
**Caminhadas Determinísticas em Redes
Complexas Aplicadas em Visão
Computacional**

Wesley Nunes Gonçalves

SERVIÇO DE PÓS-GRADUAÇÃO DO ICMC-USP

Data de Depósito:

Assinatura: _____

Caminhadas Determinísticas em Redes Complexas Aplicadas em Visão Computacional

Wesley Nunes Gonçalves

Orientador: *Prof. Dr. Odemir Martinez Bruno*

Dissertação apresentada ao Instituto de Ciências Matemáticas e de Computação - ICMC-USP, para o Exame de Qualificação, como parte dos requisitos necessários à obtenção do título de Mestre em Ciências da Computação e Matemática Computacional.

USP - São Carlos
Janeiro/2010

Aos meus pais, Osmar e Marlene

Agradecimentos

Aos meus pais Osmar e Marlene, e aos meus irmãos Diogo e Pâmela, pelo incentivo e amor que manifestaram ao longo da minha vida.

Ao meu orientador prof. Odemir Martinez Bruno, por participar ativamente na minha formação acadêmica, pela orientação e amizade.

A Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), pela concessão da bolsa de Mestrado.

Aos amigos de pós-graduação pelos momentos de alegria, discussões e trocas de conhecimentos.

Ao Instituto de Ciências Matemáticas e de Computação (ICMC-USP), aos professores e funcionários e a todos que, direta ou indiretamente, colaboraram neste trabalho.

Resumo

As redes complexas têm recebido um crescente interesse nas mais diversas áreas do conhecimento. Esse crescimento se deve principalmente a sua flexibilidade em modelar e simular estruturas topológicas que aparecem em nosso cotidiano. Na maioria das vezes, a caracterização das redes complexas é baseada em medidas básicas, como média dos graus, graus hierárquicos, coeficiente de aglomeração, entre outras. Muitas das medidas propostas são correlacionadas, implicando em redundância. Este trabalho propõe o uso das caminhadas determinísticas do turista como uma medida de representação robusta e eficiente de redes complexas. Nesta medida, caminhadas são iniciadas por exploradores que partem de um dos vértices da rede e em seguida, informações são extraídas sobre essas caminhadas. Experimentos foram realizados em redes complexas artificiais e em redes modelando imagens de textura. No reconhecimento de redes artificiais, o método proposto foi aplicado em quatro modelos de redes complexas teóricos: redes aleatórias, pequeno-mundo, livre de escala e geográficas. No reconhecimento de textura, o método foi avaliado em bancos de texturas sintéticas e reais (texturas de folhas de plantas). Em ambas as aplicações, o método alcançou excelentes resultados comparados com o estado da arte.

Palavras-chave: Caminhadas determinísticas; Redes complexas; Visão Computacional; Reconhecimento de textura;

Abstract

Complex networks have received a growing interest in several areas of knowledge. This growth is mainly due to its flexibility in modeling and simulating topological structures that appear in our daily life. In most cases, complex networks characterization are based on basic measurements such as average degree, hierarchical degree, clustering coefficient, among others. Many of the measures are correlated, resulting in redundancy. This dissertation proposes the use of deterministic walks as a robust and efficient complex network measurement. In this measurement, walks are initiated by explorers starting from each vertex and then, informations are extracted on these walks. Experiments were performed on artificial complex networks and network modeling texture images. In artificial network recognition, the proposed method was applied to four theoretical complex network models: random, small-world, free-scale and geographical networks. In texture recognition, the method was evaluated in synthetic and real (texture of leaves) databases. In both applications, the method achieved excellent results compared with the state of the art methods.

Keywords: Deterministic walks; Complex networks; Computer Vision; Texture recognition;

Sumário

Sumário	xiv
Lista de Figuras	xviii
Lista de Tabelas	xx
1 Introdução	1
1.1 Contextualização e Motivação	1
1.2 Objetivos	3
1.3 Organização do Texto	4
2 Redes Complexas	7
2.1 Redes Aleatórias	10
2.2 Redes Pequeno Mundo	10
2.3 Redes Livres de Escala	11
2.3.1 Redes Geográficas	12
2.4 Medidas	13
2.4.1 Conectividade	14
2.4.2 Coeficiente de Aglomeração	14
2.4.3 Grau Hierárquico	15
3 Caminhadas Determinísticas	17
3.1 Estratégia para Detecção de Atratores	20
3.2 Dinâmica da Caminhada Determinística do Turista	22
3.3 Critérios de Desempate	23
3.3.1 Estratégia Primeira	25
3.3.2 Estratégia Aleatória	25
3.3.3 Estratégia Visibilidade	26
3.4 Extração de Informações	27
3.4.1 Região	27
3.4.2 Histograma	27
3.4.3 Propriedades Estatísticas	28

3.5	Complexidade Computacional	29
3.6	DTW em Redes Complexas	30
4	Reconhecimento de Redes Complexas Artificiais	33
4.1	Experimentos	33
4.1.1	Avaliação dos Parâmetros da DTW	34
4.1.2	Comparação com Medidas Tradicionais	39
4.2	Conclusão	40
5	Reconhecimento de Textura	41
5.1	Modelagem de Imagens com Redes Complexas	41
5.1.1	Redes Complexas em Imagens	42
5.1.2	Vetor de Características	45
5.2	Experimentos e Resultados	48
5.2.1	Análise dos Parâmetros do Método	49
5.2.2	Comparação com Métodos da Literatura	55
5.2.3	Tolerância a Ruídos	61
5.3	Conclusão	64
6	Conclusão	67
6.1	Contribuições	68
6.2	Trabalhos Futuros	68
6.3	Trabalhos Gerados	69
	Referências	84

Lista de Figuras

2.1	Rede aleatória de Erdős e Rényi. Distribuição do grau para redes complexas construídas com 1000 vértices e probabilidade $p = 0.02$.	11
2.2	Modelo de rede complexa pequeno-mundo. Distribuição do grau para redes complexas construídas com 1000 vértices, número de vizinhos $k = 11$ e probabilidade $p = 0.02$.	12
2.3	Rede complexa livre de escala. Distribuição do grau para redes complexas construídas com 1000 vértices e número de aresta $m = 10$.	13
2.4	Rede complexa geográfica. Distribuição do grau para redes complexas construídas com 1000 vértices, escala $\lambda = 0.03$ e espaço dimensional $N \times N, N = 500$.	14
2.5	Exemplo do cálculo do coeficiente de aglomeração. Neste caso $c_1 = \frac{2}{20} = 0.1$	15
2.6	Exemplo do cálculo do nível hierárquico, $k_i^1 = 4$ (arestas em laranja) e $k_i^2 = 13$ (arestas em verde). Note que o grau hierárquico k_i^1 é igual à conectividade do vértice.	16
3.1	Exemplo de uma caminhada do turista. As cidades são representadas pelas esferas. As cidades pertencentes à parte transiente e ao atrator estão em laranja e verde, respectivamente.	18
3.2	Distribuição conjunta do tempo de transiente t e período p com $N = 1000$ e $d = 2$ para dois valores de μ . O aspecto geral de ambos os as distribuições são similares.	19
3.3	Exemplos de caminhadas do Turista. A parte em cinza representa o transiente enquanto a parte listrada representa o atrator.	21
3.4	Lista L utilizada para manter a caminhada.	21
3.5	Comparação entre diferentes partes de uma lista. Uma cidade (cinza) repetida ao menos três vezes na lista pode configurar a existência de seqüências de cidades repetidas.	22

3.6	Lista duplamente encadeada associando cada cidade da caminhada mantida em L	22
3.7	Atratores gerados pelo turista em uma imagem de textura utilizando a distância mínima e diferentes valores de μ . (a) Imagem Original; (b) $\mu = 1$; (c) $\mu = 3$; (d) $\mu = 7$	23
3.8	Atratores gerados pelo turista em uma imagem de textura utilizando a distância máxima e diferentes valores de μ . (a) Imagem Original; (b) $\mu = 1$; (c) $\mu = 3$; (d) $\mu = 7$	23
3.9	Gráfico de probabilidades de ocorrência de empates na caminhada determinística do turista. Número de cidades vizinhas X Probabilidade de empate.	24
3.10	Média do tamanho das caminhadas para diferentes valores de memórias μ e estratégias de desempate aleatória e primeira.	25
3.11	Exemplo da resolução do empate para a estratégia Visibilidade para um turista caminhando em direção da mínima distância. Um empate ocorre entre as cidades 2 e 4 no primeiro nível. Um novo empate ocorre no segundo nível entre as cidades 6 e 10. O desempate ocorre no terceiro nível, optando-se então sair da cidade 1 e ir para a cidade 4.	26
3.12	Ilustração dos valores considerados da distribuição conjunta para a estratégia região com parâmetros a e b	28
3.13	Média do tamanho das caminhadas para diferentes valores de memórias μ	30
4.1	Primeira coluna: Distribuição Conjunta para cada modelo de rede complexa gerada com $N = 1000$ e $\langle k \rangle = 20$. Segunda coluna: Distribuição com uma visão 2D.	35
4.2	Discriminantes do PCA para 4000 redes derivadas dos modelos aleatório, geográfico, pequeno mundo e livre de escala. As redes foram geradas com $N = 1000$ e $\langle k \rangle = 20$ e a aplicação do turista com regras de movimentação min e max	36
4.3	Discriminantes do PCA para os modelos de rede complexa gerada com $N = 1000$ e $\langle k \rangle = 20$ e a aplicação do turista com diferentes memórias.	37
4.4	Discriminantes do PCA para os modelos de rede complexa gerada com $N = 1000$ e $\langle k \rangle = 20$ e a aplicação do turista com combinação das memórias 0, 1, 2, 3, 4 e 5.	38
5.1	Imagem modelada em uma rede complexa utilizando $r = \sqrt{2}$	43
5.2	Transformações ψ com diferentes valores t em uma rede regular com $r = \sqrt{2}$. (a) $t_1 <$ (b) $t_2 <$ (c) t_3	44

5.3	Transformações ψ com diferentes valores t em uma rede regular com $r = \sqrt{8}$. (a) $t_1 <$ (b) $t_2 <$ (c) t_3	44
5.4	Distribuição conjunta obtida pela caminhada determinística do turista em redes regulares transformadas por diferentes limiares. Da esquerda para direita: $t = 10$; $t = 90$ e $t = 170$	45
5.5	Método aplicado em problemas de visão computacional.	47
5.6	Exemplo de algumas classes de texturas de Brodatz utilizadas nos experimentos. Cada imagem possui 200×200 pixels e 256 níveis de cinza.	48
5.7	Exemplo das texturas das folhas de plantas utilizadas nos experimentos. Cada coluna representa uma diferente classe, enquanto as linhas representam a variação de cada classe.	49
5.8	Gráfico do desempenho no reconhecimento de textura. Porcentagem de classificação correta em função do raio r em $\sqrt{\cdot}$ e limiar inicial t_0	50
5.9	(a) Resultados comparativos para diferentes raios ($\sqrt{\cdot}$), (b) Limiar t_0 X Porcentagem de Classificação Correta, (c) Limiar t_{inc} X Porcentagem de Classificação Correta e (d) Limiar t_f X Porcentagem de Classificação Correta.	52
5.10	Visibilidade do turista X porcentagem de classificação correta para a estratégia de desempate.	54
5.11	Avaliação dos parâmetros das estratégias de extração de informações (a) Parâmetro a e b para estratégia Região, (b) Quantidade de valores considerados do histograma, (c) Estatísticas.	56
5.12	Matrizes de confusão para o banco de imagens Brodatz. (a) Descritores de Fourier. (b) Matriz de co-ocorrência.	58
5.13	Matrizes de confusão para o banco de imagens Brodatz. (a) Filtros de Gabor. (b) Método proposto.	59
5.14	Exemplo de algumas imagens de texturas de Brodatz com ruídos utilizadas nos experimentos. Cada linha representa os ruídos com densidades $d = [0,01 \ 0,05 \ 0,07 \ 0,1 \ 0,3 \ 0,5]$ e as colunas representam diferentes classes.	62
5.15	Exemplo de algumas imagens de texturas de folhas com ruídos utilizadas nos experimentos. Cada coluna representa os ruídos com densidades $d = [0,01 \ 0,05 \ 0,07 \ 0,1 \ 0,3 \ 0,5]$ e as colunas representam diferentes classes.	63
5.16	Imagem com ruído modelada por uma rede complexa.	65
1	(a) Imagem Original; (b) Distribuição de probabilidade conjunta do tempo de transiente e período do atrator para $\mu = 5$	71

2	Atratores gerados pelo turista em uma imagem gerada aleatoriamente utilizando a mínima distância e diferentes valores de μ : (a) Imagem original; (b) $\mu = 1$; (c) $\mu = 3$; (d) $\mu = 7$	72
3	Atratores gerados pelo turista em uma imagem gerada aleatoriamente utilizando a máxima distância e diferentes valores de μ : (a) Imagem original; (b) $\mu = 1$; (c) $\mu = 3$; (d) $\mu = 7$	72
4	Atratores gerados pelo turista em uma imagem de textura utilizando a mínima distância e diferentes valores de μ . Atratores são misturados com textura para melhor compreensão: (a) Imagem original; (b) $\mu = 1$; (c) $\mu = 3$; (d) $\mu = 7$	72
5	Atratores gerados pelo turista em uma imagem de textura utilizando a máxima distância e diferentes valores de μ . Atratores são misturados com textura para melhor compreensão: (a) Imagem original; (b) $\mu = 1$; (c) $\mu = 3$; (d) $\mu = 7$	73
6	Histograma do tamanho do transiente para $\mu = 4$	73

Lista de Tabelas

4.1	Porcentagem de classificação correta para $\psi_{\mu,din}$ utilizando diferentes valores de μ e regras de movimentação na caminhada determinística do turista.	34
4.2	Porcentagem de classificação correta para φ combinando diferentes valores de μ no banco de imagens Brodatz.	38
4.3	Comparação dos resultados para diferentes medidas extraídas das redes complexas.	39
4.4	Comparação dos resultados para a combinação entre todas as medidas.	39
4.5	Matriz de confusão para as caminhadas determinísticas do turista.	40
4.6	Matriz de confusão para as medidas tradicionais.	40
4.7	Matriz de confusão para as caminhadas determinísticas do turista em conjunto com as medidas tradicionais.	40
5.1	Análise conjunta entre limiar inicial t_0 e raio r	51
5.2	Porcentagem de classificação correta para $\psi_{\mu,din}$ utilizando diferentes valores de μ e dinâmica din na caminhada do turista. . . .	53
5.3	Porcentagem de classificação correta para φ combinando diferentes valores de μ	53
5.4	Comparação entre as estratégias de desempate.	54
5.5	Comparação entre diferentes estratégias para extração de informações da distribuição conjunta.	55
5.6	Comparação dos resultados para diferentes métodos de caracterização de texturas no banco de imagens Brodatz.	58
5.7	Comparação dos resultados para diferentes métodos na caracterização de texturas no banco de imagens de folhas de plantas. . .	60
5.8	Teste t-pareado para os métodos no banco de imagens de Brodatz.	61
5.9	Teste t-pareado para os métodos no banco de imagens de folhas.	61

5.10	Comparação dos resultados para o banco de Brodatz com ruídos. Resultados são apresentados em % e desvio padrão entre parênteses.	64
5.11	Comparação dos resultados para o banco de Brodatz com ruídos. Resultados são apresentados em % e desvio padrão entre parênteses.	64
5.12	Comparação dos resultados para o banco de folhas com ruídos. Resultados são apresentados em % e desvio padrão entre parênteses.	64
5.13	Comparação dos resultados para o banco de folhas com ruídos. Resultados são apresentados em % e desvio padrão entre parênteses.	65

Introdução

1.1 *Contextualização e Motivação*

A visão computacional é uma área que tem crescido consideravelmente nos últimos anos, principalmente em aplicações em imagens médicas (Solaiman et al., 1998; Gonçalves et al., 2009), inspeção industrial (Stivanello and Gomes, 2006) e sistemas de segurança (Tisse et al., 2002). Esse crescimento se deve, sobretudo, a evolução da visão artificial e seus algoritmos. As técnicas desenvolvidas nessa área são diretamente aplicáveis em problemas reais, auxiliando os seres humanos na resolução de tarefas repetitivas e envolvendo um alto risco para realizá-las. Apesar da extraordinária capacidade de um ser humano em resolver problemas através da visão, em alguns casos, os resultados obtidos por um ser humano podem ser influenciados pela fadiga ou distração durante a realização de uma tarefa. Por outro lado, a resolução de tarefas utilizando visão computacional é particularmente apropriada para registrar variáveis que um ser humano é incapaz de estimar com exatidão, como as informações espaciais de distância percorrida, velocidade, área, volume, entre outras. A visão computacional tem como objetivo construir sistemas artificiais para interpretar imagens, simulando o processo de visão através da transformação de uma imagem em uma estrutura de dados simplificada. O estado da arte dessa área tem como principal característica a não existência de um algoritmo genérico que simule o processo de visão, mas é uma área com pesquisas bastante promissoras envolvendo diversas teorias e métodos, como as redes complexas.

Redes complexas, que englobam o formalismo dos grafos e a análise de

mecânica estatística, têm recebido um crescente interesse nas mais diversas áreas do conhecimento. Esse crescimento se deve principalmente a sua flexibilidade em modelar e simular estruturas topológicas que aparecem em nosso cotidiano. Devido à aquisição de dados em larga escala, o foco das pesquisas tem se transferido da análise de componentes individuais para a análise de redes com milhares ou milhões de componentes interligados (Newman, 2003), o que reforça a utilização das redes complexas. As redes complexas recebem esse nome, pois apresentam algumas propriedades topológicas não triviais, como a distribuição do grau dos vértices em leis de potência.

Diversas análises com as redes complexas são realizadas em duas etapas. A modelagem do problema em uma rede complexa e a extração de características topológicas. Como demonstrado na literatura, diversas redes presentes no mundo real possuem propriedades topológicas em comum, i.e. medidas de graus similares. Um problema que tem recebido uma considerável atenção é a classificação dessas redes de acordo com as medidas topológicas extraídas (Chalumeau et al., 2007; Backes et al., 2009a) em grupo de modelos já estabelecidos e com diversas propriedades conhecidas. Entre as aplicações estão a identificação de diferentes modelos teóricos de redes complexas (e.g. redes aleatórias (Erdős and Rényi, 1959; Solomonoff and Rapoport, 1951), pequeno-mundo (Watts and Strogatz, 1998; Watts, 1999), livre de escala (Barabasi and Albert, 1999) e redes geográficas (Gastner and Newman, 2004; Hayashi, 2006)) e a aplicação das redes complexas em reconhecimento de padrões (Chalumeau et al., 2007; Backes et al., 2009a; Gonçalves et al., 2009).

Para a classificação robusta das redes complexas, medidas representativas precisam ser extraídas. O principal problema está na escolha do conjunto de medidas adequadas em um determinado problema. Atualmente, diversas medidas foram propostas na literatura, como média dos graus, graus hierárquicos, coeficiente de aglomeração, entre outras. O artigo (da F. Costa et al., 2007) apresenta uma boa discussão sobre diferentes medidas em redes complexas. Muitas das medidas propostas são correlacionadas, implicando em redundância. Uma alternativa a essa redundância, embora isso não garanta resultados ótimos, é o uso de métodos estatísticos para selecionar um conjunto de medidas adequadas (e.g. análise discriminante linear e análise de componentes principais). Devido a essas dificuldades, o que se tem feito atualmente é basear-se no conhecimento de um especialista sobre o problema para selecionar um conjunto de medidas adequadas (da F. Costa et al., 2007).

Este trabalho propõe o uso das caminhadas determinísticas como uma métrica de representação robusta e eficiente de redes complexas. Nesta métrica, caminhadas são iniciadas por exploradores que partem de um dos vértices da rede e em seguida, informações são extraídas sobre essas caminhadas. As

caminhadas seguem a regra determinística, onde cada caminhante vai para o vértice mais próximo que não tenha sido visitado nos últimos μ passos. A regra dessa caminhada, denominada caminhada determinística do turista, pode ser modificada para ir em direção ao vértice mais próximo ou mais distante do vértice atual. Entretanto, uma vez escolhida a regra de movimentação, essa não deve ser alterada durante a caminhada. Seguindo esses procedimentos, uma otimização local é alcançada em uma estrutura da rede e padrões apresentados nessa rede são identificados. Apesar da simplicidade do modelo da caminhada determinística, a dinâmica se torna complexa devido à combinação das informações das caminhadas iniciadas em cada vértice e com isso, os resultados obtidos não são triviais.

A principal motivação para o desenvolvimento desse projeto foi a necessidade de explorar e avaliar as redes complexas e caminhadas determinísticas. Devido à flexibilidade das redes complexas em modelar problemas, a combinação entre ambas as teorias fornece um método genérico para reconhecimento de redes, sendo aplicável em diversas áreas, como reconhecimento de padrões e visão computacional. Embora muitas pesquisas estejam sendo desenvolvidas nessas áreas, ainda existe um amplo campo de pesquisa inexplorado a ser desenvolvido quanto à utilização das redes complexas e caminhadas determinísticas, desde a utilização desses métodos em aplicações de reconhecimento de padrões e visão computacional à validação quando aplicados em imagens.

1.2 *Objetivos*

Este trabalho teve como principal objetivo desenvolver pesquisas no âmbito teórico e prático das caminhadas determinísticas do turista em redes complexas, contribuindo com o desenvolvimento desses conceitos em diversas áreas e propondo uma nova medida de caracterização de redes complexas.

Em termos teóricos, o objetivo principal foi propor e avaliar um modelo que combine ambas as teorias. Nesse modelo, as caminhadas determinísticas são utilizadas como uma métrica de caracterização de redes complexas. Como objetivos secundários, avaliar metodologias para a resolução de algumas lacunas das caminhadas determinísticas. Mais precisamente, avaliar critérios de desempate nas caminhadas determinísticas do Turista (quando é equiprovável o explorador seguir em dois ou mais caminhos diferentes). Outro objetivo secundário foi avaliar estratégias de extração de informações sobre um conjunto de caminhadas. Atualmente essa extração é realizada em função de apenas uma pequena parte da distribuição conjunta do tempo de transiente e do período. A região da distribuição é selecionada empiricamente e o restante das informações é descartado. O objetivo foi avaliar outras formas de extração de

informações sem desconsiderar partes da distribuição. Por fim, em termos teóricos, propor uma maneira de representar uma imagem com a metodologia das redes complexas, considerando que as ligações entre esses vértices sejam realizadas de acordo com propriedades visuais presentes na imagem.

Em termos práticos, implementar e avaliar as caminhadas determinísticas do turista em classificação de redes complexas artificiais e em problemas de visão computacional. As redes artificiais foram geradas com base em quatro modelos bem conhecidos na literatura, enquanto os problemas de visão computacional abordados foram: reconhecimento de texturas sintéticas (Backes et al., 2009b; Chalumeau et al., 2007) e classificação de espécies de plantas baseado em textura (de Oliveira Plotze and Bruno, 2009; Bruno et al., 2008a). Por fim, comparar a metodologia desenvolvida com os principais métodos citados na literatura que resolveram esses problemas eficientemente.

Resumidamente, os objetivos específicos desse trabalho foram:

- Propor formalmente as caminhadas determinísticas em redes complexas.
- Avaliar critérios de desempate para as caminhadas determinísticas do turista.
- Avaliar estratégias de extração de informações sobre um conjunto de caminhadas.
- Aplicar e avaliar o método em redes complexas artificiais.
- Comparar os resultados com medidas tradicionais de redes complexas.
- Propor uma representação de imagens com redes complexas.
- Aplicar o método em problemas de visão computacional.
- Comparar o método com o estado da arte em cada problema.

1.3 Organização do Texto

O texto é descrito como a seguir. O Capítulo 2 aborda os conceitos das redes complexas em conjunto com os trabalhos correlatos sobre esse tema. Além disso, alguns modelos de redes complexas e medidas tradicionais são apresentados. No Capítulo 3, as caminhadas determinísticas são apresentadas tanto sob o ponto de vista teórico, com um método para extrair padrões em um ambiente regular, como em termos de aplicações, com alguns trabalhos correlatos dessa teoria aplicados em reconhecimento de padrões. Ainda neste Capítulo, um modelo matemático para as caminhadas determinísticas

do turista em redes complexas é apresentado. O Capítulo 4 apresenta os experimentos no reconhecimento de redes complexas artificiais, comparando os resultados com medidas tradicionais. Os experimentos no reconhecimento de texturas são apresentados no Capítulo 5. O método proposto para representação de imagens com redes complexas e a extração de informações com caminhadas determinísticas em redes complexas é apresentado. Nos experimentos, dois bancos de texturas foram utilizados na classificação, sendo os resultados comparados com o estado da arte em análise de textura. Por fim, no Capítulo 6, as conclusões e as contribuições são discutidas e alguns trabalhos futuros são sugeridos.

Redes Complexas

Redes complexas é uma teoria recente que engloba o formalismo dos grafos e a análise de mecânica estatística. Pesquisas em redes complexas estão estritamente ligadas com as pesquisas em grafos. Na verdade, a distinção entre grafos e redes complexas é sutil. O que realmente difere é a abordagem adotada. Por um lado, os grafos são bastante tradicionais na ciência da computação e matemática e, por outro lado, redes complexas emergem da física e usam a abordagem estatística para fazer a análise de seus componentes.

As redes complexas são representadas através de grafos compostos por um conjunto de vértices interligados por um conjunto de arestas. O grafo é descrito por $G(V, E)$, onde $V = \{v_1, v_2, \dots, v_N\}$ é o conjunto de vértices, $E = \{e_1, e_2, \dots, e_M\}$ é o conjunto de arestas, N e M são o número de vértices e arestas, respectivamente. Computacionalmente, os grafos podem ser implementados através de uma matriz de adjacência A . Cada componente a_{ij} da matriz armazena o valor 1 se existir uma aresta ligando o vértice i ao vértice j ou armazena o valor 0 caso contrário. Grafos podem também possuir arestas com pesos. Neste caso, uma matriz W é utilizada e cada elemento w_{ij} representa o peso da aresta que liga o vértice i ao vértice j .

As redes complexas geralmente possuem um conjunto de vértices e arestas compostos por um grande número de componentes e propriedades topológicas particulares, como a distribuição do grau dos vértices (Boccaletti et al., 2006), dada pelas Equações 2.1 e 2.2.

$$D(j) = \sum_{i=1}^N \begin{cases} 1, & \text{se } j == k_i \\ 0, & \text{caso contrário} \end{cases} \quad (2.1)$$

$$k_i = \sum_{j=1}^N a_{ij} \quad (2.2)$$

onde k_i é o grau do vértice i , e D a distribuição dos graus.

Uma das utilidades da distribuição dos graus dos vértices é sua capacidade em diferenciar diversos modelos de redes e decidir a sua natureza (aleatória ou seguindo determinada lei de formação). Outra característica importante em redes complexas é o caminho mínimo médio (Boccaletti et al., 2006). Essa característica está relacionada com a estrutura global de uma rede. Para a definição dessa característica, é preciso definir o caminho mínimo. Um caminho mínimo d_{ij} é o número mínimo de arestas que interligam dois vértices i e j . O caminho mínimo pode ser calculado por diversos algoritmos propostos na literatura, como o algoritmo de Dijkstra (Zhan and Noon, 1998). Dado essas definições, o caminho mínimo médio de um vértice i é dado pela distância média dos caminhos mínimos que interligam o vértice i a todos os outros vértices (Equação 2.3). Caso não exista um caminho ligando dois vértices, $d_{ij} = N$.

$$l_i = \frac{1}{N-1} \sum_{j=1, j \neq i}^N d_{ij} \quad (2.3)$$

Os primeiros trabalhos, utilizando redes complexas, buscaram definir os conceitos para caracterizar a topologia de redes do mundo real (Newman, 2003). Com isso, foi possível identificar princípios e propriedades estatísticas em comum na maioria das redes reais avaliadas. Devido a sua capacidade de generalização e caráter multidisciplinar, as redes têm sido aplicadas nos mais variados problemas, como redes sociais (Barabasi et al., 2002; Newman et al., 2002), redes tecnológicas (Sen et al., 2003) e redes biológicas (West et al., 1999).

Em (Barabasi et al., 2002), as redes sociais entre autores e co-autores de artigos científicos foram avaliadas. As redes foram construídas com os dados disponíveis na base eletrônica de relevantes jornais de publicações. Neste trabalho, a estrutura e a dinâmica deste sistema complexo foram estudadas, concluindo-se que a rede apresentou características de uma rede livre de escala, a média dos graus dos vértices é incrementada com o tempo e com isso a distância entre dois vértices quaisquer é decrementada. O estudo para prevenir a proliferação de vírus de computador por e-mails foi realizado em (Newman et al., 2002). A estrutura da rede, formada pelos endereços de e-mail e os contatos de várias pessoas, foi empiricamente investigada. Uma estratégia para prevenção é remover vértices altamente conectados, identificados por medidas extraídas da rede complexa. Esses vértices representam os responsáveis

por grande parte da epidemia de vírus de computador.

