r_D = 2\sqrt{2}$ foram realizados recentemente com a base de texturas Brodatz são apresentados na seção seguinte. Nesse caso, o vetor de característica tem tamanho 398 (192 para cada dinâmica, mais 7 estatísticas para cada dinâmica).

Outro parâmetro relevante é o tamanho da memória μ . Na literatura, valores menores de μ demonstraram ser mais descritivos. Backes *et al.* (BACKES et al., 2010) e Gonçalves *et al.* (GONÇALVES; BRUNO, 2013b) usam valor máximo de $\mu = 4$, enquanto Gonçalves *et al.* usa um valor máximo de $\mu = 5$ e (CAMPITELI et al., 2006) usam valor máximo $\mu = 8$. Resultados para tamanhos de memória maiores que $\mu = 8$ não aprimoram substancialmente as taxas de classificação correta para compensar o número adicional de dimensões e maior tempo de processamento exigidos, e podem até prejudicar taxas de classificação. Nos experimentos, usamos tamanhos de memória μ no intervalo $[1, 8]$, e comparamos resultados com μ nos intervalos $[1, 4]$ e $[5, 8]$.

3.3 Caminhadas determinísticas tridimensionais

Diversos dos trabalhos mais relevantes de pesquisa em caracterização de texturas estáticas apresentam, como mostramos, versões estendidas para três dimensões, análogas ao caso bidimensional, para que os métodos sejam aplicados a texturas dinâmicas. Nesta seção, proporemos uma extensão do nosso método de caminhadas determinísticas baseadas em redes complexas e histograma de direções para texturas dinâmicas.

Zhao e Pietikäinen (ZHAO; PIETIKAINEN, 2007) estenderam o método LBP para a terceira dimensão para descrição de texturas dinâmicas, com aplicação bem-sucedida em reconhecimento de expressões faciais. Isso foi realizado, como visto no capítulo anterior, com duas propostas distintas; a primeira, chamada de VLBP, considerando de uma única vez uma vizinhança volumétrica do pixel nas três dimensões, e a segunda, o LBP-TOP, dividindo a aplicação do método em três planos. Andrearczy *et al.* (ANDREARCZYK; WHELAN, 2017) e Gonçalves *et al.* (GONÇALVES; BRUNO, 2013b) propõem abordagens para descrição de texturas dinâmicas baseadas na análise da textura em três planos ortogonais. Tal abordagem alcança bons resultados aplicando três caminhadas bidimensionais a cada pixel da sequência de vídeo. Para isso, as três dimensões, $[X, Y, T]$ (coordenadas espaciais e coordenada temporal, respectivamente), são divididas em três pares, $[X, Y]$, $[X, T]$ e $[Y, T]$, formando três planos ortogonais como no caso do LBP-TOP, e os vetores de características resultantes de cada plano são concatenados. De acordo com nosso conhecimento, resultados experimentais que envolvam caminhadas considerando a vizinhança do pixel nos 3 planos simultaneamente não estão disponíveis na literatura.

No caso de nosso trabalho, duas abordagens distintas são propostas para a extensão

para texturas dinâmicas de nosso método de caminhadas determinísticas: uma delas considerando as 3 dimensões $[X, Y, T]$, analisando a vizinhança tridimensional para geração de um único descritor e a outra dividindo as 3 dimensões em 3 planos ortogonais $[X, Y]$, $[X, T]$ e $[Y, T]$ como no trabalho das caminhadas em texturas dinâmicas e no LBP-TOP.

A primeira proposta é que o método seja aplicado considerando a vizinhança do pixel nas três dimensões $[X, Y, T]$ como candidatos para escolha do próximo passo da caminhada. Chamaremos este método de caminhadas volumétricas, devido à analogia com o VLBP. Assim, concentra-se a informação espaço-temporal em apenas um vetor. Nossa hipótese é de que a execução destas caminhadas determinísticas possa gerar bons classificadores com o tamanho do vetor de características reduzido em relação à abordagem de 3 planos ortogonais.

A adaptação mais importante para a adequação do método das caminhadas proposto está no novo conjunto de vizinhos do pixel. A vizinhança de 8 (*8-neighborhood*), por exemplo, é gerada no caso 2D pelo raio $r_D = \sqrt{2}$. Se considerarmos o caso volumétrico e não planar, a vizinhança deve ser adequada para incluir vizinhos das 3 dimensões. Zhao e Pietikäinen mostraram que há relação entre o aumento do raio e o aumento da eficácia da descrição de texturas dinâmicas para raios não maiores que 3, mas a relação cúbica de aumento de vizinhança em relação ao raio faz com que raios maiores sejam inviáveis (ZHAO; PIETIKAINEN, 2007), e Backes *et al.* mostraram a eficácia das caminhadas bidimensionais com raios menores que 2 (BACKES *et al.*, 2010). Mostraremos também a eficácia de raios $r_D = \sqrt{2}$ em nossos experimentos sobre texturas estáticas. Considerando estes resultados, a Figura 14 mostra um exemplo de vizinhança com raio $r_D = \sqrt{3}$, um valor de raio que apresenta simetria radial nas 3 dimensões enquanto mantém o número de vizinhos manejável. A Figura também mostra os índices que denotam a ordem de comparação dos vizinhos usados no método, escolhida arbitrariamente. Nos casos em que há empate entre 2 pixels pelo critério de decisão de passos da dinâmica, opta-se pelo vizinho de menor índice, como no caso estático. Dessa forma, em um vídeo V , um pixel de coordenada $V(x, y, t)$ terá o conjunto de vizinhos $\eta(x, y, t)$, conforme a Equação (26). Nesse caso, um passo de caminhada pode levar a caminhada de um quadro para o quadro seguinte ou anterior no vídeo.

$$\eta(x, y, t) = \{(x', y', t') | \sqrt{(x - x')^2 + (y - y')^2 + (t - t')^2} \leq \sqrt{3}\} \quad (26)$$

Uma vantagem das caminhadas volumétricas de nossa proposta em relação ao VLBP é que o aumento do número de vizinhos possíveis, no caso das caminhadas, acarreta em um aumento linear do vetor de características. No caso do VLBP, por conta da natureza combinacional da codificação LBP, o vetor de características aumenta exponencialmente

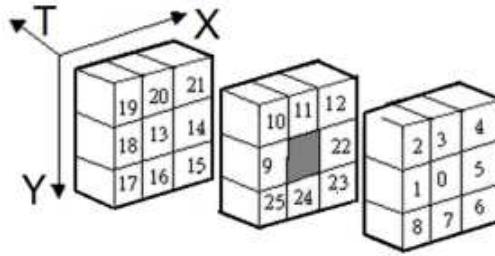


Figura 14 – Vizinhança tridimensional de pixel, considerando 26 vizinhos. Adaptado de <https://clouard.users.greyc.fr/Pandore/c++doc/html>. (Acessado em 03/04/2018).

com o aumento dos vizinhos. Um vetor com 14 vizinhos possíveis, nas 3 dimensões, teria $2^{14} = 16384$ atributos, uma dimensionalidade alta para o vetor de características.

Considerando-se as três dimensões para a caminhada determinística parcialmente auto-repelente, a proposta de caminhadas volumétricas consiste da execução do método das caminhadas sobre valores dos *pixels* dos quadros do vídeo, e também sobre uma rede complexa gerada a partir dos quadros do vídeo, em uma versão espaço-temporal do método descrito na seção 2.2. Descreveremos agora nossa proposta para geração da rede complexa considerando as 3 dimensões dos dados. A geração da imagem transformada D_t , baseada nos graus da rede G_t , nesse caso, sofre algumas alterações. A função $\omega(v_{x,y,t}, v_{x',y',t'})$ de dissimilaridade entre dos vértices agora é função de 3 coordenadas e é dada pela Equação 27. O conjunto de arestas E , antes da aplicação do limiar é dado de acordo com a Equação 28. A seleção do subconjunto E^* em função da dissimilaridade ω e do limiar t permanece inalterado em relação ao apresentado na Equação 29.

$$\omega(v_{x,y,t}, v_{x',y',t'}) = (x - x')^2 + (y - y')^2 + (t - t')^2 + r_G^2 \frac{|V(x, y, t) - V(x', y', t')|}{L} \quad (27)$$

$$E = \left\{ (v_{x,y,t}, v_{x',y',t'}) \in V \times V \mid \sqrt{(x - x')^2 + (y - y')^2 + (t - t')^2} \leq r_G \right\} \quad (28)$$

$$E^* = \{e \in E \mid w(e) \leq t\} \quad (29)$$

A Figura 15 apresenta um resumo do método das caminhadas volumétricas proposto. Cabe notar as semelhanças com o método de caminhadas estáticas que apresentamos na Figura 9. No caso dinâmico, assim como no caso estático, nossa proposta difere da literatura no uso da transformação prévia da imagem e na informação de direção de passos extraída da caminhada para criação dos histogramas que compõem o vetor de características. As diferenças principais entre as propostas estática e dinâmica estão na

natureza tridimensional dos dados de entrada, e portanto do grafo gerado, e na vizinhança considerada para cada passo de caminhada.

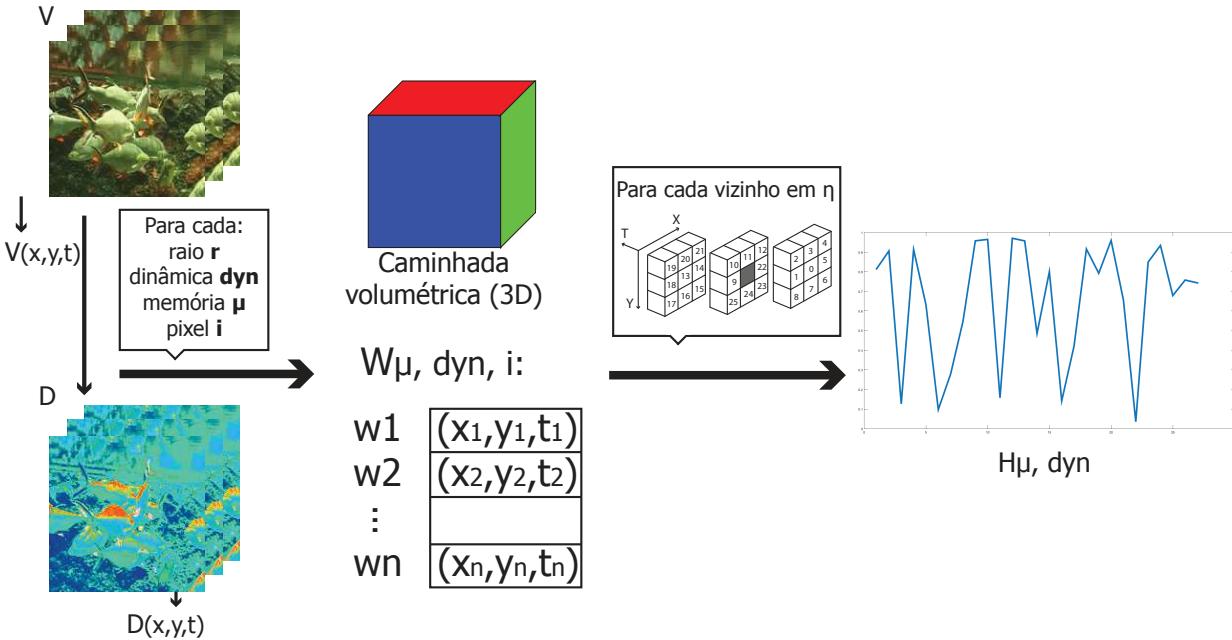


Figura 15 – Sumário do método de caminhadas volumétricas. A partir da transformação baseada em redes complexas dos quadros do vídeo V em quadros D , as caminhadas volumétricas são aplicadas em cada pixel (cada passo de caminhada tem 3 coordenadas), gerando por fim o histograma de direções de passos de caminhada descritor de V .

Gerando-se descritores de seções de vídeos representativos de uma textura (um subconjunto de quadros) em uma base de vídeos de texturas dinâmicas, é possível verificar se os descritores da proposta são eficazes.

Também propomos a execução do método de caminhadas considerando os 3 planos ortogonais, pareando as coordenadas $[X, Y]$, $[X, T]$ e $[Y, T]$, de forma análoga à estratégia aplicada nos métodos LBP-TOP e de Gonçalves *et al.*. É comum considerar o vídeo como um conjunto de quadros bidimensionais $[X, Y]$ em sequência ao longo do eixo T , mas da mesma forma é possível considerar o vídeo como um conjunto de quadros $[X, T]$ ao longo do eixo Y e $[Y, T]$ ao longo do eixo X (ZHAO; PIETIKAINEN, 2007). A consideração destes 3 pontos de vista captura no vetor de características, além do comportamento espacial da textura, as interações das dimensões espaciais com a dimensão temporal, um aspecto fundamental das texturas dinâmicas. Chamaremos esta proposta para caminhadas em texturas dinâmicas de Caminhadas TOP (da sigla para *Three Orthogonal Planes*). A aplicação, neste caso, é semelhante ao caso bidimensional; a diferença principal está na geração do vetor de características. Neste caso, o vídeo é considerado como um conjunto de imagens bidimensionais para as quais um único vetor de características será gerado. Os histogramas gerados para cada quadro do vídeo são somados elemento a elemento, com-

pondendo um histograma final $H_{plano} = H_1 + H_2 + \dots + H_N$ descritor da sequência de quadros completa. Este processo é realizado 3 vezes, e em cada uma das vezes a análise do vídeo se dá sob a perspectiva de um dos planos $[X, Y]$, $[X, T]$ e $[Y, T]$. Para compor o vetor de características completo F_V do vídeo, os histogramas para cada plano são concatenados, como mostra a Equação 30.

$$F_V = \{H_{[X,Y]}, H_{[X,T]}, H_{[Y,T]}\} \quad (30)$$

A Figura 16 mostra um sumário entre a metodologia baseada em caminhadas TOP e o método de caminhadas volumétricas. Uma comparação entre esta figura e a Figura 15 denota as semelhanças e diferenças principais entre os métodos. É possível ver como o método volumétrico consiste de caminhadas que consideram as 3 dimensões do vídeo simultaneamente, enquanto o método TOP considera cada um dos três planos ortogonais individualmente, e o vetor de características é formado pela concatenação do vetores de características obtidos para cada plano.

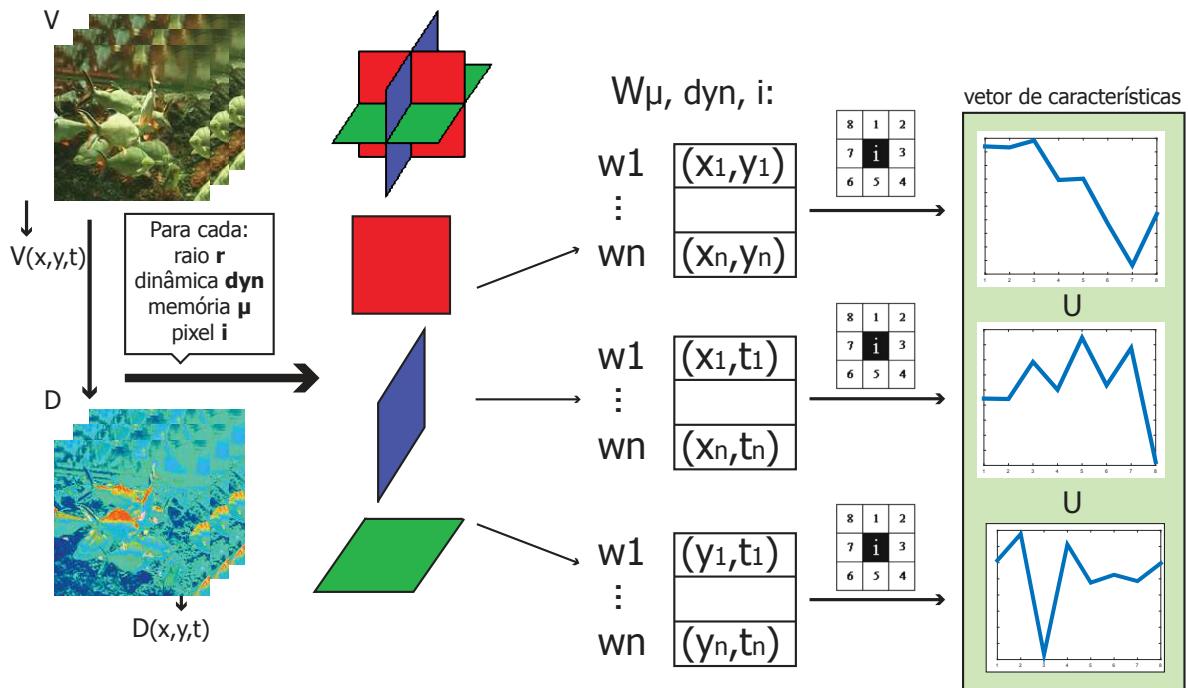


Figura 16 – Sumário do método de caminhadas TOP (três planos ortogonais). A partir da transformação baseada em redes complexas dos quadros do vídeo V em quadros D , caminhadas são aplicadas em cada plano, gerando um histograma de direções de passos de caminhada para cada plano, que concatenados compõem o descritor de V .

3.3.1 Complexidade computacional

A análise da complexidade computacional do método proposto no caso espaço-temporal é significativa, já que a componente temporal introduz uma nova dimensão de *pixels* que devem ser considerados, o que multiplica o espaço a ser computado. A complexidade desse algoritmo pode ser melhor compreendida como uma extensão do método bidimensional, para texturas estáticas, já analisado anteriormente. Consideremos, primeiramente, a proposta de caminhadas volumétricas. Seja V um vídeo de tamanho $N \times N \times N$, isto é, um vídeo composto por N quadros de resolução $N \times N$ *pixels*. Da mesma forma que no caso bidimensional, haverá uma caminhada que se inicia em cada *pixel*, de modo que se o tamanho de uma caminhada W é $|W|$. Assumindo um valor médio para $|W|$, temos que a complexidade computacional da execução do método em um vídeo inteiro é $O(N^3|W|)$. É importante notar que como neste caso a vizinhança é maior, é menos provável que um determinado passo leve a caminhada ao laço infinito que é seu critério de parada. Uma análise empírica do tamanho médio de caminhadas $|W|$ para o caso tridimensional volumétrico é apresentada no capítulo seguinte, mostrando como as caminhadas volumétricas são em média bem maiores.

No caso das caminhadas em três planos ortogonais, realiza-se 3 caminhadas sobre cada vídeo, sob 3 perspectivas diferentes, cada uma considerando seções do vídeo paralelas a um dos planos ortogonais. Assim, a complexidade é de ordem $O(3(N^3|W|))$. O fator de multiplicação 3 é uma constante de menor magnitude, se comparado aos outros fatores da equação, e portanto pode ser suprimido. Resta que a complexidade do algoritmo de caminhadas no caso de três planos ortogonais é de ordem igual à do caso volumétrico, $O(N^3|W|)$, mas na prática ele demora mais (as caminhadas TOP são de menor tamanho que as volumétricas, mas o método TOP requer o triplo de caminhadas). A análise dos tamanhos de caminhadas $|W|$, neste caso, é a mesma que para o caso bidimensional.

3.4 Padrões singulares de campos vetoriais aplicados a texturas dinâmicas

A seguir, apresentamos nossa proposta de aplicação do método de extração de padrões singulares de campos vetoriais apresentado por Liu e Ribeiro (LIU; RIBEIRO, 2012) como uma ferramenta para caracterização de texturas dinâmicas.

O detector de padrões singulares, como apresentado no capítulo anterior, baseia-se em identificar componentes recorrentes de campos vetoriais baseado em um dicionário de campos fundamentais. A ideia de usar um dicionário de componentes de textura não é nova, tendo ampla aplicação na descrição de texturas estáticas, sendo que os elementos do dicionários são com frequência chamados de *textons* (VARMA; ZISSERMAN, 2005) (ZHU et al., 2005), um termo usado para descrever elementos atômicos da percepção

visual humana (JULESZ, 1981). No caso de texturas dinâmicas, para descrever o aspecto temporal da textura através de um dicionário, é necessário um dicionário que capture variações temporais da imagem. É exatamente estas variações temporais que os padrões singulares de campos vetoriais denotam.

Propomos um método para compilar estatisticamente pontos de interesse ocorrentes em campos vetoriais, obtidos do dicionário de padrões singulares apresentados no trabalho de Liu e Ribeiro, para gerar um descritor que caracteriza uma textura dinâmica. A proposta se aproveita da periodicidade temporal e espacial das texturas dinâmicas, que sugere que os padrões detectados sejam recorrentes e que esta característica pode ser salientada através de uma abordagem estatística. Esta análise de distribuição de características locais, na qual se baseia nossa proposta, também é chamada de *bag-of-features*, e tem sido considerada extremamente efetiva para descrição de imagens (AGARWAL; AWAN; ROTH, 2004) (YANG et al., 2007) e, especificamente, texturas (LEUNG; MALIK, 2001) (NOWAK; JURIE; TRIGGS, 2006) (ZHANG et al., 2007) (XU et al., 2012). Uma propriedade essencial dos padrões singulares propostos por Liu e Ribeiro é sua invariância a rotação e escala. Atributos que conferem uma importante flexibilidade de aplicações ao reconhecimento de texturas.

O primeiro passo de nossa proposta é a computação do fluxo óptico para todos os pares de quadros consecutivos do vídeo por um método como Lucas-Kanade (LUCAS; KANADE et al., 1981) ou Horn-Schunck (HORN; SCHUNCK, 1981). Os dois métodos são bastante distintos; o Lucas-Kanade apresenta uma abordagem local e o Horn-Schunck uma abordagem global à geração do fluxo. O método Horn-Schunck, por sua vez, gera campos vetoriais mais densos, o que é desejável em aplicações de extração de pontos e interesse em campos vetoriais (CORPETTI; MEMIN; PÉREZ, 2003). Optamos pelo método Lucas-Kanade porque ele é mais robusto a ruído e acarreta em uma menor suavização das variações entre os quadros de entrada que o método Horn-Schunck (BRUHN; WEICKERT; SCHNÖRR, 2005). A suavização do método Horn-Schunck pode ser desejável em algumas aplicações, mas no caso de padrões locais como texturas dinâmicas é útil capturar tais detalhes, e qualquer suavização pode prejudicar a caracterização de texturas (PIETIKÄINEN; OJALA; XU, 2000). Há um compromisso natural que mantém os dois métodos relevantes ao longo das décadas de uso.

De um vídeo de n quadros, portanto, extraí-se $n - 1$ campos vetoriais de fluxo óptico. A seguir, decompõe-se cada fluxo óptico em um vetor de componentes descritivos (LIU; RIBEIRO, 2012). O resultado disso é uma sequência de coeficientes a_k dos k padrões singulares encontrados naquele fluxo óptico (vide seção 2.4). Exploraremos a utilização dos vetores de coeficientes de cada padrão singular como uma assinatura para todo o campo vetorial.

As Figuras 17 e 18 mostram exemplos dos passos de execução do método de detecções de padrões singulares em campos vetoriais. É possível observar como os padrões singulares

detectados, após a aplicação do limiar de energia singular, descrevem as regiões de maior movimento do quadro, ignorando o fundo imóvel. Os campos vetoriais usados foram extraídos de vídeos de texturas dinâmicas da base de dados DynTex (PÉTERI; FAZEKAS; HUISKES, 2010) usando uma implementação do algoritmo de fluxo óptico de Lucas-Kanade.