Em redes tecnológicas, foram estudadas as propriedades estruturais das linhas de ferrovias (Sen et al., 2003). As estações de trem foram representadas como vértices na rede complexa. A inserção de arestas foi executada da seguinte maneira. Dado dois vértices selecionados de forma aleatória, uma aresta era inserida se pelo menos um trem parasse em ambas as estações representadas pelos dois vértices. Após a análise desses dados, a rede complexa apresentou propriedades das redes pequeno-mundo, que possui alta conectividade entre os vértices como principal característica.

Embora as redes complexas tenham sido aplicadas com sucesso nestes exemplos, ainda existe um amplo campo de pesquisa a ser desenvolvido em diversas áreas, como a visão computacional e reconhecimento de padrões. Nestas áreas, as redes complexas ainda estão inexploradas e poucas pesquisas são encontradas na literatura (da Fontoura Costa, 2004; Backes et al., 2007; Silva and Zhao, 2007; Backes et al., 2009a).

Em (Backes et al., 2007) foi apresentado um método para calcular a aproximação poligonal de contornos de objetos. O contorno foi modelado como uma rede complexa pequeno-mundo e a aproximação poligonal foi obtida através do menor ciclo encontrado na rede. O método apresentou bons resultados, demonstrando superioridade sobre a aproximação por curvatura. A rede complexa pequeno-mundo também foi utilizada em (Backes et al., 2009a) para caracterizar a forma de objetos. Após a modelagem da forma com uma rede complexa, algumas medidas baseadas no grau da rede são calculadas e utilizadas no reconhecimento de espécies de folhas. O método apresentou tolerância a ruído e invariância à escala e rotação.

A clusterização de pixels em imagens através da identificação de comunidades em redes complexas foi realizada em (Silva and Zhao, 2007). A grande vantagem dessa abordagem está na estrutura topológica das redes, pois são obtidas não somente propriedades estatísticas, mas também propriedades baseadas nas interações entre os elementos. Os resultados são provenientes de imagens artificiais e imagens reais. Uma interessante análise de textura foi realizada em (Chalumeau et al., 2007) e (Gonçalves et al., 2009). Nessa abordagem, a imagem é modelada como uma rede complexa e a média do grau dos vértices é extraída para caracterizar a rede, e conseqüentemente as diferentes texturas nas imagens. A metodologia foi comparada com os filtros de Gabor e matriz de co-ocorrência e obteve resultados superiores. A metodologia dos artigos acima também foi aplicada no reconhecimento de faces (Gonçalves et al.,). A metodologia mostrou-se invariante a rotação, obtendo resultados superiores quando comparados com os métodos tradicionais da literatura.

As pesquisas em redes complexas evoluíram principalmente após os con-

ceitos dos modelos de redes pequeno-mundo (Watts and Strogatz, 1998) e redes livres de escala (Barabasi and Albert, 1999). Além desses dois modelos, são apresentadas a seguir, as redes aleatórias (Barabasi and Albert, 1999) e as redes geográficas (Albert et al., 2004; Hayashi, 2005).

2.1 Redes Aleatórias

As redes aleatórias foram originalmente propostas por Erdős e Rényi (Erdős and Rényi, 1959) e Solomonoff e Rapoport (Solomonoff and Rapoport, 1951) com o intuito de estudar os grafos com base na teoria das probabilidades. Esse modelo é um dos modelos mais simples de rede complexa. Inicialmente, o conjunto com N vértices estão desconectados e as $\frac{N(N-1)}{2}$ arestas possíveis são adicionadas com probabilidade fixa p . Com isso, todas as arestas possuem a mesma probabilidade de ocorrerem, sendo a estrutura da rede homogênea. A probabilidade p pode assumir valores entre o intervalo $[0, 1]$. Quando $p = 0$, a rede não possui aresta, quando $p = 1$, a rede possui todas as possíveis arestas (completamente conectada). Dependendo da probabilidade p escolhida, é possível que vértices ou subgrafos estejam desconectados na rede.

A maioria das propriedades dos grafos aleatórios é obtida quando se aplica o limite em grafos grandes. Com isso, quando $N \rightarrow \infty$, a distribuição do grau dos vértices se aproxima de uma distribuição de Poisson, com média dada por $\langle k \rangle = p(N - 1)$ e probabilidade p_k de um vértice da rede possuir um grau k dada por:

$$p_k = \binom{N}{k} p^k (1-p)^{N-k} \equiv \frac{\langle k \rangle^k e^{-\langle k \rangle}}{k!} \quad (2.4)$$

Na Figura 2.1 é ilustrada a distribuição do grau dos vértices para redes aleatórias construídas com 1000 vértices e probabilidade $p = 0.02$.

2.2 Redes Pequeno Mundo

Os modelos utilizando redes aleatórias foram originalmente desenvolvidos para buscar uma base teórica, sem interesse em aplicações reais. Watts e Strogatz, estudando as redes aleatórias em aplicações, concluíram que muitas redes do mundo real apresentam vértices altamente conectados. Devido a essa propriedade, um vértice pode ser alcançado pelos demais vértices através de um pequeno conjunto de arestas.

Outra propriedade interessante dessas redes é a presença de muitos caminhos fechados de tamanho igual a três, i.e. se uma pessoa A é amiga de duas outras pessoas B e C , existe uma alta probabilidade de duas pessoas

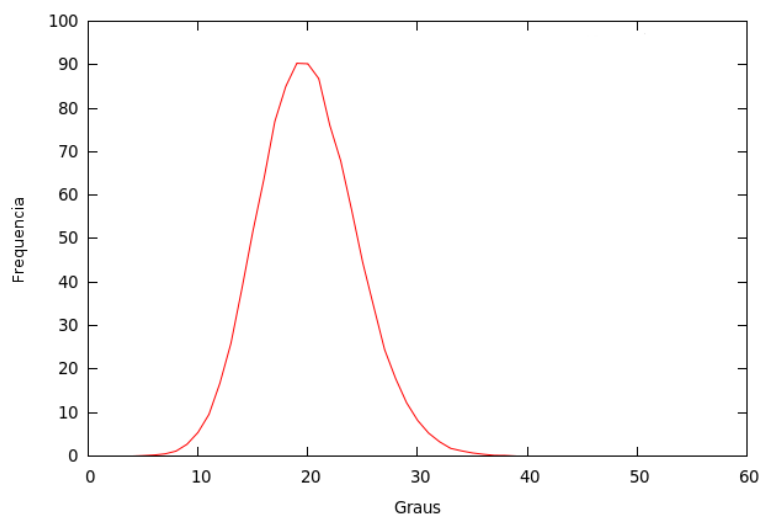


Figura 2.1: Rede aleatória de Erdős e Rényi. Distribuição do grau para redes complexas construídas com 1000 vértices e probabilidade $p = 0.02$.

B e C serem amigas, formando um triângulo. Com isso, um modelo de rede complexa que não possui arestas adicionadas de maneira completamente aleatória, denominada rede pequeno-mundo, foi proposta (Watts, 1999; Watts and Strogatz, 1998).

A rede pequeno-mundo é construída com N vértices conectados de forma regular com os k vizinhos mais próximos, sendo $N \gg k \gg \log N \gg 1$. Após a obtenção da rede regular, cada aresta é reconectada com probabilidade p . A probabilidade p introduz a aleatoriedade ao modelo, quando $p = 0$ a rede é completamente regular e quando $p = 1$ a rede é aleatória. O modelo pequeno-mundo está localizado entre a completa regularidade e aleatoriedade.

A distribuição do grau dos vértices pode ser modelada por uma distribuição de Poisson com um pico em $\langle k \rangle = 2k$, indicando características das redes aleatórias. A principal característica das redes pequeno-mundo é o baixo valor do caminho mínimo médio. Na Figura 2.2 é ilustrada a distribuição do grau para as redes pequeno mundo.

2.3 Redes Livres de Escala

Barabási e Albert mostraram que a distribuição do grau de muitas outras redes do mundo real é uma distribuição desigual, propondo um novo modelo chamado redes livres de escala (Barabasi and Albert, 1999). Nesse modelo, as conexões dos vértices não são aleatórias, mas seguindo o tipo livre de escala na forma $P(k) \sim k^{-\gamma}$, com $\gamma = 3$ para $N \rightarrow \infty$.

As distribuições geradas de forma aleatória apresentam uma escala descrita, geralmente, por uma média e um desvio padrão. Entretanto, diversos

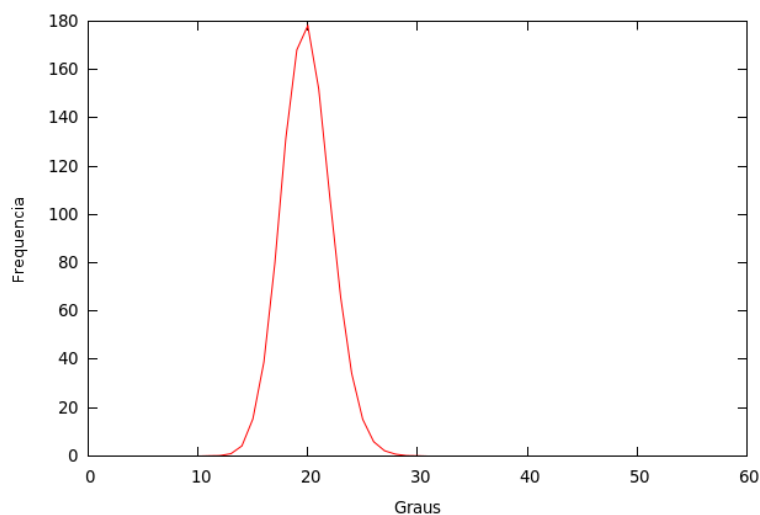


Figura 2.2: Modelo de rede complexa pequeno-mundo. Distribuição do grau para redes complexas construídas com 1000 vértices, número de vizinhos $k = 11$ e probabilidade $p = 0.02$.

eventos do mundo real não possuem uma média e desvio padrão, possuindo invariância a escala. As redes livres de escala apresentam características de uma lei de potência com curva continuamente decrescente, sem uma moda característica, sendo descrita por um único expoente.

Essas redes são construídas através de duas etapas. Inicialmente, a rede é iniciada com um pequeno número de vértices N_0 e em seguida são inseridos os $N - N_0$ vértices. Para cada novo vértice i , $m \leq N_0$ arestas são inseridas conectando o novo vértice a algum vértice j presente na rede. O vértice j é escolhido proporcionalmente a probabilidade do seu grau, conforme a Equação 2.5.

$$P_{i \rightarrow j} = \frac{k_j}{\sum_{j=0}^M k_j} \quad (2.5)$$

onde M é o número atual de vértices na rede complexa.

Nos modelos anteriores (aleatório e pequeno-mundo) os vértices são conectados de maneira homogênea com média e desvio padrão, enquanto que, nesse modelo os vértices com maior conectividade possuem probabilidade maior de receber conexões. Um exemplo da distribuição de graus desse tipo de rede é ilustrado na Figura 2.3.

2.3.1 Redes Geográficas

A grande maioria das redes complexas modela problemas em um espaço abstrato, onde a posição espacial de cada vértice não é relevante. Entretanto, existem algumas redes denominadas redes geográficas ou espaciais, onde a posição do vértice torna-se essencial, influenciando na topologia e evolução

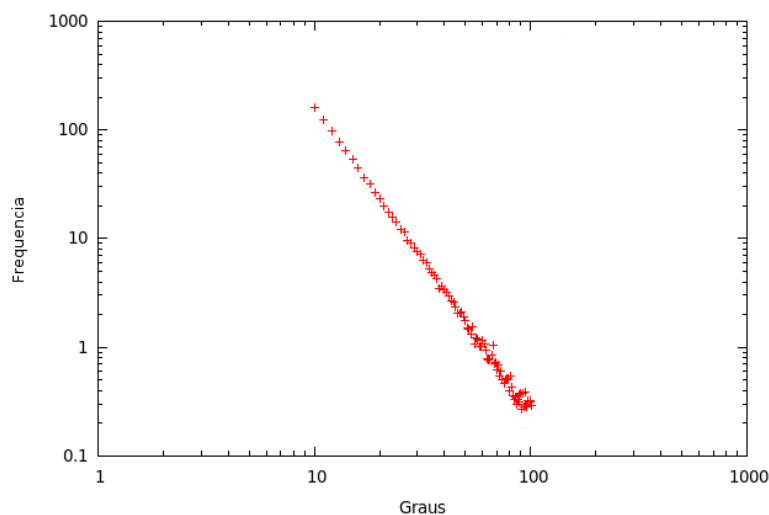


Figura 2.3: Rede complexa livre de escala. Distribuição do grau para redes complexas construídas com 1000 vértices e número de aresta $m = 10$.

da rede.

Essas redes geralmente são construídas seguindo a topologia espacial da rede real a ser modelada, podendo ser influenciadas pela distância entre os vértices, limites geográficos, etc. Para gerar uma rede artificial, os vértices são distribuídos de forma aleatória em um espaço 2D e as arestas são inseridas com base em uma probabilidade que decai proporcionalmente a distância entre os vértices dada por:

$$P_{i \rightarrow j} = e^{-\lambda d_{ij}} \quad (2.6)$$

onde λ fixa a escala das arestas e d_{ij} é a distância espacial entre os vértices i e j .

A Figura 2.4 apresenta a distribuição do grau da rede. Esse tipo de rede segue uma distribuição de Poisson como observado nas redes aleatórias e pequeno-mundo.

2.4 Medidas

As redes complexas podem ser classificadas nos modelos apresentados anteriormente através de diversas medidas. A escolha dessas medidas é essencial, pois medidas que retornem valores redundantes ou estatisticamente similares para diferentes classes podem influenciar na classificação. Nas subseções a seguir são descritas as medidas mais utilizadas relacionadas à conectividade, coeficiente de aglomeração e grau hierárquico.

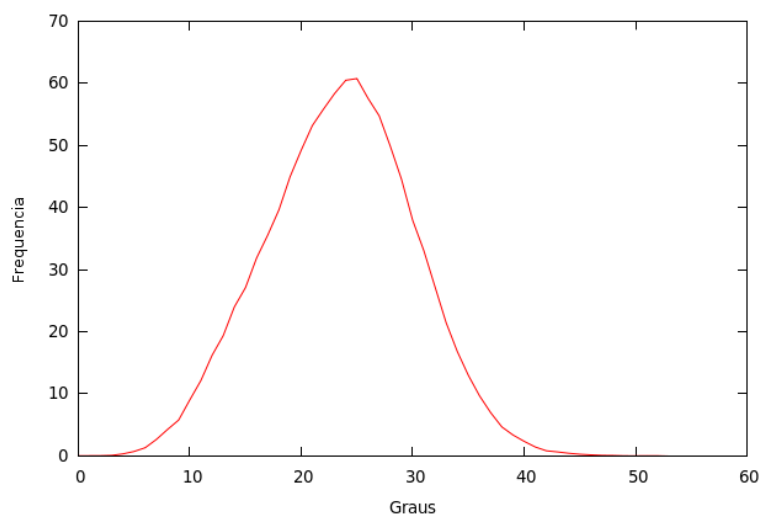


Figura 2.4: Rede complexa geográfica. Distribuição do grau para redes complexas construídas com 1000 vértices, escala $\lambda = 0.03$ e espaço dimensional $N \times N$, $N = 500$.

2.4.1 Conectividade

As redes complexas se diferenciam dos grafos por possuírem um conjunto de vértices e arestas compostos por um grande número de componentes e propriedades topográficas particulares, como o número de conexões entre os vértices (Boccaletti et al., 2006). O grau k_i de um vértice é dado pelo número de arestas incidentes a ele, calculado através da Equação 2.7.

$$k_i = \sum_{j=1}^N a_{ij} \quad (2.7)$$

O grau com pesos kw_i é calculado de maneira similar, mas somando-se o peso das arestas da matriz de pesos W , Equação 2.8.

$$kw_i = \sum_{j=1}^N w_{ij} \quad (2.8)$$

Essa medida é interessante, pois através da distribuição do grau dos vértices é possível diferenciar diversos modelos de redes complexas e decidir a sua natureza (aleatória ou seguindo determinada lei de formação) de diversas redes.

2.4.2 Coeficiente de Aglomeração

O coeficiente de aglomeração de um vértice mede a relação entre o número de arestas dos seus vizinhos e o número máximo de arestas possíveis entre eles. Essa medida pode ser analisada como a probabilidade de dois vértices

vizinhos serem conectados.

Através dessa medida, é possível mensurar a conectividade dos vizinhos de um vértice, uma medida interessante quando se pretende quantificar texturas em imagens, pois a textura está relacionada com a interação de um *pixel* e de seus vizinhos através da repetição de algum padrão. Com a utilização da matriz de adjacência, essa medida é calculada através da Equação 2.9.

$$c_i = \frac{\sum_{j=1}^N \sum_{k=1}^N a_{ij} a_{jk} a_{ki}}{k_i(k_i - 1)} \quad (2.9)$$

O coeficiente de aglomeração com pesos é calculado de maneira similar através da Equação 2.10.

$$cw_i = \frac{1}{kw_i(k_i - 1)} \sum_{j=1}^N \sum_{k=1}^N \frac{w_{ij} + w_{ik}}{2} a_{ij} a_{jk} a_{ki} \quad (2.10)$$

Na Figura 2.5, um exemplo do cálculo do coeficiente de aglomeração é apresentado. O coeficiente de aglomeração para o vértice 1 é dado por $c_1 = \frac{2}{20} = 0.1$.

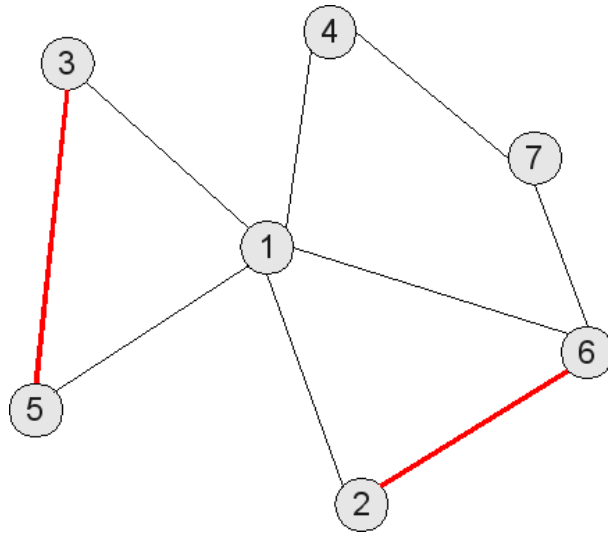


Figura 2.5: Exemplo do cálculo do coeficiente de aglomeração. Neste caso $c_1 = \frac{2}{20} = 0.1$

2.4.3 Grau Hierárquico

As redes complexas podem ser caracterizadas de maneira hierárquica e basicamente todas as medidas podem ser quantificadas através da hierarquização. O nível hierárquico é definido sobre uma dilatação $\delta(g)$, resultando em um subgrafo que contém os vértices de g mais os vértices conectados a cada vértice de g . Com isso, a d -dilatação de um grafo g é a aplicação $\delta(g)$ por d vezes como demonstrado na Equação 2.11.

$$\delta_d(g) = \underbrace{\delta(\delta(\dots(g)\dots))}_d \quad (2.11)$$

O d -anel de um subgrafo g , definido como $R_d(g)$, é composto pelos vértices e arestas de acordo com a Equação 2.12.

$$R_d(g) = V(\delta_d(g)) - V(\delta_{d-1}(g)) \quad (2.12)$$

onde $V(\cdot)$ é conjunto de vértices e $-$ é a operação de diferença de conjuntos. As arestas são obtidas pelas conexões entre os vértices resultantes da operação.

O grau hierárquico k_i^d é definido através do número de arestas que conecta os anéis $R_d(i)$ e $R_{d-1}(i)$ sobre todos os vértices i presentes na rede. De forma análoga, o grau hierárquico com pesos kw_i^d é definido como a soma dos pesos das arestas que conectam os anéis em questão. O exemplo de cálculo do grau hierárquico 1 e 2 é apresentado na Figura 2.6.

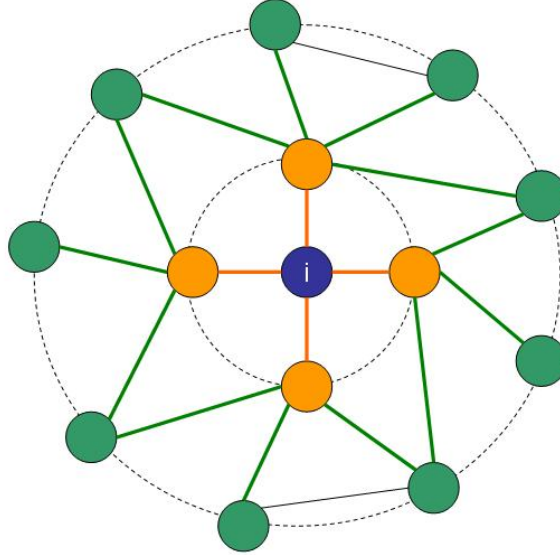


Figura 2.6: Exemplo do cálculo do nível hierárquico, $k_i^1 = 4$ (arestas em laranja) e $k_i^2 = 13$ (arestas em verde). Note que o grau hierárquico k_i^1 é igual à conectividade do vértice.

Caminhadas Determinísticas

As caminhadas aleatórias em ambientes regulares e desordenados têm sido freqüentemente exploradas e possui uma grande quantidade de aplicações (Metzler and Klafter, 2000; Derrida, 1997). Apesar de não serem exaustivamente estudadas, as caminhadas determinísticas em ambientes regulares (Gale et al., 1995; Bunimovich and Troubetzkoy, 1992) e desordenados (Bunimovich, 2004) também apresentam diversos resultados interessantes. Neste trabalho, uma caminhada determinística *self-avoiding*, denominada caminhada determinística do turista (*deterministic tourist walk* - DTW) é explorada (Campitelli et al., 2006b; Backes et al., 2006).

Essa caminhada determinística pode ser visualizada como um viajante que pode visitar N pontos distribuídos em um mapa de d dimensões. Considere esses pontos como cidades e um turista movimentando-se entre elas seguindo a regra determinística: vá para a cidade mais próxima que não tenha sido visitada nos últimos μ passos. A memória de tamanho μ é vista como o tempo necessário para uma cidade já visitada tornar-se atrativa novamente a caminhada. Através dessa memória μ é possível proibir que a caminhada se intersecte, limitada em uma janela $\tau = \mu - 1$. O turista se movimenta com base em uma tabela de vizinhança. A distância entre as cidade é simplesmente uma maneira de ranquear seus vizinhos, tornando o método invariante a transformações de escala (Campitelli et al., 2006b), pois se a escala for incrementada, todas as distâncias entre as cidades também serão incrementadas proporcionalmente, não tendo assim, influência na caminhada do turista. Essa característica permite a aplicação do método em diversos problemas, como a análise de textura (Backes et al., 2009b). Embora a caminhada possua regras simples, os resultados obtidos apresentam uma dinâmica bastante complexa.

A caminhada (conjunto de cidades) realizada pelo turista possui uma parte transitente de tamanho t e uma parte final composta por um atrator com período $p \geq \mu + 1$. O tempo transitente é o número de cidades que o turista visita, desde a cidade inicial até a primeira cidade do atrator. O atrator é um conjunto de cidades em que o turista é aprisionado, sendo o período do atrator, a cardinalidade desse conjunto. Após ingressar em um atrator, as mesmas cidades sempre serão visitadas, independente do tamanho da caminhada, como pode ser visto na Figura 3.1. Existem alguns casos onde o turista não é capaz de encontrar um atrator, sendo a caminhada composta apenas do tempo de transitente t . Um problema freqüentemente encontrado é a existência de duas ou mais direções que satisfaçam a regra de movimentação da caminhada. Neste caso de empate, a primeira cidade que satisfaz a regra é selecionada, preservando a natureza determinística (Campiteli et al., 2006b).

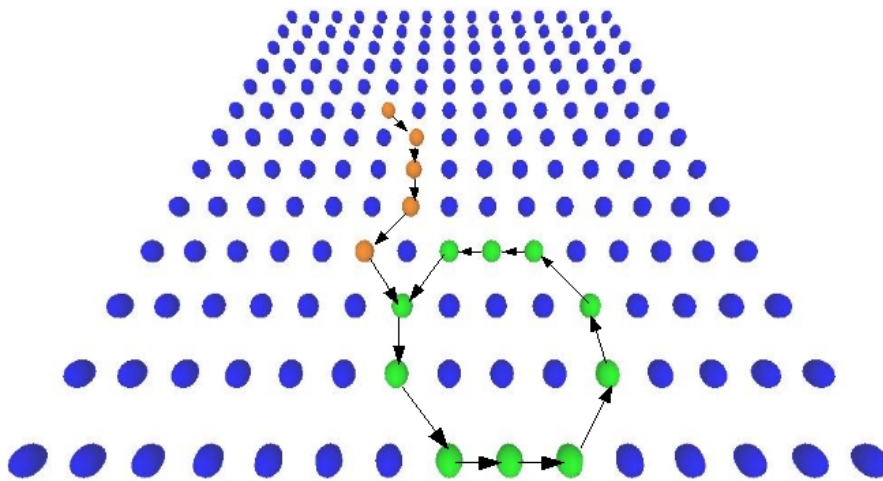


Figura 3.1: Exemplo de uma caminhada do turista. As cidades são representadas pelas esferas. As cidades pertencentes à parte transitente e ao atrator estão em laranja e verde, respectivamente.

Para cada situação inicial, a caminhada do turista produz uma trajetória diferente. O comportamento do turista depende estritamente da configuração do conjunto de cidades, da memória e da cidade inicial. Neste sentido, considera-se cada cidade do mapa como uma condição inicial de uma caminhada. Com isso, se o ambiente em que o turista será aplicado possuir N cidades, existirão N caminhadas diferentes iniciadas em cada uma das cidades. Ambos, tempo de transitente e período do atrator de cada caminhada podem ser combinados em uma distribuição conjunta $S_{\mu,d}^{(N)}(t,p)$. Cada posição da distribuição conjunta $S_{\mu,d}^{(N)}(t,p)$ armazena a quantidade de caminhadas que obtiveram tempo de transitente igual a t e período do atrator igual a p , de acordo com a Equação 3.1. A Figura 3.2 apresenta a distribuição conjunta para $\mu = 9$ e $\mu = 25$. Abaixo alguns casos especiais da distribuição conjunta

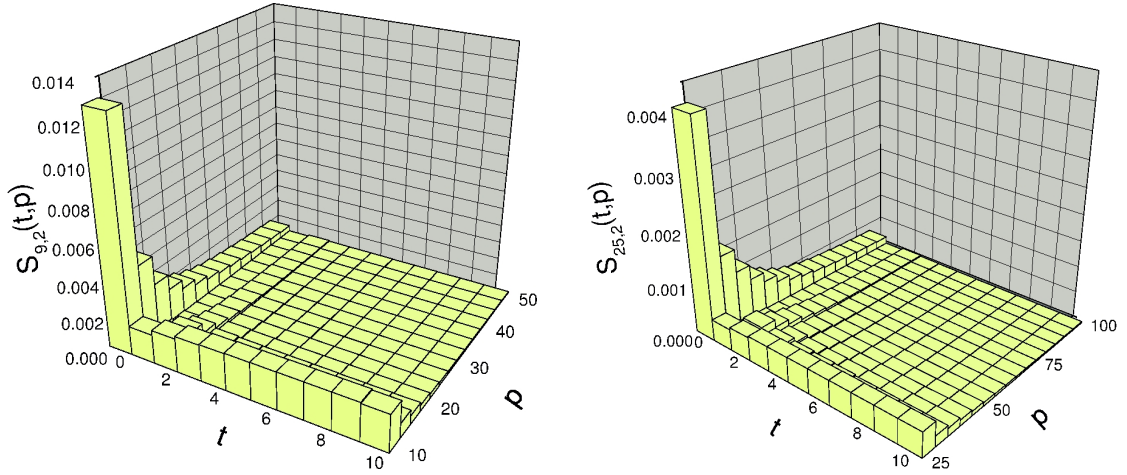


Figura 3.2: Distribuição conjunta do tempo de transiente t e período p com $N = 1000$ e $d = 2$ para dois valores de μ . O aspecto geral de ambas as distribuições são similares.

obtidos para diferentes valores de μ são discutidos.

$$S_{\mu,d}^{(N)}(t,p) = \frac{1}{N} \sum_{i=1}^N \begin{cases} 1, & \text{se } t_i = t, p_i = p \\ 0, & \text{caso contrário} \end{cases} \quad (3.1)$$

onde t_i e p_i são o transiente e o período da caminhada, respectivamente.

O caso mais trivial considera $\mu = 0$, pois o turista possui uma memória nula. Neste caso, o turista permanece na mesma cidade e a trajetória tem transiente nulo e um atrator de período $p = 1$. Apesar da sua simplicidade, a DTW com $\mu = 0$ torna-se interessante porque é a mais simples situação da caminhada do turista estocástica (Risau-Gusman et al., 2003). A distribuição conjunta do transiente e do período é simplesmente dada por:

$$S_{0,d}^{(N)}(t,p) = \delta_{t,0} \delta_{p,1}, \quad (3.2)$$

onde $\delta_{i,j}$ é o delta de Kronecker.

Para a caminhada do turista com $\mu = 1$, o turista armazena em sua memória somente a cidade em que ele está. Assim, a cada passo de tempo, ele deve deixar a cidade atual e ir para o vizinho mais próximo, pois não se lembra de quaisquer cidades visitadas anteriormente. Esta regra não conduz à exploração do ambiente, uma vez que depois de um pequeno tempo de transiente, o turista é apanhado por uma dupla de vizinhos mais próximos entre si. A distribuição conjunta do tempo de transiente e período pode ser analiticamente obtida para $N \gg 1$ (Terçariol and Martinez, 2005):

$$S_{1,d}^{(\infty)}(t,p) = \frac{\Gamma(1 + I_d^{-1})(t + I_d^{-1})}{\Gamma(t + p + I_d^{-1})} \delta_{p,2}, \quad (3.3)$$

onde $\Gamma(z)$ é a função gamma e $I_d = I_{\frac{1}{4}}(\frac{1}{2}, \frac{(d+1)}{2})$ é a função beta normalizada e incompleta.

Um interessante fenômeno ocorre quando altos valores de μ são considerados. Neste caso, a distribuição do atrator não possui um pico em $p_{min} = \mu + 1$, mas apresenta uma completa gama de atratores com períodos $p \geq p_{min}$, com uma possível lei de decaimento (Lima et al., 2001; Stanley and Buldyrev, 2001; Kinouchi et al., 2001).

As caminhadas determinísticas do turista são capazes de extrair propriedades para uma grande variedade de problemas. Entretanto, por ser um método relativamente novo, existem poucos estudos sobre ele. Em (Campitelli et al., 2006b), um estudo inicial com o período do atrator e o tempo transiente foi realizado para analisar e caracterizar texturas de imagens em níveis de cinza. As caminhadas determinísticas apresentaram certa robustez na extração de atributos, capturando a organização dos pixels em diversas escalas. Os experimentos foram realizados em texturas do banco de imagens Brodatz, mostrando um grande potencial na análise de texturas.

Em (Campitelli et al., 2006a), um algoritmo de reconhecimento de padrões baseado na caminhada determinística do turista foi apresentado. O algoritmo proposto permite uma representação hierárquica dos dados de maneira eficiente. Os atratores são identificados com base no ranqueamento dos vizinhos, sendo esses atratores afetados pela memória μ . Nesse algoritmo, a memória é vista como um parâmetro hierárquico no qual os clusters são identificados e utilizados na classificação.

As caminhadas determinísticas do turista também foram utilizadas na análise topológica de dicionários construídos como uma rede complexa (Kinouchi et al., 2001). As palavras em um dicionário estão interligadas pelo seu significado e, segundo os resultados obtidos, a distância média entre dois sinônimos é aproximadamente igual ao logarítmico de N com um alto grau de agrupamento entre os sinônimos. Outro importante resultado é o efeito pequenomundo apresentado pelas redes complexas construídas através de sinônimos.

3.1 *Estratégia para Detecção de Atratores*

Um dos principais problemas envolvidos no estudo da caminhada determinística do turista é a detecção do atrator. Um atrator é um ciclo de período p que existe no final da trajetória do turista. O atrator é caracterizado por iniciar e terminar na mesma cidade, de onde o turista é incapaz de escapar (Campitelli et al., 2006b; Terçariol and Martinez, 2005; Lima et al., 2001).

Durante a caminhada, uma cidade pode ser revisitada sem configurar um atrator, o que dificulta a detecção de um atrator. Devido à memória do turista

μ , que indica as cidades visitadas nos últimos μ passos e que não podem ser visitadas no tempo presente, permite que alguns passos da caminhada sejam repetidos sem caracterizar um atrator. Esta propriedade possibilita construir caminhadas mais sofisticadas, entretanto aumenta-se a dificuldade em detectar um atrator, assim demandando uma metodologia mais sofisticada. A Figura 3.3 apresenta alguns exemplos de caminhadas.

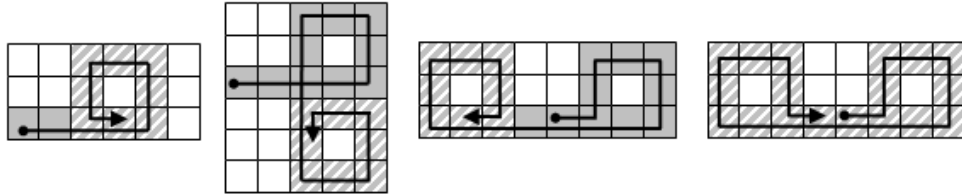


Figura 3.3: Exemplos de caminhadas do Turista. A parte em cinza representa o transiente enquanto a parte listrada representa o atrator.

Repetidas cidades, sem configurar um atrator, demanda que diferentes partes da caminhada sejam comparadas para identificar uma seção de cidades repetidas. Em (Backes et al., 2009b) foi proposto um método baseado em uma lista L que mantém todas as cidades visitadas. Essa lista permite verificar a existência de seções repetidas e conseqüentemente, a presença de um atrator na caminhada. A Figura 3.4 ilustra uma lista para uma caminhada do turista.

1	2	3	4	5	6	7	8	9	10	11
C1	C2	C3	C4	C5	C6	C3	C4	C5	C6	C3

Figura 3.4: Lista L utilizada para manter a caminhada.

A comparação de seções em uma lista consome um alto custo computacional. A cada novo passo do turista, é necessário checar se existe uma seqüência de cidades $p \geq \mu + 1$ repetida ao longo da lista. Uma alternativa é utilizar a busca exaustiva para checar se a cidade atual do turista já foi visitada três vezes e em seguida checar as seções entre essa cidade. Sendo o atrator definido por uma seqüência de cidades repetidas ao longo da caminhada e que esta seqüência inicia e termina na mesma cidade, ao menos três repetições de uma cidade são necessárias para caracterizar um atrator na caminhada (Figura 3.5). Esta condição permite aperfeiçoar a busca na lista L e executar a detecção do atrator de forma simples e rápida.

Para checar a repetição de cidades na caminhada, uma nova lista é associada a cada cidade e ela é responsável por manter a posição que a cidade aparece na lista L (Figura 3.6). Esta lista auxiliar permite checar rapidamente a existência de um conjunto de cidades repetidas, bem como, o número de

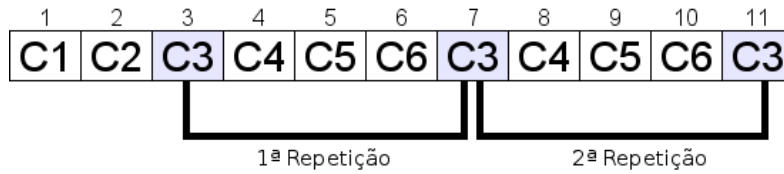


Figura 3.5: Comparação entre diferentes partes de uma lista. Uma cidade (cinza) repetida ao menos três vezes na lista pode configurar a existência de seqüências de cidades repetidas.

repetições de cada cidade, reduzindo o custo computacional envolvido na detecção do atrator.

C1	1		
C2	2		
C3	3	7	11
C4	4	8	
C5	5	9	
C6	6	10	

Figura 3.6: Lista duplamente encadeada associando cada cidade da caminhada mantida em L .

3.2 Dinâmica da Caminhada Determinística do Turista

A caminhada do turista aplicada em um mapa com cidades geradas de forma aleatória é caracterizada por um comportamento caótico, pois não há uma correlação entre as cidades vizinhas e com isso, o turista muda de direção com a mesma freqüência em que o ambiente muda. Entretanto, ambientes regulares apresentam um contexto que correlacionam os vizinhos (e.g. os pixels de uma imagem apresentam níveis de cinza próximos para regiões de um objeto). Esta correlação entre vizinhos influencia na escolha da direção do turista ao longo de sua caminhada.

Uma importante questão sobre a caminhada do turista está na regra de movimentação adotada pelo turista. Quando o turista inicia sua caminhada, ele decide ir para a cidade que está mais próxima, mais distante ou qualquer outra regra de movimentação. Uma vez escolhida a regra de movimentação,

esta não deve ser mudada ao longo da caminhada. Como resultado, as caminhadas guiadas por diferentes regras (e.g. maior ou menor distância) geram atratores com padrões distintos para um mesmo ambiente. As caminhadas do turista guiadas para a menor distância são apropriadas para encontrar atratores em regiões que apresentam alta homogeneidade. Por outro lado, as caminhadas guiadas para a maior distância encontram atratores localizados em regiões de baixa homogeneidade, como em regiões de bordas em uma imagem.

A utilização de diferentes regras de movimentação reflete em diferentes distribuições conjuntas do atrator e tempo transiente. Este comportamento permite usar informações de diferentes fontes na caracterização de um ambiente, como pode ser visto nas Figuras 3.7 e 3.8 para atratores gerados para a regra de movimentação guiada para a menor e maior distância, respectivamente.

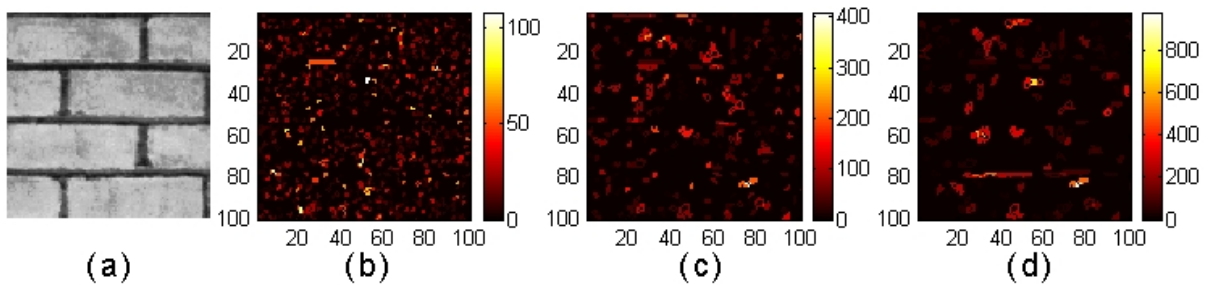


Figura 3.7: Atratores gerados pelo turista em uma imagem de textura utilizando a distância mínima e diferentes valores de μ . (a) Imagem Original; (b) $\mu = 1$; (c) $\mu = 3$; (d) $\mu = 7$.

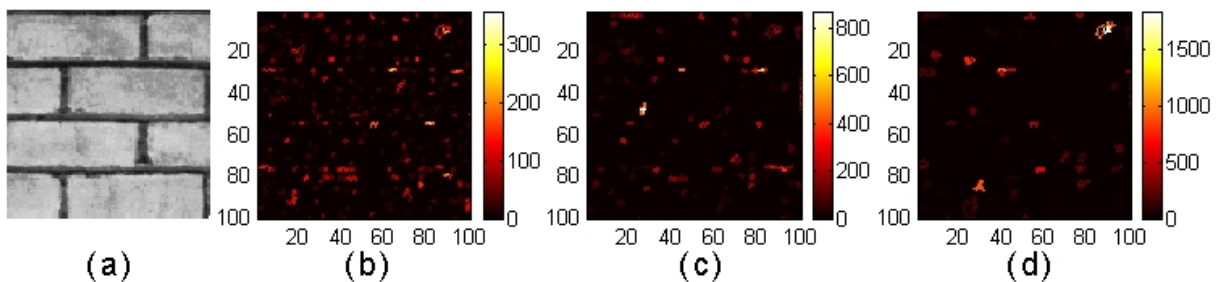


Figura 3.8: Atratores gerados pelo turista em uma imagem de textura utilizando a distância máxima e diferentes valores de μ . (a) Imagem Original; (b) $\mu = 1$; (c) $\mu = 3$; (d) $\mu = 7$.

3.3 Critérios de Desempate

Atualmente, uma das principais questões não resolvidas na caminhada do turista é o critério de desempate (Campitelli et al., 2006b; Backes et al., 2006).

O problema do desempate consiste em decidir a próxima cidade que o turista irá visitar, dado um conjunto de cidades vizinhas igualmente próximas da cidade atual. Certamente, se existe somente uma cidade vizinha mais próxima, essa deve ser seguida. Entretanto, se duas ou mais cidades vizinhas estão à mesma distância e ambas são as mais próximas, o turista pode decidir incorretamente por uma cidade, sendo que a outra cidade seria mais apropriada para encontrar um atrator e representar o padrão daquela região. A solução ideal para esse problema é avaliar todas as caminhadas possíveis e então decidir qual cidade a ser seguida. Entretanto, essa solução é computacionalmente inviável, já que as cidades empatadas podem crescer exponencialmente a cada passo.

Para demonstrar a importância da resolução desse problema, as probabilidades de empates estimadas em um conjunto de 1100 imagens de textura são apresentadas no gráfico da Figura 3.9. Como pode ser observado, ao aumentar o número de cidades vizinhas do turista, o número de empates também aumenta. Para cidades com 4 e 56 cidades vizinhas, a probabilidade de empate a cada instante é de 11.34% e 27.06%, respectivamente.

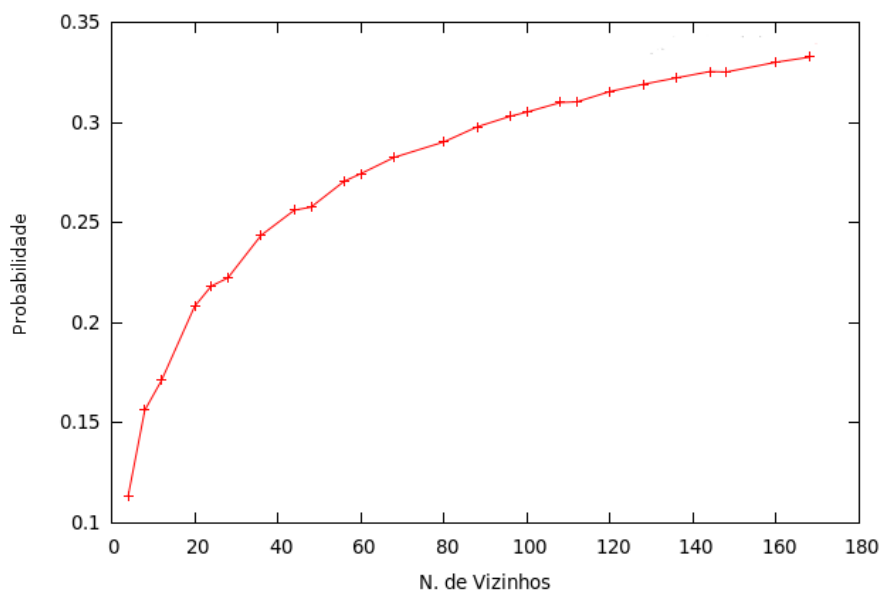


Figura 3.9: Gráfico de probabilidades de ocorrência de empates na caminhada determinística do turista. Número de cidades vizinhas X Probabilidade de empate.

Para a resolução desse problema, três estratégias são apresentadas nas próximas seções.

3.3.1 Estratégia Primeira

Dado que cada cidade possui um índice único, a estratégia primeira seleciona a cidade com menor índice entre as cidades empatadas (Campitelli et al., 2006b; Backes et al., 2006). Essa estratégia não é influenciada pela maneira como as cidades são percorridas (e.g. sentido horário e anti-horário), pois os índices são fixos e não mudam com o tempo. Com o uso dessa estratégia, percorrendo as cidades no sentido horário, a caminhada do turista obteve resultados satisfatórios em imagens (Campitelli et al., 2006b; Backes et al., 2006; Campitelli et al., 2006a).

3.3.2 Estratégia Aleatória

A estratégia aleatória é sugerida com base nos interessantes resultados obtidos pelas caminhadas aleatórias (Metzler and Klafter, 2000; Derrida, 1997). A ideia dessa estratégia é selecionar uma cidade de forma aleatória, dado um conjunto de cidades empatadas em um determinado instante. Todas as cidades empatadas possuem a mesma probabilidade de serem selecionadas.

Embora a regra deixe de ser completamente determinística, essa estratégia é interessante, pois o turista consegue caminhar por mais tempo na imagem antes de ser aprisionado em um atrator, devido ao fator aleatório. Na Figura 3.10, o tamanho da caminhada é apresentado para as estratégias primeira e aleatória. Podemos notar que com a estratégia aleatória, as caminhadas são maiores e conseqüentemente, o ambiente é mais explorado.

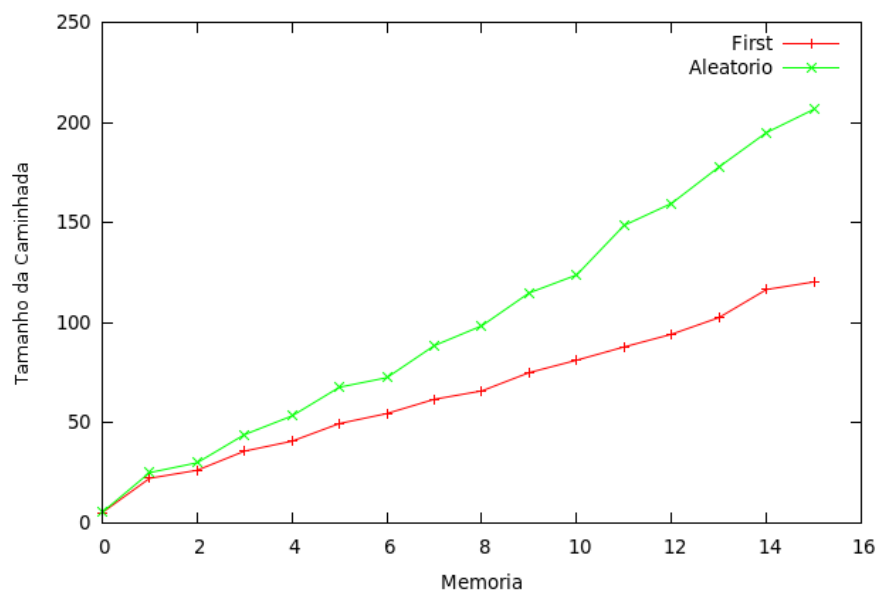


Figura 3.10: Média do tamanho das caminhadas para diferentes valores de memórias μ e estratégias de desempate aleatória e primeira.

3.3.3 Estratégia Visibilidade

A estratégia Visibilidade utiliza as informações da dinâmica do turista. Dado o conjunto de n cidades empatadas, a idéia da estratégia é iniciar n caminhadas do turista simultâneas em cada cidade empatada. Essas caminhadas simultâneas serão iteradas até que não haja mais cidades empatadas, levando em consideração a dinâmica utilizada na caminhada do turista.

Pode ser que as caminhadas simultâneas se prolonguem por muito tempo, o que tornaria o método ineficiente. Com isso, se as caminhadas possuírem um tamanho (número de cidades percorridas) maior que um determinado limiar, as trajetórias são interrompidas e uma cidade é selecionada através da estratégia aleatória.

Embora nessa estratégia ainda possa ocorrer empates devido ao tamanho máximo escolhido, a probabilidade de isso ocorrer é menor e número de cidades empatadas no final das caminhadas simultâneas também é reduzido. Essa estratégia pode ser vista como um turista com maior visibilidade dos caminhos a serem percorridos até um determinado raio (limiar das caminhadas simultâneas) e então decidir qual cidade percorrer de forma mais precisa.

Para ilustrar essa estratégia de desempate, considere a Figura 3.11. O turista, caminhando em direção a mínima distância, encontra um empate entre os vértices 2 e 4. Em seguida, os vizinhos dos vértices empatados são analisados. Um novo empate ocorre entre os vértices 6 e 10. No terceiro nível, o turista consegue resolver o empate, optando então, em sair do vértice 1 e ir ao vértice 4.

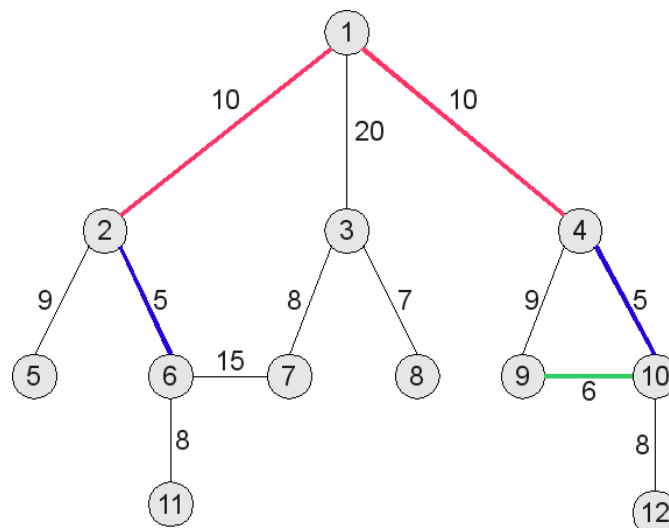


Figura 3.11: Exemplo da resolução do empate para a estratégia Visibilidade para um turista caminhando em direção da mínima distância. Um empate ocorre entre as cidades 2 e 4 no primeiro nível. Um novo empate ocorre no segundo nível entre as cidades 6 e 10. O desempate ocorre no terceiro nível, optando-se então sair da cidade 1 e ir para a cidade 4.

3.4 Extração de Informações

A grande quantidade de informações resultante da aplicação da caminhada do turista em um ambiente é representada pela distribuição conjunta do tempo de transiente e período. Para a utilização da caminhada do turista de forma eficiente, informações relevantes precisam ser extraídas dessa distribuição.

As estratégias de extração de informações são utilizadas para construir um vetor de características $\psi_{\mu,din}$ a partir da distribuição conjunta construída com memória μ e dinâmica de movimentação din . Esse vetor $\psi_{\mu,din}$ é utilizado para caracterizar o ambiente onde o turista foi executado. Nas subseções abaixo três estratégias são descritas.

3.4.1 Região

Nessa estratégia, utilizada em trabalhos com a caminhada do turista (Campiteli et al., 2006b; Backes et al., 2006), o vetor de características $\psi_{\mu,din}$ é construído com os valores de uma região da distribuição conjunta. Essa região é empiricamente escolhida através dos parâmetros a (número máximo do tempo de transiente) e b (número máximo de períodos a serem considerados). O vetor de característica é dado pela Equação 3.4 para uma distribuição conjunta construída com memória μ em um ambiente de duas dimensões. Um ilustração dos valores considerados da distribuição conjunta, dado os parâmetros a e b , é apresentada na Figura 3.12.

$$\psi_{\mu,din} = [S_{2,\mu}^{(N)}(0 : a, \mu + 1), S_{2,\mu}^{(N)}(0 : a, \mu + 2), \dots, S_{2,\mu}^{(N)}(0 : a, \mu + b)] \quad (3.4)$$

Embora a região seja empiricamente escolhida, os resultados obtidos para valores pequenos de a e b são interessantes, como $a = 4$ e $b = 5$ em (Backes et al., 2006). Esses resultados indicam que as informações relevantes estão no início da distribuição conjunta.

3.4.2 Histograma

Essa estratégia constrói um histograma da caminhada do turista $h_{\mu,din}(t+p)$. Este histograma representa o número de caminhadas que possui tamanho igual $(t+p)$ em sua distribuição conjunta, onde t e p representam o tempo transiente e período, respectivamente. O número total de características n são selecionadas do histograma da caminhada para compor o vetor de característica $\psi_{\mu,din}$. Como não existem períodos menores que $\mu + 1$, o primeiro descritor selecionado possui esse tamanho. O vetor de característica, construído com

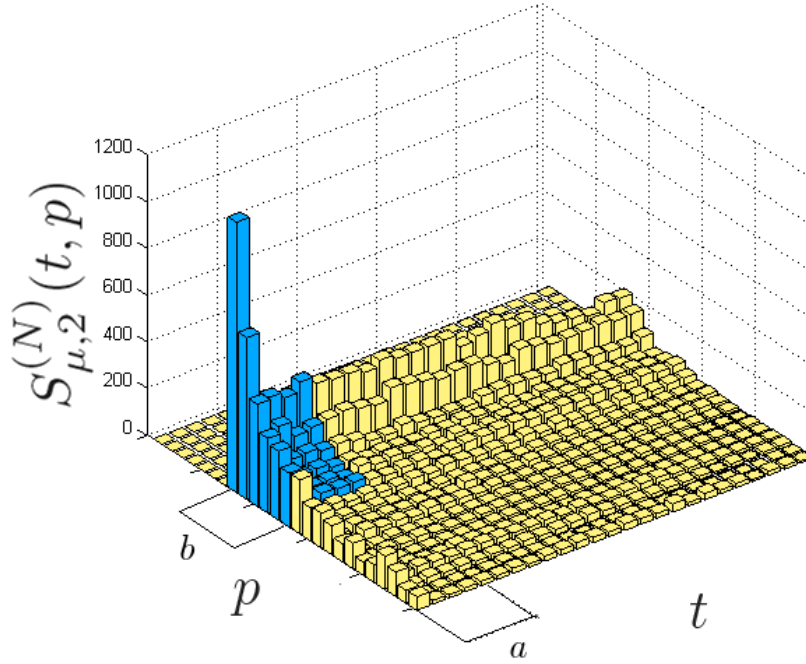


Figura 3.12: Ilustração dos valores considerados da distribuição conjunta para a estratégia região com parâmetros a e b .

essa estratégia, é dado na Equação 3.5.

$$\psi_{\mu,din} = [h_{\mu,din}(\mu + 1), h_{\mu,din}(\mu + 2), \dots, h_{\mu,din}(t + p), \dots, h_{\mu,din}(\mu + n)]$$

$$\begin{aligned} h_{\mu,din}(\mu + 1) &= S_{\mu,2}^{(N)}(0, \mu + 1) \\ h_{\mu,din}(\mu + 2) &= S_{\mu,2}^{(N)}(0, \mu + 2) + S_{\mu,2}^{(N)}(1, \mu + 1) \\ h_{\mu,din}(\mu + 3) &= S_{\mu,2}^{(N)}(0, \mu + 3) + S_{\mu,2}^{(N)}(1, \mu + 2) + S_{\mu,2}^{(N)}(2, \mu + 1) \end{aligned} \quad (3.5)$$

Com essa estratégia, mais informações da distribuição conjunta são inseridas no vetor $\psi_{\mu,din}$ com um número menor de características, comparado com a estratégia descrita anteriormente. Embora essa estratégia e a anterior apresentem resultados interessantes, estas não consideram todas as informações presentes na distribuição conjunta.

3.4.3 Propriedades Estatísticas

Diversas características têm sido propostas para extrair informações de uma distribuição 2D (Haralick et al., 1973; Weszka et al., 1976). Essas características estatísticas são interessantes, pois consideram as informações da distribuição com um todo. Em (Haralick et al., 1973), 14 características são propostas, entretanto, somente algumas são utilizadas, como em (Weszka et al., 1976) e (Connors and Harlow, 1980) onde foram utilizadas somente 4 ($\phi_1, \phi_2, \phi_3, \phi_4$) e 5 ($\phi_1, \phi_2, \phi_3, \phi_5, \phi_6$) características, respectivamente. Seguindo

estes trabalhos, na Equação 3.6 são descritas essas características.

$$\begin{aligned}
\phi_1 &= \text{Energia} = \sum_{i,j} S_{2,\mu}^{(N)}(i,j)^2 \\
\phi_2 &= \text{Entropia} = - \sum_{i,j} S_{2,\mu}^{(N)}(i,j) \log S_{2,\mu}^{(N)}(i,j) \\
\phi_3 &= \text{Inércia} = \sum_{i,j} (i-j)^2 S_{2,\mu}^{(N)}(i,j) \\
\phi_4 &= \text{Correlação de Haralik} = \frac{\sum_{i,j} (ij) S_{2,\mu}^{(N)}(i,j) - \mu_x \mu_y}{\sigma_i \sigma_j} \\
\phi_5 &= \text{Correlação} = - \sum_{i,j} \frac{(i-\mu_i)(j-\mu_j) S_{2,\mu}^{(N)}(i,j)}{\sigma_i \sigma_j} \\
\phi_6 &= \text{Momento da diferença inversa} = \sum_{i,j} \frac{1}{1+(i-j)^2} S_{2,\mu}^{(N)}(i,j)
\end{aligned} \tag{3.6}$$

onde μ_x, μ_y, σ_x e σ_y são as médias e desvio-padrão da soma das linhas e colunas da distribuição conjunta, respectivamente.