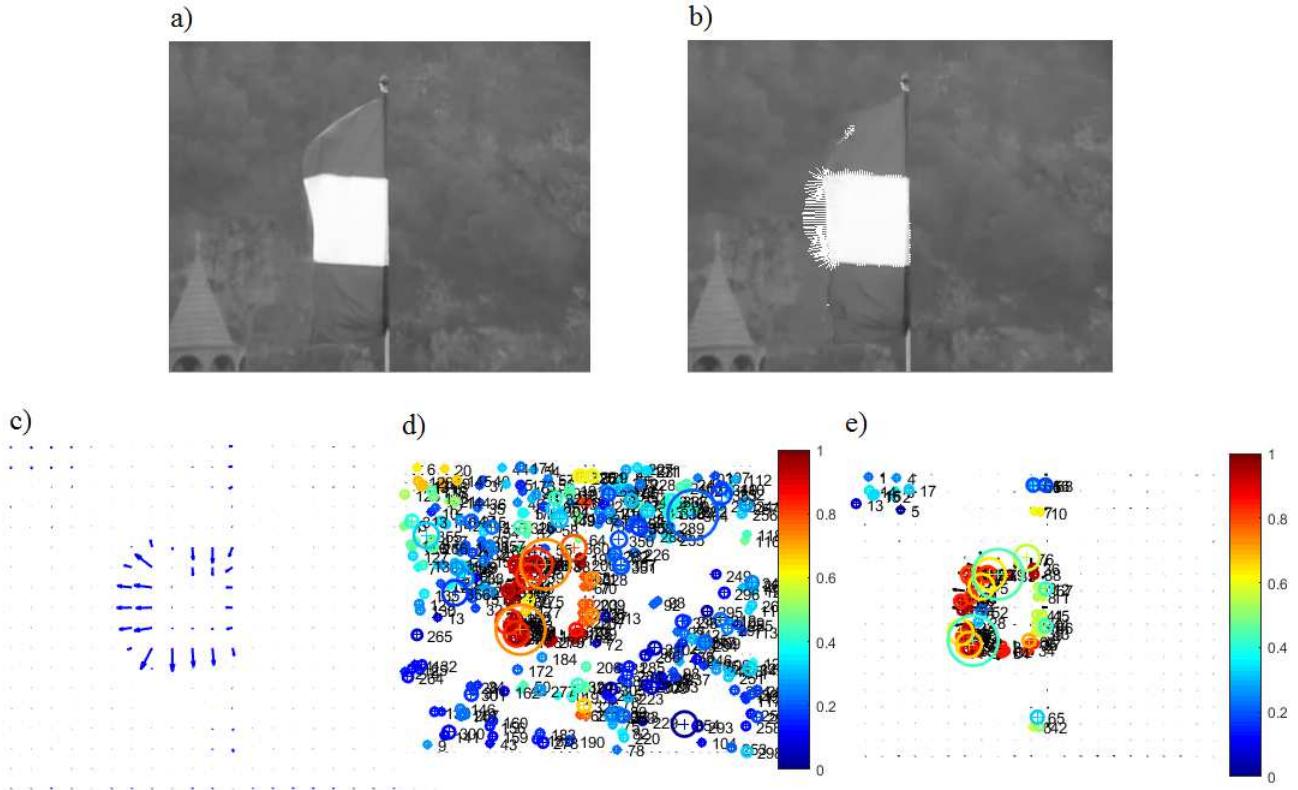


Figura 17 – Exemplo de execução do método de detecção de padrões singulares (LIU; RIBEIRO, 2012) sobre vídeo de uma bandeira. a) ilustra o quadro original, b) apresenta linhas onde o fluxo óptico é acentuado entre os quadros, c) apresenta os vetores do fluxo óptico, d) ilustra os padrões detectados pelo método; o tamanho do círculo representa magnitude e a cor representa a magnitude em logaritmo da energia singular do padrão e e) apresenta o subconjunto de padrões mais fortes (de maior energia singular).

Cada fluxo óptico da imagem apresenta um conjunto de padrões singulares. Cada padrão singular é descrito por um vetor de coeficientes a_k . A hipótese que fundamenta nossa proposta neste caso é que, quando se trata de texturas dinâmicas, a periodicidade temporal promove a repetição de padrões ao longo do vídeo. Assim, os descritores destes padrões também recorrem espacialmente ao longo da imagem e temporalmente entre quadros diferentes. Dessa maneira, existe similaridade entre os conjuntos de padrões singulares obtidos dos fluxos ópticos de pares distintos de quadros do vídeo da textura dinâmica. Adicionalmente, a periodicidade espacial indica que dentro de um fluxo óptico

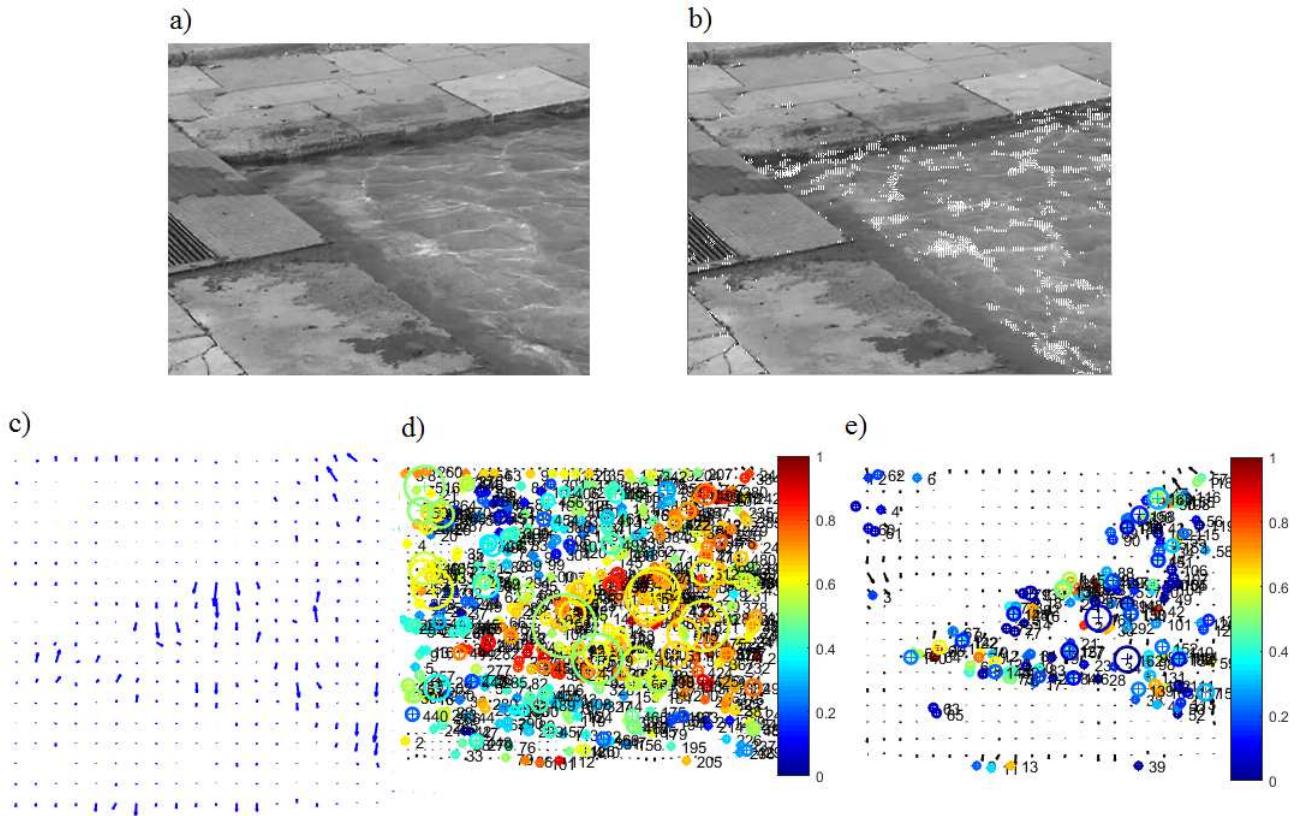


Figura 18 – Exemplo de execução do método de detecção de padrões singulares (LIU; RIBEIRO, 2012) sobre vídeo de uma piscina. a) ilustra o quadro original, b) apresenta linhas onde o fluxo óptico é acentuado entre os quadros, c) apresenta os vetores do fluxo óptico, d) ilustra os padrões detectados pelo método; o tamanho do círculo representa magnitude e a cor representa a magnitude em logaritmo da energia singular do padrão e e) apresenta o subconjunto de padrões mais fortes (de maior energia singular).

representante de um único quadro há repetição de padrões singulares. Para realçar tais padrões repetitivos característicos de cada textura dinâmica, nossa proposta é que sejam organizados em um histograma, que servirá como descritor de um fluxo óptico completo, ou de um conjunto de fluxos ópticos. Metodologias baseadas em *bag-of-features* abrem precedente para o uso de histogramas de pontos de interesse para descrição de padrões repetitivos como texturas dinâmicas, como no trabalho de Xu *et al.* (XU *et al.*, 2012), onde os pontos de interesse gerados por uma extração de características baseada no método SIFT foram compilados em um histograma para composição do vetor de características. O método dos padrões singulares é recente e não há na literatura uma abordagem *bag-of-features* para descrição de um campo complexo a partir dos componentes locais.

Para a construção do vetor de características a partir da *bag-of-features*, apresentaremos duas propostas na seção a seguir.

3.4.1 Vetor de características por *bag-of-features*

A origem do método *bag-of-features* pode ser traçada para os trabalhos de Sivic *et al.* (SIVIC; ZISSEMAN, 2003) e Csurka *et al.* (CSURKA et al., 2004). O método é inspirado em um método originalmente usado para caracterização de textos, de forma que descritores estatísticos similares são com frequência chamados de *bag-of-visual-words*, ou *bag-of-keypoints*.

A geração do descritor *bag-of-features* consiste em reunir um conjunto de características, ou pontos de interesse, em uma única estrutura que é mais descriptiva do todo que qualquer ponto de interesse local isolado. A construção típica de uma *bag-of-features* envolve os seguintes passos, em ordem (O'HARA; DRAPER, 2011):

- **Computação dos padrões de treinamento:** A computação dos pontos de interesse locais (como SIFT ou padrões singulares) para um conjunto de dados de treinamento.
- **Quantização:** A quantização ou clusterização dos descritores, através, por exemplo, do algoritmo k-means (DUDA; HART; STORK, 2012) (XU; TIAN, 2015). Este passo é por vezes chamado de construção do "vocabulário";
- **Agrupamento:** A extração de pontos de interesse para um novo elemento e seu agrupamento nos clusters por proximidade.
- **Geração do histograma:** A construção da *bag-of-features* através da construção de um histograma de ocorrência de padrões obtidos para o elemento descrito em cada um dos *clusters*.

No histograma resultante do processo, cada *bin* representa o número de características agrupadas em cada *cluster* do elemento descrito. Este histograma é um vetor de características que utiliza informações locais, agrupando-as para gerar uma descrição global. Levando em consideração o reduzido tamanho do descritor em relação ao volume de dados que ele representa, o *bag-of-features* tem alto poder de caracterização (AGGARWAL; AGRAWAL, 2012).

No caso da *bag-of-features* aplicada aos descritores de padrões singulares, considere um conjunto de treinamento de n campos vetoriais $F_{treino} = \{F_1, F_2, \dots, F_n\}$. Destes campos vetoriais de treinamento extrai-se amostras para construção do vocabulário, de forma que chamaremos de $X_{F_{treino}}$ o conjunto de padrões singulares computados para F_{treino} , e cada um dos padrões singulares individuais $x \in X_{F_{treino}}$. Cada x possui seu respectivo vetor de k coeficientes descritores $\mathbf{a}_x = (a_1, a_2, \dots, a_k)$. A quantização consiste em dividir o conjunto de padrões $X_{F_{treino}}$ em M *clusters*, cujos centroides são dados por C_1, C_2, \dots, C_M . Note que cada centroide tem dimensão k , o mesmo número de elementos de cada vetor \mathbf{a}_x . Todo este processo está sumarizado na Figura 19.



Figura 19 – Resumo da geração do vocabulário na metodologia *bag-of-features* aplicada aos padrões singulares de campos vetoriais.

Para gerar o descritor de um novo campo vetorial F_{novo} , computa-se seus padrões singulares $X_{F_{novo}}$. Cada padrão singular x é agrupado em um dos M clusters, por proximidade ao centroide C_M . A Equação 31 mostra a escolha do cluster C_x para um padrão singular x .

$$C_x = \arg \min_{1 \leq m \leq M} |a_x - C_m|, x \in X_{F_{novo}} \quad (31)$$

A seguir, construímos o histograma $H_{F_{novo}}$ que descreve o conjunto de padrões singulares $X_{F_{novo}}$, incrementando o *bin* correspondente a cada C_x , para todos os $x \in X_{F_{novo}}$, como na Equação 32.

$$H_{F_{novo}}(m) = \sum_{m=1}^M \delta(C_x, m), \forall x \in X_{F_{novo}} \quad (32)$$

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

$$\delta(j, i) = \begin{cases} 1, & j = i \\ 0, & j \neq i \end{cases} \quad (33)$$

O processo de agrupamento de pontos para uma nova amostra, gerando um histograma de ocorrências de pontos em cada cluster, está descrito na Figura 20. o processo é repetido para múltiplos campos vetoriais de uma mesma textura dinâmica, para geração de um descritor completo.

O histograma resultante pode ser descrito mais profundamente pelo cálculo do conjunto de estatísticas St , que é adicionado ao vetor de características. Assim, o vetor de características descritor de um campo vetorial F , dado por ψ_F , é apresentado na Equação 34.

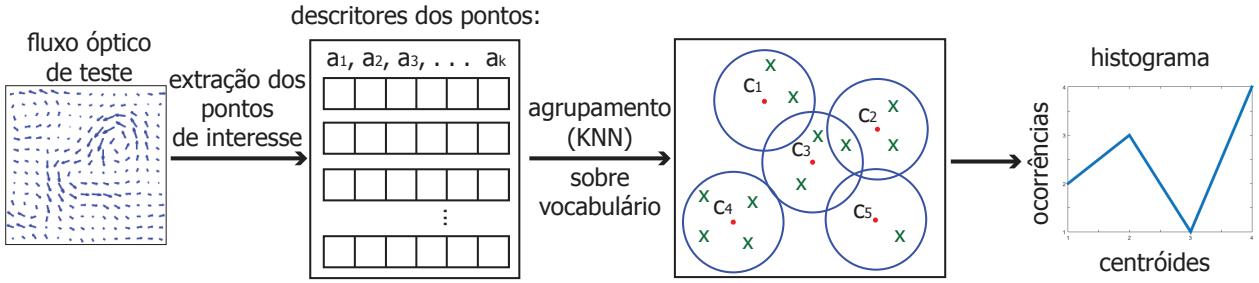


Figura 20 – Agrupamento dos pontos de interesse, ou *features*, adquiridos pelo detector de padrões singulares no vocabulário gerado da *bag-of-features*, de onde é obtido o histograma de ocorrências.

$$\psi_F = \{H_F, St_{H_F}\} \quad (34)$$

O número de *clusters* gerados no passo de quantização é um informação de grande importância na geração da *bag-of-features*. Na sessão 4.5.1 discutiremos as duas heurísticas usadas para decisão do número de *clusters*.

Na seção a seguir, apresentamos uma nova proposta para a construção de uma estrutura similar à *bag-of-features*, mas que oferece uma solução de aplicação particular ao problema dos padrões singulares, já que considera as propriedades específicas dos descritores locais de padrões singulares de campos vetoriais. A abordagem dispensa a quantização.

3.4.2 Vetor de características por histograma de coeficientes a_k

Considere um campo vetorial F . Considere ainda cada um dos padrões singulares $x \in F$, e seu respectivo vetor de coeficientes descritores $\mathbf{a}_x = (a_1, a_2, \dots, a_k)$. Note que, como visto no capítulo anterior, cada coeficiente a_k do vetor a_x denota a representatividade de determinado padrão singular do dicionário no padrão real x . Devido à periodicidade espaço-temporal da textura dinâmica, agrupar os valores de a_k em um histograma pode conter informações descritivas a respeito da ocorrência de cada elemento do dicionário em cada textura dinâmica. Todo padrão real x tem um a_x , mas seus valores diferem substancialmente.

Para representar as ocorrências de valores de a_k , propomos a construção de k histogramas $H_{F,k}$ onde, para cada valor de k , avaliamos os valores de a_k para todos os $x \in X_F$ padrões singulares de F e os organizamos nos n bins de $H_{F,k}$, cada bin representando o número de ocorrências de a_k dos padrões singulares de F em cada intervalo A , de acordo com a Equação 35. Assim, para cada valor de k em um campo vetorial F é gerado um histograma.

$$H_{F,k}(i) = \sum_{i=1}^n \mathbf{1}_A(\mathbf{a}_x(k)), \quad \forall x \in F, \quad (35)$$

onde $\mathbf{1}_A(x)$ é a função indicadora:

$$\mathbf{1}_A(x) = \begin{cases} 0, & x \notin A \\ 1, & x \in A \end{cases} \quad (36)$$

A proposta para geração do vetor de características a partir de um histograma de valores de coeficientes a_k está sumarizada na Figura 21. Note como um histograma é gerado para cada coeficiente a_k , e cada histograma descreve a presença de um padrão do dicionário na composição do conjunto de fluxos ópticos da textura dinâmica descrita.

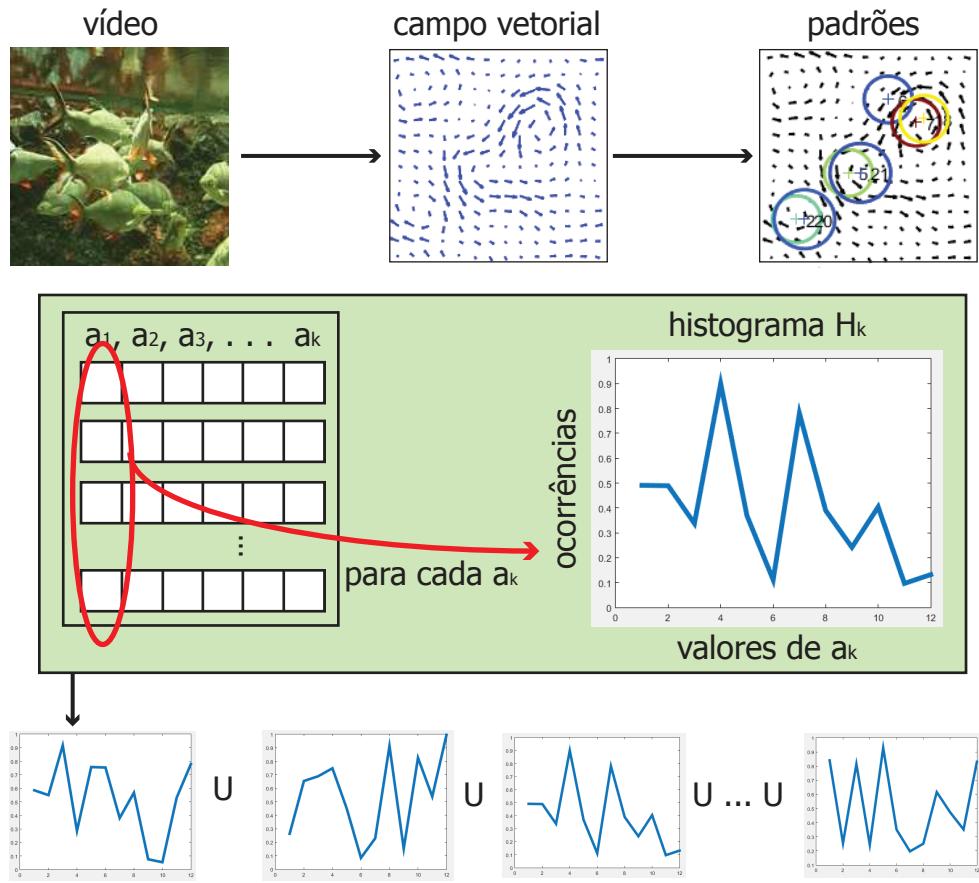


Figura 21 – Sumário da geração do vetor de características por histogramas de coeficientes a_k . Neste método os pontos de interesse obtidos do fluxo óptico tem os coeficientes de seus vetores agrupados em histogramas, cada histograma agrupando as ocorrências de valores para um coeficiente a_k .

O conjunto de estatísticas St é calculado para cada histograma H_k e acrescentado na composição do descriptor completo do campo vetorial, já que as métricas contribuem com

a descrição da topologia dos histogramas. Realizamos a concatenação dos histogramas de diferentes valores de k e suas respectivas estatísticas, formando para F um descritor completo ψ_F , como mostra a Equação 37:

$$\psi_F = \{H_{F,1}, St_{H_{F,1}}, H_{F,2}, St_{H_{F,2}}, \dots, H_{F,k}, St_{H_{F,k}}\} \quad (37)$$

O número n de *bins* e o tamanho dos *bins* representam pelos valores dos intervalos A (da Equação 36) são parâmetros importantes para evidenciar a distribuição dos dados no histograma. Os intervalos entre os *bins* devem prover uma distribuição adequada de ocorrências. Na prática, usaremos tamanhos de *bins* que variam exponencialmente, já que os valores de coeficientes a_k variam nessa proporção, como mostraremos no próximo capítulo.

Qualquer que seja a estratégia para geração do histograma descritor dos fluxos ópticos, para descrever um vídeo V de q quadros, do qual podem ser extraídos $q - 1$ fluxos ópticos (ou campos vetoriais) F , geramos k histogramas $H_{F,k}$ para cada quadro de V . Em seguida os histogramas de mesmo k de diferentes quadros são somados *bin* a *bin*, compondo $H_{V,k}$, o histograma descritor do vídeo V , como indicado na Equação 38.

$$H_{V,k}(i) = \sum_{i=1}^n H_{F,k}(i) \quad (38)$$

O descritor final de V se dá pela concatenação dos histogramas dos vídeos e suas estatísticas St concatenadas, da forma apresentada na Equação 39.

$$\psi_V = \{H_{V,1}, St_{H_{V,1}}, H_{V,2}, St_{H_{V,2}}, \dots, H_{V,k}, St_{H_{V,k}}\} \quad (39)$$

No próximo capítulo validamos esta metodologia aplicada à classificação de vídeos. Mostraremos assim como a proposta é adequada para reconhecimento de vídeos de texturas dinâmicas similares.

3.4.3 Complexidade computacional

O primeiro passo de nossa proposta é o cálculo do fluxo óptico. Apresentamos o algoritmo Lucas-Kanade no capítulo anterior, concluindo que, para um par de imagens de resolução $N \times N$, sua complexidade é $O(n^2N^2 + n^3)$. Como n em geral é bem menor

que N^2 , podemos eliminá-lo da notação como um termo de menor ordem, resultando em $O(N^2)$.