Por fim, o vetor de característica $\psi_{\mu,din}$ para essa estratégia é composto pelas características descritas acima, calculadas sobre a distribuição conjunta (Equação 3.7).

$$\psi_{\mu,din} = [\phi_1, \phi_2, \phi_3, \phi_4, \phi_5, \phi_6] \tag{3.7}$$

3.5 Complexidade Computacional

As caminhadas determinística do turista são executadas para cada cidade do ambiente, como por exemplo, para cada pixel de uma imagem. Considere que o turista será executado em um ambiente com N cidades. Cada caminhada do turista consiste de uma parte transiente de tamanho t e um atrator de tamanho $p \geq \mu + 1$, se ele existir. Em casos onde a atrator não é encontrado, o turista caminha até um determinado tamanho, geralmente igual ao número de componentes, por exemplo, $t = N$ e $p = 0$.

A complexidade computacional da caminhada do turista é determinada pelo número de cidades e tamanho de cada caminhada, $O(N(t + p))$. O melhor caso do algoritmo é alcançado quando todas as caminhadas iniciam em um atrator, ou seja, transiente nulo ($t = 0$) e o atrator apresenta tamanho mínimo ($p = 1$).

O tamanho do atrator depende do contexto do ambiente e da memória μ . Para $\mu = 0$ e considerando um contexto oportuno, o tamanho do atrator é minimizado para $p = 1$, o qual lida com complexidade $O(N)$ no melhor caso. O pior caso é alcançado quando nenhum atrator é encontrado durante a caminhada. Neste caso, para qualquer tamanho de memória μ , a caminhada do turista apresenta uma tamanho $t + p = N$, o qual leva a uma complexidade

$O(N^2)$. É importante enfatizar que o pior caso é muito raro, acontecendo em uma configuração muito específica do ambiente. Mesmo uma imagem gerada de forma aleatória não produz este caso especial de caminhada.

A Figura 3.13 apresenta a média do tamanho de caminhada para diferentes memórias μ . Para valores de memórias $\mu = 11$, a média do tamanho da caminhada é igual a 55 passos. Este resultado está bem próximo da complexidade do turista no melhor caso, $O(N)$ para $N \gg 55$, o que geralmente ocorre nas aplicações em imagens. Este é um excelente resultado quando comparado com a complexidade de alguns métodos de reconhecimento de texturas como filtros de Gabor ($O(N \log N)$, devido à transformada de Fourier) e matrizes de co-ocorrência ($O(N)$), considerando N a quantidade de pixels.

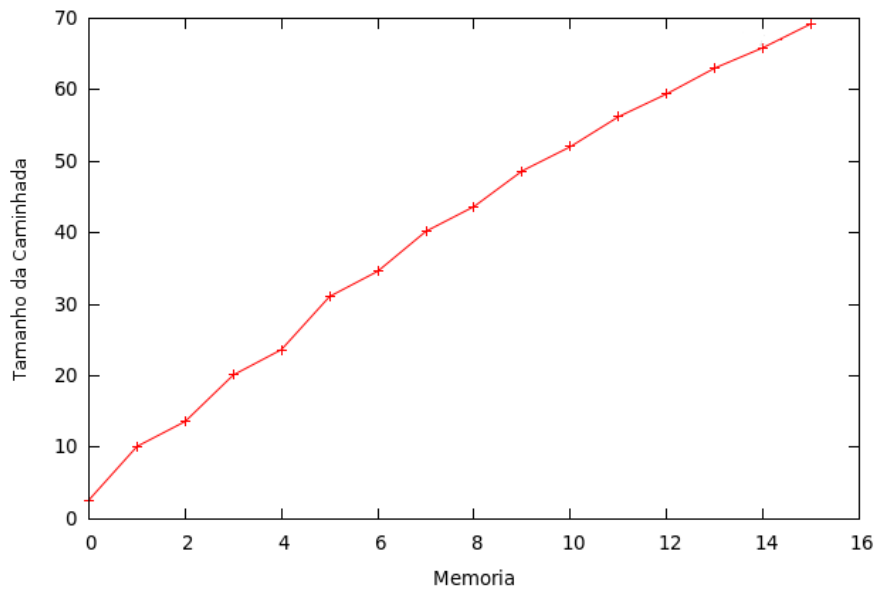


Figura 3.13: Média do tamanho das caminhadas para diferentes valores de memórias μ .

3.6 DTW em Redes Complexas

Nesta seção, uma descrição matemática das caminhadas determinísticas do turista em redes complexas é apresentada. Dado uma rede complexa representada por um conjunto de n vértices $V = \{v_1, \dots, v_n\}$ e um conjunto de arestas $W = \{w_{v_i, v_j} | v_i, v_j \in V, w_{v_i, v_j} \in \mathbb{R}\}$, onde cada componente w_{v_i, v_j} representa o peso da aresta que interliga o vértice v_i ao vértice v_j . Considere ainda, $\eta(v) = \{k \in V | \exists w_{kv}\}$ um conjunto tal que todos os vértices do conjunto são vizinhos do vértice v e $C = \{c_0, c_1, \dots, c_i | c_i \in V\}$ uma lista de vértices que armazena a caminhada do turista.

O turista inicia sua caminhada em um vértice v_0 selecionada previamente (Equação 3.8).

$$c_0 = v_0 \quad (3.8)$$

O turista segue a regra: vá para o vértice mais próximo que não tenha sido visitado nos últimos μ passos. Em redes complexas, a movimentação é completamente realizada com base na memória (Equação 3.9) e no conjunto W de pesos das arestas. A memória M_i é um subconjunto composto pelos últimos μ vértices da caminhada C .

$$M_i = \{X \subset C \mid x = \bigcup_{k=i-\mu}^i c_k\} \quad (3.9)$$

A caminhada é iterada através das Equações 3.10 e 3.11. Na equação 3.10, F é um conjunto de vértices tal que esses vértices são os mais próximos do vértice atual c_i e que não estão na memória M_i . F é definido como um conjunto, pois podem ocorrer empates entre os vértices. Dependendo da dinâmica de movimentação, a cada passo o turista deixa o vértice atual e vai para o vértice mais próximo (definido pelo $\arg \min$) ou mais distante (trocando-se $\arg \min$ por $\arg \max$ na Equação 3.10).

$$F = \{v \in V \mid \arg \min_{v \in \eta(c_i)} w_{v,c_i} \mid v \notin M_i\} \quad (3.10)$$

Na Equação 3.11, o possível empate dos vértices contidos em F é resolvido. Caso F tenha somente um elemento, esse é escolhido como o próximo passo do turista. Caso contrário, o empate é resolvido por uma função ϕ que retorna somente um vértice. A função pode retornar um vértice aleatoriamente ou executar alguma outra estratégia mais sofisticada, como aumentar a visibilidade do turista.

$$c_{i+1} = \begin{cases} f_1, & n(F) = 1 \\ \phi(F), & \text{caso contrário} \end{cases} \quad (3.11)$$

Em um determinado instante, o turista que está caminhando sobre os vértices do grafo é aprisionado em um atrator com período p e tempo transiente t . A caminhada pode ser iterada até um determinado tempo e em seguida buscar-se o atrator, ou a cada passo de tempo, realizar a busca pelo atrator e encerrar a caminhada caso seja encontrado. A detecção do atrator é definida pela Equação 3.12, considerando $\varphi(v)$ como o índice do vértice v , i.e. $\varphi(v_1) = 1, \varphi(v_2) = 2$. Caso $Sum_{p,it} = 0$, um atrator com período p e tempo de transiente $t = it - 1$ foram identificados. Caso $F = \emptyset$, o turista não encontrou qualquer atrator, e com isso $p = 0$ e $t = i$.

$$\text{Fim da Caminhada} \leftarrow \begin{cases} F = \emptyset \\ \exists \text{Sum}_{p,it} = 0, \quad 0 \leq p, it \leq i \end{cases} \quad (3.12)$$

$$\text{Sum}_{p,it} = \sum_{j=0}^{p-1} |\varphi(t_{it+j}) - \varphi(t_{it+j+p})|$$

Reconhecimento de Redes Complexas Artificiais

Diversas redes do mundo real descritas na literatura apresentam propriedades estruturais e topológicas em comum. A classificação dessas redes em um dos modelos de redes complexas conhecido e com diversas propriedades estabelecidas é uma aplicação importante (da F. Costa et al., 2007). A classificação tem sido realizada com medidas básicas, como média dos graus dos vértices, caminho mínimo, etc.

Entretanto, somente uma medida não possui o poder de discriminação, sendo necessário um conjunto de várias medidas básicas. Devido à grande maioria das medidas serem redundantes, o que tem se feito atualmente é basear-se no conhecimento de um especialista sobre o problema para selecionar um conjunto de medidas adequadas (da F. Costa et al., 2007). Este capítulo descreve os experimentos realizados na classificação de redes complexas artificiais com as caminhadas determinísticas do turista - DTW. Por fim, os resultados são comparados com as medidas básicas comumente utilizadas na literatura.

4.1 Experimentos

Para avaliar a DTW como uma medida para representação de redes complexas, os experimentos foram realizados em um conjunto com 40000 redes artificiais. Esse conjunto foi gerado com base em quatro modelos de redes complexas: rede aleatória, pequeno mundo, geográfica e livre de escala. As redes complexas artificiais foram construídas com pesos das arestas aleatórios

variando de 0 a 255, número de vértices variando de $N = 100$ até $N = 1000$ com incremento de 100 e com média do grau variando de $\langle k \rangle = 2$ até $\langle k \rangle = 20$ com incremento de 2.

As caminhadas determinísticas do turista foram aplicadas nas redes complexas descritas acima e a distribuição conjunta foi obtida. Seguindo os procedimentos descritos na Ref. (Backes et al., 2009b), um total de $n = 4$ descritores do histograma foram considerados na extração de informações da distribuição conjunta, pois algumas análises estatísticas revelaram que as informações relevantes da distribuição conjunta se concentram em poucos elementos. Os vetores de assinatura φ extraídos para cada rede complexa foram classificados através do classificador K-neighbor nearest (Mitchell, 1997), com $k = 1$, em uma estratégia cross-validation com 10 pastas.

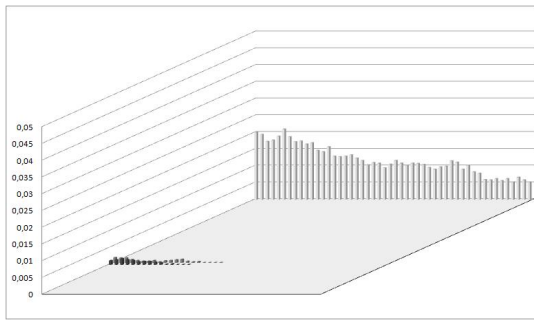
4.1.1 Avaliação dos Parâmetros da DTW

Na Figura 4.1, a distribuição conjunta da caminhada determinística do turista para cada modelo de rede complexa é ilustrada. Nesta figura, a DTW com memória $\mu = 5$ e regra de movimentação em direção a menor distância foi aplicado nos modelos de redes complexas gerados com $N = 1000$ e $\langle k \rangle = 20$. É possível notar que, as distribuições apresentam padrões distintos e, portanto essas informações podem ser utilizadas para caracterizar cada modelo de rede complexa.

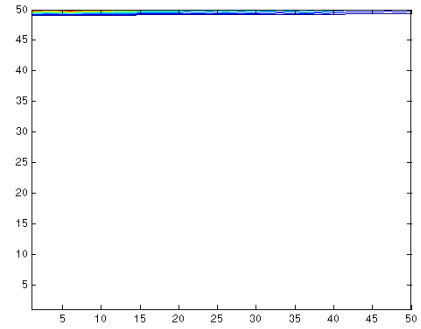
A Tabela 4.1 apresenta os resultados da classificação das redes para diferentes valores de μ e regra de movimentação din da caminhada determinística do turista. Na maioria das vezes, os melhores resultados são obtidos quando a regra de movimentação do turista é em direção à menor distância quando comparadas com a máxima distância. Neste caso, os atratores do turista são formados em regiões homogêneas, i.e., regiões onde o peso das arestas entre os vértices são baixos. A combinação entre vetores de características com as regras de movimentação min e max ($[Min Max]$) incrementou a porcentagem de classificação correta. Isso ocorre porque o novo vetor de assinatura apresenta ambas as informações locais e globais da rede complexa, fornecendo assim uma ferramenta mais poderosa para caracterizar redes complexas.

	Memória (μ)					
	1	2	3	4	5	6
<i>Min</i>	74,31	87,27	73,86	64,45	61,59	58,45
<i>Max</i>	76,20	81,26	70,98	63,21	58,94	55,76
$[Min Max]$	84,07	93,02	84,03	74,84	68,37	61,09

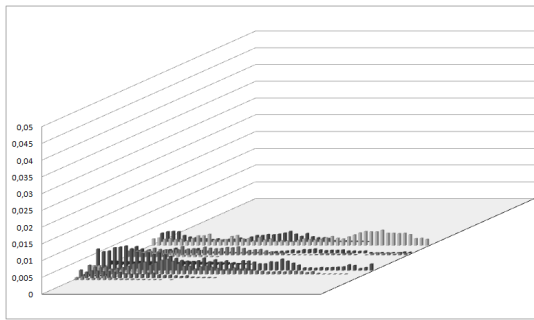
Tabela 4.1: Porcentagem de classificação correta para $\psi_{\mu, din}$ utilizando diferentes valores de μ e regras de movimentação na caminhada determinística do turista.



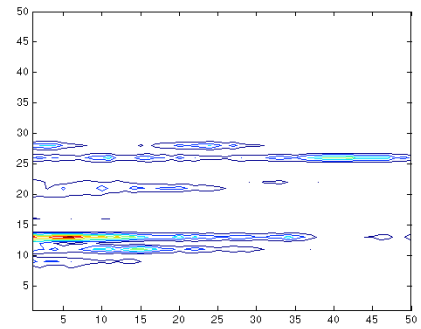
(a) Random Graph.



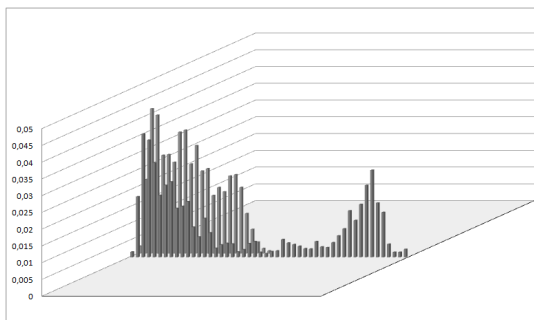
(b) Random Graph.



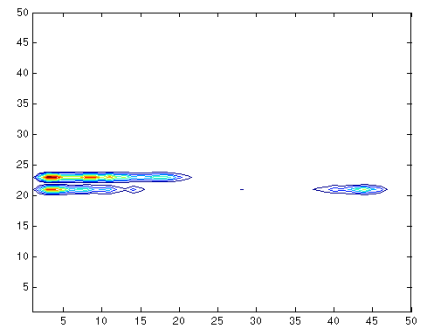
(c) Geographical Network.



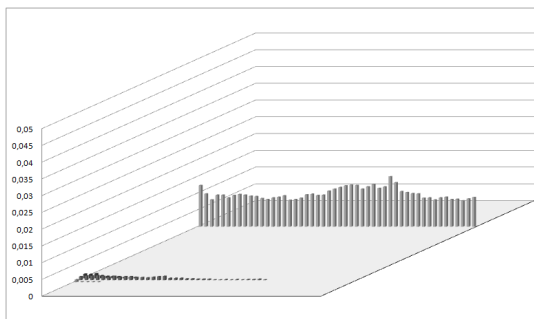
(d) Geographical Network.



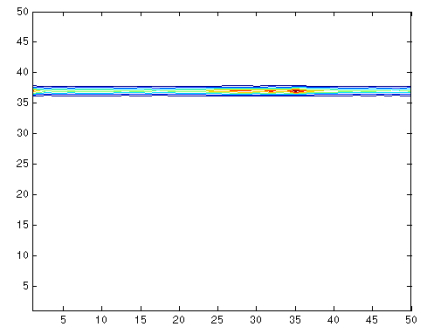
(e) Scale-free.



(f) Scale-free.



(g) Small World.



(h) Small World.

Figura 4.1: Primeira coluna: Distribuição Conjunta para cada modelo de rede complexa gerada com $N = 1000$ e $\langle k \rangle = 20$. Segunda coluna: Distribuição com uma visão 2D.

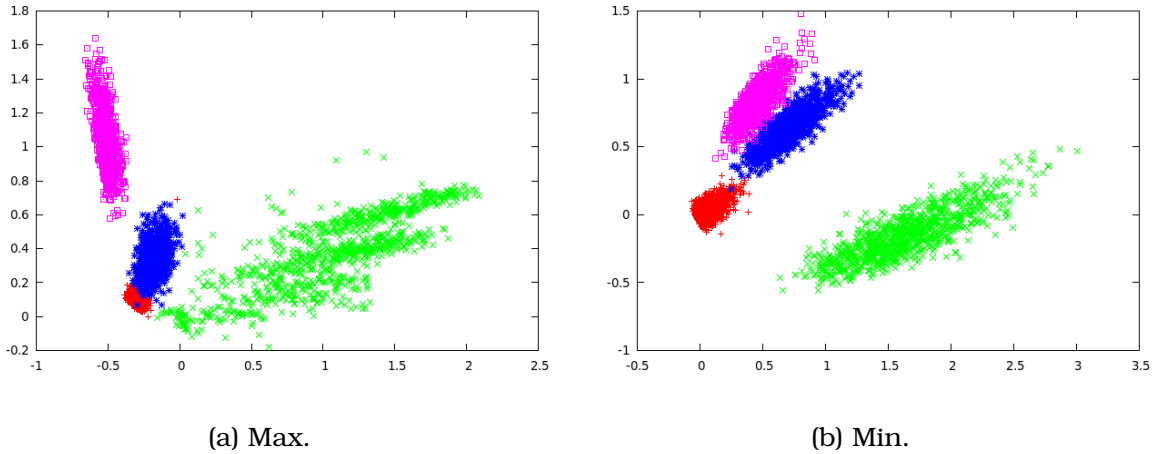
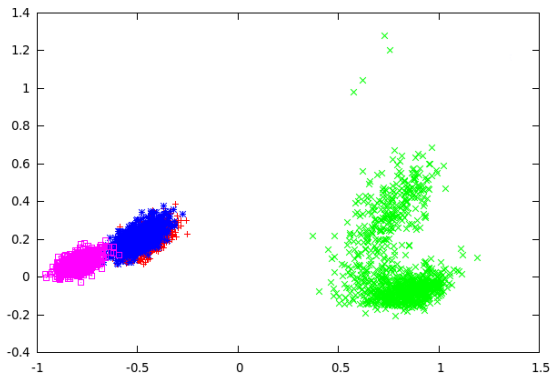


Figura 4.2: Discriminantes do PCA para 4000 redes derivadas dos modelos aleatório, geográfico, pequeno mundo e livre de escala. As redes foram geradas com $N = 1000$ e $\langle k \rangle = 20$ e a aplicação do turista com regras de movimentação *min* e *max*.

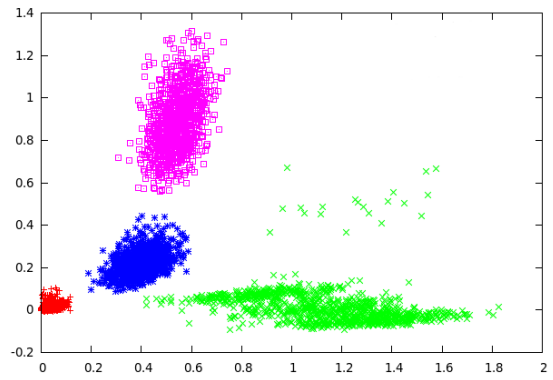
A Figura 4.2 apresenta os dois primeiros discriminantes após a aplicação do método de análise de componentes principais (principal component analysis - PCA) (Fukunaga, 1990) para as regras de movimentação *Min* e *Max*. Como pode ser visto, a separação dos modelos de rede é melhor visualizada quando se utiliza a regra *Min*.

Os resultados da Tabela 4.1 também mostraram que, para ambas as regras de movimentação e suas combinações, a porcentagem de classificação correta é decrementada ao mesmo passo que a memória é incrementada, com exceção da $\mu = 1$ que é um caso trivial da caminhada determinística do turista. A explicação para este resultado lida com o fato de que, se a memória é incrementada, o turista possui mais dificuldade de encontrar um atrator na rede complexa (Backes et al., 2009b). Com isso, o tamanho da caminhada é incrementado e as longas caminhadas podem conduzir a uma armadilha, casos onde o turista não consegue encontrar um atrator, pois todos os seus vizinhos estão na memória. Por outro lado, os pequenos valores de memória μ permitem uma melhor análise local da estrutura da rede, refletindo em uma maior porcentagem de classificação correta. Como ilustração, a Figura 4.3 apresenta os dois primeiros discriminantes do PCA para diferentes memórias.

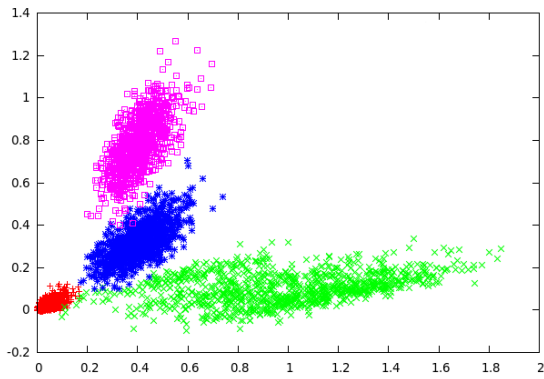
A Tabela 4.2 apresenta os resultados quando múltiplos valores de μ são considerados. Este método diminui a importância individual para cada valor de μ e permite que o turista realize caminhadas de diferentes tamanhos, proporcionando assim uma caracterização mais robusta. Por fim, na Figura 4.4, a análise de componentes principais foi aplicada na estratégia com múltiplos valores de μ (1, 2, 3, 4, 5, 6) para redes complexas geradas com $N = 1000$ e $\langle k \rangle = 20$.



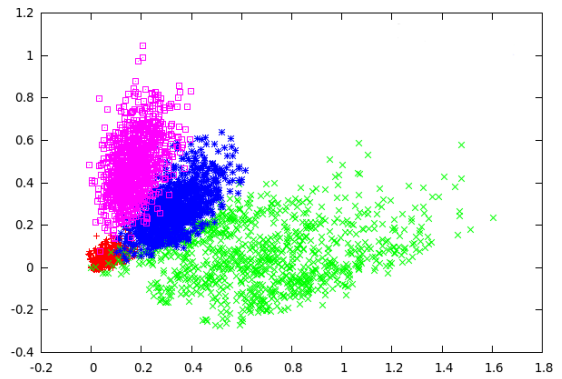
(a) $\mu = 1.$



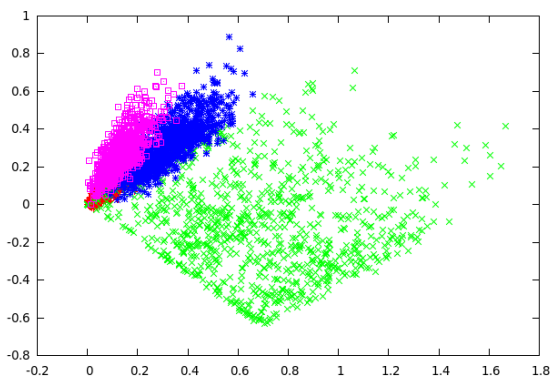
(b) $\mu = 2.$



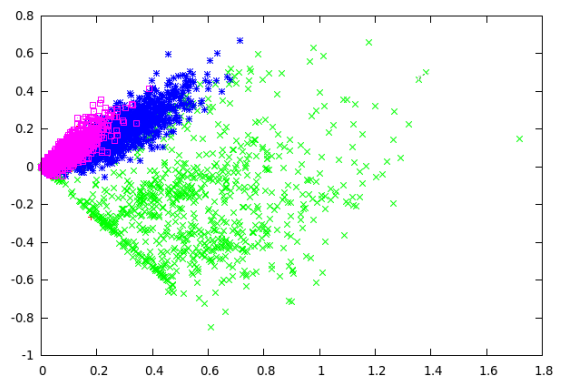
(c) $\mu = 3.$



(d) $\mu = 4.$



(e) $\mu = 5.$



(f) $\mu = 6.$

Figura 4.3: Discriminantes do PCA para os modelos de rede complexa gerada com $N = 1000$ e $\langle k \rangle = 20$ e a aplicação do turista com diferentes memórias.

Memórias (μ)	<i>min</i>	<i>max</i>	[<i>min max</i>]
{1, 2}	95,49	94,23	98,11
{1, 2, 3}	95,65	94,47	98,16
{1, 2, 3, 4}	95,73	94,68	98,19
{1, 2, 3, 4, 5}	95,66	94,82	98,21
{1, 2, 3, 4, 5, 6}	95,65	94,83	98,22

Tabela 4.2: Porcentagem de classificação correta para φ combinando diferentes valores de μ no banco de imagens Brodatz.

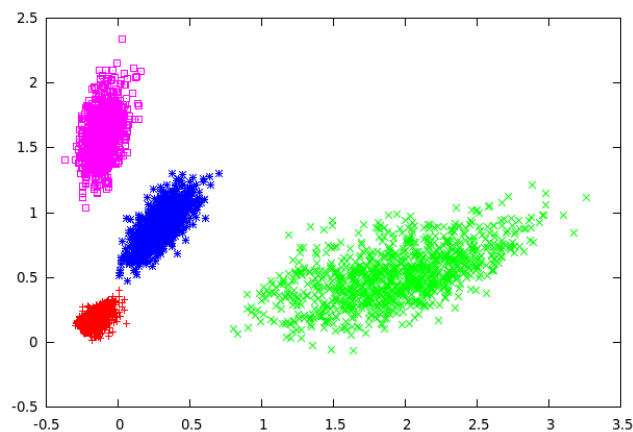


Figura 4.4: Discriminantes do PCA para os modelos de rede complexa gerada com $N = 1000$ e $\langle k \rangle = 20$ e a aplicação do turista com combinação das memórias 0, 1, 2, 3, 4 e 5.

4.1.2 Comparação com Medidas Tradicionais

Os resultados obtidos pela caminhada do turista na caracterização são comparados com as medidas tradicionais na Tabela 4.3 para o grupo de redes complexas geradas de forma artificial. As medidas tradicionais são calculadas para cada vértice da rede complexa e a média e desvio padrão são obtidos. Portanto, cada medida tradicional é representada por um vetor de 2 dimensões correspondentes à média e desvio padrão.

Para compreender a regra das características no reconhecimento de padrões, três classificadores foram utilizados: K-Vizinhos mais próximos (*k-nearest neighbor* - Knn) (Mitchell, 1997), *naive bayes* (Zhang, 2004) e árvores de decisão (Breiman et al., 1984). A caminhada do turista supera as medidas tradicionais com uma taxa de classificação correta de 98,22% enquanto as melhores taxas foram de 78,32% e 71,87%, obtidas respectivamente por coeficiente de clusterização e correlação de Pearson.

Medida	Knn	Naive Bayes	Árvores de Decisão
Grau	66,72	25,64	67,77
Grau Hier. 2	54,67	28,45	64,47
Grau Hier. 3	54,14	32,17	64,33
Grau com Pesos	43,49	25,63	54,72
Grau Hier. 2 com Pesos	51,98	28,50	62,61
Grau Hier. 3 com Pesos	53,27	32,11	63,51
Coefficiente de Clusterização	68,13	69,19	78,32
Correlação de Pearson	67,02	58,96	71,87
DTW	98,22	80,12	98,24

Tabela 4.3: Comparação dos resultados para diferentes medidas extraídas das redes complexas.

Na Tabela 4.4, as medidas foram concatenadas e utilizadas na classificação nos três classificadores. Quando todas as medidas tradicionais são utilizadas na classificação, os resultados obtidos somente pela caminhada do turista são equivalentes (99,18% contra 98,22%). Quando ambas medidas tradicionais e DTW são utilizadas, a classificação alcançou uma taxa de classificação correta de 99,96%.