A complexidade do método de extração de padrões singulares depende de diversos parâmetros. Dado um campo vetorial F de tamanho $N \times N$ e um dicionário de n padrões de campo base, cada um de tamanho $B \times B$, o campo vetorial é projetado sobre os n padrões base por correlação cruzada, como uma "janela deslizante". Isso é repetido para cada oitava (análise multi-escala). Outro fator relevante é a ordem da aproximação, que sumarizamos no valor $k \in K$. Assim, k possui $|K|$ valores possíveis. Para cada valor de k um polinômio é somado à aproximação. Assim, podemos dizer que a complexidade computacional do algoritmo é proporcional a $O(nN^2B^2|K|)\log_2N^2$. O logaritmo aparece aplicado ao tamanho da imagem, N^2 , devido à consideração de que cada oitava é uma sub-amostragem da oitava anterior que diminui de tamanho de modo quadrático.

Dado que os valores n , N_0 e $|K|$ são via de regra significativamente menores que o tamanho N^2 do campo vetorial (os experimentos de Liu e Ribeiro (LIU; RIBEIRO, 2012) usam $n = 18$, $N_0 = 4$ e $\Phi = [-4, 4]$, ou seja, $|K| = 9$), e com o pressuposto seguro de que o B^2 também é em geral muito menor que N^2 , podemos afirmar que a complexidade do algoritmo depende principalmente do tamanho do fluxo original e dos fluxos base e é da ordem $O(N^2)$.

Nossa proposta envolve a construção de um histograma de padrões singulares, o que significa que a complexidade computacional depende do número de padrões adquiridos a partir de um campo vetorial. O número de padrões é difícil de estimar e está intrinsecamente ligado às características do campo vetorial. Considerando um vídeo V de tamanho $N \times N \times N$ do qual são extraídos P padrões singulares, temos que, em nossa proposta baseada no *bag-of-features* tradicional, cada um dos P padrões deve ser agrupado em um dos M *clusters*, o que depende de uma comparação dos padrões com cada centroide. Cada padrão é um vetor de coeficientes de $k \in K$ dimensões, e portanto o agrupamento tem complexidade $O(PKM)$. No caso de nossa proposta de histograma de valores de a_k , a construção do histograma envolve o percurso dos P padrões e a distribuição dos valores dos coeficientes a_k em histogramas de N *bins*. Portanto, temos que esse passo do algoritmo tem complexidade $O(PKN)$. Generalizando para os dois casos, se chamarmos de n o número de *bins* do histograma, ambos os casos apresentam complexidade $O(PKn)$.

Para um vídeo V de $N \times N \times N$, serão gerados N campos vetoriais $N \times N$. Combinando a complexidade dos 3 algoritmos, temos que a complexidade do método proposto é da ordem de $O(N^2 + N^2N + Mn)$. Considerando que o número de padrões dicionário n e o número de agrupamentos M são valores relativamente pequenos, como veremos no próximo capítulo, a complexidade do algoritmo todo pode ser considerada $O(N^3)$. Cabe o adendo de que essa simplificação apresenta apenas a variável de maior ordem, e ao longo da análise excluímos coeficientes e termos de menor ordem que, enquanto não afetam a ordem de complexidade do método, contribuem com o aumento do tempo de execução do

algoritmo.

CAPÍTULO 4

Experimentos e Análise dos Resultados

4.1 Bases de dados de textura

Para validação do método de caminhadas proposto para texturas estáticas, assinaturas foram computadas a partir de imagens e aplicadas à classificação de textura. Optamos por bases de dados difundidas e popularmente usadas em outros métodos na literatura, para fins de comparação. A Figura 22 traz um exemplo de amostras da base de dados Brodatz (BRODATZ, 1966), uma das bases de texturas usadas em conjunto com o método relativo a texturas estáticas, enquanto a Figura 23 traz um exemplo de amostras da base de dados Dyntex, a base de texturas dinâmicas que utilizamos com a proposta de caminhadas tridimensionais sobre vídeos e com a proposta de padrões singulares de fluxos ópticos.

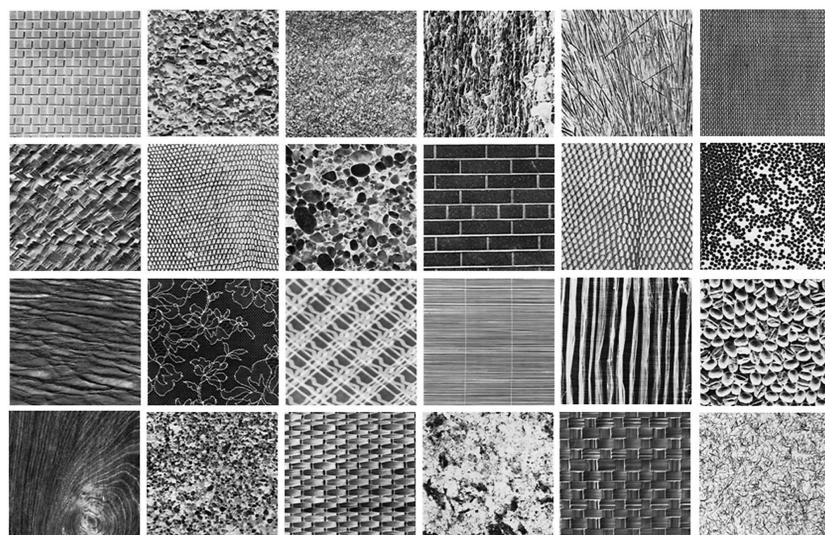


Figura 22 – Exemplos de imagens da base de texturas Brodatz.

A Tabela 1 mostrada a seguir lista as bases de dados usadas na validação da metodologia proposta. As seis bases de dados de texturas estáticas usadas foram: o álbum de texturas Brodatz (BRODATZ, 1966), com 1776 imagens de 111 classes diferentes, a base de texturas coloridas VisTex (VISION..., 2009), com 864 imagens em 54 classes, a suíte



Figura 23 – Exemplos de quadros extraídos de vídeos da base de texturas dinâmicas Dyntex.

Outex_TC_00013 (OJALA et al., 2002), com 1360 imagens em 68 classes, a UIUC Texture Database (UIUCTex) (LAZEBNIK; SCHMID; PONCE, 2005), com 1000 imagens em 25 classes, a base de dados de Reflectância e Textura Columbia-Utrecht (CURET) (DANA et al., 1999), com 5612 imagens em 61 classes e a base de dados de Texturas sob Variações de Iluminação e Escala (KTH-TIPS) (HAYMAN; CAPUTO; EKLUNDH, 2004), com 810 imagens em 10 classes. A base de texturas dinâmicas Dynamic Texture Database (Dyntex) (PÉTERI; FAZEKAS; HUISKES, 2010), por sua vez, possui 679 amostras de vídeo (cada amostra possui pelo menos 250 quadros) distribuídas em 36 classes. Cabe notar que, no caso da Dyntex, o número de classes pode variar pois não há um agrupamento padrão das amostras.

Tabela 1 – Sumário das bases de dados usadas na validação dos métodos da proposta.

Bases de Imagens	Amostras	Classes	Tipo
Brodatz (BRODATZ, 1966)	1776	111	Estática
Vistex (VISION..., 2009)	864	54	Estática
Outex (OJALA et al., 2002)	1360	68	Estática
UIUCTex (LAZEBNIK; SCHMID; PONCE, 2005)	1000	25	Estática
KTH_TIPS (HAYMAN; CAPUTO; EKLUNDH, 2004)	810	10	Estática
CURET (DANA et al., 1999)	5612	61	Estática
Dyntex (PÉTERI; FAZEKAS; HUISKES, 2010)	679	36	Dinâmica

4.2 Análise da Rede Complexa

Nesta seção faremos uma breve análise das características topológicas da rede complexa usada neste trabalho. Há modelos clássicos propostos para descrever tipos de redes recorrentes (COSTA et al., 2011), então é útil identificar a rede da aplicação proposta e verificar se é possível classificá-la como pertencente a um dos modelos da literatura. Um dado importante de uma rede complexa é a conectividade de seus vértices, visualizável através de sua distribuição de graus. Ela é de especial importância neste trabalho pois usamos a matriz de graus médios da rede como objeto das caminhadas. Para esta análise, foram selecionadas 8 amostras, cada uma obtida de uma classe de texturas com propriedades altamente distintas da base Brodatz, e um limiar $t = 0.1$, escolhido para fins de melhor visualização. A Figura 24 mostra o histograma de graus para cada uma dessas classes, nas escalas linear e logarítmica.

A versão logarítmica da distribuição é apresentada para verificarmos se a conectividade da rede usada neste trabalho e em Backes *et al.* (BACKES; CASANOVA; BRUNO, 2013) segue a lei de potência, isto é, se o histograma de distribuição de graus segue uma regra do tipo $y = ax^k$, onde x representa cada grau, y representa a frequência de dado grau, e a e k são constantes (GUERRIERO, 2012). Redes cuja distribuição de grau seguem a lei de potência são chamadas de redes livres de escala (*scale-free networks*), e são bastante frequentes em sistemas naturais e fractais (SONG; HAVLIN; MAKSE, 2005) e apresentam propriedades interessantes (BARABÁSI; ALBERT, 1999). Na Figura 24, é possível observar que este não é o caso das redes geradas a partir da base Brodatz; se fosse, a distribuição de graus $\log \times \log$ aproximaria uma reta ($y = kx + a$). O que podemos observar, ao invés disso, é que a distribuição linear apresenta grande variação em algumas classes como na classe 001, mas em outras classes se aproxima de uma distribuição Gaussiana. A distribuição Gaussiana não é frequente em redes complexas (BULLMORE; SPORNS, 2009) e a redução da ocorrência de graus de valor alto ocorre notavelmente em redes como as do modelo de Pequeno Mundo, onde as vértices possuem probabilidades de estarem ligados por arestas dadas por sua proximidade. Concluímos que esta rede, quando aplicada a texturas, sendo um grafo geográfico ou espacial, pode ser categorizada como uma rede de Pequeno Mundo.

O comportamento Gaussiano da distribuição de graus é independente do histograma de intensidades da imagem original. Isto é notável quando analisamos os histogramas de cada uma das 8 imagens, que são perceptivelmente bastante distintos entre si, como mostra a Figura 25. Concluímos ainda que, para nosso método, a organização espacial dos graus é mais importante que a distribuição para a discriminação das texturas, já que esta é similar para determinadas classes.

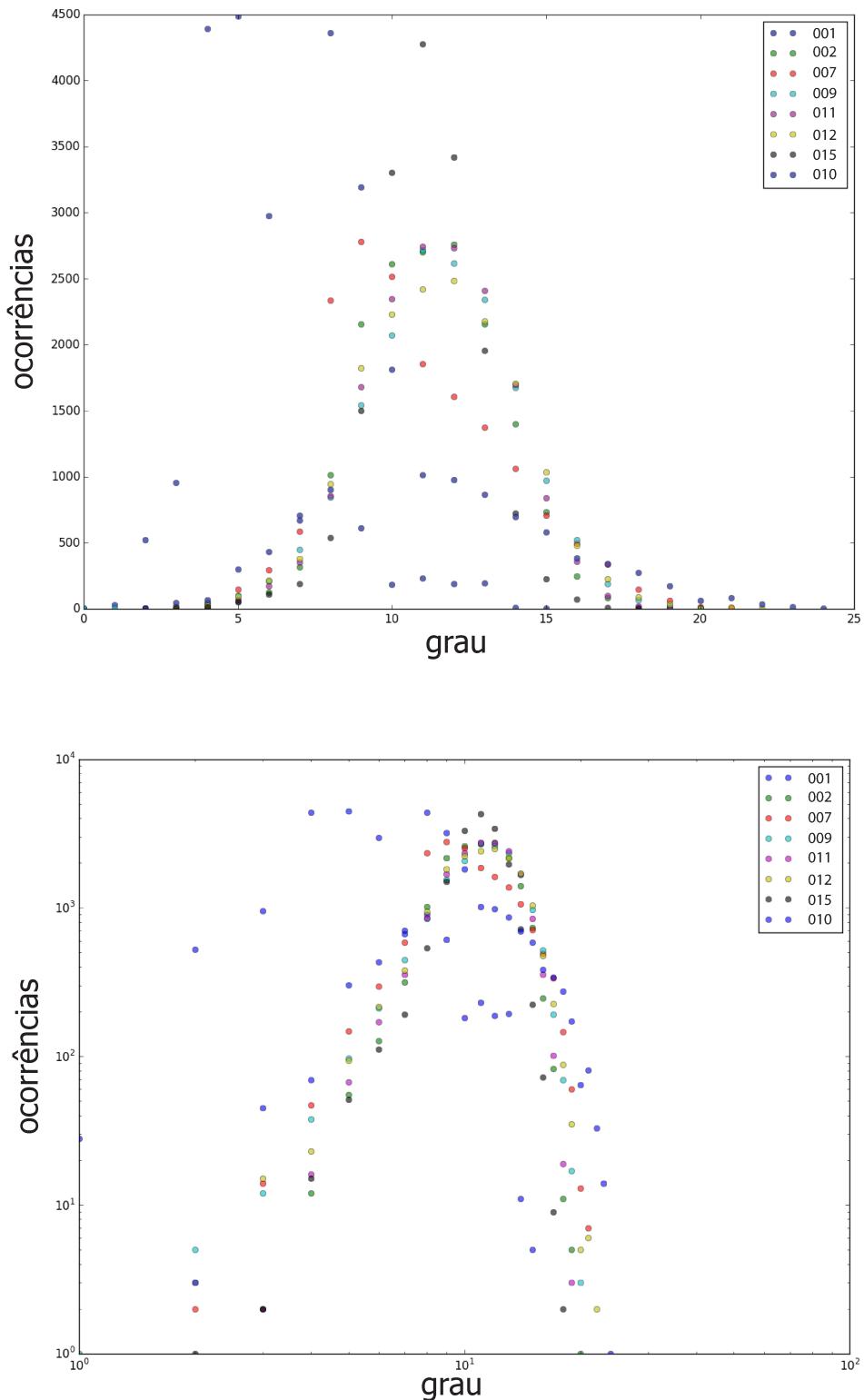


Figura 24 – Histograma linear e $\log \times \log$ de 8 amostras de classes da base Brodatz. A análise mostra que a distribuição não segue uma lei de potência, mas em alguns casos aproxima uma Gaussiana.

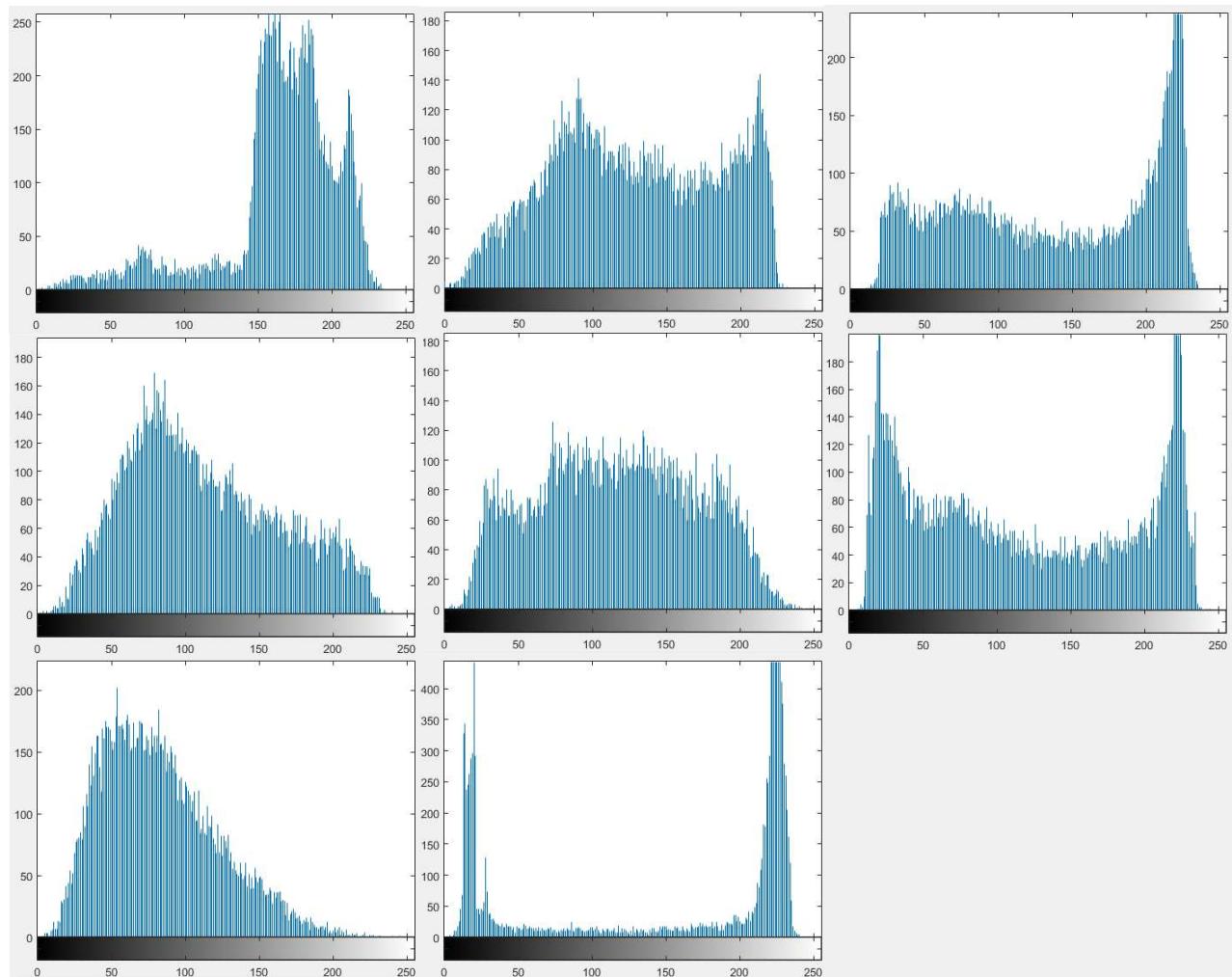


Figura 25 – Histogramas de ocorrências de valores de intensidade de pixels das amostras da Brodatz usadas na análise de distribuição de graus.

4.3 Experimentos do método das caminhadas estáticas

4.3.1 Parâmetros dos experimentos

Os experimentos deste capítulo foram implementados no ambiente MATLAB 2015b, com as partes do código mais dispendiosas em termos de tempo de processamento sendo implementadas em C/C++ compilados como funções MEX para interface com o código MATLAB.

O mesmo conjunto de parâmetros foram usados para todos os experimentos a seguir, relativos à proposta de método para caminhadas em texturas estáticas usando redes complexas. Quando aplicável, o raio da rede complexa foi estabelecido como $r_G = 3$, o que implica que o grau máximo para cada nó $v_{x,y}$ é de 28. Além disso o intervalo S para limiares t usado é $[0.005, 0.020..., 0.530]$, com incrementos de 0.015, o que significa um total de 36 grafos D_t cuja média gera um grafo D para cada imagem. Incrementos menores aumentam o número de limiares a serem avaliados sem melhorias significativas nas redes, o que sugere que o limiar está pequeno, e o intervalos maiores geram redes efetivamente menos discriminantes (BACKES; CASANOVA; BRUNO, 2013), possivelmente por causa de *overfitting*, ou seja, alto viés no dilema viés-variância do classificador ao conjunto de treinamento, causado por dimensionalidade desnecessariamente alta (DUDA; HART; STORK, 2012). Para as caminhadas, o raio de vizinhança do pixel foi estabelecido como $r_D = \sqrt{2}$, totalizando 8 vizinhos, valor que ofereceu bons resultados e mantém viável o número de dimensões no vetor de características, já que o tamanho do vetor de características é proporcional ao número de vizinhos possíveis. Para melhores resultados, é importante combinar diversos valores de memória, mas valores de memória altos aumentam o custo computacional sem melhorar as taxas de classificação correta (GONÇALVES; BRUNO, 2013a) (BACKES et al., 2010). O intervalo que apresentou memórias com resultados mais descriptivos foi o intervalo $[1, 8]$. Considerando o intervalo de $[1, 8]$, grupos de $\Theta = 4$ memórias foram usadas para construir o vetor de características, suficientes para permitir a comparação de dois intervalos de tamanhos de memória, $[1, 4]$ e $[5, 8]$.

O classificador usado em todos os testes foi o *Linear Discriminant Analysis* (LDA) (DUDA; HART; STORK, 2012) (FRIEDMAN; HASTIE; TIBSHIRANI, 2001) (RIPLEY, 2007), uma técnica que inclui um importante passo de redução de dimensionalidade, usando um esquema de validação cruzada *leave-one-out* para avaliação individual de cada imagem das bases. O método *leave-one-out* baseia-se no treinamento do classificador com toda a base, com exceção de uma das observações, que será classificada pelo modelo treinado, assim validando o classificador. Esse processo é repetido para cada observação da base, cada vez treinando o classificador com toda a base de dados menos uma das instâncias observadas. A exclusão da observação a ser classificada do processo de treina-

mento é importante para que não haja viés do classificador em relação aos dados a serem classificados. Todas as bases usadas tem o mesmo número de amostras por classe.

Para ambas as dinâmicas de caminhadas existem 8 tamanhos possíveis de memória μ , cada um com 8 *bins* (um *bin* para cada vizinho possível), num total de 128 dimensões. Considerando que St consiste de 7 valores computados para cada dinâmica, o vetor de características completo possui um total de 142 dimensões, sendo $2 \times 8 \times 8$ das caminhadas e 2×7 das estatísticas dos histogramas. Em testes em que só metade dos tamanhos de memória foram usados, o vetor de características totaliza 78 dimensões ($2 \times 4 \times 8 + 2 \times 7$), e quando apenas uma das dinâmicas é usada o vetor totaliza 71 dimensões ($8 \times 8 + 7$).

Por fim, neste capítulo, sempre que nos referirmos à taxa de acertos, estamos nos referindo à acurácia Acc do método em termos percentuais, dada pela fórmula 40, em que t é o número de classificações corretas e n é o total de classificações, a soma das classificações corretas com as incorretas.

$$Acc = 100 \frac{t}{n} \quad (40)$$

4.3.2 Resultados e discussão

Este primeiro experimento avalia a importância de ordenar o vetor de características para levar em conta variações como rotação dos padrões da textura. Foram ordenados todos os sub-vetores correspondentes a cada valor de memória μ e cada dinâmica dyn diferentes. A Tabela 2 mostra os resultados de classificação com e sem a ordenação do vetor. Com esses resultados, é possível evidenciar que as taxas de classificação são melhores quando o vetor de características ordenado é acrescido ao vetor original, para todos os intervalos de memória testados. Porém, é importante perceber que a melhora é maior quando os descritores de diferentes dinâmicas de caminhada não são combinados. Dinâmicas $\{min\}$ e $\{max\}$ exploram a imagem de forma diferente, mas ambas estão sujeitas à ordem de visitação dos vizinhos. Assim, pequenas variações na imagem podem afetar a vizinhança de um pixel, causando diferenças nas posições dos descritores propostos, o que é compensado pela ordenação.

Tabela 2 – Comparação de taxas de classificação na base Brodatz com e sem ordenação do vetor de características. Colunas assinaladas com * indicam resultados obtidos para vetores ordenados.

Dinâmica	Taxa de acertos (%)					
	[1,8]	[1,4]	[5,8]	[1,8]*	[1,4]*	[5,8]*
min	93,29	87,33	68,97	96,05	93,91	95,21
max	93,52	82,93	82,15	96,67	94,76	95,35
min \cup max	98,36	94,20	92,28	98,36	96,62	97,18

Também avaliamos o efeito de usar partes do vetor de características como descritores, avaliando o desempenho individual e agrupado das partes para as duas dinâmicas, e diferentes valores e intervalos de memória. É possível observar na Figura 26 que as memórias isoladas apresentaram resultados razoáveis de classificação na base Brodatz, dado o reduzido número de dimensões usado. Também cabe apontar que valores menores de memória obtiveram, em geral, melhor taxa de acerto, enquanto em relação ao uso das dinâmicas, a dinâmica *max* obteve resultados melhores em todos os casos exceto quando $\mu = 3$. O uso de ambas as dinâmicas superou o uso das dinâmicas isoladas de forma clara.

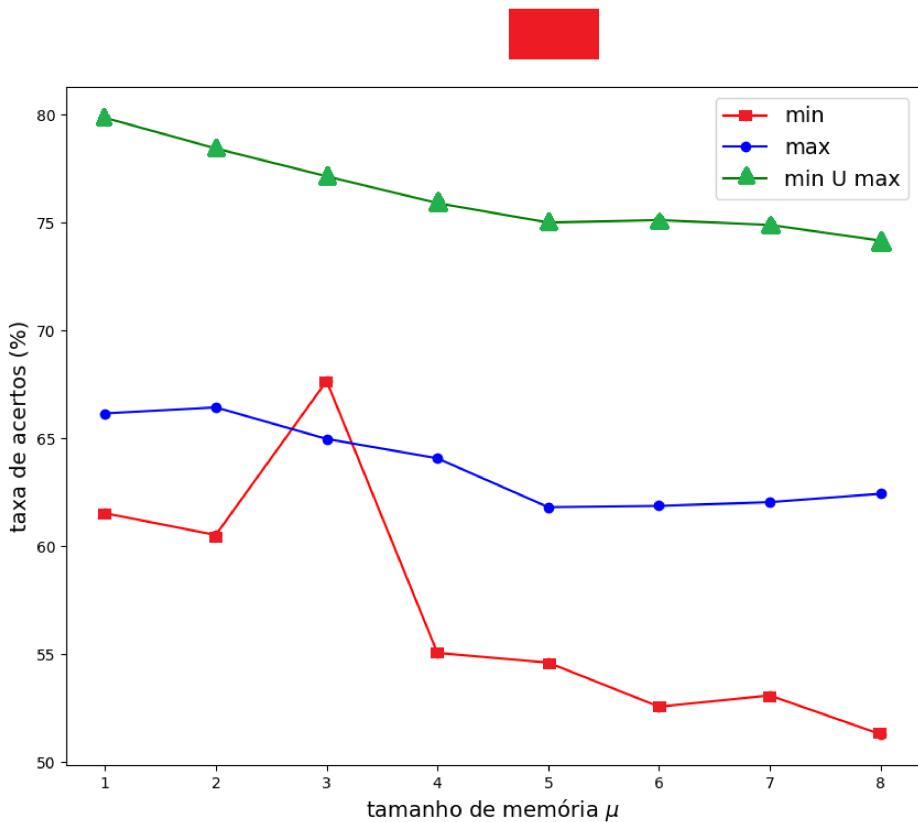


Figura 26 – Taxas de acerto para valores de memória μ isolados na base Brodatz, considerando as 3 combinações de dinâmicas.

A Tabela 3 apresenta os resultados da classificação adquiridos sobre a base de texturas Brodatz para conjuntos de valores de memória, comparando taxas de classificação para histogramas de direções de caminhada ordenados, nos três intervalos de memória listados, e em três combinações de dinâmicas: apenas a dinâmica $\{max\}$, apenas dinâmica $\{min\}$ e dinâmicas $\{min\} \cup \{max\}$ concatenadas. Além disso, também verificamos o efeito de incluir ou não ou vetor de características as estatísticas *St* dos histogramas.

É possível perceber que o agrupamento de ambas as dinâmicas ($\{min\} \cup \{max\}$) resultou em taxas superiores de acerto, o que era esperado dado o potencial de cada dinâmica

Tabela 3 – Resultados de classificação na base Brodatz para dinâmicas isoladas e combinadas, com e sem as estatísticas St , e apenas com as estatísticas St .

Dinâmica	Taxa de acertos (%)						
	[1,8]	[1,4]	[5,8]	[1,8]+ St	[1,4]+ St	[5,8]+ St	St
min	96,05	93,91	95,21	96,50	94,53	95,60	54,28
max	96,67	94,76	95,35	97,18	95,94	95,32	60,59
minUmax	98,36	96,62	97,18	98,36	97,35	97,97	77,25

de evidenciar elementos distintos das texturas (propriedade descrita na análise da Seção 3.2). A inclusão das estatísticas St por sua vez, também melhorou os resultados enquanto acrescentou um número relativamente pequeno de dimensões ao vetor de características (14 no total, como citado anteriormente). Até mesmo quando usadas isoladamente, as estatísticas St produziram bons resultados, se considerarmos sua pequena dimensionalidade. Portanto, todos os experimentos de classificação de texturas comparativos com o estado-da-arte no restante do capítulo foram realizados usando ambas as dinâmicas e a inclusão das estatísticas St aos descritores.

Em relação às memórias, os melhores resultados são provenientes do conjunto completo de memórias $\mu \in [1, 8]$, embora os resultados para os intervalos de memória menores tenham sido próximos aos do intervalo completo.

As classificações incorretas geralmente ocorrem em imagens visualmente muito similares. Isso é o caso especialmente em certas classes que apresentam variação significativa entre as amostras da mesma classe. A Figura 27 apresenta um exemplo de matriz de confusão para a base Brodatz que destaca dois pares de classes frequentemente confundidas. As amostras das classes onde houve maior confusão do classificador podem ser verificadas como visualmente semelhantes.

Os experimentos na Tabela 4 apresentam os resultados do método para todos os conjuntos de dados usados, apresentando ainda os resultados para os três intervalos de memória usados. A tendência sugerida pelos resultados da base Brodatz se mantém, com os melhores resultados sendo obtidos para o intervalo de memória $\mu \in [1, 8]$ em todas as bases de texturas exceto a KTH_TIPS. Os resultados para o intervalo [1, 4] tiveram melhor desempenho que o intervalo [5, 8] em todas as bases com exceção das texturas da base Outex. Mantém-se ainda a tendência de que os intervalos menores apresentam resultados inferiores, porém muito próximos aos do intervalo completo. É preciso levar em consideração que os vetores de características dos intervalos de memória menores apresentam metade da dimensão do intervalo completo. Assim, optamos por usar o intervalo [1, 4] para as próximas comparações do método com métodos do estado-da-arte.

Estes resultados concedem versatilidade à nossa proposta: tamanhos de memória menores são significativamente mais rápidos de computar, devido às menores restrições que eles impõem à formação de atratores, o que resulta em caminhadas com menor número de

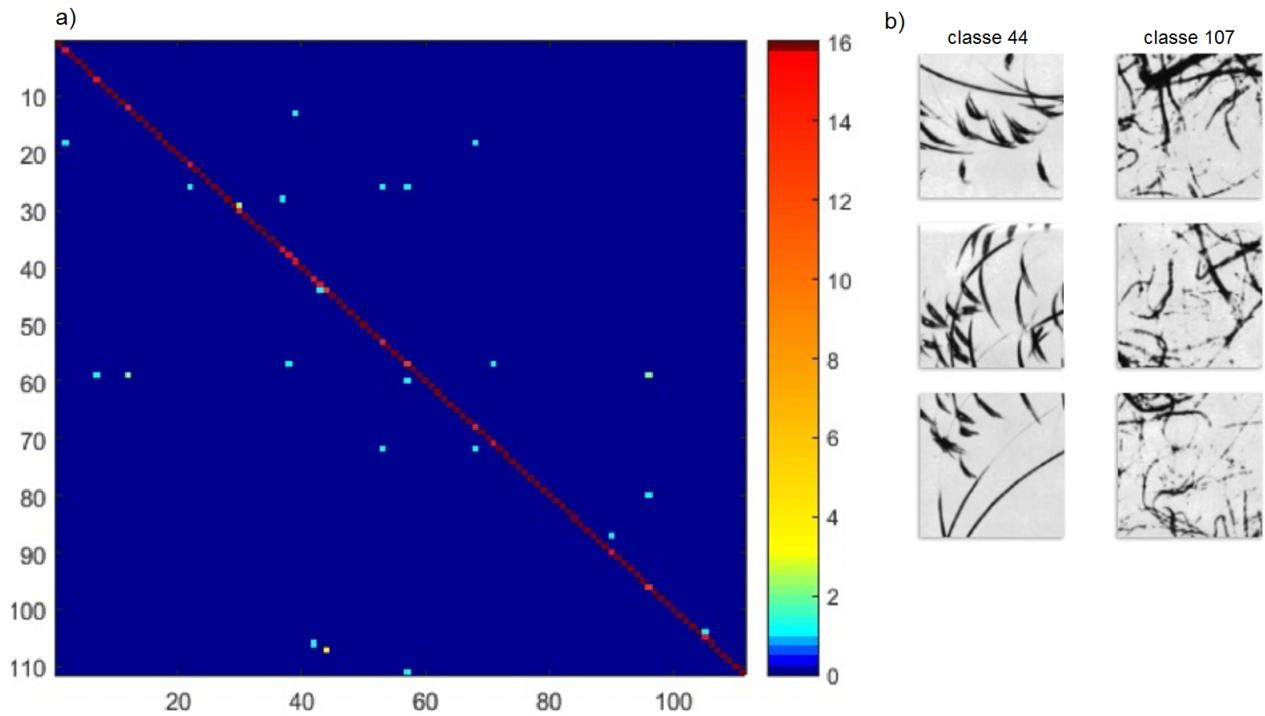


Figura 27 – Exemplo de matriz de confusão para caminhadas na base Brodatz. Ao lado dela há amostras dos pares de classes que causaram maior confusão ao classificador, as classes 44 e 107 (4 classificações incorretas).

Tabela 4 – Comparação das taxas de classificação para os três diferentes conjuntos de tamanhos de memória. Melhores resultados em negrito.

Bases de texturas	Taxa de acertos (%)		
	[1,8]	[1,4]	[5,8]
Brodatz	98,36	98,25	97,97
Vistex	98,03	97,80	96,87
Outex	97,79	97,28	97,43
UIUCTex	98,30	98,10	97,10
KTH_TIPS	99,51	99,88	99,14
CUReT	98,81	98,45	97,72

passos. Se uma computação mais eficiente é desejável, é possível usar intervalos menores, enquanto que se houver necessidade de alta precisão na classificação, o intervalo completo oferece melhorias marginais mas consistentes na taxa de classificação.

A Figura 28 mostra a relação entre os tamanhos de memória usados e os tamanhos médios de caminhadas, medidos em números de passos, quando as caminhadas são aplicadas sobre a base de dados Brodatz. Os tamanhos de caminhadas foram verificados para uma imagem completa de cada uma das 111 classes da base, com a média e o desvio padrão apresentados na imagem. Ainda que haja alta variação nos números de passos para valores de memória altos, como demonstram as barras verticais de desvio padrão, é possível observar o crescimento do tamanho médio de caminhada $|W|$ conforme o crescimento do

tamanho de memória μ . Os valores altos de desvio padrão são um comportamento esperado, considerando que os tamanhos de caminhadas são um fator discriminante poderoso entre caminhadas sobre diferentes imagens. É também possível observar que a relação entre μ e $|W|$ tem comportamento aproximadamente linear. A regressão linear resulta na fórmula $|W| = 2,93\mu + 8,52$, e o coeficiente de correlação linear entre as variáveis é $r = 0,995$, o que denota uma correlação positiva alta. Em contraste com os valores de memória, a dinâmica de caminhada escolhida (*min* ou *max*) não causa diferenças estatisticamente significativas no número de passos das caminhadas.

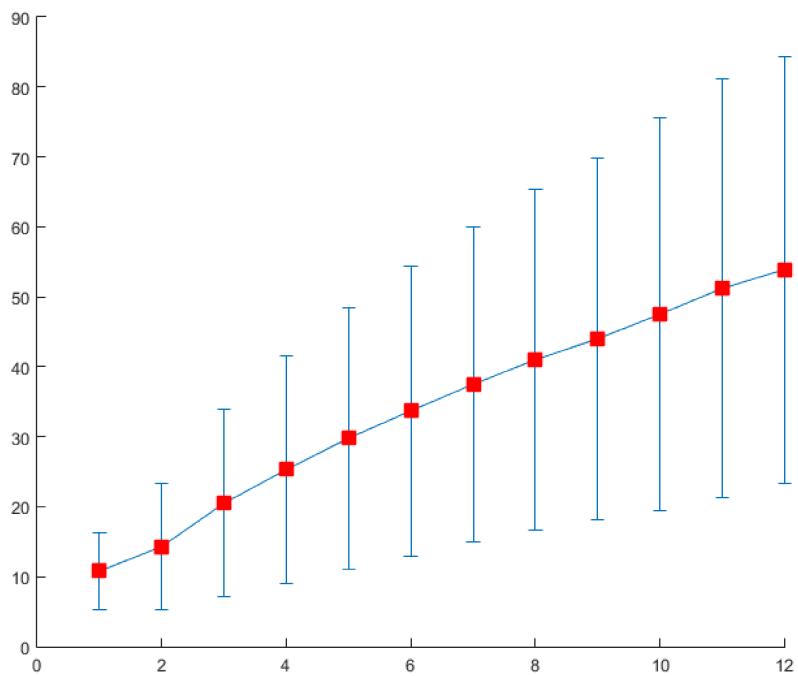


Figura 28 – Tamanhos médios de caminhadas para cada tamanho de memória μ , na base Brodatz. As barras verticais indicam o desvio padrão.

Uma comparação da abordagem proposta com métodos tradicionais da literatura para análise de texturas é apresentada na Tabela 5. A tabela também exibe o tamanho do vetor de características de cada método, que é um atributo importante. Diversos resultados apresentados nesta tabela foram obtidos com a utilização do método LDA como classificador e os mesmos parâmetros. Nos outros casos, extraímos do artigo citado o melhor resultado obtido. A abordagem proposta neste trabalho obteve o melhor resultado de classificação para a base de texturas Brodatz (98,25% e 97,28%, respectivamente), enquanto obteve o segundo melhor resultado para a base Outex, inferior apenas ao MRELBP, a versão mais eficaz do LBP segundo a exaustiva comparação de Liu *et al.* (LIU et al., 2016a). Cabe apontar que o MRELBP possui descritores com dimensão substancialmente maior que o método proposto. No caso da base Vistex o método proposto obteve 97,80%, que é inferior a 5 das abordagens comparadas, mas tem taxa de acerto apenas 1,85% menor que a do melhor resultado (GLDM, com 99,65%).

Tabela 5 – Comparação do método proposto com outros métodos da literatura. Métodos das 8 primeiras linhas foram executados com a mesma implementação do LDA. Melhores resultados em negrito.

Método	Dimensão	Taxa de acertos (%)		
		Brodatz	Outex	Vistex
Filtros 1 ^a Ordem (MATERKA; STRZELECKI et al., 1998)	5	34,29	52,86	50,11
Filtros de Gabor (IDRISSA; ACHEROY, 2002)	64	84,85	80,00	91,66
DCT (NG; TAN; KITTLER, 1992)	8	82,20	75,22	83,91
Lacunaridade (DONG, 2000)	16	69,14	64,33	78,24
Fourier Radial (AZENCOTT; WANG; YOUNES, 1997)	63	87,66	93,45	91,89
GLDM (KIM; PARK, 1999)	60	97,29	92,20	99,65
Matrizes de co-ocorrência (HARALICK, 1979)	16	80,29	80,95	82,40
Redes Complexas (BACKES; CASANOVA; BRUNO, 2013)	108	95,27	86,76	98,03
Wavelets (CHANG; KUO, 1993)	36	85,64	78,45	89,69
Florindo et al. (FLORINDO; LANDINI; BRUNO, 2016)	242/142	92,74	-	95,60
Casanova et al. (CASANOVA et al., 2016)	105/78	-	95,00	99,31
CLBP (GUO; ZHANG; ZHANG, 2010)	648	95,32	85,80	98,03
LBP ($r = \{1, 2, 3\}$) (PIETIKÄINEN; ZHAO, 2015)	54	97,57	92,64	99,18
MRELBP (LIU et al., 2016b)	800	-	99,82	-
Método proposto	78	98,25	97,28	97,80

Na Tabela 6, comparamos os resultados do método proposto em outras bases de dados com outros métodos de alta eficácia de classificação da literatura recente que utilizam essas mesmas bases. Estes métodos usam outros classificadores que não o LDA, como K-NN ou Redes Neurais. Para essa comparação usamos os resultados de nosso método já previamente apresentados na Tabela 4. O método proposto tem melhor desempenho de classificação que os outros métodos publicados para os conjuntos de dados mostrados. A comparação também corrobora a observação de Hossain *et al.* de que a base UIUCTex é a mais desafiadora das bases de texturas populares devido à sua alta variação interna em cada classe (HOSSAIN; SERIKAWA, 2013), já que a base apresentou regularmente os piores resultados para a maioria dos métodos.

Tabela 6 – Comparação do método proposto com métodos da literatura sobre as bases KTH_TIPS, UIUCTex e CUReT (melhores resultados em negrito).

Método	Taxa de acertos (%)		
	KTH_TIPS	UIUCTex	CUReT
Zhang et. al (ZHANG et al., 2015)	98,86 \pm 1,12	-	94,44 \pm 1,13
Zhao et. al (ZHAO et al., 2013)	-	93,31	96,27
Perea & Carlsson (PEREA; CARLSSON, 2014)	94,77 \pm 1,3	91,23 \pm 1,13	95,66 \pm 0,45
Barley et. al (BARLEY; TOWN, 2014)	96,00	90,60	-
Lasmar et. al (LASMAR; BAUSSARD; CHENADEC, 2015)	-	81,70	-
Serra et. al (SERRA; GRANA; CUCCHIARA, 2014)	98,62	-	-
Backes et. al (BACKES; CASANOVA; BRUNO, 2013)	90,37	86,50	84,32
Goncalves et. al (GONÇALVES et al., 2016)	94,63	-	-
Ribas et. al (RIBAS et al., 2015)	-	90,90	-
Método Proposto	99,88	98,10	98,45

Por fim, apresentamos na Tabela 7 os resultados experimentais obtidos pelo método

de caminhadas sobre a base de texturas Brodatz usando um raio r_D maior, $r_D = 2\sqrt{2}$. Nesse caso, o total de vizinhos possíveis será de 24 em vez de 8 com o raio $r_D = \sqrt{2}$. Nesse caso, cada histograma de direções de passos de caminhadas que compõe o vetor de características triplica de tamanho. É possível observar que, mesmo com o número bem maior de descritores no vetor de características, as taxas de acerto do método são comparáveis às obtidas quando um raio de tamanho menor é usado.

Tabela 7 – Resultados para caminhadas sobre a base de dados Brodatz com raio $r_D = 2\sqrt{2}$. Melhores resultados em negrito

Dinâmicas	Taxa de Acertos (%)					
	[1,8]	[1,4]	[5,8]	[1,8]+St	[1,4]+St	[5,8]+St
min	94,70	91,89	89,80	95,32	93,01	92,11
max	92,39	88,00	86,48	92,39	88,96	88,96
min \cup max	98,64	95,83	95,77	98,19	96,90	96,45

A Tabela 8 evidencia essa comparação repetindo lado a lado alguns dos resultados das Tabelas 2 e 7. O uso de raio de maior valor gera melhores resultados quando só a dinâmica *min* é usada, mas é inferior nos outros dois casos. Portanto, essa comparação mostra como o aumento do raio não aperfeiçoa o poder descritivo do método, e na verdade pode até impactar negativamente as taxas de classificações corretas enquanto aumenta o tempo de processamento e a dimensionalidade do vetor, uma conclusão semelhante à obtida pelos autores do método LBP (OJALA; PIETIKAINEN; MAENPAA, 2002).

Tabela 8 – Resultados de classificação, na base Brodatz com o raio $r_D = 2\sqrt{2}$ e $r_D = \sqrt{2}$.

Raio	Taxa de acertos (%)		
	min	max	min \cup max
$r_D = 2\sqrt{2}$	95,32	92,39	98,19
$r_D = \sqrt{2}$	93,29	93,52	98,36

Como um último experimento relativo a texturas estáticas, validamos isoladamente o descritor composto do histograma de direção de passos proposto. Para avaliar o descritor de forma isolada, realizamos caminhadas na imagem original, sem utilizar neste caso a matriz de graus D , em que os valores dos *pixels* são os graus médios dos nós da rede complexa. Em vez disso, a caminhada foi aplicada considerando os valores de intensidade dos *pixels* da imagem original I . Dessa maneira, podemos comparar nossos resultados ao descritor baseado em tamanhos de caminhada da literatura (BACKES et al., 2010), que tem a mesma complexidade computacional e também é aplicado diretamente sobre os valores de intensidade da imagem. A Tabela 9 mostra os resultados de classificação para a base de texturas Brodatz, com o intervalo de memórias de [1, 8].

Os resultados demonstram o potencial classificatório do método de caminhadas proposto mesmo sem considerar a transformação da imagem. Ainda assim, uma comparação

Tabela 9 – Resultados para o descritor de histograma de direções de passos aplicados sobre imagens I sem transformação, para cada dinâmica (melhores resultados e negrito).

Datasets	Descritor de caminhadas original (%)			Método proposto (%)		
	min	max	min \cup max	min	max	min \cup max
Brodatz	70,94	87,04	93,24	88,45	93,13	97,46
Vistex	78,47	89,00	96,29	94,57	96,64	99,76
Outex	69,70	78,89	88,23	88,67	90,51	97,13

destes resultados com os resultados da Tabela 2 torna possível observar que a realização das caminhadas propostas sobre a transformada de graus médios baseada em redes complexas proposta D oferece resultados superiores à aplicação das caminhadas propostas sobre a imagem não transformada I . A transformação de fato melhora os resultados nos casos analisados. Isso evidencia a eficácia da transformação baseada em redes complexas na descrição da vizinhança de cada pixel. A Tabela 10 repete os resultados relevantes das Tabelas 2 e 9, exibindo os resultados com e sem o uso da rede complexa lado a lado para evidenciar essa comparação.

Tabela 10 – Resultados de classificação, na base Brodatz com intervalo de memória [1, 8], das caminhadas sobre a imagem original I e sobre a transformada D .

Métodos	Taxa de acertos (%)		
	min	max	min \cup max
Caminhadas sobre I	88,45	93,13	97,46
Caminhadas sobre D	93,29	93,52	98,36

4.4 Experimentos com caminhadas em texturas dinâmicas

A proposta de caminhadas aplicada a texturas dinâmicas foi testada na base de dados DynTex. Há diversos desafios relativos a esta base de dados, entre eles vídeos com movimento de câmera e vídeos periódicos ao longo do tempo mas com pouca ou nenhuma periodicidade espacial (HOSSAIN; SERIKAWA, 2013). Outro fator que representa um obstáculo expressivo é a maior carga de processamento trazida pelo número maior de *pixels* que um vídeo possui em relação a uma imagem, devido ao acréscimo da dimensão temporal. Para contornar estas dificuldades, selecionamos um sub-conjunto de vídeos da base. Separamos manualmente os vídeos em pares de vídeos da mesma classe, considerando apenas vídeos sem movimentação de câmera, resultando em uma base de 82 pares de vídeos (164 vídeos ao todo). Os pares de vídeos representam uma mesma cena, mas em momentos diferentes e em geral sob perspectivas diferentes.