Medida	Knn	Naive Bayes	Árvores de Decisão
Medidas Tradicionais	99,18	80,96	99,10
DTW	98,22	80,12	98,24
Medidas Tradicionais e DTW	99,96	81,54	99,92

Tabela 4.4: Comparação dos resultados para a combinação entre todas as medidas.

Nas Tabelas 4.5, 4.6 e 4.7 são apresentadas respectivamente as matrizes de confusão para a classificação utilizando caminhadas determinísticas do

turista, medidas tradicionais e combinação entre DTW e medidas tradicionais.

	Aleatória	Livre de Escala	Geográfica	Pequeno Mundo
Aleatória	9554	103	326	17
Livre de Escala	33	9950	13	4
Geográfica	109	21	9828	42
Pequeno Mundo	16	3	25	9956

Tabela 4.5: Matriz de confusão para as caminhadas determinísticas do turista.

	Aleatória	Livre de Escala	Geográfica	Pequeno Mundo
Aleatória	9820	52	113	15
Livre de Escala	7	9969	20	4
Geográfica	43	12	9913	32
Pequeno Mundo	10	8	12	9970

Tabela 4.6: Matriz de confusão para as medidas tradicionais.

	Aleatória	Livre de Escala	Geográfica	Pequeno Mundo
Aleatória	9991	3	4	2
Livre de Escala	0	9998	2	0
Geográfica	2	0	9996	2
Pequeno Mundo	0	0	1	9999

Tabela 4.7: Matriz de confusão para as caminhadas determinísticas do turista em conjunto com as medidas tradicionais.

4.2 Conclusão

Os resultados apresentados acima, corroboram o uso das caminhadas determinísticas do turista como uma métrica robusta para redes complexas. Com a utilização de somente essa métrica, os resultados foram similares aos resultados obtidos por um conjunto de métricas tradicionais. A caracterização das redes pelas caminhadas determinísticas do turista pode auxiliar o uso das redes complexas em diversas áreas, desde a Física através da classificação dos modelos de redes complexas teóricos à visão computacional, reconhecimento de padrões e inteligência artificial.

Reconhecimento de Textura

Textura é uma das mais importantes propriedades da imagem e utilizada em aplicações de visão computacional. A textura está presente na maioria dos objetos e pode ser facilmente entendida pelos sistemas de visão natural como um padrão repetido sobre uma superfície em sua forma exata ou com pequenas variações. Além disso, padrão de textura está diretamente relacionada com as propriedades físicas das superfícies dos objetos, o que o torna um excelente descritor regional.

A literatura apresenta uma vasta quantidade de abordagens para análise de textura. Algumas abordagens são baseadas no estudo das organizações dos pixels em termos de distâncias e direções (Haralick, 1979). Outras abordagens avaliam a complexidade da iteração entre pixels em uma região (Kasparis et al., 2001; Backes and Bruno, 2008). Existe ainda a possibilidade de representar a textura por meio de informações espectrais (Azencott et al., 1997; Jain and Farrokhnia, 1991).

Inicialmente neste capítulo, o método proposto para modelar imagens com redes complexas é descrito. Em seguida, os experimentos executados para avaliar o método proposto no reconhecimento de texturas e compará-lo com os principais métodos da literatura são apresentados.

5.1 Modelagem de Imagens com Redes Complexas

Nas seções abaixo, um método para representação de imagens com redes complexas e a extração de um vetor características com as caminhadas determinísticas do turista é proposto para ser aplicado em problemas de visão computacional. Inicialmente, uma imagem é modelada como um grafo regu-

lar e transformações de eliminação de arestas são aplicadas sobre o grafo, transformando-o em uma rede complexa pequeno-mundo. Em seguida, diversas caminhadas determinísticas do turista são realizadas na rede complexa e então, uma distribuição conjunta do tempo de transiente e período do atrator é construída. Por fim, um vetor de assinatura é construído a partir da distribuição conjunta, sendo esse vetor de assinatura composto por caminhadas realizadas com diferentes memórias. Após a construção do vetor de assinatura para diversas imagens, algoritmos de aprendizagem de máquina podem ser utilizados para classificar as imagens em diferentes classes.

Desta maneira, o método pode ser separado em dois estágios: modelagem da rede complexa e extração de características com as caminhadas determinísticas do turista. As seções abaixo detalham estes dois estágios.

5.1.1 Redes Complexas em Imagens

Para modelar uma imagem como uma rede $G = (V, E)$, cada pixel é mapeado em um vértice do conjunto V . Para simplificar a tarefa de implementação, cada vértice v pode ser rotulado com um valor determinado por $labelv = y + (x - 1) * altura$. Este rótulo pode ser facilmente transformado nas coordenadas dos pixels por $x = \lfloor (labelv - 1) / altura \rfloor + 1$ e $y = ((labelv - 1) \% altura) + 1$, onde $1 \leq x \leq largura$, $1 \leq y \leq altura$ e $\lfloor \cdot \rfloor$ é o operador que retorna o inteiro inferior.

As arestas da rede são definidas pela distância Euclidiana. Dado dois vértices i e j , relacionados aos pixels p_i e p_j . Os vértices são conectados se a distância espacial entre os pixels p_i e p_j é menor que um dado raio r e o peso é definido pela diferença entre a intensidade dos pixels (veja Equações 5.1 e 5.2). O raio r limita as conexões espacialmente distantes na imagem. A Figura 5.1 apresenta uma imagem modelada em uma rede complexa utilizando raio $r = \sqrt{2}$.

$$a_{ij} = \begin{cases} 1, & \text{se } dist(p_i, p_j) \leq r \\ 0, & \text{caso contrário} \end{cases} \quad (5.1)$$

$$w_{ij} = \begin{cases} |I(p_i) - I(p_j)|, & \text{se } dist(p_i, p_j) \leq r \\ 0, & \text{caso contrário} \end{cases} \quad (5.2)$$

onde a_{ij} é a aresta e w_{ij} é a aresta com pesos, $I(x)$ é a intensidade (níveis de cinza) do pixel x , $dist$ é a distância Euclidiana e r é o raio de vizinhança.

A rede obtida é um grafo regular, pois os vértices possuem o mesmo número de conexões (desconsiderando os pixels dos cantos da imagem). O próximo passo é aplicar um conjunto de operações para transformar a rede obtida até o momento em uma rede complexa com uma rica estrutura e topologia. Neste trabalho, é aplicado um simples método para transformar um grafo

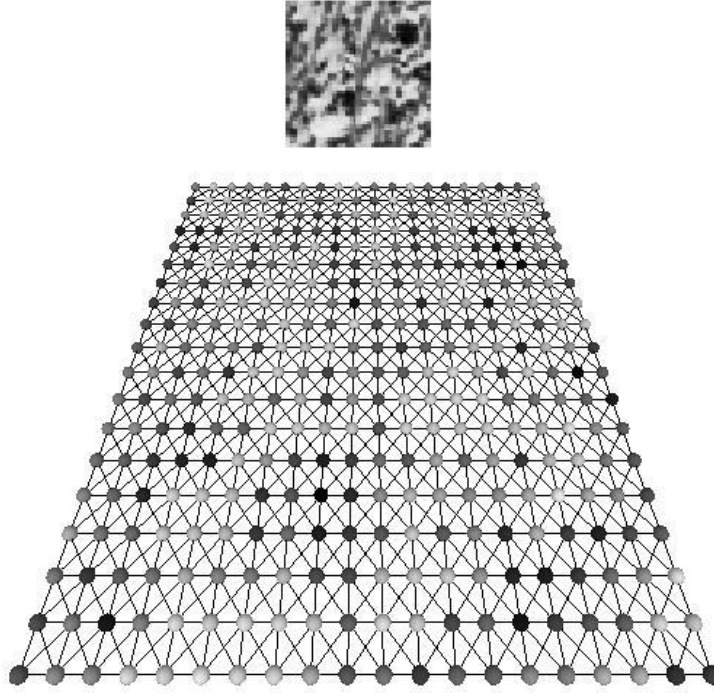


Figura 5.1: Imagem modelada em uma rede complexa utilizando $r = \sqrt{2}$.

regular em uma rede complexa pequeno-mundo. A transformação, definida como $\psi(t, G)$, consiste na seleção de aresta de acordo com o peso. A seleção de arestas é realizada aplicando-se um limiar t , o qual pode ser variado de $0 \leq t \leq \max(W)$, com o intuito de preservar somente as arestas com pesos menores que esse limiar (veja Equações 5.3 e 5.4). As Figuras 5.2 e 5.3 ilustram três transformações $\psi(t, G)$, com $t_1 < t_2 < t_3$ para redes complexas construídas com $r = \sqrt{2}$ e $r = \sqrt{8}$, respectivamente.

$$a_{ij} = \begin{cases} 1, & \text{se } w_{ij} \leq t \\ 0, & \text{caso contrário} \end{cases} \quad (5.3)$$

$$w_{ij} = \begin{cases} w_{ij}, & \text{se } w_{ij} \leq t \\ 0, & \text{caso contrário} \end{cases} \quad (5.4)$$

A transformada $\psi(t, G)$ pode ser vista como uma análise multi-escala da rede, pois para cada valor de t , a rede original é transformada em uma rede com escala t . Para cada escala (valor de t) é obtida uma rede que apresenta diferentes propriedades e revela a estrutura e topologia relacionada a sua escala. Para pequenos valores de t , a rede fornece informações de detalhes na imagem, representando melhor pequenos conjuntos de pixels ou regiões (Figura 5.2 (a)). À medida que t é incrementado, a rede representa melhor as informações globais da imagem. Para altos valores de t , a rede pode fornecer informações de bordas (Figura 5.2 (c)). A proposta do método é combinar

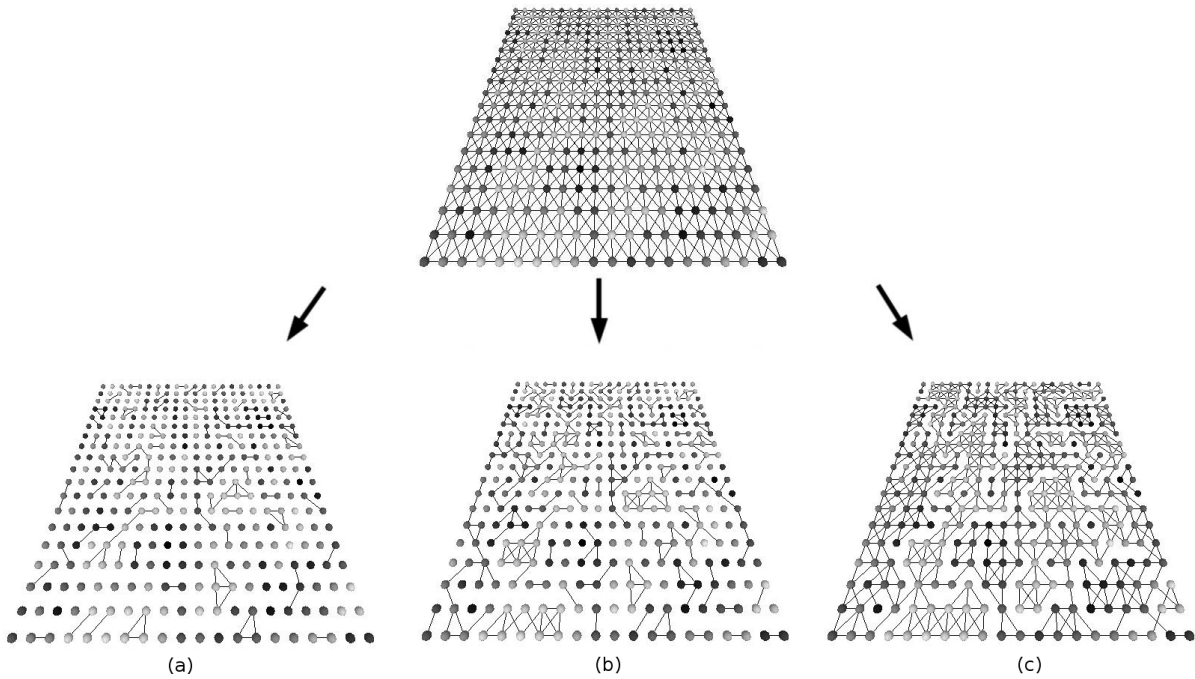


Figura 5.2: Transformações ψ com diferentes valores t em uma rede regular com $r = \sqrt{2}$. (a) $t_1 <$ (b) $t_2 <$ (c) t_3

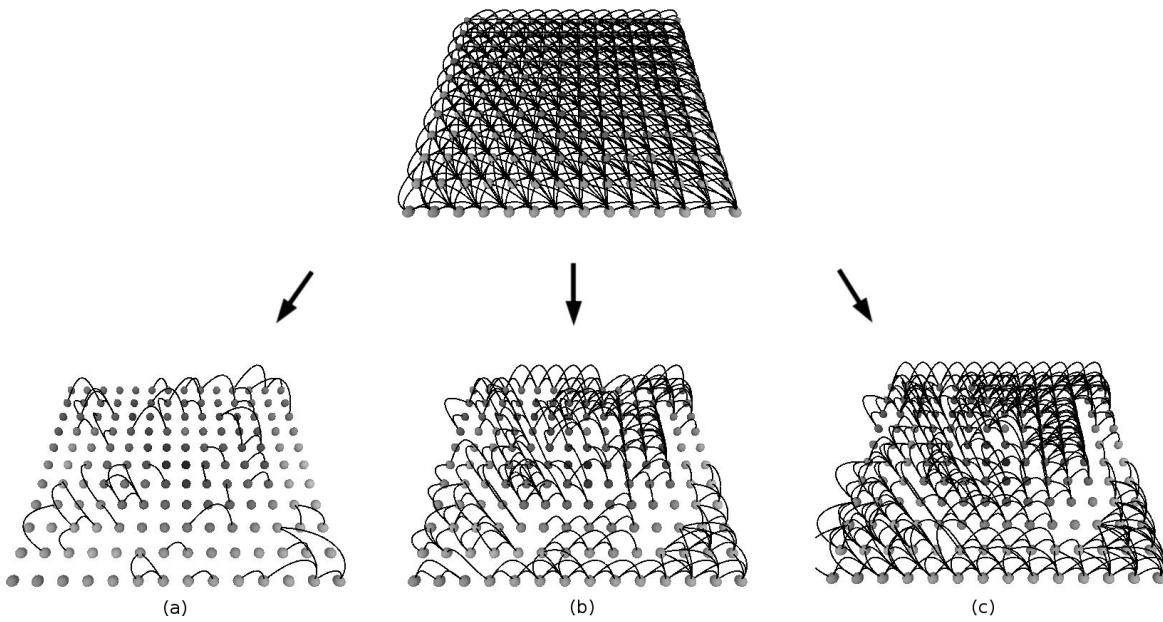


Figura 5.3: Transformações ψ com diferentes valores t em uma rede regular com $r = \sqrt{8}$. (a) $t_1 <$ (b) $t_2 <$ (c) t_3

medidas de diferentes valores de t , obtendo uma análise multi-escala.

A Figura 5.4 apresenta distribuições conjuntas computadas pela caminhada determinística do turista em diferentes configurações da rede. Para cada limiar t , diferentes configurações da rede são obtidas, resultando em diferentes dinâmicas do turista e como consequência, diferentes características da imagem são realçadas.

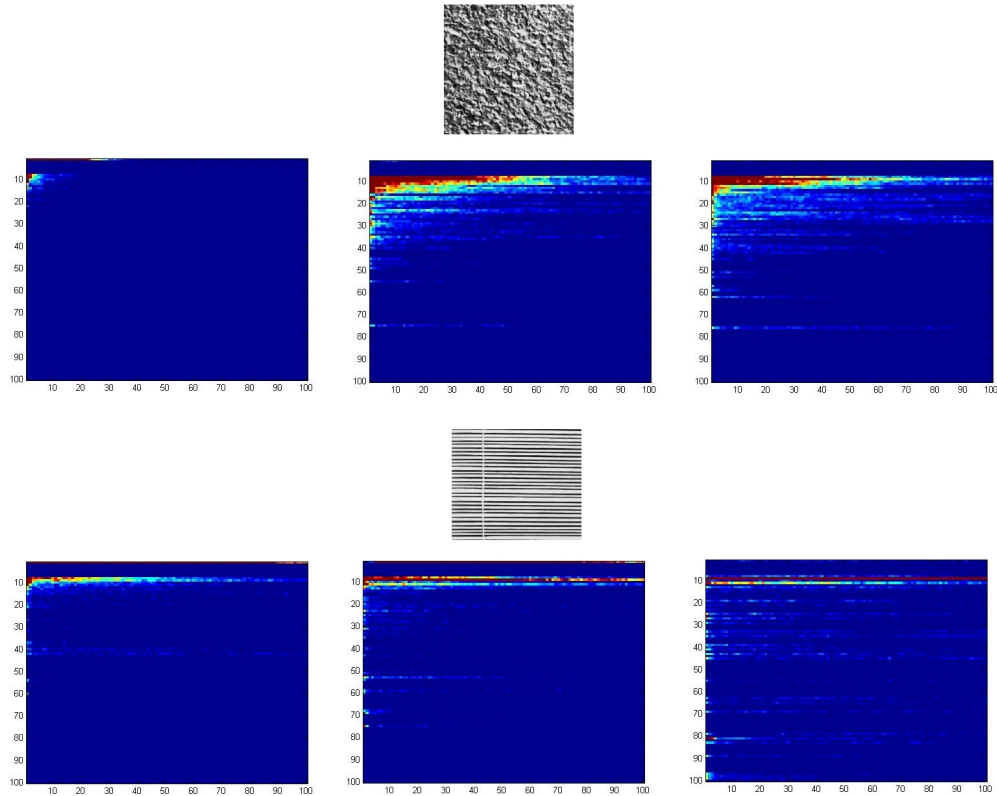


Figura 5.4: Distribuição conjunta obtida pela caminhada determinística do turista em redes regulares transformadas por diferentes limiares. Da esquerda para direita: $t = 10$; $t = 90$ e $t = 170$.

5.1.2 Vetor de Características

Após a modelagem de imagens com redes complexas é necessário caracterizá-las de forma robusta e eficiente. Essa caracterização pode ser realizada por meio das diversas medidas propostas para redes complexas (e.g. média do grau dos vértices, coeficiente de aglomeração). Entretanto, uma técnica que tem demonstrado resultados interessantes e que capturam padrões em diversas escalas simultaneamente são as caminhadas determinísticas do turista. A distribuição conjunta construída a partir das caminhadas do turista depende do valor μ e da dinâmica de movimentação do turista. Além disso, o turista é aplicado em uma rede complexa obtida através de uma transformação com um limiar t . Para capturar informações de diferentes escalas e fonte, um vetor

de assinatura φ_t considerando diferentes valores de μ e diferentes dinâmicas din é apresentado na Equação 5.5. Este vetor de assinatura é composto pela concatenação dos vetores de características $\psi_{\mu,din}$ aplicados em uma rede complexa transformada com um limiar t .

$$\varphi_t = [\psi_{\mu_1,max}, \psi_{\mu_1,min}, \psi_{\mu_2,max}, \psi_{\mu_2,min}, \dots, \psi_{\mu_M,max}, \psi_{\mu_M,min}] \quad (5.5)$$

onde $\psi_{\mu,din}$ pode ser uma das estratégias para extrair informações de uma distribuição conjunta de um turista com memória μ e dinâmica din , max é a dinâmica de movimentação onde o turista caminha para a máxima distância e min é a estratégia análoga.

A utilização de várias memórias e a movimentação máxima e mínima auxilia a caracterização dos padrões e apresenta um grande potencial como um método de classificação de imagens (Backes et al., 2006). Entretanto, o desempenho pode ser aumentado utilizando uma estratégia combinando diferentes limiares t na transformação da rede complexa, pois cada transformação apresenta informações relevantes de escalas. Os valores estão entre o intervalo $t_0 \leq t \leq t_n$ e incrementado por uma constante t_{inc} . Com isso, o vetor de assinatura φ (Equação 5.6) considera o uso de informações extraídas pelo turista em redes complexas transformadas com diferentes limiares.

$$\varphi = [\varphi_{t_0}, \varphi_{t_0+t_{inc}}, \dots, \varphi_{t_n}] \quad (5.6)$$

O método aplicado em problemas de visão computacional é sumarizado na Figura 5.5 e nos passos descritos abaixo.

Passo 1: A imagem é modelada como um grafo regular G .

Passo 2: n transformações de eliminação de arestas com diferentes limiares (t_0, \dots, t_n) são aplicados sobre o grafo G , transformando-o em n redes complexas pequeno-mundo.

Passo 3: Caminhadas do turista com diferentes memórias (μ_1, \dots, μ_M) e dinâmica min e max são aplicadas em cada rede complexa, obtendo-se as distribuições conjuntas.

Passo 4: Para cada distribuição conjunta, um vetor de característica é obtido (e.g. $\psi_{\mu_1,max}$). Em seguida esses vetores são concatenados obtendo uma representação para a rede complexa transformada com um limiar t ($\varphi_t = [\psi_{\mu_1,max}, \psi_{\mu_1,min}, \dots, \psi_{\mu_M,max}, \psi_{\mu_M,min}]$).

Passo 5: Os vetores anteriores são concatenados, obtendo-se o vetor de características final φ , contendo informações de várias escalas e fontes.

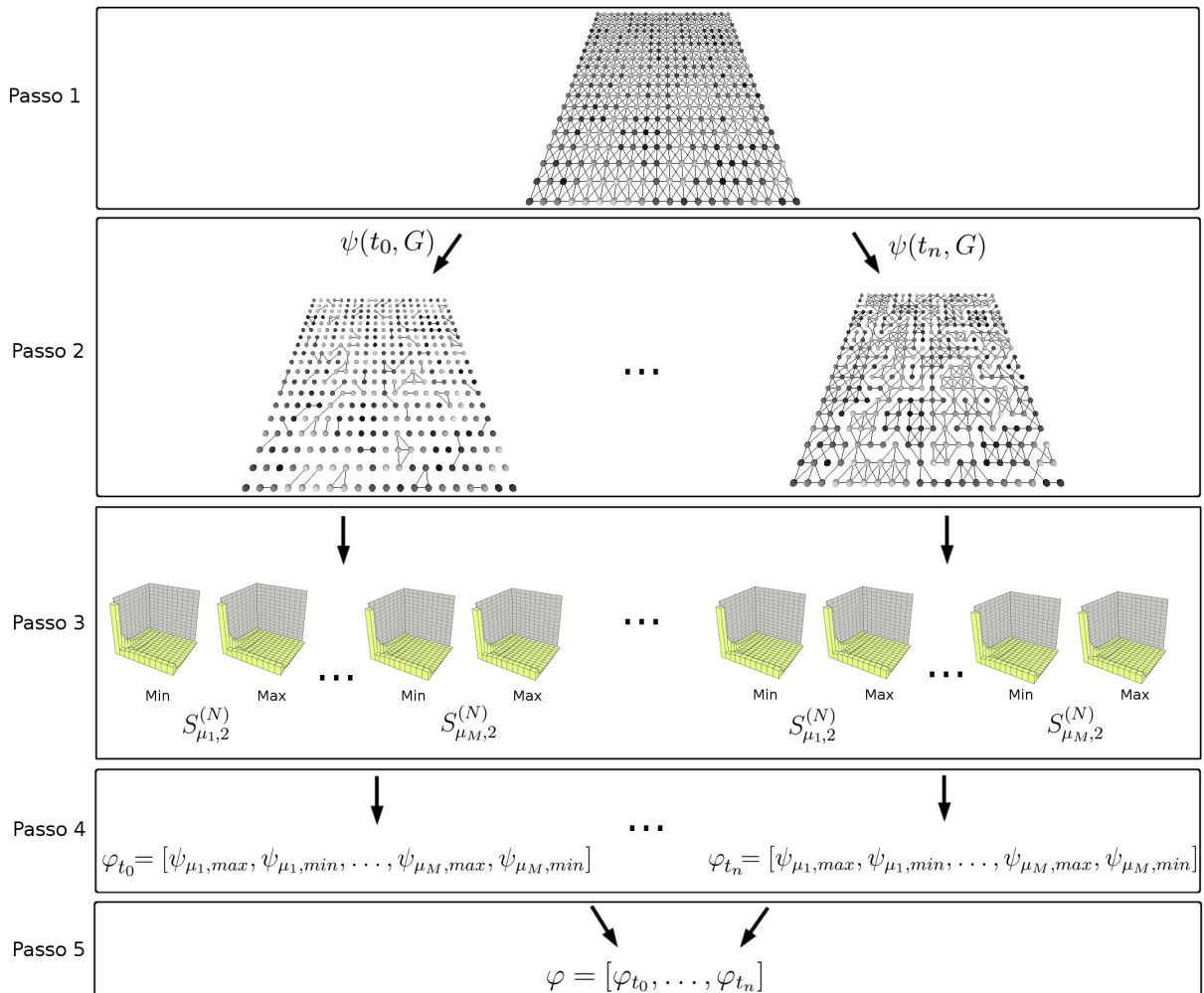


Figura 5.5: Método aplicado em problemas de visão computacional.

5.2 Experimentos e Resultados

Para avaliar a proposta do método e comparar sua eficiência com outros métodos no reconhecimento de texturas, um conjunto de experimentos foi realizado utilizando dois bancos de imagens: texturas sintéticas e texturas naturais. O primeiro banco é composto por imagens do livro de Brodatz (Brodatz, 1966), um conjunto amplamente utilizado na literatura como um *benchmark* para análise de texturas. Neste experimento, cada imagem possui 200×200 pixels com 256 níveis de cinza. Um total de 1110 imagens, divididas igualmente em 111 classes, foram utilizadas. Exemplos de algumas classes de textura são apresentados na Figura 5.6.

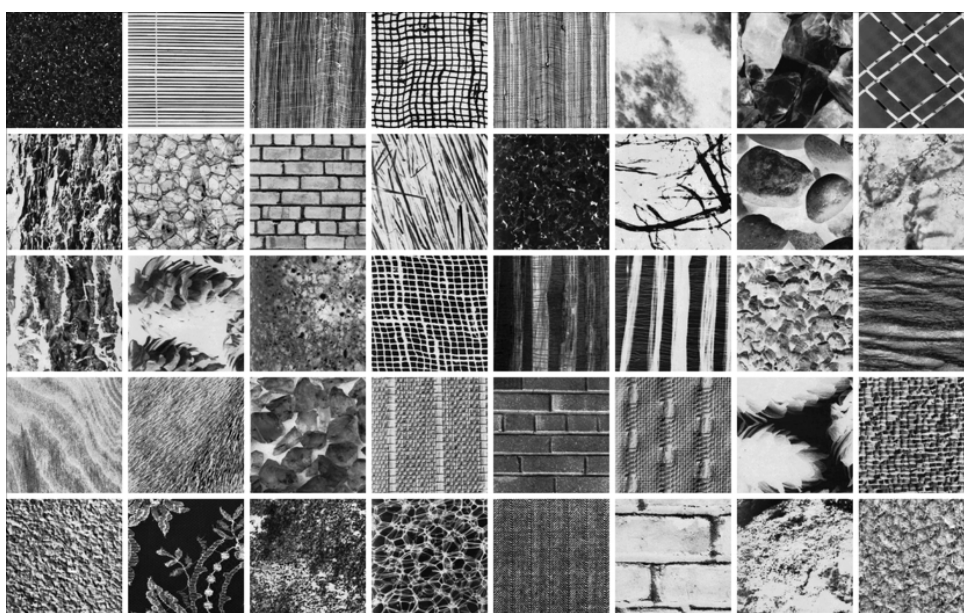


Figura 5.6: Exemplo de algumas classes de texturas de Brodatz utilizadas nos experimentos. Cada imagem possui 200×200 pixels e 256 níveis de cinza.

O segundo banco é composto por imagens de texturas extraídas de folhas de plantas de 5 diferentes espécies da flora brasileira. Para cada espécie, 10 amostras de 128×128 pixels com 256 níveis de cinza foram utilizadas para os experimentos. Exemplos de cada classe e suas variações são apresentados na Figura 5.7. A principal motivação desses experimentos é avaliar o algoritmo em problemas reais e complexos. Desenvolver uma ferramenta computacional que possa identificar plantas é um desafio científico atualmente. Identificar folhas de plantas, bem como, as texturas de folhas é uma tarefa difícil, considerando a alta variabilidade dentro da mesma classe e a similaridade entre as classes (Plotze et al., 2005; Bruno et al., 2008b).