Além disso, os vídeos foram amostrados em partes para realização dos experimentos, de forma que obtivemos 10 amostras para cada classe de vídeo (5 de cada vídeo), e cada amostra sendo composta de uma janela de *pixels* extraída do vídeo. Esta opção por incluir nos experimentos apenas vídeos com câmera estática e selecionar 10 amostras a partir de sub-conjuntos dos vídeos é a mesma abordagem de trabalhos seminais da área de classificação de texturas dinâmicas usando a base Dyntex (FAZEKAS; CHETVERIKOV, 2005) (ZHAO; PIETIKAINEN, 2007). É importante notar que não há intersecção entre as amostras de um mesmo vídeo, e as amostras foram extraídas do centro do vídeo em blocos de $20 \times 20 \times 20$ *pixels* do espaço (X, Y, T) , com diferentes valores de T .

Com a adição da terceira dimensão, no caso das caminhadas volumétricas, o valor do raio escolhido r apresenta uma relação de ordem n^3 com o número de dimensões do vetor de características. Assim, para manter uma dimensionalidade razoável no vetor de características, no caso das caminhadas volumétricas o raio escolhido para os experimentos foi $r = \sqrt{3}$. Assume-se que a distância d entre *pixels* de mesma coordenada $I(x, y)$ em imagens distintas é 1, isto é: $d(I_1(x, y), I_2(x, y)) = 1$, o que resulta em 26 vizinhos para cada pixel. Para as caminhadas TOP, os experimentos foram executados nos 3 planos e em cada um a vizinhança foi considerada dado o raio $r = \sqrt{2}$, o que resulta em 8 vizinhos para o pixel em cada plano, 24 ao todo. Optamos por usar este raio também porque foi o raio que obteve melhores resultados no LBP-TOP (ZHAO; PIETIKAINEN, 2007). Os demais parâmetros de tamanhos de memória μ , dinâmicas *dyn* e método de classificação (LDA com validação cruzada *leave-one-out*) foram os mesmos usados nos experimentos com texturas estáticas, exceto onde indicado no texto.

Uma análise do tamanho médio das caminhadas volumétricas em texturas dinâmicas indica que a relação de linearidade observada no caso das texturas estáticas se mantém verdadeira, com a regressão linear apresentando a fórmula $|W| = 7,22\mu + 4,09$, e um coeficiente positivo de linearidade igualmente alto de $r = 0,995$. A Figura 29 mostra os valores de tamanhos médios de caminhada para texturas dinâmicas da base Dyntex, para valores de μ de 1 até 12. É notável que as medições indicam que os tamanhos médios de caminhadas para o caso volumétrico são maiores: no caso estático a regressão linear da Figura 28 na forma $|W| = A\mu + B$ gerou $A = 2,93$, enquanto a regressão linear no caso dinâmico volumétrico gerou $A = 7,22$. O desvio padrão do caso volumétrico é ainda maior que no caso estático, o que denota grande variação nos tamanhos de caminhada.

Para avaliar o método de direções de caminhadas, aplicamos as caminhadas em todo o sub-conjunto de 164 vídeos da Dyntex estabelecido. Primeiramente, testamos a proposta de caminhadas volumétricas (uma única caminhada por pixel nas 3 dimensões (X, Y, T)). Avaliamos os resultados com e sem a inclusão dos valores das estatísticas de histograma *St* ao vetor de características. Avaliamos também os diferentes conjuntos de tamanhos de memória μ , nos intervalos $[1, 4]$, $[5, 8]$ e $[1, 8]$, e as dinâmicas $dyn = \{min, max\}$. A Tabela 11 mostra os resultados obtidos, com os melhores resultados para cada dinâmica

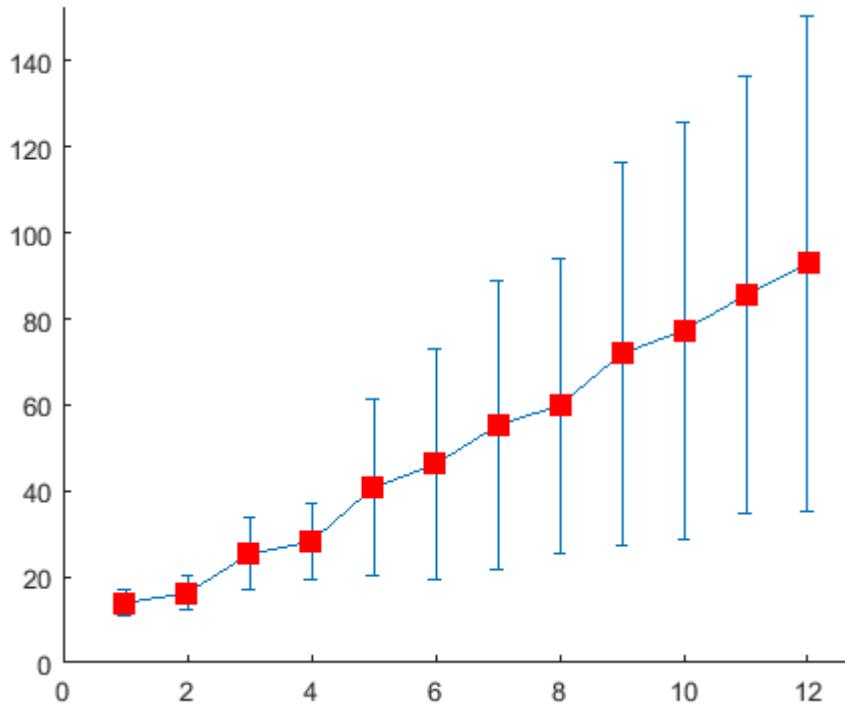


Figura 29 – Tamanhos médios de caminhadas volumétricas para diferentes tamanhos de memória μ , na base DynTex. As barras verticais indicam o desvio padrão.

destacados em negrito.

Tabela 11 – Resultados de classificação com caminhadas volumétricas (sobre as 3 dimensões (X, Y, T)), na base Dyntex, com e sem adição dos valores de estatísticas St ao descritor.

Dinâmicas	Taxa de acertos (%)						
	[1,8]	[1,4]	[5,8]	[1,8]+ <i>St</i>	[1,4]+ <i>St</i>	[5,8]+ <i>St</i>	<i>St</i>
min	62,80	66,95	58,54	71,59	69,76	62,68	35,49
max	63,17	61,71	60,49	69,76	65,00	66,23	29,76
min \cup max	76,59	79,63	77,80	80,00	82,20	80,73	57,20

As taxas de acerto inferiores às bases de texturas estáticas evidenciam o desafio de classificação na base Dyntex. Os resultados mostram que houve um equilíbrio entre as caminhadas de dinâmica *min* e *max*, enquanto experimentos com as dinâmicas concatenadas geraram nitidamente os melhores resultados. Entre os diferentes intervalos de memória, a vantagem de usar o intervalo completo de [1, 8] foi menos pronunciada; na verdade, no caso de dinâmicas *min* e *min* \cup *max* no caso em que não foram considerados os valores de *St* no vetor de características e no caso em que foram utilizadas ambas as dinâmicas com o vetor *St*, o intervalo [1, 4] gerou resultados melhores que o intervalo completo de memórias [1, 8], de forma que a decisão em relação a qual intervalo de memórias utilizar deve ser baseada em restrições de tempo de processamento e tamanho da base de dados. Este resultado corrobora as observações de Gonçalves *et al.* (GONÇALVES;

BRUNO, 2013b) para texturas dinâmicas de que parece haver um limite superior para a utilidade de tamanhos de memória maiores, especialmente considerando o alto custo computacional associado. A adição do vetor de estatísticas de histograma S_t melhorou os resultados em todos os casos. Ainda mais positivo é que o vetor S_t é computacionalmente rápido de calcular.

A análise dos vetores de características gerados pelo método de caminhadas volumétricas elicitava uma comparação entre as dimensões e as dinâmicas: no caso da base DynTex, a pouca variação entre um quadro e outro no caso de diversos vídeos provoca uma forte tendência de que o passo de distância mínima a partir de um pixel $V(x, y, t)$ ser para um quadro adjacente, nas coordenadas $V(x, y, t+1)$ ou $V(x, y, t-1)$. Em outras palavras, há uma proeminência de passos na direção da dimensão T em relação às dimensões X e Y . Isso ocorre porque a dinâmica *min* dá preferência por regiões homogêneas da imagem. Se há pouca ou nenhuma variação temporal em uma região do vídeo, as caminhadas tendem a ter este comportamento. A pouca variação entre quadros é um fenômeno presente em diversas das texturas da DynTex, já que em muitos casos o movimento é sutil ou ocupa apenas parte do quadro. Esse fenômeno já não ocorre na dinâmica *max*, que valoriza as regiões onde realmente ocorrem mudanças significativas no padrão de intensidade de *pixels* no espaço ou tempo. A Figura 30 mostra como os índices $V(x, y, t+1)$ e $V(x, y, t-1)$ apresentaram valores maiores no caso *min*, para texturas diversas, enquanto há maior equilíbrio do caso da dinâmica *max*. Curiosamente, a acurácia da proposta usando a dinâmica *min* não parece significativamente afetada por esta característica.

Os próximos resultados são relativos à aplicação do método de caminhadas volumétricas sobre a transformação do vídeo baseada em redes complexas. Para este caso, o vetor de características completo possui 430 dimensões, considerando 26 vizinhos por pixel, em 8 tamanhos de memória, 2 dinâmicas executadas e um vetor S_t para cada dinâmica ($|F_{Volume}| = (26 \times 8 + 7) \times 2$).

A Tabela 12 apresenta os resultados dos experimentos, com os melhores resultados destacados em negrito.

Tabela 12 – Resultados de classificação com caminhadas volumétricas (sobre as 3 dimensões (X, Y, T)), na base DynTex, com e sem adição do vetor de estatísticas S_t , sobre a transformação baseada em rede complexa da imagem.

Dinâmicas	Taxa de acertos (%)						
	[1,8]	[1,4]	[5,8]	[1,8]+ S_t	[1,4]+ S_t	[5,8]+ S_t	S_t
min	49,15	49,27	55,49	61,71	56,10	58,90	22,07
max	52,20	47,56	45,73	55,12	52,20	48,17	19,88
min \cup max	65,85	64,88	66,22	66,71	66,71	69,39	33,66

A observação mais evidente na Tabela 12 talvez seja que os resultados são, sem exceção, inferiores em relação aos do caso sem rede complexa. Este resultado contrasta com que

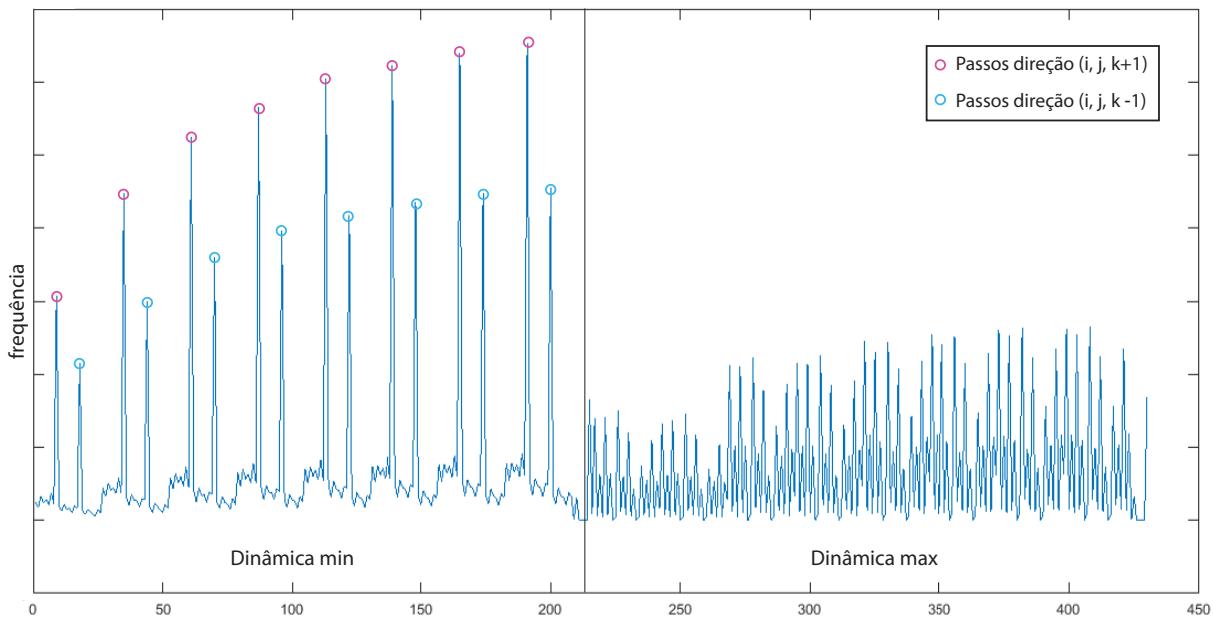


Figura 30 – Exemplo de vetor de características para caminhadas volumétricas. As caminhadas com dinâmica *min*, mostradas na metade esquerda do vetor, valorizam de forma pronunciada as direções $V(x, y, t+1)$ ou $V(x, y, t-1)$ na maioria das texturas dinâmicas da base Dyntex, quando comparadas às caminhadas com dinâmica *max*, à direita.

ocorreu com a maioria das texturas estáticas. As taxas de acerto indicam que o uso da rede complexa no caso tridimensional efetivamente prejudicou a caracterização da textura dinâmica. Possivelmente, a menor variação temporal das texturas dinâmicas torna a informação de grau menos discriminante no caso tridimensional. O método análogo ao caso bidimensional para geração da rede complexa no caso de texturas dinâmicas não parece, portanto, adequado.

Nossa segunda proposta para estender o método de caminhada bidimensional para o caso espaço-temporal é o método de caminhadas TOP. Neste caso, o vetor de características completo possui dimensão 426, dados 8 vizinhos por pixel, em 8 tamanhos de memória, 2 dinâmicas executadas sobre 3 planos e um vetor de estatísticas St para cada dinâmica em cada plano ($|F_{TOP}| = (8 \times 8 + 7) \times 2 \times 3$).

Os resultados da execução deste método nos vídeos da base Dyntex são exibidos na Tabela 13. Os melhores resultados para cada dinâmica estão destacados em negrito.

O método de caminhadas-TOP utilizando direções de passos de caminhadas nos três planos ortogonais alcançou taxas de classificação correta notavelmente superiores às das caminhadas volumétricas, com direções de passos de caminhadas nas três dimensões simultaneamente. Este também é via de regra o caso de comparações entre LBP-TOP e VLBP, por exemplo, o que indica que a análise de vizinhanças nos 3 planos é em geral

Tabela 13 – Resultados de classificação com caminhadas TOP (sobre os 3 planos ortogonais (X, Y) , (X, T) e (Y, T)), na base Dyntex, com e sem adição do vetor de estatísticas St , sobre a transformação baseada em rede complexa da imagem.

Dinâmicas	Taxa de acertos (%)						
	[1,8]	[1,4]	[5,8]	[1,8]+ <i>St</i>	[1,4]+ <i>St</i>	[5,8]+ <i>St</i>	<i>St</i>
min	85,00	84,51	73,78	87,20	87,07	84,27	63,29
max	85,12	82,56	78,17	89,02	87,56	84,76	66,34
min \cup max	93,54	91,46	89,39	94,39	92,93	93,29	81,34

superior à análise tridimensional. Novamente, observamos que os resultados melhoraram quando ambas as dinâmicas são consideradas, e as estatísticas St são poderosas dada sua dimensionalidade relativamente baixa. Os melhores resultados foram obtidos utilizando o maior vetor de características possível, dado pelos parâmetros $[1, 8] + St$.

A execução do método com o vetor de características completo (parâmetros $[1, 8] + St$) gerou a seguinte matriz de confusão. Entre as classes que causaram confusão ao classificador, apontamos alguns resultados interessantes, evidenciados na Figura 31: as classes 15 e 72, por exemplo, causaram 3 classificações incorretas. A classe 15 é um vídeo de uma máquina de lavar em funcionamento, e a classe 72 é um vídeo do círculo de pás de um moinho de vento em rotação. De fato são cenas distintas, mas há semelhança de movimento das duas texturas, já que ambas representam elementos em movimento de rotação em torno de um eixo perpendicular à perspectiva da câmera. As classes 62 e 65 também foram motivo de confusão para o classificador, com 4 classificações incorretas, mas um exame dos vídeos das classes em questão (vide figura) permite observar que os vídeos são ambos cenas de ondas na praia sob diferentes perspectivas. Por isso, o classificador mostra tendência correta em ter considerado as classes como similares. O erro foi, na verdade, diferenciar as classes ao estabelecer os valores de referência. A classe de número 10, por outro lado, foi classificada corretamente apenas para 5 das 10 amostras; é uma classe desafiadora que representa uma cena de um grupo de pelicanos e apresenta pouca ou nenhuma periodicidade temporal, com movimentos variados de penas, pescoço e bico.

Para comparação com os métodos LBP, tomamos os melhores resultados das propostas de caminhadas volumétricas e caminhadas-TOP e comparamos aos melhores resultados das extensões espaço-temporais do LBP, no caso, o $LBP - TOP_{[8,8,8]}$, isto é, LBP-TOP com vizinhança de tamanho 8 em cada plano, parâmetros que ofereceram os melhores resultados do LBP-TOP e são equivalentes aos usados no método que propusemos. Os resultados estão na Tabela 14, juntamente com as dimensões dos vetores de características. Estão apresentados os resultados da aplicação dos métodos em cada plano ortogonal individualmente, além do resultado referente ao vetor completo. Os resultados relativos ao LBP-TOP foram obtidos por experimentos executados sobre o mesmo sub-conjunto da base de dados Dyntex, utilizando a implementação do LBP-TOP disponibilizada pelos

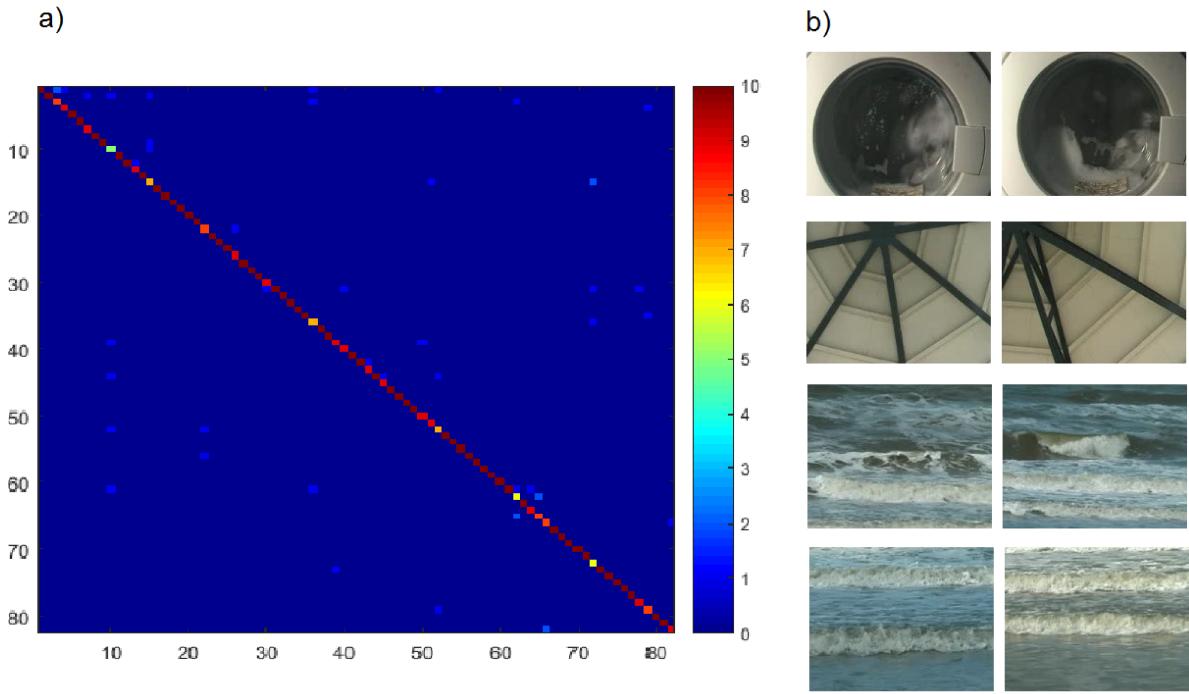


Figura 31 – a) Exemplo de matriz de confusão para caminhadas TOP na base Dyntex.
b) Amostras dos pares de classes que causaram maior confusão mútua do classificador, de cima para baixo: o par 15 e 72 (ambas texturas com movimentação circular, com 4 classificações incorretas), e o par 62 e 65 (momentos diferentes do mesmo cenário, com 3 classificações incorretas).

autores do método.

Tabela 14 – Comparação do melhor resultado do nosso método com o melhor resultado do *LBP – TOP*[8,8,8], para cada plano individualmente e para os 3 planos juntos. Melhor resultado em negrito.

Método	Nº de descritores	Taxa de acertos (%)			
		[X, Y]	[X, T]	[Y, T]	[X, Y] \cup [X, T] \cup [Y, T]
LBP-TOP	256 (1 plano)/768 (3 planos)	87,32	86,10	83,29	95,98
Caminhadas TOP	142 (1 plano)/426 (3 planos)	77,44	76,22	75,37	94,39

Podemos ver que, em ambos os métodos, os 3 diferentes planos oferecem contribuições semelhantes em termos de discriminação de texturas, de modo que o vetor de características completo, composto pelos descritores dos 3 planos, ofereceu os melhores resultados em ambos os casos, com diferença significativa em relação aos casos de planos individuais. Este resultado realça a dependência da descrição de texturas dinâmicas da análise da variação espaço-temporal da textura. Cabe lembrar que o vetor completo, no LBP-TOP e nas caminhadas TOP, apresenta vetor de características e tempo de execução 3 vezes maiores em relação às caminhadas isoladas. Podemos ainda destacar que o método LBP-TOP obteve resultados melhores, mas esta vantagem é reduzida no caso que considera o

vetor de características completo; neste caso, é importante notar que o método que propusemos apresenta um vetor de características com quase metade do tamanho do vetor do LBP-TOP.

4.5 Experimentos de descrição de texturas dinâmicas por padrões singulares de campos vetoriais

Experimentos foram realizados relativos à proposta do uso de padrões singulares em campos vetoriais aplicados a texturas dinâmicas. A primeira categoria de experimentos é relativa à abordagem *bag-of-features* aplicada aos padrões singulares. A segunda categoria de experimentos baseia-se no agrupamento estatístico não dos pontos de interesse completos mas de seus coeficientes a_k .

4.5.1 Parâmetros dos experimentos

Para ambos os métodos que propusemos relativos às estatísticas de padrões singulares de fluxos ópticos, foi utilizado o mesmo subconjunto de vídeos da base Dyntex usado para os experimentos relativos às caminhadas espaço-temporais. Em relação aos quadros utilizados, utilizamos 5 blocos de 20 quadros extraídos de cada vídeo. Como há um par de vídeos para cada classe, foram obtidas 10 amostras por classe, de modo análogo ao procedimento adotado nos experimentos com caminhadas.

Dois parâmetros importantes para a aquisição dos padrões singulares do campo vetorial são pertinentes à geração do campo vetorial, e são a área do quadro utilizada e a densidade do fluxo óptico. As duas variáveis influenciam de forma direta e significativa o tempo de processamento. Optamos por realizar os experimentos em quadros completos do vídeo, que possuem resolução 288×352 , e também em partes menores de 100×100 extraídas do centro de cada quadro. Em relação à densidade do fluxo óptico, experimentos foram realizados considerando o campo vetorial e fluxo óptico completo, com densidade igual à resolução do vídeo (onde a cada pixel da imagem I corresponde um vetor do campo vetorial V), e também com o campo vetorial após sub-amostragem (*down-sampling*) de ambas as dimensões de V pela metade. A sub-amostragem de V consiste da criação de um campo vetorial V' , composto pelos vetores de V de coordenada (x, y) tal que x e y são pares, mantendo as posições relativas dos vetores, situação em que o campo vetorial resultante V' tem tamanho final 4 vezes menor que o de V . A sub-amostragem de V em V' está representada de forma simples na Equação 41.

$$V'(x, y) = V(2x, 2y) \quad (41)$$

4.5.1.1 Histograma de valores de coeficientes de padrões singulares

Em primeiro lugar, no caso do histograma de valores de coeficientes de padrões singulares, os parâmetros foram configurados aos valores sugeridos na implementação da literatura por Liu e Ribeiro (LIU; RIBEIRO, 2012). No caso, as combinações de pares de $i \in [1, 2]$ e $k \in [-4, 4]$ oferecem um dicionário de 18 padrões básicos, com limiar mínimo de energia para elegibilidade de padrões singulares em 0.4.

Justificaremos agora nossa escolha de tamanhos variáveis de *bins*. O objetivo do histograma proposto é discriminar estatisticamente a textura com base nas ocorrências de valores para cada coeficiente a_k . Optamos por um histograma cujo tamanho aumenta exponencialmente em relação à distância do *bin* central de valor zero. Isso é porque, como mostra a Figura 32, observamos empiricamente que a ocorrência dos valores de a_k segue uma distribuição aproximadamente Gaussiana em torno do valor zero. A figura foi gerada através da análise do primeiro campo vetorial de cada vídeo dos vídeos da base Dyntex usados nos experimentos, compilando os coeficientes a_k para todos os 18 diferentes (i, k) em um histograma de *bin* de mesmo tamanho. Tal observação implica que um histograma com *bins* cujos tamanhos variam de forma logarítmica seria mais adequado para descrever as ocorrências de valores de coeficientes mais próximos de zero.

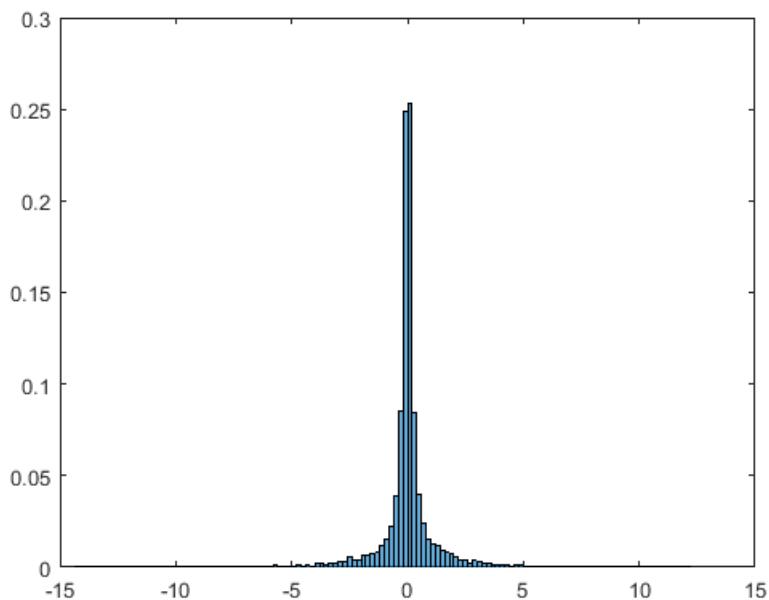


Figura 32 – O histograma de valores de a_k amostrados de múltiplos vídeos da Dyntex mostra que a distribuição de valores de a_k segue aproximadamente uma distribuição normal (Gaussiana).

O número de *bins* selecionado foi 12. Empiricamente, esse valor divide razoavelmente bem os valores de coeficientes observados de forma que eles estejam distribuídos em certa igualdade ao longo dos *bins*, sem que haja *bins* super-utilizados ou sub-utilizados. A

quantia ainda resulta em um número aceitável de dimensões para o vetor de características: há 18 histogramas de 12 *bins*, totalizando 216 dimensões. Cada histograma tem um conjunto de 7 estatísticas *St*, o que totaliza 126 estatísticas. O vetor completo possui dimensão 342.

4.5.1.2 Agrupamento de padrões em *bag-of-features*

Um parâmetro importante, este pertinente à criação do vetor de características, é o número de agrupamentos utilizados na geração da *bag-of-features* tradicional. Este parâmetro afeta diretamente o tamanho do vetor de características. Cada agrupamento é representado por um centroide, que define um *bin* da *bag-of-features*, e o número de agrupamentos não é uma decisão trivial pois a quantização ideal é muitas vezes ambígua; se o número de *clusters* for muito pequeno, agrupa-se elementos pouco similares, e se o número de *clusters* for muito grande, corre-se o risco de causar separações indesejadas entre elementos próximos (KETCHEN; SHOOK, 1996). Para decisão a respeito deste número, utilizamos duas heurísticas, a análise da evolução da soma dos erros quadrados e o número de classes do conjunto de dados.

A heurística baseada na análise da evolução dos erros quadrados é também conhecida como método do "cotovelo" (KETCHEN; SHOOK, 1996), e é um critério empírico para decisão do número de *clusters* que envolve calcular os *clusters* com um algoritmo como *k-means* para diversos valores *k*. Para cada *k*, o valor da soma dos erros quadrados dos pontos de cada agrupamento em relação ao seu centroide é calculado. O objetivo é escolher um valor *k* que promova uma boa divisão dos dados (com baixo somatório de erros quadrados dentro de um mesmo grupo), e que ao mesmo tempo não seja substancialmente melhorado por mais divisões (ou seja, um conjunto de *k* tal que a soma dos erros quadrado não é muito menor para *k* + 1, já que queremos evitar *clusters* em excesso). O cálculo da soma dos erros quadrados *SEQ_k*, para um determinado número de grupos *k*, é dado pela Equação 42, onde *C_m* é o centroide de cada *clusters* e cada *x_i* representa um dos *n* pontos dentro de um mesmo *cluster*.

$$SEQ_k = \sum_{m=1}^k \sum_{i=1}^n |x_i - C_m| \quad (42)$$

Uma análise visual do gráfico dos valores *SEQ_k* para cada valor *k*, apresentado na Figura 33, mostra a região de valores de *k* em que *k* + 1 não reduz significativamente a soma dos erros quadrados dentro dos *clusters*. Este valor fica aproximadamente no "cotovelo" da curva do gráfico. Pela análise gráfica, decidimos que um bom valor para o número de grupos é *k* = 10.

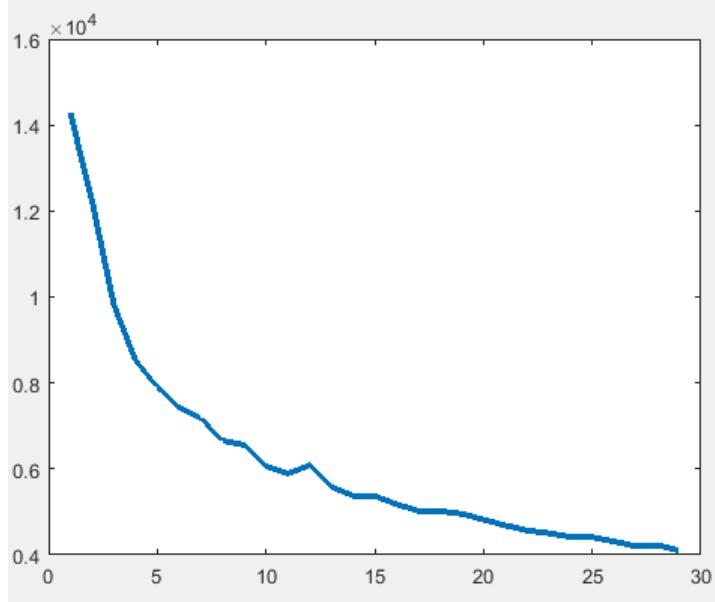


Figura 33 – Soma dos erros quadrados para cada *cluster* por número de grupos k . É possível ver o "cotovelo" da curva do gráfico, que indica um valor adequado de k .

A outra heurística para decisão sobre o número de *clusters* consiste do uso do número de classes do conjunto de dados, se for conhecido. Considerando-se que uma textura dinâmica pode ser um arranjo espaço-temporal de diversos padrões menores, não é possível saber ao certo o número de classes de fluxos reais. Por outro lado, o método detector de padrões singulares já nos providencia um vocabulário comprovadamente eficaz de padrões base de campos vetoriais. Nesse caso, como mencionamos, os parâmetros experimentais escolhidos na implementação de Liu e Ribeiro geram um dicionário de 18 padrões base, então também realizaremos experimentos com o número de agrupamentos $k = 18$.

O vetor de características do *bag-of-features* é de tamanho reduzido. No caso de $k = 10$, e com as estatísticas St , temos um vetor de dimensão 17. No caso de $k = 18$, temos com as estatísticas um vetor de dimensão 25.

4.5.2 Resultados e discussão

Para os primeiros experimentos, validaremos nossa hipótese de que há, em uma textura dinâmica, recorrência de padrões de fluxo óptico decorrentes da periodicidade espaço-temporal. Comparamos os pareamentos de pontos de interesse obtidos entre quadros de fluxo óptico obtidos de texturas dinâmicas distintas da mesma classe e de classes diferentes. Em outras palavras, buscamos verificar que diferentes fluxos ópticos de uma mesma classe de textura dinâmica apresentam padrões singulares recorrentes que podem ser identificados pela análise da ocorrência de valores dos coeficientes dos pontos de interesse de texturas dinâmicas de uma mesma classe.

Para isso, executamos o algoritmo de detecção de padrões singulares em duas amostras

de textura dinâmica. Em seguida, calculamos, por distância euclidiana, os pontos mais próximos entre os pontos das amostras. Fazemos isso para múltiplas amostras, incluindo pares de amostras da mesma classe, e entre amostras de outras classes. Se as distâncias entre os pontos de amostras da mesma classe forem em média menores que as de pontos de classes diferentes, isso é indício de que classes semelhantes geram padrões semelhantes. De fato, a análise dos padrões encontrados indica que pontos mais próximos obtidos para amostras distintas de uma mesma classe apresentam entre si distância média de $d_{iguais} = 1,87 \pm 1,66$ e pontos mais próximos de classes diferentes distam entre si, em média, em $d_{distintas} = 3,07 \pm 1,73$. Nota-se que, apesar do alto desvio, há uma tendência de que classes similares apresentem padrões singulares similares.

Os resultados dos experimentos de relativos à abordagem de histograma de valores de coeficientes e à abordagem *bag-of-features* sobre a base Dyntex são apresentados na Tabela 15. A tabela discrimina as taxas de classificação correta para os diferentes números de agrupamentos usados, além dos diferentes tamanhos de bloco e aplicação ou não de *downsampling* no fluxo óptico. Para cada combinação de parâmetros, a tabela apresenta os resultados com e sem o acréscimo das estatísticas St , calculadas para o respectivo histograma, ao vetor de características. Também experimentamos com a concatenação dos vetores de características resultantes das duas abordagens.

Um desafio particular da base Dyntex para o método de fluxos ópticos, que não ocorre no caso das caminhadas, é que algumas classes da Dyntex apresentam pouca ou nenhuma variação temporal, o que acarreta em fluxos ópticos com vetores de magnitude muito pequena. São classes da Dyntex que representam cenas como nuvens se movendo no céu ou lâmpadas LED piscando. Nos fluxos ópticos destas cenas, o algoritmo de detecção de padrões singulares tem dificuldades para detectar quaisquer padrões que ultrapassem o limiar de energia mínima para que sejam considerados pontos de interesse e não flutuações irrelevantes ou ruído. Por esse motivo, essas classes são bastante confundidas pelo classificador umas com as outras. A Figura 34 mostra exemplos de matriz de confusão com e sem 4 dessas classes problemáticas, o que evidencia como tais classes afetam a classificação, sendo notório que a maior parte dos erros de classificação se concentram nelas. Isto ocorre porque o método de padrões singulares usa fluxos ópticos, que são descritores de movimento e variações temporais da sequência de quadros. O fluxo óptico não é um método capaz de descrever variações espaciais, apenas variações no eixo temporal.

Concluímos que, em casos com fluxos ópticos de magnitude muito pequena, ambos os métodos baseados em padrões singulares de fluxo óptico propostos deve ser utilizados em conjunto com algum outro descritor que destaque a informação espacial, como as caminhadas determinísticas ou LBP-TOP. A concentração dos erros de classificação em poucas classes demonstra a limitação do método nestes casos, mas é um aspecto positivo pois é possível esperar resultados ainda melhores para experimentos que não consideram esse tipo de classe.

Tabela 15 – Taxas de classificação correta para o descritor de padrões singulares em fluxos ópticos sobre a base Dyntex. Melhor resultado em negrito.

Método	Dimensão do vetor	Taxa de acertos (%)
Histograma de coeficientes (100 × 100, com <i>downsampling</i>)		
<i>St</i>	126	69,51
Histograma	216	64,51
Histograma + <i>St</i>	342	72,44
Histograma de coeficientes (288 × 352, com <i>downsampling</i>)		
<i>St</i>	126	89,88
Histograma	216	84,63
Histograma + <i>St</i>	342	93,29
Histograma de coeficientes (100 × 100)		
<i>St</i>	126	79,15
Histograma	216	74,39
Histograma + <i>St</i>	342	82,80
Histograma de coeficientes (288 × 342)		
<i>St</i>	126	91,83
Histograma	216	87,80
Histograma + <i>St</i>	342	94,39
<i>Bag-of-features</i> (288 × 342, $k = 10$)		
<i>St</i>	7	24,76
<i>Bag-of-features</i>	10	44,27
<i>Bag-of-features</i> + <i>St</i>	17	44,74
<i>Bag-of-features</i> (288 × 342, $k = 18$)		
<i>St</i>	7	29,39
<i>Bag-of-features</i>	18	47,68
<i>Bag-of-features</i> + <i>St</i>	25	53,02
Histograma de coeficientes + <i>Bag-of-features</i> $k = 18$ (288 × 342)		
<i>St</i>	133	93,17
Histograma	234	86,83
Histograma + <i>St</i>	367	94,51

Ainda assim, observando a Tabela 15, é possível notar que os melhores resultados foram obtidos pela combinação dos dois métodos de agrupamento estatístico dos padrões singulares com taxa de acerto de 94,51%. Ainda assim, o método de histograma de valores de coeficientes a_k obteve bons resultados por si só, e a contribuição oferecida pela adição da *bag-of-features* foi pequena, melhorando a taxa de acerto em apenas 0,12%, ou 0,25% sem as classes problemáticas. Os resultados do *bag-of-features* por si só são insatisfatórios. Acreditamos que isso ocorre devido à natureza dos padrões singulares de campos vetoriais: cada ponto de interesse é formado por componentes, dadas por coeficientes a_k , que representam cada um dos padrões singulares do dicionário de padrões. A análise individual de cada coeficiente oferece uma apreciação da presença e perfil individual de cada coeficiente no campo vetorial, algo que a *bag-of-features* nem sempre é capaz de capturar. Mais uma vez é possível notar a influência positiva do conjunto de estatísticas *St* e inclusive seu potencial discriminativo quando usado isoladamente.

A análise das taxas de acerto também torna notório que o bloco de menor resolução (100 × 100) gerou resultados bastante inferiores aos do bloco de resolução máxima (288 × 352). De maneira parecida, a realização de *downsampling* no fluxo óptico, enquanto

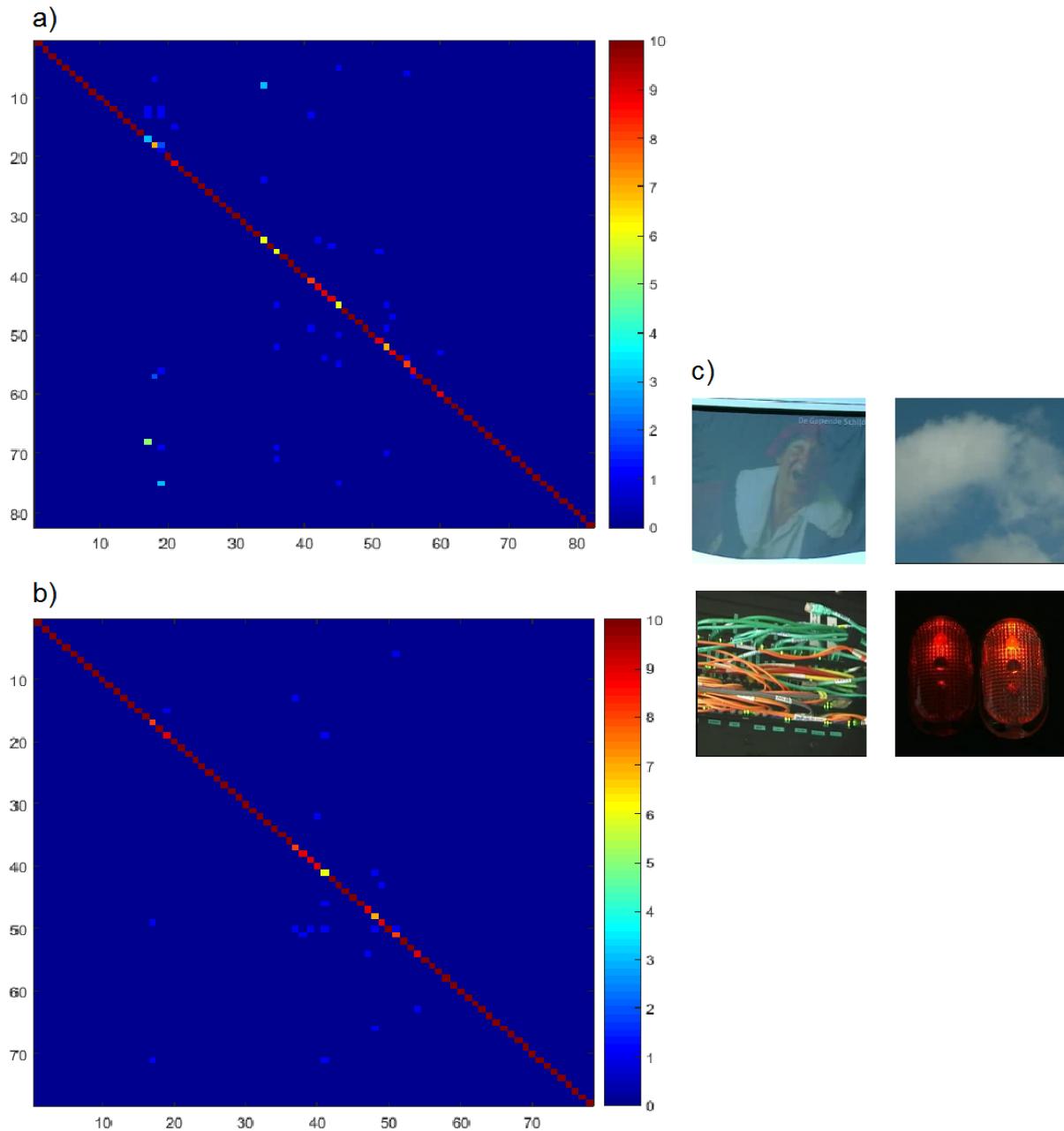


Figura 34 – Comparação de matrizes de confusão para testes com a) inclusão e b) exclusão das classes problemáticas (que possuem fluxo óptico de baixa magnitude). Em c) temos amostras destas classes (classes 17, 19, 33 e 35).

reduz substancialmente o custo computacional, também acarreta em uma redução da taxa de acertos. O motivo para a redução da taxa de acerto nos dois casos é o mesmo: o número reduzido de pontos de interesse detectados no fluxo no caso da resolução 100×100 faz com que haja uma menor quantidade de observações de forma que não é possível popular o histograma com a mesma relevância estatística que métodos que usam mais dados. Usaremos os resultados com o quadro inteiro do vídeo, sem *downsampling* e com a combinação das a abordagem estatística e histograma de valores de coeficientes para a

análise comparativa.

Compararemos o método de histogramas de coeficientes de padrões singulares (que obteve as taxas de classificação correta melhores que o método *bag-of-features*) com um método tradicional da literatura para caracterização de texturas dinâmicas, o *LBP* – *TOP*[8, 8, 8].

A Tabela 16 mostra os resultados comparativos entre o *LBP-TOP* e nossa proposta sobre a base Dyntex. Note que, como exposto anteriormente, o *LBP-TOP* é um classificador aplicado aos mesmos dados sob o ponto de vista de 3 planos, $[X, Y]$, $[X, T]$ e $[Y, T]$. Outra forma de entender o método seria dizer que a análise sobre cada plano descreve os eixos X , Y e T . O classificador proposto de padrões singulares em fluxos ópticos, por sua vez, descreve apenas variações ao longo do eixo T , ou seja, variações sobre os planos $[X, T]$ e $[Y, T]$. O fluxo óptico não é capaz de descrever as variações no plano $[X, Y]$. Se removemos classes em que há pouca variação temporal, as taxas de acerto do nosso método são aprimoradas, como mostram as taxas de classificação entre parênteses na tabela.

Os resultados da parte inferior da tabela são referentes à combinação de nosso descritor de histograma de coeficientes com os descritores do *LBP-TOP* referentes ao plano $[X, Y]$. Nesse caso, obtemos taxas de acertos superiores às do *LBP-TOP* completo, com um descritor combinado de menos dimensões. A concatenação de nosso método baseado em fluxos ópticos com um método como o *LBP*, capaz de descrever as variações espaciais que o fluxo óptico não representa é, portanto, extremamente viável.

Tabela 16 – Taxas de classificação correta para o descritor de padrões singulares em fluxos ópticos sobre a base Dyntex. Melhor resultado em negrito.