Em cada banco de imagens foram realizados dois experimentos: análise de parâmetros e comparação com outros métodos da literatura. No primeiro, uma análise dos parâmetros do método proposto foi realizada para compre-

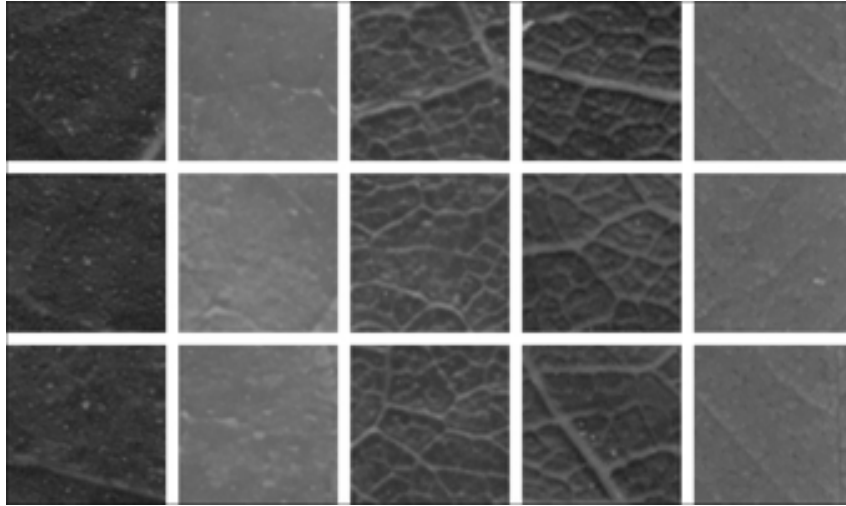


Figura 5.7: Exemplo das texturas das folhas de plantas utilizadas nos experimentos. Cada coluna representa uma diferente classe, enquanto as linhas representam a variação de cada classe.

ender o comportamento do método no reconhecimento de imagens. No segundo experimento, foram realizadas comparações entre o método proposto com algoritmos consagrados pela comunidade, como filtros de Gabor (Jain and Farrokhnia, 1991), matriz de co-ocorrência (Haralick, 1979) e descritores de Fourier (Azencott et al., 1997). Para ambos os experimentos, um vetor de características foi extraído para cada imagem do banco e um processo de classificação utilizando K vizinhos mais próximos (Mitchell, 1997) foi executado, com $k = 1$. Um classificador simples foi escolhido para evidenciar a regra das características no reconhecimento de padrões. Para os experimentos, foi adotado o esquema leave-one-out cross-validation para separação do banco em conjunto de treinamento e classificação.

5.2.1 Análise dos Parâmetros do Método

Esta seção apresenta um estudo sobre os parâmetros do método proposto baseado em seu desempenho no reconhecimento de texturas. Os parâmetros considerados para as redes complexas são: (i) a influência do raio de vizinhança para a transformação que converte uma imagem em uma rede, (ii) os limiares utilizados nas transformações multi-escala (t_0, t_{inc}, t_f). Os parâmetros do turista são: (iii) as memórias utilizadas na caminhada determinística do turista, (iv) o critério de desempate da caminhada, (v) estratégia de extração de informações da distribuição conjunta. O objetivo dessa análise é apresentar o comportamento do método e determinar o melhor conjunto de parâmetros.

Parâmetros da Rede Complexa

Os principais parâmetros da construção da rede complexa são os raios de vizinhança r e o limiar inicial t_0 . A taxa de vizinhança é o parâmetro que determina a zona de influência das conexões de um vértice da rede. Como este parâmetro determina a escala de conexões de cada vértice, ele está relacionado com a estrutura da rede e mudanças nesse parâmetro pode influenciar diretamente o desempenho no reconhecimento de imagens. Os limiares t_0, t_{inc} e t_f são utilizados na transformação ψ , o qual transforma um grafo regular em uma rede complexa e fornece uma análise multi-escala. O limiar inicial t_0 é o parâmetro responsável por limitar a escala da análise do método.

A análise de parâmetros inicia com a escolha de uma faixa de r e t_0 onde o método alcança seu desempenho ótimo. A Figura 5.8 apresenta um gráfico do desempenho em porcentagem de classificação correta para os valores de r e t_0 . É possível notar que o método alcança seus melhores resultados para $r \geq \sqrt{8}$. Uma análise mais precisa é apresentada na Tabela 5.1, onde o melhor resultado é obtido para $r = \sqrt{16}$ e $t_0 = 13$. Estes valores, podem ser combinados com os outros parâmetros para encontrar o conjunto de parâmetros adequados para o método.

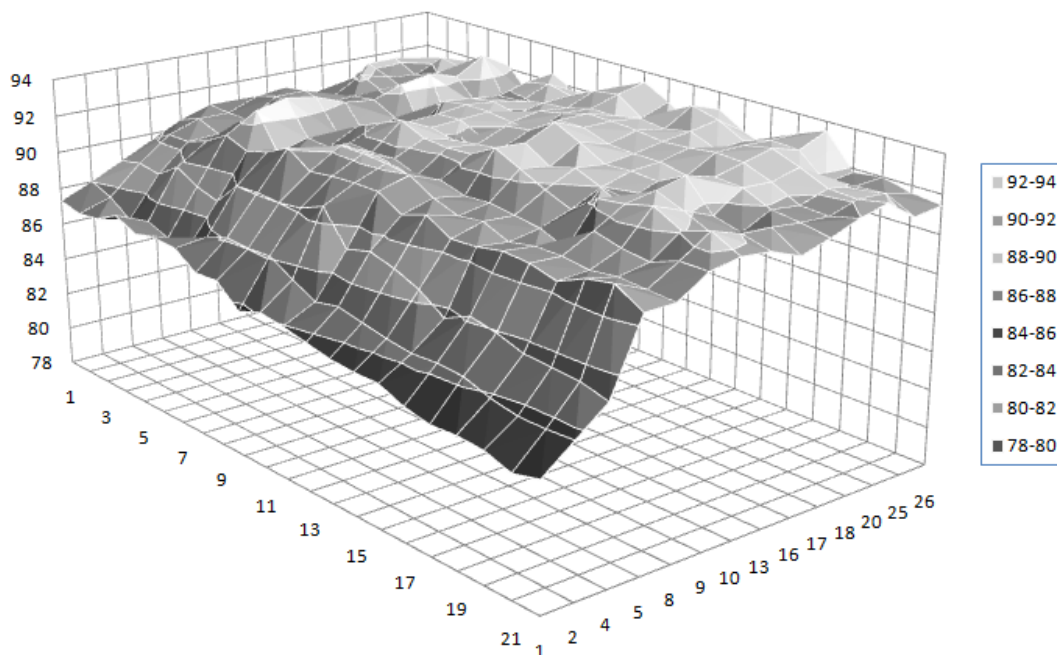


Figura 5.8: Gráfico do desempenho no reconhecimento de textura. Porcentagem de classificação correta em função do raio r em $\sqrt{\cdot}$ e limiar inicial t_0 .

A Figura 5.9 apresenta a análise para todos os parâmetros de construção da rede complexa. Para isso, os valores encontrados anteriormente foram fixados. Na Figura 5.9(a) é apresentado um gráfico do raio r versus porcentagem

Raio r ($\sqrt{\cdot}$)	Limiar Inicial t_0						
	8	9	10	13	14	15	16
8	92,25	92,16	91,98	91,26	90,81	90,99	90,45
9	91,62	91,17	91,35	90,99	91,08	90,90	90,72
10	92,34	92,25	92,34	93,06	92,43	91,80	91,26
13	91,89	92,07	92,25	92,97	92,16	92,52	91,98
16	91,62	92,70	92,88	93,69	93,06	92,34	92,43
17	92,16	92,07	93,06	93,06	92,61	92,52	92,52
18	92,43	92,70	93,06	93,06	92,97	92,79	92,88
20	92,25	92,61	92,61	92,88	92,70	92,97	92,43
25	92,88	93,06	92,16	92,43	93,60	93,24	92,52

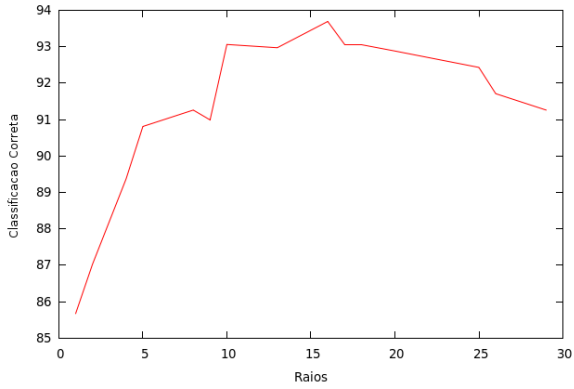
Tabela 5.1: Análise conjunta entre limiar inicial t_0 e raio r .

de classificação correta. Note que, como nos resultados anteriores, a classificação alcança seu máximo em 93,69% para $r = \sqrt{16}$. A Figura 5.9(b) apresenta um gráfico de t_0 versus porcentagem de classificação correta. A porcentagem máxima é alcançada para $t_0 = 13$. Para baixos valores de t_0 , como aqueles menores que 8, a rede pode ser altamente perturbada por ruído. Por outro lado, para altos valores de t_0 , a rede pode não apresentar os detalhes da imagem e perder informações de texturas importantes.

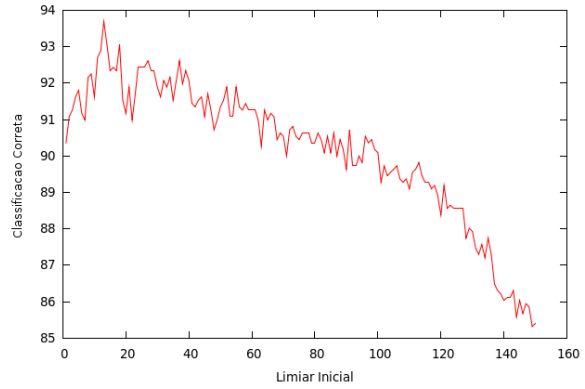
As Figuras 5.9(c) e 5.9(d) apresentam respectivamente t_{inc} e t_f versus porcentagem de classificação correta. Para apresentar esses gráficos, foi utilizado $t_f = t_0 + x * t_{inc}$, onde x é o número de vezes de t_{inc} . Esta função faz com que t_{inc} e t_f sejam valores correlacionados. Na primeira análise, foi utilizado $x = 3$ (veja Figura 5.9(d)). O gráfico mostra que o melhor resultado foi alcançado para altos valores de t_{inc} . Isto era esperado, pois para altos valores de t_{inc} o método pode explorar uma melhor faixa de escalas na análise multi-escala. Na Figura 5.9(d), o gráfico mostra o número de limiares, dado por x o qual determina t_f , versus a porcentagem de classificação correta. Note que a curva obtém seu máximo quando $x = 3$. Com estes quatro gráficos, nós podemos determinar o conjunto adequado de parâmetros para a construção da rede complexa, o qual é $r = \sqrt{16}$, $t_0 = 13$, $t_{inc} = 59$ e $t_f = t_0 + 3 * t_{inc}$.

Parâmetros da Caminhada Determinística do Turista

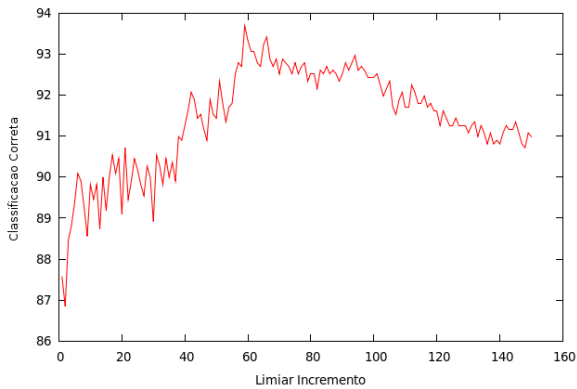
Com os parâmetros das redes complexas multi-escala determinados, os experimentos apresentados a seguir foram realizados para avaliar os parâmetros da caminhada determinística do turista. A Tabela 5.2 apresenta os resultados para diferentes valores de memória μ da caminhada do turista. Nós notamos que a caminhada do turista apresenta melhores resultados quando a dinâmica do turista é em direção à máxima distância, comparadas com a mínima distância. Neste caso, os atratores do turista são formados em regiões heterogêneas, i.e., regiões onde a distância entre dois vértices é elevada, o que



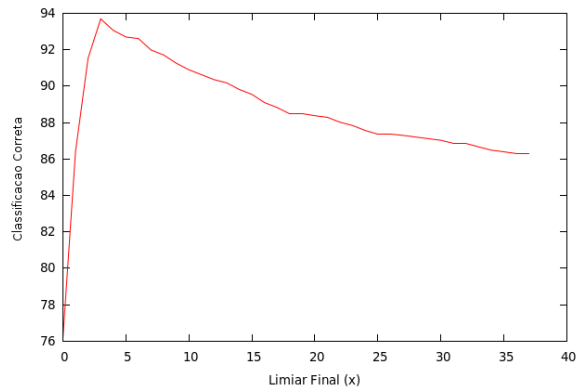
(a)



(b)



(c)



(d)

Figura 5.9: (a) Resultados comparativos para diferentes raios ($\sqrt{\cdot}$), (b) Limiar t_0 X Porcentagem de Classificação Correta, (c) Limiar t_{inc} X Porcentagem de Classificação Correta e (d) Limiar t_f X Porcentagem de Classificação Correta.

caracteriza a presença de contornos na imagem. A combinação entre vetores de características com as dinâmicas mínima e máxima ($[Min, Max]$) incrementou a porcentagem de classificação correta. Isso ocorre porque o novo vetor de assinatura apresenta ambas as informações homogêneas e heterogêneas da imagem, fornecendo assim uma ferramenta mais poderosa para análise de imagens.

	Memória (μ)				
	1	2	3	4	5
<i>Min</i>	49,55	45,41	44,14	38,74	36,76
<i>Max</i>	90,33	85,22	83,33	79,46	79,46
$[Min, Max]$	93,69	89,10	87,75	84,59	84,23

Tabela 5.2: Porcentagem de classificação correta para $\psi_{\mu, din}$ utilizando diferentes valores de μ e dinâmica din na caminhada do turista.

Os resultados também mostraram que, para ambas as regras de movimentação máxima e mínima e suas combinações, a porcentagem de classificação correta diminui ao mesmo passo que a memória é incrementada. A explicação para este resultado está relacionada com o fato de que, se a memória é incrementada, o turista possui mais dificuldade em encontrar um atrator na rede complexa (Backes et al., 2009b). Com isso, o tamanho da caminhada é incrementado e as longas caminhadas podem conduzir a uma armadilha (caso onde o turista não consegue encontrar um atrator). Caso o turista não encontre um atrator, este fato pode influenciar na sua distribuição conjunta e afetar a classificação das amostras. Por outro lado, os pequenos valores de memória μ permitem uma melhor análise local da textura da imagem, refletindo em uma maior porcentagem de classificação correta.

A Tabela 5.3 apresenta os resultados quando múltiplos valores de μ são considerados. Como nos resultados individuais para cada memória, a dinâmica em direção a máxima distância apresentou melhores resultados quando comparado com a dinâmica em direção a mínima distância. Para memórias $\{1, 2, 3\}$, o método obteve 94,32%. A utilização de múltiplas memórias diminui a importância individual de cada valor de μ , proporcionando assim uma classificação de imagens mais robusta.

Memórias (μ)	<i>Min</i>	<i>Max</i>	$[Min, Max]$
$\{1\}$	49,55	89,33	93,69
$\{1, 2\}$	64,86	89,41	93,96
$\{1, 2, 3\}$	67,93	90,78	94,32
$\{1, 2, 3, 4\}$	67,30	91,10	94,23
$\{1, 2, 3, 4, 5\}$	67,66	90,96	94,05

Tabela 5.3: Porcentagem de classificação correta para φ combinando diferentes valores de μ .

Uma questão importante na caminhada do turista é o critério de desempate (Campiteli et al., 2006b; Backes et al., 2006). Este problema consiste em decidir a próxima cidade que o turista irá visitar, dado um conjunto de cidades vizinhas igualmente próximas da cidade atual. As três estratégias avaliadas foram: primeira, aleatória e visibilidade. A estratégia primeira escolhe a cidade com base no seu índice. A cidade com menor índice entre as cidades empatadas é escolhida. Essa estratégia mantém a característica determinística do método. A estratégia aleatória sorteia uma das cidades empatadas. Por fim, a estratégia visibilidade aumenta a visibilidade do turista e então decide qual o melhor caminho a seguir. Essas estratégias foram descritas no Capítulo 3.

A última estratégia precisa ser avaliada com relação ao tamanho da visibilidade do turista. Na Figura 5.10, a visibilidade do turista é avaliada. Caso, a caminhada ainda esteja empatada no final, a estratégia aleatória é aplicada. Para visibilidade composta por $n = 3$, o turista alcança 93,48% de classificação correta.

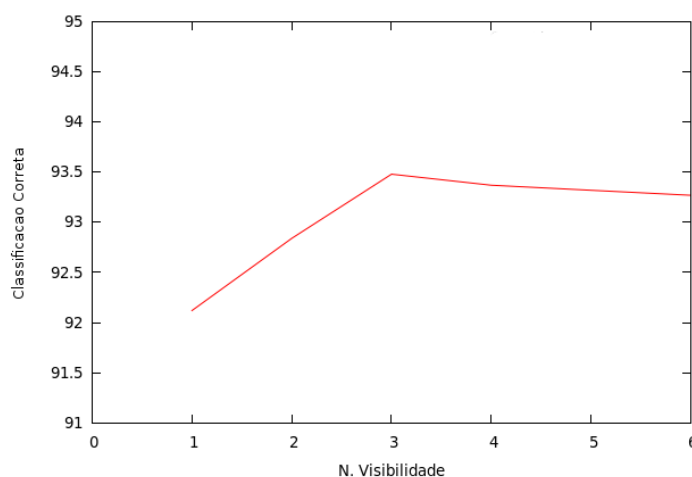


Figura 5.10: Visibilidade do turista X porcentagem de classificação correta para a estratégia de desempate.

Na Tabela 5.4 são apresentados os resultados obtidos para diferentes critérios de desempate. A estratégia primeira obteve o melhor desempenho comparado com as estratégias aleatórias e visibilidade. Embora o desempenho entre todas as estratégias seja parecido.

Critérios	Classificação (%)	Desvio Padrão
Primeira	94,32	1,99
Aleatória	92,67	2,36
Visibilidade	93,48	2,27

Tabela 5.4: Comparação entre as estratégias de desempate.

Por fim, experimentos para determinação dos parâmetros do método proposto foram realizados nas estratégias de extração de informações da dis-

tribuição conjunta. A distribuição conjunta é o resultado da aplicação da caminhada determinística do turista em um ambiente (e.g. rede complexa, imagem). Para a utilização da caminhada determinística do turista de forma eficiente, informações relevantes precisam ser extraídas dessa distribuição. As estratégias consideradas foram: região, histograma e estatística, descritas no Capítulo 3.

A avaliação dos parâmetros de cada estratégia é apresentado na Figura 5.11. A Figura 5.11(a) apresenta a avaliação dos parâmetros a e b responsáveis pelo tamanho da região da distribuição conjunta a ser considerada. Os melhores resultados são obtidos para pequenos valores de a e b , o que sugere que as informações relevantes da distribuição conjunta estão concentrados no início da distribuição. Para valores $a = 1$ e $b = 2$, a estratégia alcançou seu melhor desempenho de 93,51%. A avaliação do número de valores n da estratégia histograma é apresentado na Figura 5.11(b). Como pode ser visto, os melhores resultados são obtidos para os primeiros valores do histograma ($n \leq 8$), novamente sugerindo que as informações relevantes da distribuição estão no seu início. Para valores $n \geq 9$, os resultados obtidos decaíram acentuadamente e em seguida se mantiveram praticamente constante. O melhor resultado foi de 94,86% utilizando $n = 6$. Por fim, os resultados das estatísticas calculadas da distribuição são apresentados na Figura 5.11(c). Cada estatística foi avaliada separadamente, tendo a energia obtido o melhor resultado. Além disso, foram avaliadas todas as estatísticas em conjunto, obtendo 88,20%, e considerando as 4 melhores estatísticas (entropia, energia, homogeneidade e variância inversa), obtendo 91,98%.

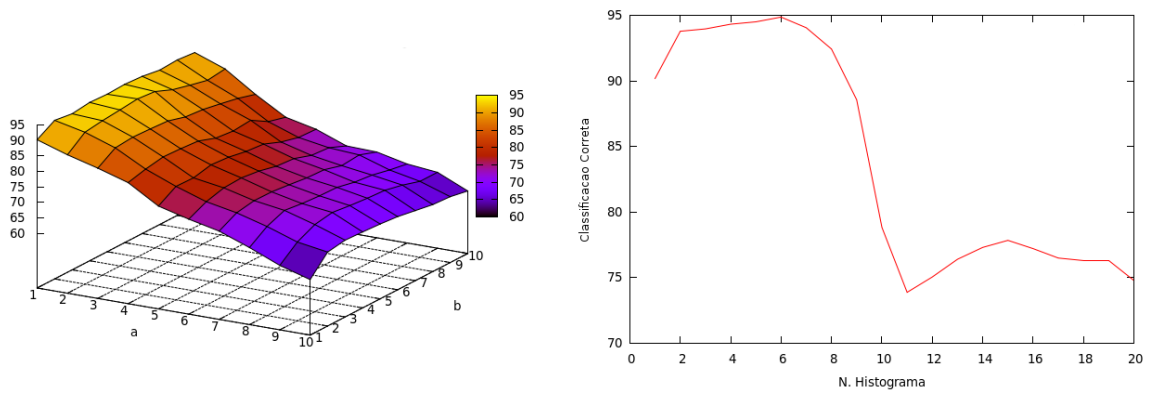
Na Tabela 5.5 são apresentados os resultados comparativos finais para diferentes estratégias de extração de informações descritas no Capítulo 3. A estratégia Histograma obteve a melhor porcentagem de classificação comparadas com as demais.

Estratégias	Classificação (%)	Desvio Padrão
Estatística	91,98	2,27
Região	93,51	1,84
Histograma	94,86	1,94

Tabela 5.5: Comparação entre diferentes estratégias para extração de informações da distribuição conjunta.

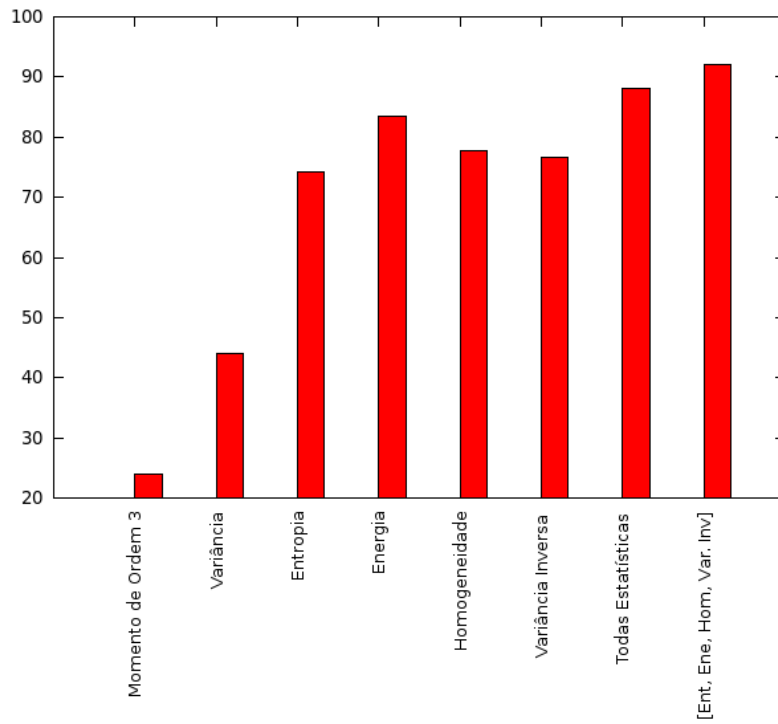
5.2.2 Comparação com Métodos da Literatura

Para uma análise mais completa do método proposto, comparações com métodos tradicionais de análise de textura foram realizadas. Os seguintes métodos foram considerados na comparação:



(a)

(b)



(c)

Figura 5.11: Avaliação dos parâmetros das estratégias de extração de informações (a) Parâmetro a e b para estratégia Região, (b) Quantidade de valores considerados do histograma, (c) Estatísticas.

Descritores de Fourier: nos experimentos, estes descritores consiste em um vetor de características com a energia dos 99 coeficientes mais significativos da transformada de Fourier. Cada coeficiente corresponde a soma dos valores absolutos do espectro dentro de uma distância radial do centro (depois de executar a operação *shifting*) de uma transformação bi-dimensional. Mais informações sobre os descritores de Fourier podem ser encontradas em (Azen-cott et al., 1997).

Matrizes de Co-ocorrência: estas matrizes representam a distribuição de probabilidade conjunta entre pares de pixels a uma dada distância e direção. Nos experimentos, distâncias de 1 a 5 pixels com ângulos de -45° , 0° , 45° , 90° , em uma versão não simétrica foram considerados. Energia, entropia, correlação, momento da diferença inversa - IDM e inércia foram computadas das matrizes resultantes (Haralick, 1979).

Filtros de Gabor: este método convolui uma imagem por uma família de filtros de Gabor (i.e., diferentes escalas e orientações da mesma configuração original). Cada filtro de Gabor é definido como uma função bi-dimensional gaussiana modulada por uma senoide orientada em uma determinada frequência e direção. Para os experimentos, uma família de 16 filtros (4 rotações e 4 escalas), com frequência variando de 0,01 a 0,30, foram utilizadas. Informações adicionais podem ser encontradas em Refs. (Jain and Farrokhnia, 1991; Manjunath and Ma, 1996; Daugman and Downing, 1995; Idrissa and Acheroy, 2002).

Caminhadas Determinística do Turista em Redes Complexas: Seguindo os experimentos realizados acima, o método proposto foi aplicado com os seguintes parâmetros: $r = \sqrt{16}$, $t_0 = 13$, $t_i = 59$, $x = 3$ para as redes complexas e $\mu = \{1, 2, 3\}$, desempate Primeira e extração de informações pela estratégia Histograma com $n = 6$.

Os resultados dos métodos de análise de texturas descritos na literatura e do método proposto são comparados na Tabela 5.6 utilizando o banco de imagens Brodatz. O método proposto supera os métodos tradicionais (como os descritores de Fourier, matrizes de co-ocorrência e filtros de Gabor). A caminhada do turista em redes complexas alcançou 94,86% contra uma porcentagem de 88,99% obtida pelos filtros de Gabor, o segundo melhor nos experimentos. Esses resultados mostram um grande potencial para análise e classificação de imagens utilizando a combinação entre as redes complexas e a caminhada do turista.

As matrizes de confusão para cada método são apresentadas nas Figuras 5.12 e 5.13. As matrizes são representadas por imagens, onde cada pixel corresponde a uma posição da matriz. Quanto menos pontos fora da diagonal, melhor o resultado obtido pelo método.

Método	Classificação (%)	Desvio Padrão
Fourier	70,67	3,39
Matriz de Co-ocorrência	86,49	2,69
Filtros de Gabor	88,99	2,03
Caminhada do Turista em Redes	94,86	1,94

Tabela 5.6: Comparação dos resultados para diferentes métodos de caracterização de texturas no banco de imagens Brodatz.