Método	Dimensão do vetor	Taxa de acertos (%)
Histograma de coeficientes (288 × 342)		
<i>St</i>	126	91,83
Histograma	216	87,80
Histograma + <i>St</i>	342	94,39
Resultados comparativos <i>LBP-TOP</i> (288 × 342)		
<i>LBP-TOP</i> $[X, Y]$	256	87,32
<i>LBP-TOP</i> $[X, Y] + [X, T] + [Y, T]$	768	95,98
Histograma de coeficientes + <i>LBP-TOP</i> $[X, Y]$ (288 × 342)		
<i>St</i>	382	97,44
Histograma	472	96,34
Histograma + <i>St</i>	598	96,34

CAPÍTULO 5

Conclusão

5.1 Principais Contribuições

O trabalho realizado e os métodos propostos tiram inspiração de métodos do estado-da-arte da caracterização de texturas estáticas e dinâmicas, enquanto acrescentam novas abordagens descritivas. Os diversos experimentos realizados em bases abrangentes, desafadoras e populares na literatura ofereceram resultados que permitem avaliar os métodos; tanto para texturas estáticas quanto para dinâmicas, os métodos obtiveram resultados de classificação bastante positivos quando comparados ao estado-da-arte. Os resultados obtidos, até mesmo os inferiores aos do estado-da-arte, oferecem observações importantes em relação ao método e à natureza do problema e dos dados, que aprofundam a compreensão do problema de classificação de texturas e podem embasar trabalhos futuros. Entre as principais contribuições do trabalho desenvolvido, podemos citar que este trabalho:

- ❑ Obtém sucesso em usar estatísticas de direções de passos de caminhadas determinísticas para extrair informações com poder descritivo que permitem atingir taxas de classificações corretas superiores às de métodos do estado-da-arte da literatura, com um vetor de características de dimensão viável e menor que em muitos métodos da literatura;
- ❑ Demonstra o potencial de uma abordagem baseada em redes complexas como uma transformação de imagem que enfatiza a informação local, aprimorando o potencial discriminativo dos métodos descritores de textura propostos em diversos casos;
- ❑ Apresenta e valida uma extensão do método baseado em redes complexas e duas extensões das caminhadas determinísticas para texturas dinâmicas, discutindo sua viabilidade;
- ❑ Apresenta e demonstra o poder descritivo de abordagens estatísticas para uso de um método extrator de padrões singulares locais de campos vetoriais na geração de descritores globais para texturas dinâmicas.

5.2 Trabalhos Futuros

O trabalho realizado oferece diversas possíveis novas direções para pesquisa no futuro próximo. Destacaremos a seguir algumas das que consideramos mais promissoras.

5.2.1 Melhor invariância a rotação e escala

A aplicação do método foi realizada sobre bases de texturas sem aplicação específica. Há bases de dados focadas em aplicações práticas em que a classificação de texturas estáticas ou dinâmicas pode ser aplicada, como sensoreamento remoto (CHENG; HAN; LU, 2017), análise de imagens médicas (MARCUS et al., 2010), reconhecimento de expressões faciais (LYONS; GYOBA; BUDYNEK,) e reconhecimento de multidões (ALI; SHAH, 2007).

5.2.2 Melhor invariância a rotação e escala

A ordenação dos vetores de características podem garantir certa resistência a variações de rotação, como visto nos experimentos em bases que apresentam imagens rotacionadas, porém ao custo de informação espacial característica do padrão local. As soluções usadas nos métodos LBP, SIFT e padrões singulares em campos vetoriais são análogas; qualquer método que busca invariância a rotação precisa ser capaz de reverter mudanças de orientação para um alvo invariante, seja atribuindo uma orientação reversível ao descritor ou detectando padrões que são versões reorientadas uns dos outros. Baseado nestas soluções, um aprimoramento possível é atribuir uma orientação reproduzível a cada vizinhança de pixel. O gradiente, por exemplo, é uma função que denota a direção de variação do sinal, e pode ser usada para atribuir orientações locais a uma imagem, e é dada pelo vetor bidimensional da equação 43 (GONZALES; WOODS; EDDINS, 2004).

$$\nabla I = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \frac{\partial I}{\partial x} \\ \frac{\partial I}{\partial y} \end{bmatrix} \quad (43)$$

O método SIFT usa o gradiente para estabelecer as orientações principais de seus pontos de interesse e alcançar robustez contra variações de rotação, na forma de um histograma de gradientes de cada pixel ao redor do ponto (LOWE, 2004). O gradiente indica a direção da maior variação de intensidade do pixel em relação à sua vizinhança. A investigação da eficácia dos valores de gradiente como fator de eliminação de sensibilidade a rotação é um caminho de interesse para a pesquisa.

Não abordamos neste trabalho o desafio de invariância a escala de texturas. Devido à natureza do histograma de direções, quaisquer variações de escala prejudicam substancialmente a classificação. Uma adaptação de nosso método para análise multi-escala

de texturas amostradas em diversas oitavas é uma solução que foi usada em alguns dos trabalhos relacionados citados nesta tese (OJALA; PIETIKÄINEN; MÄENPÄÄ, 2002) (LOWE, 2004) (LIU; RIBEIRO, 2012) e sua viabilidade pode ser investigada.

5.2.3 Avaliação da evolução dos limiares

A complexidade do método das caminhadas, em diversos casos práticos com requisitos temporais, torna impraticável a realização de múltiplas caminhadas para todos os valores de limiares usados ($|S|=36$). A opção pelo uso de caminhadas sobre graus médios contorna estes altos custos de desempenho. Ainda assim, Backes *et al.* (BACKES; CASANOVA; BRUNO, 2013) e nossa pesquisa em Ribeiro *et al.* (RIBEIRO et al., 2015) mostram com clareza que a evolução da rede mediante variações de limiar carrega informação discriminativa sobre a textura. Consideramos como um tópico de interesse investigar maneiras de integrar a variação de limiares da rede com nossa proposta de caminhadas. Esta via de pesquisa pode ser combinada com a seleção automática de limiares para a rede proposta em Ribeiro et al..

5.2.4 Integração com LBP

O método LBP é, como apresentado, um dos mais eficazes métodos para classificação de texturas dinâmicas. Por sua vez, o método das caminhadas e suas variantes, nas quais se baseou grande parte de nossa proposta, tem obtido resultados significativos. O método LBP é uma descrição mais precisa dos padrões locais, porém mais sensível a variações no padrão, enquanto o método das caminhadas é robusto a variações de intensidade entre pixels de valores dos pixels. Seria interessante investigar as possibilidades de integração dos dois métodos.

A matriz de valores LBP pode ser tratada como uma transformação da imagem. Assim, destacamos o potencial da realização de caminhadas sobre os valores da imagem transformada com o LBP, já que a codificação LBP do pixel substitui seu valor de intensidade por um valor que carrega uma comparação relativa entre o valor original do pixel e o de sua vizinhança. Este trabalho demonstrou o potencial de caminhadas sobre transformações de imagens baseadas em vizinhanças, já que quando a caminhada foi aplicada sobre a transformação da imagem baseada em redes complexas os resultados de classificação de texturas estáticas foram superiores aos da aplicação das caminhadas sobre a imagem original.

5.2.5 Análise fractal

Recentemente, métodos de análise da dimensão fractal para descrição de texturas tem recebido atenção (LAHMIRI, 2016) (XU et al., 2015) (RIBAS et al., 2015) (QUAN et

al., 2014) (GONÇALVES; BRUNO, 2013a) (FLORINDO; BRUNO, 2012). Uma propriedade significativa da análise fractal é sua relativa robustez a variações de rotação e escala (VARMA; GARG, 2007) (XU; JI; FERMÜLLER, 2009). A dimensão fractal de determinado elemento se relaciona com sua complexidade em termos de ocupação espacial e autossimilaridade (FLORINDO; BRUNO, 2012) e é um poderoso descritor de geometrias irregulares. Há trabalhos na literatura que realizaram com sucesso a combinação entre medidas de dimensão fractal e agentes de caminhada (GONÇALVES; BRUNO, 2013a) (GONÇALVES; MACHADO; BRUNO, 2014) e, portanto, é uma área de pesquisa promissora.

5.2.6 Padrões singulares com *Three Orthogonal Planes*

O descritor baseado em padrões singulares locais mostrou-se eficaz como um descritor global de texturas dinâmicas. Apesar disso, o método teve dificuldades para descrever texturas que apresentam pouca movimentação, já que os campos vetoriais de baixa magnitude não geram padrões singulares locais em quantidade ou qualidade suficientes. Isto ocorreu porque o método é baseado na descrição de movimento e flutuações em um único eixo (no caso, o eixo temporal). Seguindo a ideia da abordagem TOP usada no LBP e na proposta de caminhadas da tese, é de interesse verificar o potencial descritivo do método dos padrões singulares sobre fluxos ópticos calculados sobre os eixos espaciais do vídeo (X e Y). A eficácia das abordagens TOP faz com que haja boas expectativas a respeito desta linha de pesquisa.

5.3 Contribuições em Produção Bibliográfica

A seguir listamos os trabalhos apresentados e publicações que foram geradas a partir de pesquisas relacionadas a este trabalho:

- ❑ Couto, L. N., Backes, A. R. and Barcelos, C. A. Z. (2017, October). Texture characterization via deterministic walks' direction histogram applied to a complex network-based image transformation. *In Pattern Recognition Letters*, vol.97, 2017 (pp. 77-83). Elsevier.
- ❑ Couto, L. N., Ribeiro, T. P., Backes, A. R. and Barcelos, C. A. Z. (2015, September). Texture characterization via improved deterministic walks on image-generated complex network. *In Image Processing (ICIP), 2015 IEEE International Conference on* (pp. 4416-4420). IEEE.
- ❑ Couto, L. N. and Barcelos, C. A. Z. Caracterização e classificação de texturas dinâmicas através de caminhadas de determinísticas. *In XIV Semana da Matemática*.

XIV SEMAT e IV SEMEST, 2014. Nota: A publicação recebeu menção honrosa da coordenação do evento.

- Ribeiro, T. P., Couto, L. N., Backes, A. R., and Barcelos, C. A. Z. (2015). Texture Characterization via Automatic Threshold Selection on Image-Generated Complex Network. *In Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications* (pp. 468-476). Springer International Publishing.
- Oliveira, G. D., Cunha, M., Couto, L. N. (2015) Extração de características e pareamento de imagens com SIFT simplificado para aplicação em localização robótica. *In II FACOM TechWeek 2015*.
- Couto, L. N. and Barcelos, C. A. Z. (2018). Singular Patterns in Optical Flows as Dynamic Texture Descriptors. *In Image Processing (ICIP), 2018 IEEE International Conference on* (submetido).

Referências

- AGARWAL, S.; AWAN, A.; ROTH, D. Learning to detect objects in images via a sparse, part-based representation. **IEEE transactions on pattern analysis and machine intelligence**, IEEE, v. 26, n. 11, p. 1475–1490, 2004.
- AGGARWAL, N.; AGRAWAL, R. First and second order statistics features for classification of magnetic resonance brain images. **Journal of Signal and Information Processing**, v. 3, n. 02, p. 146, 2012. Disponível em: <<https://doi.org/10.4236/jsip.2012.32019>>.
- ALI, S.; SHAH, M. A lagrangian particle dynamics approach for crowd flow segmentation and stability analysis. In: IEEE. **Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on**. 2007. p. 1–6. Disponível em: <<https://doi.org/10.1109/CVPR.2007.382977>>.
- ANDREARCZYK, V.; WHELAN, P. F. Convolutional neural network on three orthogonal planes for dynamic texture classification. **arXiv preprint arXiv:1703.05530**, 2017.
- AZENCOTT, R.; WANG, J.-P.; YOUNES, L. Texture classification using windowed fourier filters. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 19, n. 2, p. 148–153, 1997. Disponível em: <<https://doi.org/10.1109/34.574796>>.
- BACKES, A. R. et al. Deterministic tourist walks as an image analysis methodology based. In: SPRINGER. **Iberoamerican Congress on Pattern Recognition**. 2006. p. 784–793. Disponível em: <https://doi.org/10.1007/11892755_81>.
- BACKES, A. R.; CASANOVA, D.; BRUNO, O. M. A complex network-based approach for texture analysis. In: **Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications**. Springer Berlin Heidelberg, 2010, (Lecture Notes in Computer Science, v. 6419). p. 354–361. Disponível em: <https://doi.org/10.1007/978-3-642-16687-7_48>.
- BACKES, A. R.; CASANOVA, D.; BRUNO, O. M. Texture analysis and classification: A complex network-based approach. **Information Sciences**, Elsevier, v. 219, p. 168–180, 2013. Disponível em: <<https://doi.org/10.1016/j.ins.2012.07.003>>.
- BACKES, A. R. et al. Texture analysis and classification using deterministic tourist walk. **Pattern Recognition**, v. 43, n. 3, p. 685 – 694, 2010. Disponível em: <<https://doi.org/10.1016/j.patcog.2009.07.017>>.

- BAKER, S.; MATTHEWS, I. Lucas-kanade 20 years on: A unifying framework. **International journal of computer vision**, Springer, v. 56, n. 3, p. 221–255, 2004.
- BARABASI, A.-L. Linked: How everything is connected to everything else and what it means. Plume, 2003.
- BARABÁSI, A.-L.; ALBERT, R. Emergence of scaling in random networks. **Science**, American Association for the Advancement of Science, v. 286, n. 5439, p. 509–512, 1999. Disponível em: <<https://doi.org/10.1126/science.286.5439.509>>.
- BARLEY, A.; TOWN, C. Combinations of feature descriptors for texture image classification. **Journal of Data Analysis and Information Processing**, v. 2, p. 67–76, 2014. Disponível em: <<https://doi.org/10.4236/jdaip.2014.23009>>.
- BASU, S. et al. Deep neural networks for texture classification—a theoretical analysis. **Neural Networks**, Elsevier, v. 97, p. 173–182, 2018. Disponível em: <<https://doi.org/10.1016/j.neunet.2017.10.001>>.
- BAY, H.; TUYTELAARS, T.; GOOL, L. V. Surf: Speeded up robust features. In: **Computer vision–ECCV 2006**. Springer, 2006. p. 404–417. Disponível em: <https://doi.org/10.1007/11744023_32>.
- BELONGIE, S. et al. Color-and texture-based image segmentation using em and its application to content-based image retrieval. In: **IEEE. Computer Vision, 1998. Sixth International Conference on**. 1998. p. 675–682.
- BLAS, M. et al. Fast color/texture segmentation for outdoor robots. In: **IEEE. Intelligent Robots and Systems, 2008. IROS 2008. IEEE/RSJ International Conference on**. 2008. p. 4078–4085.
- BOUREAU, Y.-L. et al. Learning mid-level features for recognition. In: **IEEE. Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on**. 2010. p. 2559–2566. Disponível em: <<https://doi.org/10.1109/CVPR.2010.5539963>>.
- BRODATZ, P. **Textures: a photographic album for artists and designers**. : Dover Publications, 1966. (Dover pictorial archives).
- BRUHN, A.; WEICKERT, J.; SCHNÖRR, C. Lucas/kanade meets horn/schunck: Combining local and global optic flow methods. **International Journal of Computer Vision**, Springer, v. 61, n. 3, p. 211–231, 2005. Disponível em: <<https://doi.org/10.1023/B:VISI.0000045324.43199.43>>.
- BULLMORE, E.; SPORNS, O. Complex brain networks: graph theoretical analysis of structural and functional systems. **Nature reviews. Neuroscience**, Nature Publishing Group, v. 10, n. 3, p. 186, 2009.
- CAMPITELI, M. G. et al. Deterministic walks as an algorithm of pattern recognition. **Phys. Rev. E**, American Physical Society, v. 74, p. 026703, 2006.
- CASANOVA, D. et al. Texture analysis using fractal descriptors estimated by the mutual interference of color channels. **Information Sciences**, Elsevier, v. 346, p. 58–72, 2016. Disponível em: <<https://doi.org/10.1016/j.ins.2016.01.077>>.

- CASTELLANO, G. et al. Texture analysis of medical images. **Clinical radiology**, Elsevier, v. 59, n. 12, p. 1061–1069, 2004.
- CHANG, T.; KUO, C. J. Texture analysis and classification with tree-structured wavelet transform. **IEEE Transactions on Image Processing**, v. 2, n. 4, p. 429–441, 1993. Disponível em: <<https://doi.org/10.1109/83.242353>>.
- CHAO, H.; GU, Y.; NAPOLITANO, M. A survey of optical flow techniques for robotics navigation applications. **Journal of Intelligent & Robotic Systems**, Springer, v. 73, n. 1-4, p. 361–372, 2014. Disponível em: <<https://doi.org/10.1007/s10846-013-9923-6>>.
- CHAUDHRY, R. et al. Histograms of oriented optical flow and binet-cauchy kernels on nonlinear dynamical systems for the recognition of human actions. In: **IEEE. computer vision and pattern recognition, 2009. CVPR 2009. IEEE Conference on**. 2009. p. 1932–1939. Disponível em: <<https://doi.org/10.1109/CVPR.2009.5206821>>.
- CHENG, G.; HAN, J.; LU, X. Remote sensing image scene classification: benchmark and state of the art. **Proceedings of the IEEE**, IEEE, v. 105, n. 10, p. 1865–1883, 2017. Disponível em: <<https://doi.org/10.1109/JPROC.2017.2675998>>.
- CHEVREKOV, D.; PÉTERI, R. A brief survey of dynamic texture description and recognition. In: **Computer Recognition Systems**. Springer, 2005. p. 17–26. Disponível em: <https://doi.org/10.1007/3-540-32390-2_2>.
- CORPETTI, T.; MEMIN, E.; PÉREZ, P. Extraction of singular points from dense motion fields: an analytic approach. **Journal of mathematical imaging and vision**, Springer, v. 19, n. 3, p. 175–198, 2003.
- COSTA, L. d. F. et al. Analyzing and modeling real-world phenomena with complex networks: a survey of applications. **Advances in Physics**, Taylor & Francis, v. 60, n. 3, p. 329–412, 2011.
- 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. Disponível em: <<https://doi.org/10.1080/00018730601170527>>.
- COUTO, C. M. V.; COMIN, C. H.; COSTA, L. da F. Effects of threshold on the topology of gene co-expression networks. **Molecular BioSystems**, Royal Society of Chemistry, v. 13, n. 10, p. 2024–2035, 2017.
- COUTO, L. **Sistema para localização robótica de veículos autônomos baseado em visão computacional por pontos de referência**. Dissertação (Mestrado) — Universidade de São Paulo, Instituto de Ciências Matemáticas e de Computação, 2012.
- COUTO, L. N.; BACKES, A. R.; BARCELOS, C. A. Z. Texture characterization via deterministic walks' direction histogram applied to a complex network-based image transformation. **Pattern Recognition Letters**, Elsevier, v. 97, p. 77–83, 2017. Disponível em: <<https://doi.org/10.1016/j.patrec.2017.07.013>>.
- COUTO, L. N. et al. Texture characterization via improved deterministic walks on image-generated complex network. In: **IEEE. Image Processing (ICIP), 2015 IEEE International Conference on**. 2015. p. 4416–4420. Disponível em: <<https://doi.org/10.1109/ICIP.2015.7351641>>.

- CSURKA, G. et al. Visual categorization with bags of keypoints. In: PRAGUE. **Workshop on statistical learning in computer vision, ECCV**. 2004. v. 1, n. 1-22, p. 1-2.
- DANA, K. J. et al. Reflectance and texture of real-world surfaces. **ACM Transactions on Graphics**, ACM, New York, NY, USA, v. 18, n. 1, p. 1-34, 1999. ISSN 0730-0301.
- DERPANIS, K. G.; WILDES, R. Spacetime texture representation and recognition based on a spatiotemporal orientation analysis. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, IEEE, v. 34, n. 6, p. 1193-1205, 2012. Disponível em: <<https://doi.org/10.1109/TPAMI.2011.221>>.
- DONG, P. Test of a new lacunarity estimation method for image texture analysis. **International Journal of Remote Sensing**, v. 21, n. 17, p. 3369-3373, 2000. Disponível em: <<https://doi.org/10.1080/014311600750019985>>.
- DORETTO, G. et al. Dynamic textures. **International Journal of Computer Vision**, Springer, v. 51, n. 2, p. 91-109, 2003. Disponível em: <<https://doi.org/10.1023/A:1021669406132>>.
- DUDA, R. O.; HART, P. E.; STORK, D. G. **Pattern classification**. : John Wiley & Sons, 2012.
- EFROS, A. A.; LEUNG, T. K. Texture synthesis by non-parametric sampling. In: **Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on**. 1999. v. 2, p. 1033-1038 vol.2. Disponível em: <<https://doi.org/10.1109/ICCV.1999.790383>>.
- FAN, L. et al. Singular points detection based on zero-pole model in fingerprint images. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, IEEE, v. 30, n. 6, p. 929-940, 2008. Disponível em: <<https://doi.org/10.1109/TPAMI.2008.31>>.
- FAZEKAS, S.; CHETVERIKOV, D. Normal versus complete flow in dynamic texture recognition: a comparative study. In: **International workshop on texture analysis and synthesis**. 2005. p. 37-42.
- FAZEKAS, S.; CHETVERIKOV, D. Dynamic texture recognition using optical flow features and temporal periodicity. In: **IEEE. Content-Based Multimedia Indexing, 2007. CBMI'07. International Workshop on**. 2007. p. 25-32. Disponível em: <<https://doi.org/10.1109/CBMI.2007.385388>>.
- FLORINDO, J. B.; BRUNO, O. M. Fractal descriptors based on fourier spectrum applied to texture analysis. **Physica A: statistical Mechanics and its Applications**, Elsevier, v. 391, n. 20, p. 4909-4922, 2012.
- FLORINDO, J. B.; LANDINI, G.; BRUNO, O. M. Three-dimensional connectivity index for texture recognition. **Pattern Recognition Letters**, Elsevier, v. 84, p. 239-244, 2016. Disponível em: <<https://doi.org/10.1016/j.patrec.2016.09.013>>.
- FRIEDMAN, J.; HASTIE, T.; TIBSHIRANI, R. **The elements of statistical learning**. : Springer series in statistics New York, 2001. v. 1.