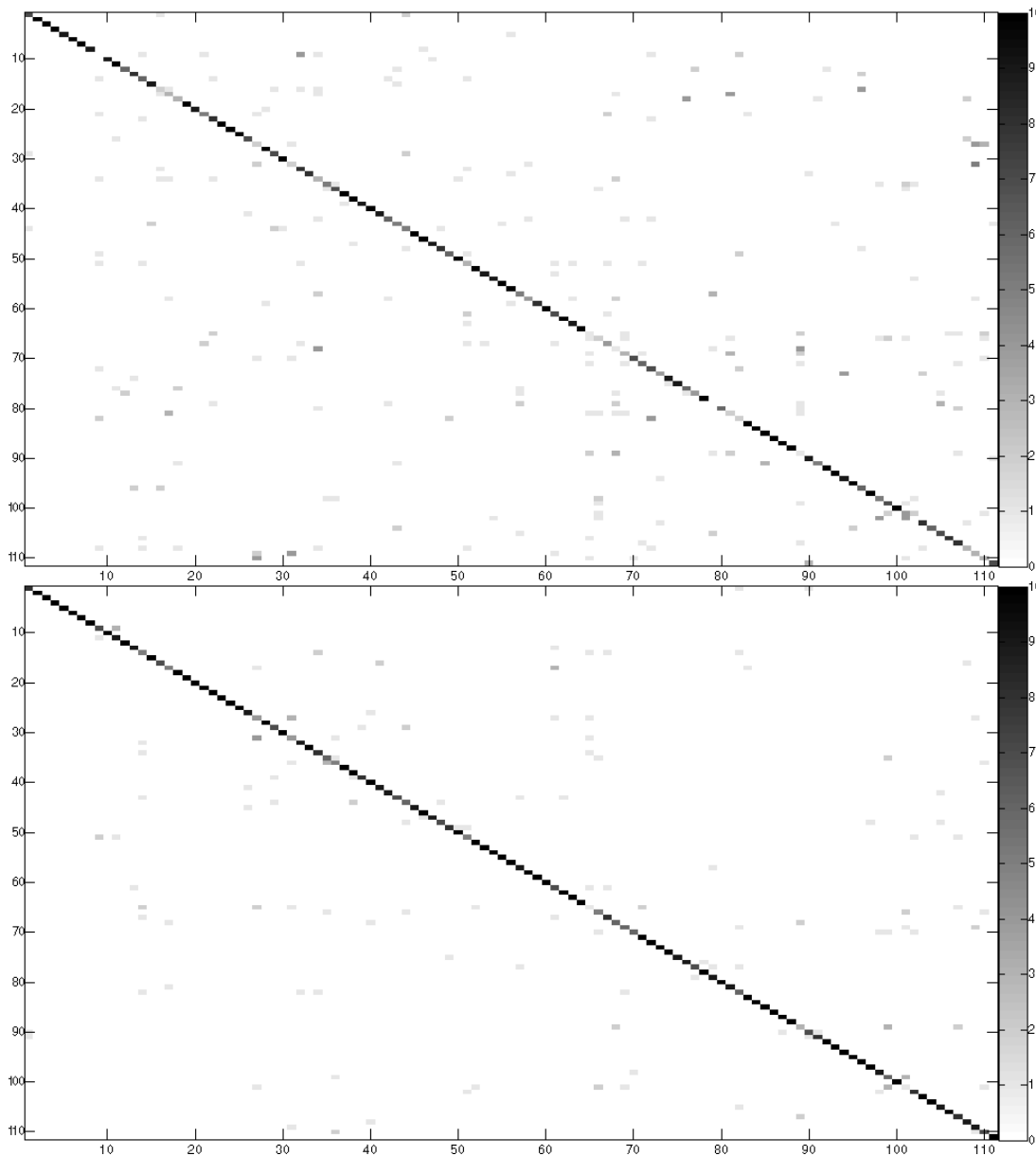


Figura 5.12: Matrizes de confusão para o banco de imagens Brodatz. (a) Descritores de Fourier. (b) Matriz de co-ocorrência.

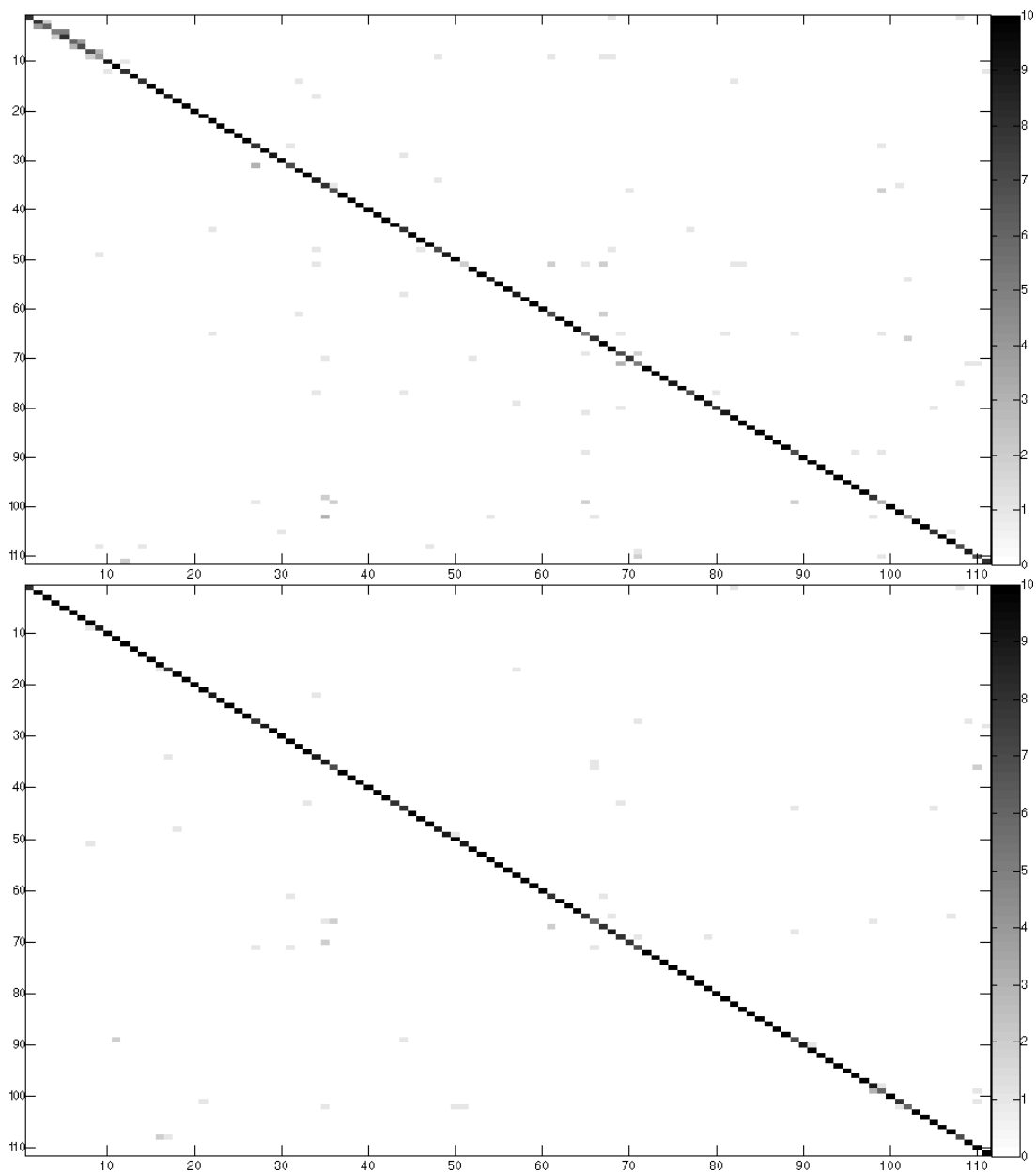


Figura 5.13: Matrizes de confusão para o banco de imagens Brodatz. (a) Filtros de Gabor. (b) Método proposto.

Os resultados apresentados acima são baseados em um banco de imagens sintéticas (banco Brodatz). Os resultados comparativos para o banco de imagens de folhas reais são apresentados na Tabela 5.7. Como nos resultados para texturas sintéticas, a caminhada do turista em redes complexas mostraram-se superior aos métodos tradicionais. É importante enfatizar que a classificação de texturas de folhas de plantas é uma tarefa difícil devido à pequena variação entre as classes (Figura 5.7). Estes resultados mostram que a caminhada do turista em redes complexas é eficiente na caracterização de padrões de texturas similares, obtendo resultados superiores quando comparados com os métodos tradicionais.

Método	Classificação (%)	Desvio Padrão
Fourier	80,20	18,96
Filtros de Gabor	85,00	13,74
Matriz de co-ocorrência	92,40	9,55
Caminhada do Turista em Redes	98,00	6,03

Tabela 5.7: Comparação dos resultados para diferentes métodos na caracterização de texturas no banco de imagens de folhas de plantas.

Nas Tabelas 5.8 e 5.9, o teste t-pareado com significância $\alpha = 0,05$ (Nadeau and Bengio, 2003) foi executado com o intuito de comparar estatisticamente os métodos no reconhecimento de texturas. Nas linhas estão diferentes algoritmos de classificação (todos os algoritmos estão disponíveis no aplicativo Weka (Eibe et al., 1999) e foram executados com os parâmetros padrões da ferramenta). Nas colunas estão os diferentes métodos de extração de características de texturas. A base do teste estatístico é o método proposto. Nessa tabela, são apresentados a porcentagem de classificação correta e um símbolo que indica se o método foi estatisticamente igual (=), superior (+) ou inferior (-) ao método base. Por fim, na última linha da tabela, é apresentado o número de vezes que cada método foi estatisticamente igual, superior ou inferior ao método proposto.

Na Tabela 5.8 é apresentado o teste estatístico para o banco de imagens de Brodatz. Comparando com os filtros de Gabor, Gabor foi estatisticamente inferior ao método proposto com os classificadores Knn (Mitchell, 1997), Naive Bayes (Zhang, 2004) e máquinas de vetor suporte (support vector machines - SVM) (Meyer et al., 2003), por outro lado, os filtros de Gabor foram estatisticamente equivalentes com os classificadores BayesNet (Su and Zhang, 2006), árvore de decisão (Breiman et al., 1984) e Redes Neurais (Arbib, 1995). Considerando todos os classificadores, o método proposto foi 15 vezes estatisticamente superior, 2 vezes equivalente e 1 vez inferior.

Na Tabela 5.9 é apresentado o teste estatístico para o banco de imagens de Brodatz. Embora o método tenha obtido uma porcentagem de classificação

	Método proposto	Fourier	Gabor	Matrizes
Knn	94,86	70,67 (-)	88,99 (-)	86,49 (-)
Naive Bayes	89,70	53,45 (-)	86,41 (-)	83,71 (-)
BayesNet	85,36	47,89 (-)	83,76 (=)	79,14 (-)
Árvore de Decisão (J48)	69,96	51,59 (-)	68,29 (=)	76,31 (+)
SVM	93,53	57,89 (-)	84,47 (-)	68,19 (-)
Redes Neurais	98,00	75,27 (-)	91,28 (=)	91,09 (=)
	(+/=/-)	(0/0/6)	(0/3/3)	(1/1/4)

Tabela 5.8: Teste t-pareado para os métodos no banco de imagens de Brodatz.

correta superior na maioria dos casos, o método proposto foi 7 vezes estatisticamente superior, 10 vezes equivalente e 1 vez inferior aos métodos da literatura.

	Método proposto	Fourier	Gabor	Matrizes
Knn	98,00	80,20 (-)	85,00 (-)	92,40 (=)
Naive Bayes	88,80	63,40 (-)	83,20 (=)	88,40 (=)
BayesNet	75,40	58,40 (=)	66,20 (=)	73,00 (+)
Árvore de Decisão (J48)	78,60	55,20 (-)	76,60 (=)	80,20 (=)
SVM	86,20	69,00 (-)	84,20 (=)	73,20 (-)
Redes Neurais	98,00	83,40 (-)	88,60 (=)	96,00 (=)
	(+/=/-)	(0/1/5)	(0/5/1)	(1/4/1)

Tabela 5.9: Teste t-pareado para os métodos no banco de imagens de folhas.

5.2.3 Tolerância a Ruídos

Uma característica importante para os métodos de reconhecimento de imagens é a sua tolerância a ruídos. Esses ruídos são provenientes de várias fontes, como variação nas condições de iluminação e qualidade do sensor do dispositivo de captura. Para avaliar a tolerância dos métodos, experimentos foram realizados em um banco modificado. Ruídos *salt and pepper* com densidades $d = [0,01 \ 0,05 \ 0,07 \ 0,1 \ 0,3 \ 0,5]$ foram adicionados para cada imagem dos bancos de Brodatz e de folhas de plantas. Os ruídos com densidade d afetam aproximadamente $d * w * h$ dos pixels da imagem, onde w e h são a largura e altura da imagem respectivamente. Exemplos de imagens acrescidas de ruídos são ilustradas nas Figuras 5.14 para o banco de imagens Brodatz e na Figura 5.15 para o banco de imagens de folhas.

Os resultados obtidos no banco de imagens Brodatz com ruídos são apresentados na Tabela 5.10 para $d = [0,01 \ 0,05 \ 0,07]$ e na Tabela 5.11 para $d = [0,1 \ 0,3 \ 0,5]$. Para ruídos com densidades baixas (Tabela 5.10), o método proposto é superior aos métodos tradicionais, embora todos os métodos apresentem alguma tolerância aos ruídos. Entretanto, quando a densidade é incrementada e a textura na imagem passa a ser mínima (Tabela 5.11), os resultados

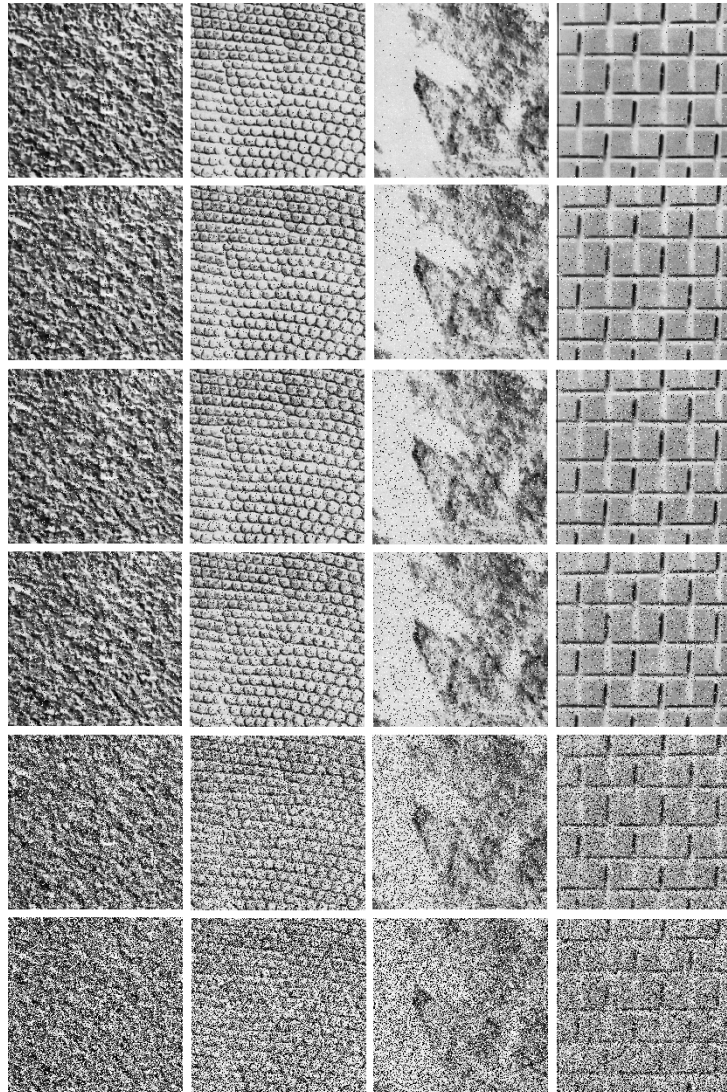


Figura 5.14: Exemplo de algumas imagens de texturas de Brodatz com ruídos utilizadas nos experimentos. Cada linha representa os ruídos com densidades $d = [0,01 \ 0,05 \ 0,07 \ 0,1 \ 0,3 \ 0,5]$ e as colunas representam diferentes classes.

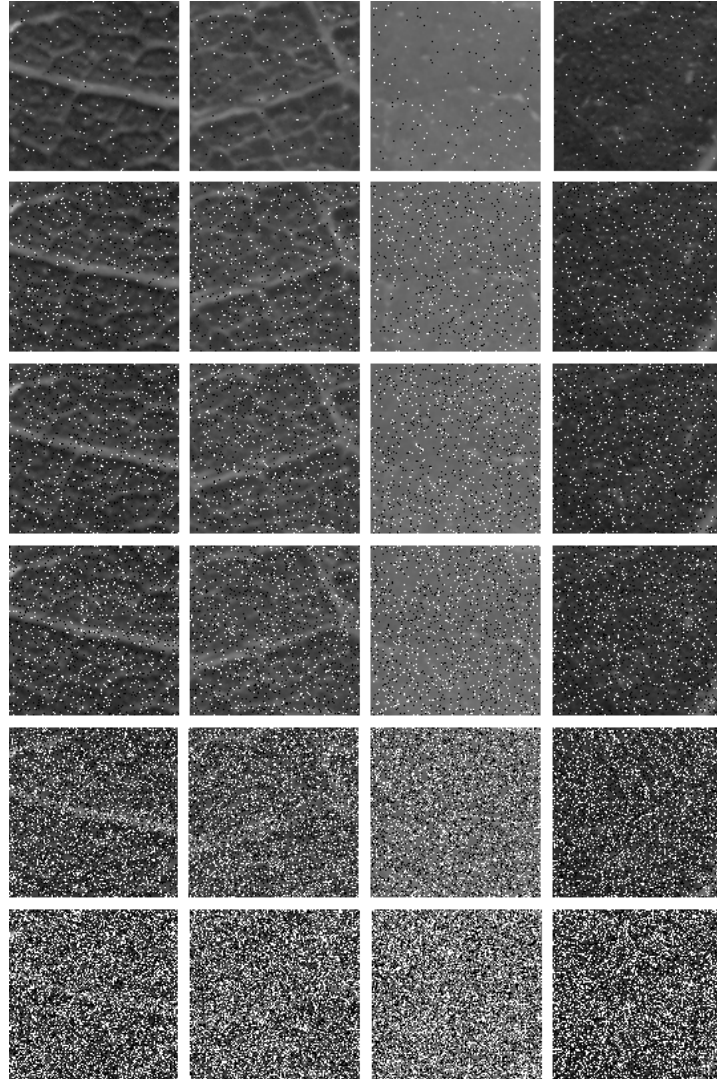


Figura 5.15: Exemplo de algumas imagens de texturas de folhas com ruídos utilizadas nos experimentos. Cada coluna representa os ruídos com densidades $d = [0,01 \ 0,05 \ 0,07 \ 0,1 \ 0,3 \ 0,5]$ e as colunas representam diferentes classes.

alcançados pelos métodos tradicionais decaem rapidamente, enquanto que o método proposto ainda fornece uma porcentagem de classificação de 88,95% contra 82,31% obtido pelos filtros de gabor para $d = 0,5$. Estes resultados mostram que o método proposto é mais eficiente para a classificação de padrões mesmo na presença de alta densidade de ruído.

Método	$d = 0,01$	$d = 0,05$	$d = 0,07$
Fourier	71,15(3,61)	72,28(3,27)	71,68(3,33)
Filtros de Gabor	88,86(1,93)	88,59(1,97)	88,53(2,11)
Matriz de co-ocorrência	86,29(2,67)	86,46(2,54)	86,15(2,50)
Caminhada do Turista em Redes	92,88(2,11)	92,07(2,67)	91,17(2,46)

Tabela 5.10: Comparação dos resultados para o banco de Brodatz com ruídos. Resultados são apresentados em % e desvio padrão entre parênteses.

Método	$d = 0,1$	$d = 0,3$	$d = 0,5$
Fourier	72,14(3,21)	68,59(3,40)	62,42(4,07)
Filtros de Gabor	88,28(2,09)	87,95(2,33)	82,31(2,71)
Matriz de co-ocorrência	86,29(2,66)	83,86(2,68)	80,51(2,48)
Caminhada do Turista em Redes	91,35(2,36)	91,08(2,53)	88,95(2,95)

Tabela 5.11: Comparação dos resultados para o banco de Brodatz com ruídos. Resultados são apresentados em % e desvio padrão entre parênteses.

Nas Tabelas 5.12 e 5.13 são apresentados os resultados para o banco de imagens de folhas. Como nos resultados anteriores, o método proposto obteve melhores resultados quando comparados com os métodos tradicionais. Para baixas densidades, como $d = 0,07$, o método proposto obteve 96,40% contra 82% obtido pelos descritores de Fourier. Para altas densidades, como $d = 0,5$, o método proposto obteve 85,20% contra 81,60% obtido pelas matrizes de co-ocorrência. A vantagem do método está na topologia da rede complexa que modela uma imagem com ruído. Como pode ser visto no exemplo da Figura 5.16, os pixels do ruído são desconsiderados na estrutura da rede complexa.

Método	$d = 0,01$	$d = 0,05$	$d = 0,07$
Fourier	68,00(19,69)	88,20(13,06)	82,80(15,31)
Filtros de Gabor	70,80(18,51)	73,20(17,34)	68,20(18,66)
Matriz de co-ocorrência	79,20(13,90)	88,20(13,66)	76,20(16,74)
Caminhada do Turista em Redes	98,00(6,03)	96,60(9,01)	96,40(9,16)

Tabela 5.12: Comparação dos resultados para o banco de folhas com ruídos. Resultados são apresentados em % e desvio padrão entre parênteses.

5.3 Conclusão

Inicialmente, o método proposto foi avaliado com relação aos seus parâmetros. Podemos notar que, as transformações realizadas nas redes complexas

Método	$d=0,1$	$d=0,3$	$d=0,5$
Fourier	78,00(16,64)	74,80(18,77)	75,20(18,67)
Filtros de Gabor	73,60(18,61)	42,40(21,33)	59,60(21,27)
Matriz de co-ocorrência	83,80(12,29)	82,80(16,09)	81,60(17,45)
Caminhada do Turista em Redes	94,40(10,67)	86,40(15,01)	85,20(16,24)

Tabela 5.13: Comparação dos resultados para o banco de folhas com ruídos. Resultados são apresentados em % e desvio padrão entre parênteses.

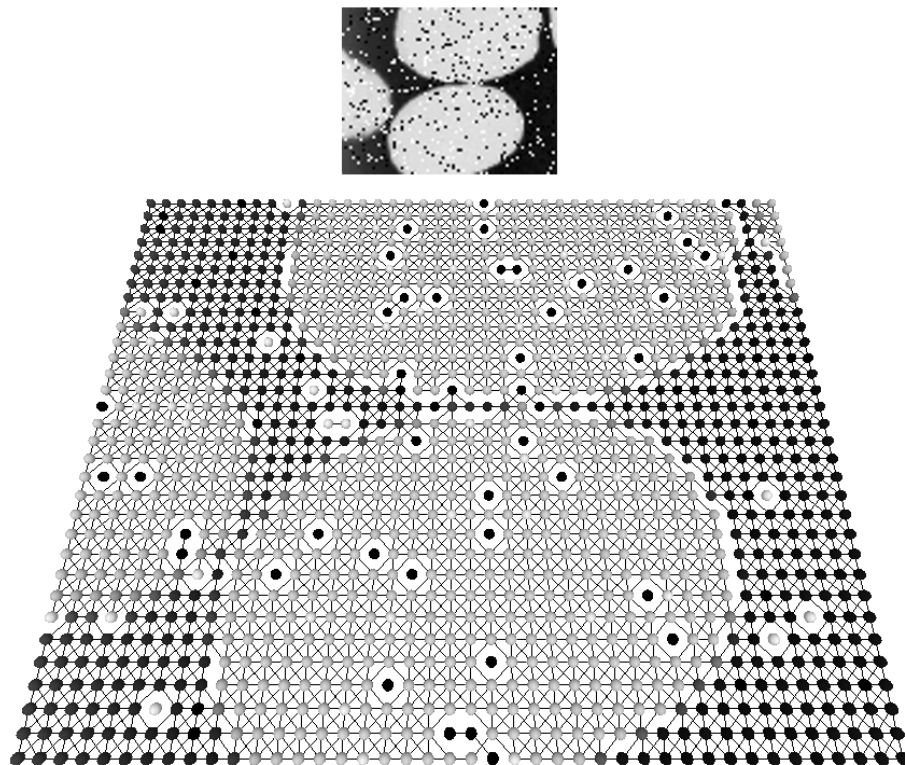


Figura 5.16: Imagem com ruído modelada por uma rede complexa.

auxiliam a caracterização das texturas e para os parâmetros da caminhada determinística do turista, a combinação entre diferentes dinâmicas e memórias incrementam a porcentagem de classificação correta. Nos experimentos, o método proposto para reconhecimento de texturas apresentou robustez tanto no reconhecimento de texturas sem ruído quanto com altas densidade de ruído. Comparado com os métodos da literatura, o método supera os métodos tradicionais em ambos os bancos de texturas sintéticas e naturais (folhas de plantas).

Conclusão

As redes complexas possuem um alto nível de modelagem que abrangem desde as interações entre células até as mais complexas redes sociais. Na maioria das vezes, o estudo dos elementos da rede não é suficiente para determinar o comportamento modelado. Com isso, a extração de informações topológicas e estruturais dessas redes é de extrema importância para aplicações reais.

Neste contexto, este trabalho propôs o uso das caminhadas determinísticas do turista para a caracterização das redes complexas. Uma caminhada é iniciada em cada elemento da rede e então, informações são extraídas dessas caminhadas. Algumas lacunas apresentadas pelas caminhadas determinísticas do turista foram estudadas e novas estratégias para a resolução desses problemas foram propostas e avaliadas.

O foco principal em termos de aplicações foi a classificação de redes complexas artificiais geradas com base em quatro modelos teóricos. Os resultados obtidos foram superiores comparados com medidas tradicionais. O método também foi avaliado no reconhecimento de texturas sintéticas e naturais. As texturas sintéticas foram extraídas do banco de Brodatz enquanto que as texturas naturais foram extraídas de folhas de plantas. Em ambos os casos, o reconhecimento de texturas, o método apresentou excelentes resultados, superando os métodos estado da arte em reconhecimento de textura, como Gabor, matriz de co-ocorrência e descritores de Fourier.

Com este trabalho pode-se concluir que as caminhadas determinísticas do turista podem ser utilizadas como uma medida robusta na geração de características em redes complexas. Essas características quando classificadas com um algoritmo de aprendizagem de máquina apresentaram bons resultados

para as aplicações estudadas neste trabalho.

6.1 Contribuições

As principais contribuições deste trabalho foram:

- Nova medida para representação de redes complexas: Um método matemático para a aplicação das caminhadas determinísticas do turista em redes complexas foi apresentado. Os resultados obtidos em redes complexas artificiais foram superiores comparados com as medidas tradicionais.
- Representação multi-escala de imagens com redes complexas: Uma nova representação multi-escala de imagens com redes complexas foi proposta e avaliada em reconhecimento de textura. Essa representação pode ser aplicada em diversos outros problemas de visão computacional e ainda pode-se utilizar outras medidas além das caminhadas do turista.
- Novo método para reconhecimento de textura: A combinação entre a representação multi-escala de imagens e a extração de características com as caminhadas do turista mostrou-se eficaz no reconhecimento de textura, superando inclusive os métodos estado da arte descritos na literatura.
- Lacunas das caminhadas determinísticas do turista: Experimentos foram realizados para avaliar estratégias de desempate e extração de informações da distribuição conjunta que, atualmente, não eram resolvidas pelo método.

6.2 Trabalhos Futuros

Os interessantes resultados obtidos pelo método sugerem que novos trabalhos podem ser gerados. A caminhada determinística do turista é denominada uma caminhada *self-avoiding*. Alguns vértices que foram visitados recentemente não podem ser visitados novamente durante um determinado tempo, ou seja, a caminhada é baseada em uma memória recente. Uma alternativa a esse tipo de memória é o uso de uma memória baseada em frequência, onde seriam evitados os vértices que mais fizeram parte da caminhada, sem qualquer relação com o tempo em que eles foram visitados.

Em cada caminhada, o período e o tempo de transiente são utilizados para construção da distribuição conjunta. Além disso, muitas outras informações

podem ser extraídas dessas caminhadas, como por exemplo, o esforço realizado pelo turista descrito como a somatória dos pesos de cada aresta. Outra maneira para extrair informações dessas redes é a comparação entre os atratores encontrados na rede. Essa comparação pode ser realizada através de um histograma entre a distância mínima entre todos os atratores. Um método que utiliza um subconjunto de vértices é apresentado em (Antiqueira and da Fontoura Costa, 2008), podendo este subconjunto ser representado pelos atratores.

O método proposto pode ser estendido e ser aplicado em imagens coloridas. A avaliação e combinação de diversos espaços de cores na construção dos pesos da rede complexa pode ser uma linha a ser seguida. O método pode ainda ser aplicado em outros problemas de visão computacional, como o reconhecimento de texturas em movimento. Texturas em movimento têm recebido um crescente interesse na comunidade de visão computacional. Exemplos de texturas em movimento são fogo, fumaça, multidão caminhando, ondas do mar, entre outros. Para isso, a rede complexa seria construída conectando os pixels de dois frames consecutivos e em seguida, aplicar-se o turista para a extração de um vetor de características.

6.3 *Trabalhos Gerados*

Os artigos gerados deste trabalho são listados a seguir. Uma descrição detalhada pode ser encontrado no Apêndice B.