- GAETANO, R.; SCARPA, G.; POGGI, G. Hierarchical texture-based segmentation of multiresolution remote-sensing images. **IEEE Transactions on Geoscience and Remote Sensing**, IEEE, v. 47, n. 7, p. 2129–2141, 2009. Disponível em: <<https://doi.org/10.1109/TGRS.2008.2010708>>.
- GHIDONI, S.; CIELNIAK, G.; MENEGATTI, E. Texture-based crowd detection and localisation. In: **Intelligent Autonomous Systems 12**. Springer, 2013. p. 725–736. Disponível em: <https://doi.org/10.1007/978-3-642-33926-4_69>.
- GONÇALVES, W. N. et al. Texture descriptor based on partially self-avoiding deterministic walker on networks. **Expert Systems with Applications**, Elsevier, v. 39, n. 15, p. 11818–11829, 2012. Disponível em: <<https://doi.org/10.1016/j.eswa.2012.01.094>>.
- GONÇALVES, W. N.; BRUNO, O. M. Combining fractal and deterministic walkers for texture analysis and classification. **Pattern Recognition**, Elsevier, v. 46, n. 11, p. 2953–2968, 2013. Disponível em: <<https://doi.org/10.1016/j.patcog.2013.03.012>>.
- GONÇALVES, W. N.; BRUNO, O. M. Dynamic texture analysis and segmentation using deterministic partially self-avoiding walks. **Expert Systems with Applications**, Elsevier, v. 40, n. 11, p. 4283–4300, 2013. Disponível em: <<https://doi.org/10.1016/j.eswa.2012.12.092>>.
- GONÇALVES, W. N.; MACHADO, B. B.; BRUNO, O. M. Spatiotemporal gabor filters: a new method for dynamic texture recognition. **arXiv preprint arXiv:1201.3612**, 2012.
- GONÇALVES, W. N.; MACHADO, B. B.; BRUNO, O. M. Texture descriptor combining fractal dimension and artificial crawlers. **Physica A: Statistical Mechanics and its Applications**, Elsevier, v. 395, p. 358–370, 2014.
- GONÇALVES, W. N. et al. Texture recognition based on diffusion in networks. **Information Sciences**, Elsevier, v. 364, p. 51–71, 2016. Disponível em: <<https://doi.org/10.1016/j.ins.2016.04.052>>.
- GONZALES, R. C.; WOODS, R. E.; EDDINS, S. L. **Digital image processing using MATLAB**. : Pearson Prentice Hall, 2004.
- GOPALAKRISHNAN, V.; HU, Y.; RAJAN, D. Random walks on graphs to model saliency in images. In: **IEEE. Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on**. 2009. p. 1698–1705. Disponível em: <<https://doi.org/10.1109/CVPR.2009.5206767>>.
- GRADY, L. Random walks for image segmentation. **Pattern Analysis and Machine Intelligence, IEEE Transactions on**, IEEE, v. 28, n. 11, p. 1768–1783, 2006.
- GUERRIERO, V. Power law distribution: Method of multi-scale inferential statistics. **Journal of Modern Mathematics Frontier**, Science and Engineering Publishing Company, v. 1, n. 1, p. 21–28, 2012.
- GUO, Y. et al. Video texture synthesis with multi-frame lbp-top and diffeomorphic growth model. **IEEE Transactions on Image Processing**, IEEE, v. 22, n. 10, p. 3879–3891, 2013. Disponível em: <<https://doi.org/10.1109/TIP.2013.2263148>>.

GUO, Z.; ZHANG, L.; ZHANG, D. A completed modeling of local binary pattern operator for texture classification. **IEEE Transactions on Image Processing**, IEEE, v. 19, n. 6, p. 1657–1663, 2010. Disponível em: <<https://doi.org/10.1109/TIP.2010.2044957>>.

HÁJEK, M. **Texture analysis for magnetic resonance imaging**. : Texture Analysis Magn Resona, 2006.

HARALICK, R. M. Statistical and structural approaches to texture. **Proceedings of the IEEE**, v. 67, n. 5, p. 786–804, 1979. ISSN 0018-9219.

HARALICK, R. M.; SHANMUGAM, K.; DINSTEIN, I. H. Textural features for image classification. **Systems, Man and Cybernetics, IEEE Transactions on**, Ieee, n. 6, p. 610–621, 1973.

HAYKIN, S.; VEEN, B. V. **Signals and systems**. : John Wiley & Sons, 2007.

HAYMAN, E.; CAPUTO, B.; EKLUNDH, J. On the significance of real-world conditions for material classification. In: **Computer Vision - ECCV 2004**. Springer Berlin Heidelberg, 2004, (Lecture Notes in Computer Science, v. 3024). p. 253–266. ISBN 978-3-540-21981-1. Disponível em: <https://doi.org/10.1007/978-3-540-24673-2_21>.

HIREMATH, P. S.; PUJARI, J. Content based image retrieval using color, texture and shape features. In: **Advanced Computing and Communications, 2007. ADCOM 2007. International Conference on**. 2007. p. 780–784.

HOLME, P.; SARAMÄKI, J. Temporal networks. **Physics reports**, Elsevier, v. 519, n. 3, p. 97–125, 2012. Disponível em: <<https://doi.org/10.1016/j.physrep.2012.03.001>>.

HORN, B. K.; SCHUNCK, B. G. Determining optical flow. In: INTERNATIONAL SOCIETY FOR OPTICS AND PHOTONICS. **1981 Technical symposium east**. 1981. p. 319–331.

HOSSAIN, S.; SERIKAWA, S. Texture databases—a comprehensive survey. **pattern recognition letters**, Elsevier, v. 34, n. 15, p. 2007–2022, 2013.

HUANG, D. et al. Local binary patterns and its application to facial image analysis: a survey. **IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)**, IEEE, v. 41, n. 6, p. 765–781, 2011.

IDRISSA, M.; ACHEROY, M. Texture classification using gabor filters. **Pattern Recognition Letters**, v. 23, n. 9, p. 1095–1102, 2002. Disponível em: <[https://doi.org/10.1016/S0167-8655\(02\)00056-9](https://doi.org/10.1016/S0167-8655(02)00056-9)>.

JAIN, A. K.; FARROKHNA, F. Unsupervised texture segmentation using gabor filters. **Pattern recognition**, Elsevier, v. 24, n. 12, p. 1167–1186, 1991.

JÉGOU, H.; DOUZE, M.; SCHMID, C. Improving bag-of-features for large scale image search. **International journal of computer vision**, Springer, v. 87, n. 3, p. 316–336, 2010.

JIANG, M.; MACHIRAJU, R.; THOMPSON, D. Detection and visualization of. **The Visualization Handbook**, Academic Press, v. 295, 2005.

- JULESZ, B. Experiments in the visual perception of texture. **Scientific American**, v. 232, p. 34–43, 1975. Disponível em: <<https://doi.org/10.1038/scientificamerican0475-34>>.
- JULESZ, B. Textons, the elements of texture perception, and their interactions. **Nature**, Nature Publishing Group, v. 290, n. 5802, p. 91, 1981.
- KASSNER, A.; THORNHILL, R. Texture analysis: a review of neurologic mr imaging applications. **American Journal of Neuroradiology**, Am Soc Neuroradiology, v. 31, n. 5, p. 809–816, 2010. Disponível em: <<https://doi.org/10.3174/ajnr.A2061>>.
- KATO, H.; HARADA, T. Image reconstruction from bag-of-visual-words. In: **IEEE. Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on**. 2014. p. 955–962. Disponível em: <<https://doi.org/10.1109/CVPR.2014.127>>.
- KETCHEN, D. J. J.; SHOOK, C. L. The application of cluster analysis in strategic management research: an analysis and critique. **Strategic management journal**, JSTOR, p. 441–458, 1996.
- KIHL, O.; TREMBLAIS, B.; AUGEREAU, B. Multivariate orthogonal polynomials to extract singular points. In: **IEEE. Image Processing, 2008. ICIP 2008. 15th IEEE International Conference on**. 2008. p. 857–860. Disponível em: <<https://doi.org/10.1109/ICIP.2008.4711890>>.
- KIM, J.-K.; PARK, H. W. Statistical textural features for detection of microcalcifications in digitized mammograms. **IEEE Transactions on Medical Imaging**, v. 18, n. 3, p. 231–238, 1999. Disponível em: <<https://doi.org/10.1109/42.764896>>.
- LAHMIRI, S. Image characterization by fractal descriptors in variational mode decomposition domain: application to brain magnetic resonance. **Physica A: Statistical Mechanics and its Applications**, Elsevier, v. 456, p. 235–243, 2016.
- LANDY, M. S.; GRAHAM, N. 73 visual perception of texture. **The visual neurosciences**, MIT press, v. 1, p. 1106, 2004.
- LASMAR, M.; BAUSSARD, A.; CHENADEC, G. L. Asymmetric power distribution model of wavelet subbands for texture classification. **Pattern Recognition Letters**, v. 52, n. 0, p. 1 – 8, 2015. Disponível em: <<https://doi.org/10.1016/j.patrec.2014.08.004>>.
- LAWS, K. I. **Textured image segmentation**. Tese (Doutorado) — University of South California, 1980. Disponível em: <<https://doi.org/10.21236/ADA083283>>.
- 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.
- LAZEBNIK, S.; SCHMID, C.; PONCE, J. Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories. In: **IEEE. Computer vision and pattern recognition, 2006 IEEE computer society conference on**. 2006. v. 2, p. 2169–2178.

LEUNG, T.; MALIK, J. Representing and recognizing the visual appearance of materials using three-dimensional textons. **International journal of computer vision**, Springer, v. 43, n. 1, p. 29–44, 2001.

LI, J.; YAU, W.-Y.; WANG, H. Constrained nonlinear models of fingerprint orientations with prediction. **Pattern Recognition**, Elsevier, v. 39, n. 1, p. 102–114, 2006. Disponível em: <<https://doi.org/10.1016/j.patcog.2005.08.010>>.

LI, Y. et al. Automatic image stitching using sift. In: **IEEE. Audio, Language and Image Processing, 2008. ICALIP 2008. International Conference on**. 2008. p. 568–571.

LIU, L. et al. Evaluation of lbp and deep texture descriptors with a new robustness benchmark. In: SPRINGER. **European Conference on Computer Vision**. 2016. p. 69–86. Disponível em: <https://doi.org/10.1007/978-3-319-46487-9_5>.

LIU, L. et al. Median robust extended local binary pattern for texture classification. **IEEE Transactions on Image Processing**, IEEE, v. 25, n. 3, p. 1368–1381, 2016. Disponível em: <<https://doi.org/10.1109/TIP.2016.2522378>>.

LIU, Q. et al. Fingerprint singular points extraction based on orientation tensor field and laurent series. **Journal of Central South University**, Springer, v. 21, n. 5, p. 1927–1934, 2014. Disponível em: <<https://doi.org/10.1007/s11771-014-2139-5>>.

LIU, W.; RIBEIRO, E. Scale and rotation invariant detection of singular patterns in vector flow fields. In: SPRINGER. **SSPR/SPR**. 2010. p. 522–531. Disponível em: <https://doi.org/10.1007/978-3-642-14980-1_51>.

LIU, W.; RIBEIRO, E. Detecting singular patterns in 2d vector fields using weighted laurent polynomial. **Pattern Recognition**, Elsevier, v. 45, n. 11, p. 3912–3925, 2012. Disponível em: <<https://doi.org/10.1016/j.patcog.2012.04.025>>.

LIU, Y.-J. et al. A main directional mean optical flow feature for spontaneous micro-expression recognition. **IEEE Transactions on Affective Computing**, IEEE, v. 7, n. 4, p. 299–310, 2016. Disponível em: <<https://doi.org/10.1109/TAFFC.2015.2485205>>.

LOWE, D. G. Distinctive image features from scale-invariant keypoints. **International journal of computer vision**, Springer, v. 60, n. 2, p. 91–110, 2004. Disponível em: <<https://doi.org/10.1023/B:VISI.0000029664.99615.94>>.

LUCAS, B. D.; KANADE, T. et al. An iterative image registration technique with an application to stereo vision. In: **International Joint Conference on Artificial Intelligence**. 1981. v. 81, p. 674–679.

LYONS, M. J.; GYOBA, J.; BUDYNEK, J. The japanese female facial expression (jaffe) database. In: . Disponível em: <<http://www.kasrl.org/jaffe.html>>.

MA, Y.; CISAR, P. Event detection using local binary pattern based dynamic textures. In: **IEEE. Computer Vision and Pattern Recognition Workshops, 2009. CVPR Workshops 2009. IEEE Computer Society Conference on**. 2009. p. 38–44. Disponível em: <<https://doi.org/10.1109/CVPRW.2009.5204204>>.

- MALIK, F.; BAHARUDIN, B. The statistical quantized histogram texture features analysis for image retrieval based on median and laplacian filters in the dct domain. **The International Arab Journal of Information Technology**, v. 10, n. 6, p. 1–9, 2013.
- MARCUS, D. S. et al. Open access series of imaging studies: longitudinal mri data in nondemented and demented older adults. **Journal of cognitive neuroscience**, MIT Press, v. 22, n. 12, p. 2677–2684, 2010.
- MARTINEZ, A. S.; KINOUCHI, O.; RISAU-GUSMAN, S. Exploratory behavior, trap models, and glass transitions. **Physical Review E**, APS, v. 69, n. 1, p. 017–101, 2004.
- MATERKA, A.; STRZELECKI, M. et al. Texture analysis methods—a review. **Technical university of lodz, institute of electronics, COST B11 report**, Brussels, p. 9–11, 1998.
- NANNI, L.; LUMINI, A.; BRAHNAM, S. Local binary patterns variants as texture descriptors for medical image analysis. **Artificial intelligence in medicine**, Elsevier, v. 49, n. 2, p. 117–125, 2010.
- NANNI, L.; LUMINI, A.; BRAHNAM, S. Survey on lbp based texture descriptors for image classification. **Expert Systems with Applications**, Elsevier, v. 39, n. 3, p. 3634–3641, 2012. Disponível em: <<https://doi.org/10.1016/j.eswa.2011.09.054>>.
- NG, I.; TAN, T.; KITTLER, J. On local linear transform and Gabor filter representation of texture. In: **Proceedings of IEEE International Conference on Pattern Recognition (ICPR)**. 1992. p. 627–631. Disponível em: <<https://doi.org/10.1109/ICPR.1992.202065>>.
- NOWAK, E.; JURIE, F.; TRIGGS, B. Sampling strategies for bag-of-features image classification. **Computer Vision–ECCV 2006**, Springer, p. 490–503, 2006. Disponível em: <https://doi.org/10.1007/11744085_38>.
- O'HARA, S.; DRAPER, B. A. Introduction to the bag of features paradigm for image classification and retrieval. **arXiv preprint arXiv:1101.3354**, 2011.
- OJALA, T. et al. Outex-new framework for empirical evaluation of texture analysis algorithms. In: **IEEE. Pattern Recognition, 2002. Proceedings. 16th International Conference on**. 2002. v. 1, p. 701–706. Disponível em: <<https://doi.org/10.1109/ICPR.2002.1044854>>.
- 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.
- OJALA, T.; PIETIKÄINEN, M.; MÄENPÄÄ, T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. **Pattern Analysis and Machine Intelligence, IEEE Transactions on**, IEEE, v. 24, n. 7, p. 971–987, 2002.
- PATIL, N. K. et al. Adaptive texture and color feature based color image compression. In: **IEEE. Smart Structures and Systems (ICSSS), 2013 IEEE International Conference on**. 2013. p. 82–86. Disponível em: <<https://doi.org/10.1109/ICSSS.2013.6623006>>.

- PAYET, N.; TODOROVIC, S. Scene shape from texture of objects. In: **IEEE. Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on**. 2011. p. 2017–2024. Disponível em: <<https://doi.org/10.1109/CVPR.2011.5995326>>.
- 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.
- PÉTERI, R.; FAZEKAS, S.; HUISKES, M. J. Dyntex: A comprehensive database of dynamic textures. **Pattern Recognition Letters**, Elsevier, v. 31, n. 12, p. 1627–1632, 2010. Disponível em: <<https://doi.org/10.1016/j.patrec.2010.05.009>>.
- PIETIKÄINEN, M.; OJALA, T.; XU, Z. Rotation-invariant texture classification using feature distributions. **Pattern Recognition**, Elsevier, v. 33, n. 1, p. 43–52, 2000. Disponível em: <[https://doi.org/10.1016/S0031-3203\(99\)00032-1](https://doi.org/10.1016/S0031-3203(99)00032-1)>.
- PIETIKÄINEN, M.; ZHAO, G. Two decades of local binary patterns: A survey. **Advances in Independent Component Analysis and Learning Machines**, Academic Press, p. 175–210, 2015. Disponível em: <<https://doi.org/10.1016/B978-0-12-802806-3.00009-9>>.
- PUJARI, J. D.; YAKKUNDIMATH, R.; BYADGI, A. S. Image processing based detection of fungal diseases in plants. **Procedia Computer Science**, Elsevier, v. 46, p. 1802–1808, 2015. Disponível em: <<https://doi.org/10.1016/j.procs.2015.02.137>>.
- QUAN, Y. et al. Lacunarity analysis on image patterns for texture classification. In: **Proceedings of the IEEE conference on computer vision and pattern recognition**. 2014. p. 160–167. Disponível em: <<https://doi.org/10.1109/CVPR.2014.28>>.
- RAO, A. R.; JAIN, R. C. Computerized flow field analysis: Oriented texture fields. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, IEEE Computer Society, v. 14, n. 7, p. 693–709, 1992.
- REED, T. R.; DUBUF, J. H. A review of recent texture segmentation and feature extraction techniques. **CVGIP: Image understanding**, Elsevier, v. 57, n. 3, p. 359–372, 1993.
- RIBAS, L. C.; BRUNO, O. M. Fast deterministic tourist walk for texture analysis. **arXiv preprint arXiv:1611.08624**, 2016.
- RIBAS, L. C. et al. Fractal dimension of maximum response filters applied to texture analysis. **Pattern Recognition Letters**, Elsevier, v. 65, p. 116–123, 2015. Disponível em: <<https://doi.org/10.1016/j.patrec.2015.07.030>>.
- RIBEIRO, T. P. et al. Texture characterization via automatic threshold selection on image-generated complex network. In: **Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications**. Springer, 2015. p. 468–476. Disponível em: <https://doi.org/10.1007/978-3-319-25751-8_56>.
- RIPLEY, B. D. **Pattern recognition and neural networks**. : Cambridge university press, 2007.

- SCHMID, C. Bag-of-features for category classification. **ENS/INRIA Visual Recognition and Machine Learning Summer School Lecture 25-29 July, 2011.**
- SERRA, G.; GRANA, C.; CUCCHIARA, R. Covariance of covariance features for image classification. In: **Proceedings of International Conference on Multimedia Retrieval**. New York, NY, USA: ACM, 2014. (ICMR '14), p. 411–414. ISBN 978-1-4503-2782-4. Disponível em: <<https://doi.org/10.1145/2578726.2578781>>.
- SHINZATO, P. Y.; WOLF, D. F. A road following approach using artificial neural networks combinations. **Journal of Intelligent & Robotic Systems**, Springer, v. 62, n. 3, p. 527–546, 2011. Disponível em: <<https://doi.org/10.1007/s10846-010-9463-2>>.
- SIQUEIRA, F. R. de; SCHWARTZ, W. R.; PEDRINI, H. Multi-scale gray level co-occurrence matrices co-occurrence matrices for texture description. **Neurocomputing**, Elsevier, v. 120, p. 336–345, 2013. Disponível em: <<https://doi.org/10.1016/j.neucom.2012.09.042>>.
- SIVIC, J.; ZISSERMAN, A. Video google: A text retrieval approach to object matching in videos. In: **IEEE. Computer Vision (ICCV 2003), Proceedings of the Ninth IEEE International Conference on**. 2003. p. 1470. Disponível em: <<https://doi.org/10.1109/ICCV.2003.1238663>>.
- SONG, C.; HAVLIN, S.; MAKSE, H. A. Self-similarity of complex networks. **arXiv preprint cond-mat/0503078**, 2005.
- SOUZA, J. R.; OSORIO, F. S.; WOLF, D. F. Vision-based waypoint following using templates and artificial neural networks. **Neurocomputing**, Elsevier, v. 107, p. 77–86, 2013.
- TAMURA, H.; MORI, S.; YAMAWAKI, T. Textural features corresponding to visual perception. **Systems, Man and Cybernetics, IEEE Transactions on**, IEEE, v. 8, n. 6, p. 460–473, 1978.
- TERÇARIOL, C. A. S.; MARTINEZ, A. S. Analytical results for the statistical distribution related to a memoryless deterministic walk: Dimensionality effect and mean-field models. **Physical Review E**, APS, v. 72, n. 2, p. 021–103, 2005.
- TODOROVIC, S.; AHUJA, N. Texel-based texture segmentation. In: **IEEE 12th International Conference on Computer Vision (ICCV)**. 2009. p. 841–848. Disponível em: <<https://doi.org/10.1109/ICCV.2009.5459308>>.
- TUCERYAN, M.; JAIN, A. texture analysis, the handbook of pattern recognition and computer vision , by ch. **Chen, LF. Pau, PSP Wang (ed)** Word Scientific Publishing Co, 1998.
- VARMA, M.; GARG, R. Locally invariant fractal features for statistical texture classification. In: **IEEE. Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on**. 2007. p. 1–8. Disponível em: <<https://doi.org/10.1109/ICCV.2007.4408876>>.
- VARMA, M.; ZISSERMAN, A. A statistical approach to texture classification from single images. **International journal of computer vision**, Springer, v. 62, n. 1-2, p. 61–81, 2005.

VIDAL, R.; RAVICHANDRAN, A. Optical flow estimation & segmentation of multiple moving dynamic textures. In: **IEEE. Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on.** 2005. v. 2, p. 516–521. Disponível em: <<https://doi.org/10.1109/CVPR.2005.263>>.

VISION Texture Database. 2009. <<http://vismod.media.mit.edu/vismod/imagery/VisionTexture/vistex.html>>. Accessado: 01-03-2018.

WAN, S. et al. Integrated local binary pattern texture features for classification of breast tissue imaged by optical coherence microscopy. **Medical image analysis**, Elsevier, v. 38, p. 104–116, 2017.

XU, D.; TIAN, Y. A comprehensive survey of clustering algorithms. **Annals of Data Science**, Springer, v. 2, n. 2, p. 165–193, 2015. Disponível em: <<https://doi.org/10.1007/s40745-015-0040-1>>.

XU, Y. et al. Scale-space texture description on sift-like textons. **Computer Vision and Image Understanding**, Elsevier, v. 116, n. 9, p. 999–1013, 2012. Disponível em: <<https://doi.org/10.1016/j.cviu.2012.05.003>>.

XU, Y.; JI, H.; FERMÜLLER, C. Viewpoint invariant texture description using fractal analysis. **International Journal of Computer Vision**, Springer, v. 83, n. 1, p. 85–100, 2009. Disponível em: <<https://doi.org/10.1007/s11263-009-0220-6>>.

XU, Y. et al. Classifying dynamic textures via spatiotemporal fractal analysis. **Pattern Recognition**, Elsevier, v. 48, n. 10, p. 3239–3248, 2015. Disponível em: <<https://doi.org/10.1016/j.patcog.2015.04.015>>.

YANG, J. et al. Evaluating bag-of-visual-words representations in scene classification. In: **ACM. Proceedings of the international workshop on Workshop on multimedia information retrieval.** 2007. p. 197–206. Disponível em: <<https://doi.org/10.1145/1290082.1290111>>.

ZHANG, D.; CHEN, Y. Q. Artificial life: a new approach to texture classification. **International Journal of Pattern Recognition and Artificial Intelligence**, World Scientific, v. 19, n. 03, p. 355–374, 2005. Disponível em: <<https://doi.org/10.1142/S021800140500406X>>.

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.

ZHANG, J. et al. Local features and kernels for classification of texture and object categories: A comprehensive study. **International journal of computer vision**, Springer, v. 73, n. 2, p. 213–238, 2007.

ZHAO, G.; PIETIKAINEN, M. Dynamic texture recognition using local binary patterns with an application to facial expressions. **IEEE transactions on pattern analysis and machine intelligence**, IEEE, v. 29, n. 6, p. 915–928, 2007.

ZHAO, Y. et al. Completed robust local binary pattern for texture classification. **Neurocomputing**, v. 106, n. 0, p. 68–76, 2013. ISSN 0925-2312. Disponível em: <<http://dx.doi.org/10.1016/j.neucom.2012.10.017>>.

ZHU, S.-C. et al. What are textons? **International Journal of Computer Vision**, Springer, v. 62, n. 1-2, p. 121-143, 2005. Disponível em: <<https://doi.org/10.1007/s11263-005-4638-1>>.