- GONÇALVES, W. N.; BRUNO, O. M.. Redes Complexas Aplicadas no Reconhecimento de Faces. Workshop de Visão Computacional, 2008, Bauru. (Publicado).
- GONÇALVES, W. N.; MACHADO, B. B.; BATISTA NETO, J. E. S.; BRUNO, O. M.. A Complex Network Approach to Texture Applied to Medical Image Classification. VipImage 2009 - II ECCOMAS Thematic Conferences on Computational Vision and Medical Image Processing, 2009, Porto. (Publicado).
- BACKES, A. R.; GONÇALVES, W. N.; MARTINEZ, A. S.; BRUNO, O. M.. Texture Analysis and Classification using Deterministic Tourist Walk. Pattern Recognition, 2009. (Publicado).
- GONÇALVES, W. N.; SILVA, J. A.; BRUNO, O. M.. A rotation invariant face recognition method based on complex network. Pattern Recognition (Submetido).

- GONÇALVES, W. N.; BACKES, A. R.; BRUNO, O. M.. Deterministic Tourist Walk on Complex Networks Applied to Texture Analysis and Classification (Desenvolvimento).
- GONÇALVES, W. N.; MARTINEZ, A. S.; BRUNO, O. M.. Complex Networks Characterization using Deterministic Tourist Walk (Desenvolvimento).

Apêndice A

Neste apêndice, o comportamento do turista em imagens é apresentado e discutido. Na Figura 1, a distribuição conjunta do tempo de transiente e atrator é ilustrada para uma imagem. Para isso, a DTW foi aplicada em uma rede regular sem qualquer transformação. A distribuição possui uma completa gama de atratores com períodos $p \geq p_{min}$, com uma lei de decaimento (Lima et al., 2001; Stanley and Buldyrev, 2001; Kinouchi et al., 2001).

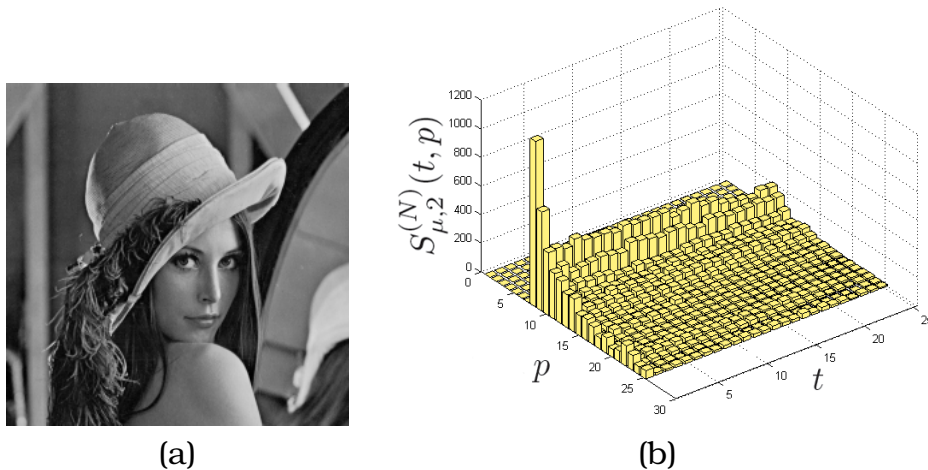


Figura 1: (a) Imagem Original; (b) Distribuição de probabilidade conjunta do tempo de transiente e período do atrator para $\mu = 5$.

A análise da DTW em imagens foi baseada em atratores gerados para diferentes memórias μ , quando a caminhada do turista é aplicada em uma imagem gerada aleatoriamente e uma imagem de textura. Ambas as imagens possuem 100×100 pixels e 256 níveis de cinza.

A memória tem uma grande influência na distribuição conjunta em imagens. Para pequenos valores de μ , um alto número de atratores é distribuído sobre a imagem. Estes atratores apresentam um simples comportamento e pequenas repetições de seções. Por outro lado, à medida que a memória aumenta, o número de atratores decrementa, pois um grande número de caminhadas tendem a encontrar o mesmo atrator. Quando a memória aumenta

ainda mais, o turista é forçado a procurar por grandes atratores. Estes novos atratores são representados em uma pequena porção na imagem, e eles também apresentam um comportamento mais complexo (Figuras 2-5). Na Figuras 2 e 3 são apresentados os atratores para uma imagem gerada de forma aleatória, enquanto nas Figuras 4 e 5 são apresentados os atratores para uma imagem real.

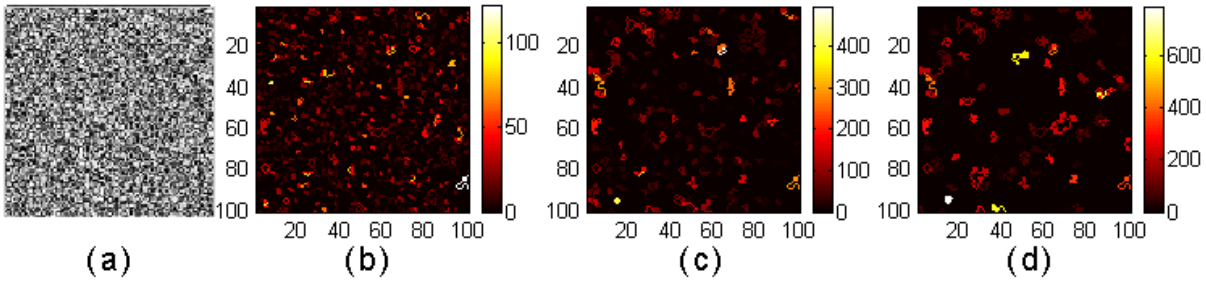


Figura 2: Atratores gerados pelo turista em uma imagem gerada aleatoriamente utilizando a mínima distância e diferentes valores de μ : (a) Imagem original; (b) $\mu = 1$; (c) $\mu = 3$; (d) $\mu = 7$.

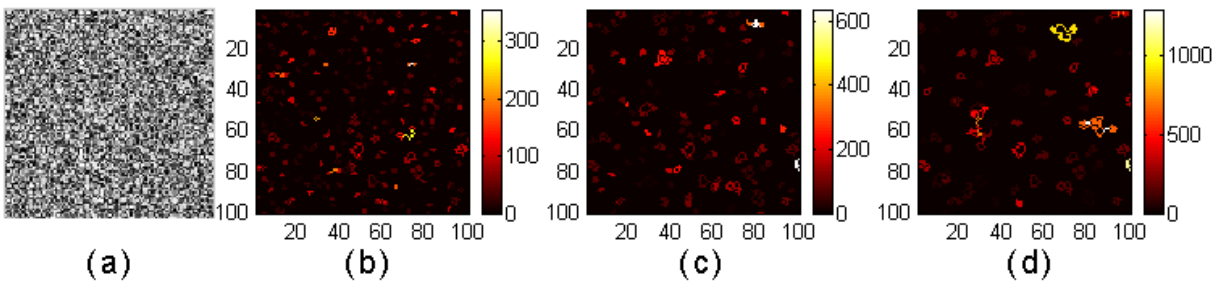


Figura 3: Atratores gerados pelo turista em uma imagem gerada aleatoriamente utilizando a máxima distância e diferentes valores de μ : (a) Imagem original; (b) $\mu = 1$; (c) $\mu = 3$; (d) $\mu = 7$.

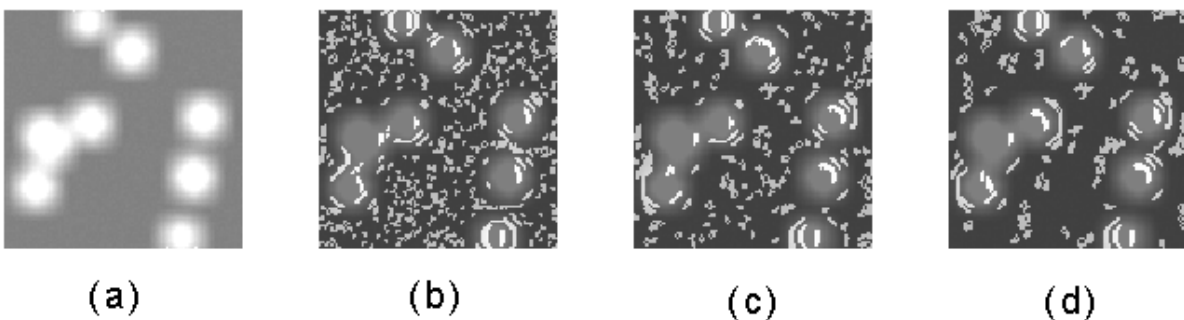


Figura 4: Atratores gerados pelo turista em uma imagem de textura utilizando a mínima distância e diferentes valores de μ . Atratores são misturados com textura para melhor compreensão: (a) Imagem original; (b) $\mu = 1$; (c) $\mu = 3$; (d) $\mu = 7$.

É possível notar também que o turista não requer uma longa caminhada para encontrar um atrator. Isto é mostrado pelo número de atratores que

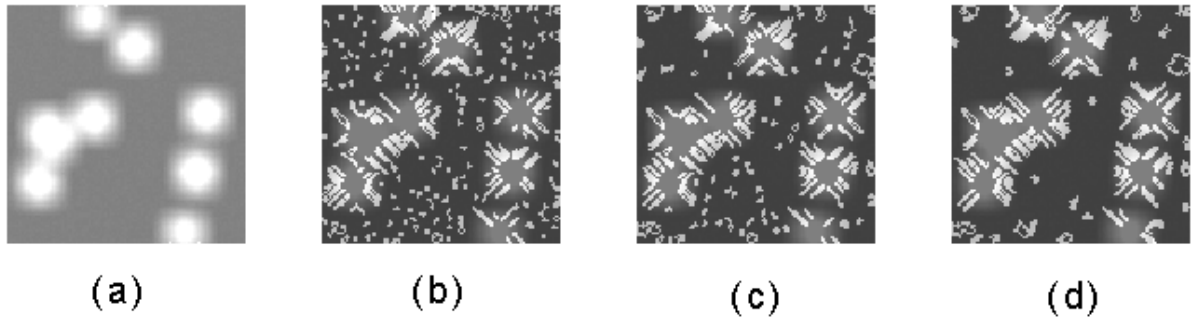


Figura 5: Atratores gerados pelo turista em uma imagem de textura utilizando a máxima distância e diferentes valores de μ . Atratores são misturados com textura para melhor compreensão: (a) Imagem original; (b) $\mu = 1$; (c) $\mu = 3$; (d) $\mu = 7$.

apresentam um transiente $t = 0$ para a DTW aplicada com memória $\mu = 4$ (Figura 6).

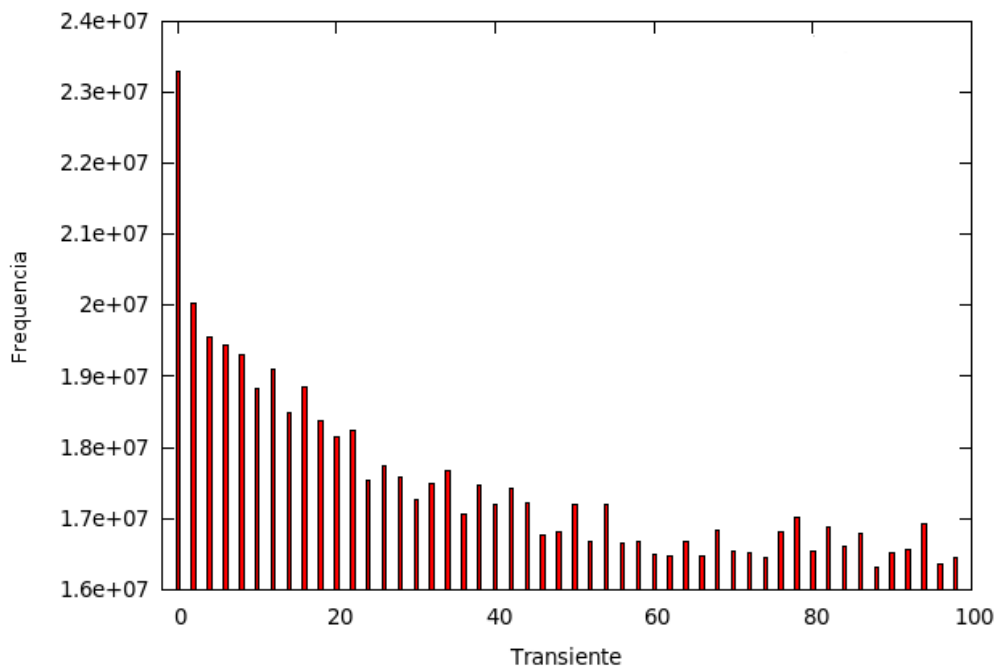


Figura 6: Histograma do tamanho do transiente para $\mu = 4$.

Em imagens geradas aleatoriamente, a caminhada do turista é caracterizada por um comportamento caótico. Este comportamento é devido ao ambiente onde o turista é aplicado (i.e. como não existe correlação entre os pixels vizinhos, o turista caminha aleatoriamente e muda sua direção com a mesma frequência que o ambiente muda (Figuras 2 e 3). Entretanto, imagens digitais apresentam um contexto visual, i.e., pixels estão correlacionados a cena, como a face de uma pessoa ou um padrão de textura. Esta correlação entre os pixels vizinhos influenciam as escolhas de direção do turista durante a sua caminhada. As Figuras 4 e 5 mostram que os atratores gerados apresentam uma correlação com o padrão visual presente na imagem.

Apêndice B

Durante o desenvolvimento deste trabalho, alguns artigos foram gerados. Abaixo, segue a listagem dos artigos publicados, submetidos e em desenvolvimento.

Redes Complexas Aplicadas no Reconhecimento de Faces.

GONÇALVES, W. N.; BRUNO, O. M.;

Workshop de Visão Computacional, 2008.

Resumo: Este artigo apresenta uma nova metodologia para o reconhecimento de faces, um importante e difícil problema que tem sido estudado pela comunidade de visão computacional e reconhecimento de padrões. A metodologia utilizada modela a imagem de uma face através de uma rede complexa e medidas são extraídas sobre essas redes para a composição do vetor de característica. Os experimentos foram conduzidos usando diferentes configurações, incluindo modificações no limiar para a modelagem da rede. A metodologia com redes foi comparada com a técnica eigenfaces, obtendo resultados promissores.

A Complex Network Approach to Texture Applied to Medical Image Classification.

GONÇALVES, W. N.; MACHADO, B. B.; BATISTA NETO, J. E. S.; BRUNO, O. M.;

VipImage 2009 - II ECCOMAS Thematic Conferences on Computational Vision and Medical Image Processing, 2009.

Resumo: Several schemes of CAD systems are being developed to classifying medical images. They are useful, because once found a suspicious lesion, a doctor consults similar images already diagnosed. The computerized analysis in addition to accelerating the clinical diagnosis can support the image analy-

sis as a second opinion in making diagnostic decisions. This paper presents a novel methodology to texture classification of medical images with complex network. Each image is modeled into a complex network. It uses the degree of vertices and a transformation on the network to compose a set of texture descriptors. The proposed method has a powerful performance to texture characterization. Experimental results show that the method is robust to noise and scale invariant. The results were compared with traditional methods of texture analysis, as co-occurrence matrices, Fourier descriptors and Gabor filters. The proposed method shows its great potential for image analysis and classification using the theory of complex networks.

Texture Analysis and Classification using Deterministic Tourist Walk.

BACKES, A. R.; GONÇALVES, W. N.; MARTINEZ, A. S.; BRUNO, O. M.;

Volume 43, Issue 3, March 2010, Pages 685-694, Pattern Recognition.

Resumo: In this paper, we present a study on a deterministic partially self-avoiding walk (tourist walk), which provides a novel method for texture feature extraction. The method is able to explore an image on all scales simultaneously. Experiments were conducted using different dynamics concerning the tourist walk. A new strategy, based on histograms, to extract information from its joint probability distribution is presented. The promising results are discussed and compared to the best-known methods for texture description reported in the literature.

A rotation invariant face recognition method based on complex network.

GONÇALVES, W. N.; SILVA, J. A.; BRUNO, O. M.;

Pattern Recognition, 2010 (Submetido)

Resumo: Face recognition is an important aspect in computer vision having applications in the security industry. It is a hard task that has been studied in depth by the science community. This paper introduces a novel graph based method for face recognition which is rotation invariant. The main idea of the approach is to model the face image onto a graph and use complex network methodology to analyze the graph and extract the features in order to proceed with face recognition. The paper presents the novel methodology in details and presents experiments comparing it with eigenfaces, fisherfaces, laplacianfaces and neighborhood preserving embedding, which are four important and state of art algorithms in face recognition. The results demonstrated that the proposed technique has more positive results than previous ones.

Complex Networks Characterization using Deterministic Tourist Walk

GONÇALVES, W. N.; MARTINEZ, A. S.; BRUNO, O. M.;

(Desenvolvimento).

Resumo: Este artigo propõe o uso das caminhadas determinísticas como uma medida de representação de redes complexas. A caminhada do turista foi aplicada na classificação dos modelos aleatórios, geográficos, pequeno-mundo e livre de escala. As redes complexas foram geradas de forma artificial e cada aresta possui um peso aleatório. Como apresentado nos resultados, com a utilização de apenas essa nova medida, foi possível obter resultados semelhantes aos obtidos com a combinação de diferentes medidas tradicionais e resultados superiores quando comparados com as medidas individualmente.

Deterministic Tourist Walk on Complex Networks Applied to Texture Analysis and Classification

GONÇALVES, W. N.; BACKES, A. R.; BRUNO, O. M.;

(Desenvolvimento).

Resumo: Esse artigo é resultado final da dissertação, apresentando uma nova abordagem para executar as caminhadas determinísticas utilizando a teoria das redes complexas para modelar as conexões entre pixels vizinhos. Originalmente, o método considera a imagem como um grafo regular, i.e. todos os vértices estão conectados com o mesmo número de vizinhos. Aplicando uma função sobre todas as arestas, é possível construir uma rede que apresenta diferentes propriedades, assim influenciando a maneira que o turista executa sua caminhada. Os resultados foram comparados com os principais métodos de reconhecimento de textura, obtendo resultados superiores tanto nas imagens originais, como nas imagens com altas densidades de ruído.

Referências

- Albert, R., Albert, I., and Nakarado, G. L. (2004). Structural vulnerability of the north american power grid. Comment: 10 pages, 4 figures, to appear in PRE. Citado na página 10.
- Antiqueira, L. and da Fontoura Costa, L. (2008). Characterization of sub-graphs relationships and distribution in complex networks. Citado na página 69.
- Arbib, M. A. (1995). *The Handbook of Brain Theory and Neural Networks*. MIT Press, Cambridge, MA, USA. Citado na página 60.
- Azencott, R., Wang, J.-P., and Younes, L. (1997). Texture classification using windowed Fourier filters. *IEEE Trans. Pattern Anal. Mach. Intell.*, 19(2):148–153. Citado nas páginas 41, 49, e 57.
- Backes, A. R. and Bruno, O. M. (2008). A new approach to estimate fractal dimension of texture images. In *ICISP*, volume 5099 of *Lecture Notes in Computer Science*, pages 136–143. Springer. Citado na página 41.
- Backes, A. R., Bruno, O. M., Campiteli, M. G., and Martinez, A. S. (2006). Deterministic tourist walks as an image analysis methodology based. In *CIARP*, pages 784–793. Citado nas páginas 17, 23, 25, 27, 46, e 54.
- Backes, A. R., Casanova, D., and Bruno, O. M. (2007). Método de aproximação poligonal de contornos utilizando redes complexas. *INFOCOMP Journal of Computer Science*, 6(2):71–80. Citado na página 9.
- Backes, A. R., Casanova, D., and Bruno, O. M. (2009a). A complex network-based approach for boundary shape analysis. *Pattern Recognition*, 42(1):54–67. Citado nas páginas 2 e 9.
- Backes, A. R., Gonçalves, W. N., and Bruno, O. M. (2009b). Texture analysis and classification using deterministic tourist walk. *Pattern Recognition*, 43:685–694. Citado nas páginas 4, 17, 21, 34, 36, e 53.

- Barabasi, A. L. and Albert, R. (1999). Emergence of scaling in random networks. *Science*, 286(5439):509–512. Citado nas páginas 2, 10, e 11.
- Barabasi, A.-L., Jeong, H., Neda, Z., Ravasz, E., Schubert, A., and Vicsek, T. (2002). Evolution of the social network of scientific collaborations. *Physica A*, 311:590–614. Citado na página 8.
- Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., and Hwang, D. U. (2006). Complex networks: Structure and dynamics. *Physics Reports*, 424(4-5):175–308. Citado nas páginas 7, 8, e 14.
- Breiman, L., Friedman, J., Olshen, R., and Stone, C. (1984). *Classification and Regression Trees*. Wadsworth and Brooks, Monterey, CA. Citado nas páginas 39 e 60.
- Brodatz, P. (1966). *Textures: A photographic album for artists and designers*. Dover Publications, New York. Citado na página 48.
- Bruno, O. M., de Oliveira Plotze, R., Falvo, M., and de Castro, M. (2008a). Fractal dimension applied to plant identification. *Inf. Sci*, 178(12):2722–2733. Citado na página 4.
- Bruno, O. M., de Oliveira Plotze, R., Falvo, M., and de Castro, M. (2008b). Fractal dimension applied to plant identification. *Inf. Sci.*, 178(12):2722–2733. Citado na página 48.
- Bunimovich, L. A. (2004). Deterministic walks in random environments. *Physica D*, 187:20–29. Citado na página 17.
- Bunimovich, L. A. and Troubetzkoy, S. E. (1992). Recurrence properties of lorentz lattice gas cellular automata. *J. Stat. Phys*, 67:289–302. Citado na página 17.
- Campiteli, M. G., Batista, P. D., and Martinez, O. K. A. S. (2006a). Deterministic walks as an algorithm of pattern recognition. *Physical Review E*, 74(2):026703. Citado nas páginas 20 e 25.
- Campiteli, M. G., Martinez, A. S., and Bruno, O. M. (2006b). An image analysis methodology based on deterministic tourist walks. In *IBERAMIA-SBIA*, pages 159–167. Citado nas páginas 17, 18, 20, 23, 25, 27, e 54.
- Chalumeau, T., L. da F Costa, L., Laligant, O., and Meriaudeau, F. (2007). Complex networks: application for texture classification. In *Eighth International Conference on Quality Control by Artificial Vision*, volume 6356. Citado nas páginas 2, 4, e 9.
- Connors, R. and Harlow, C. (1980). A theoretical comparison of texture algorithms. *TransPAMI*, 2:204–222. Citado na página 28.

- da F. Costa, L., Rodrigues, F. A., Travieso, G., and Boas, P. R. V. (2007). Characterization of complex networks: A survey of measurements. *Advances In Physics*, 56:167–242. Citado nas páginas 2 e 33.
- da Fontoura Costa, L. (2004). Complex networks, simple vision. Citado na página 9.
- Daugman, J. and Downing, C. (1995). Gabor wavelets for statistical pattern recognition. In Arbib, M. A., editor, *The Handbook of Brain Theory and Neural Networks*, pages 414–419. MIT Press, Cambridge, Massachusetts. Citado na página 57.
- de Oliveira Plotze, R. and Bruno, O. M. (2009). Automatic leaf structure biometry: Computer vision techniques and their applications in plant taxonomy. *IJPRAI*, 23(2):247–262. Citado na página 4.
- Derrida, B. (1997). From random walks to spin glasses. *Phys. D*, 107(2-4):186–198. Citado nas páginas 17 e 25.
- Eibe, I. W., Witten, I. H., Frank, E., Trigg, L., Hall, M., Holmes, G., and Cunningham, S. J. (1999). Weka: Practical machine learning tools and techniques with java implementations. Citado na página 60.
- Erdős, P. and Rényi, A. (1959). On random graphs. *Publ. Math. Debrecen*, 6:290–297. Citado nas páginas 2 e 10.
- Fukunaga, K. (1990). *Introduction to statistical pattern recognition*. Academic Press Professional, Inc., San Diego, CA, USA. Citado na página 36.
- Gale, D., Propp, J., Sutherland, S., and Troubetzkoy, S. (1995). Further travels with my ant. *Mathematical Intelligencer*, 17:48–56. Citado na página 17.
- Gastner, M. T. and Newman, M. E. J. (2004). The spatial structure of networks. Citado na página 2.
- Gonçalves, W. N., de Andrade Silva, J., and Bruno, O. M. A rotation invariant face recognition method based on complex network. *Pattern Recognition (Submetido)*. Citado na página 9.
- Gonçalves, W. N., Machado, B. B., Neto, J. E. S. B., and Bruno, O. M. (2009). A complex network approach to texture applied to medical image classification. In *VipImage 2009 - II ECCOMAS Thematic Conferences on Computational Vision and Medical Image Processing*. Citado nas páginas 1, 2, e 9.
- Haralick, R. M. (1979). Statistical and structural approaches to texture. *Proc. IEEE*, 67(5):768–804. Citado nas páginas 41, 49, e 57.

- Haralick, R. M., Shanmugam, K., and Dinstein, I. (1973). Textural features for image classification. *Systems, Man and Cybernetics, IEEE Transactions on*, 3(6):610–621. Citado na página 28.
- Hayashi, Y. (2005). A review of recent studies of geographical scale-free networks. Comment: 9 pages, 5 figures, 1 table, REVTeX. Citado na página 10.
- Hayashi, Y. (2006). A review of recent studies of geographical scale-free networks. *IPSJ Trans. Special Issue on Network Ecology*, 47:776. Citado na página 2.
- Idrissa, M. and Acheroy, M. (2002). Texture classification using gabor filters. *Pattern Recognition Letters*, 23(9):1095–1102. Citado na página 57.
- Jain, A. K. and Farrokhnia, F. (1991). Unsupervised texture segmentation using gabor filters. *Pattern Recogn.*, 24(12):1167–1186. Citado nas páginas 41, 49, e 57.
- Kasparis, T., Charalampidis, D., Georgiopoulos, M., and Rolland, J. P. (2001). Segmentation of textured images based on fractals and image filtering. *Pattern Recognition*, 34(10):1963–1973. Citado na página 41.
- Kinouchi, O., Martinez, A. S., Lima, G. F., Lourenco, G. M., and Risau-Gusman, S. (2001). Deterministic walks in random networks: an application to thesaurus graphs. Citado nas páginas 20 e 71.
- Lima, G. F., Martinez, A. S., and Kinouchi, O. (2001). Deterministic walks in random media. *Phys. Rev. Lett.*, 87(1):010603. Citado nas páginas 20 e 71.
- Manjunath, B. S. and Ma, W.-Y. (1996). Texture features for browsing and retrieval of image data. *IEEE Trans. Pattern Anal. Mach. Intell*, 18(8):837–842. Citado na página 57.
- Metzler, R. and Klafter, J. (2000). The random walk’s guide to anomalous diffusion: a fractional dynamics approach. *Physics Reports*, 339(1):1–77. Citado nas páginas 17 e 25.
- Meyer, D., Leisch, F., and Hornik, K. (2003). The support vector machine under test. *Neurocomputing*, 55(1-2):169–186. Citado na página 60.
- Mitchell, T. M. (1997). *Machine Learning*. McGraw-Hill, New York. Citado nas páginas 34, 39, 49, e 60.
- Nadeau, C. and Bengio, Y. (2003). Inference for the generalization error. *Machine Learning*, 52(3):239–281. Citado na página 60.
- Newman, M. E. J. (2003). The structure and function of complex networks. *SIAM Review*, 45:167–256. Citado nas páginas 2 e 8.

- Newman, M. E. J., Forrest, S., and Balthrop, J. (2002). Email networks and the spread of computer viruses. *Physical Review Letters*, 66(3):035101. Citado na página 8.
- Plotze, R. O., M., F., Pádua, J. G., Bernacci, L. C., Vieira, M. L. C., Oliveira, G. C. X., and Bruno, O. M. (2005). Leaf shape analysis using the multiscale Minkowski fractal dimension, a new morphometric method: a study with passiflora (passifloraceae). *Canadian Journal of Botany*, 83(3):287–301. Citado na página 48.
- Risau-Gusman, S., Martinez, A. S., and Kinouchi, O. (2003). Escaping from cycles through a glass transition. *Physica D*, 68:016104. Citado na página 19.
- Sen, P., Dasgupta, S., Chatterjee, A., Sreeram, P. A., Mukherjee, G., and Manna, S. S. (2003). Small-world properties of the indian railway network. *Physical Review E*, 67:036106. Citado nas páginas 8 e 9.
- Silva, T. C. and Zhao, L. (2007). Pixel clustering by using complex network community detection technique. In *ISDA '07: Proceedings of the Seventh International Conference on Intelligent Systems Design and Applications*, pages 925–932, Washington, DC, USA. IEEE Computer Society. Citado na página 9.
- Solaiman, B., Burdsall, B., and Roux, C. (1998). Hough transform and uncertainty handling. application to circular object detection in ultrasound medical images. In *ICIP (3)*, pages 828–831. Citado na página 1.
- Solomonoff, R. and Rapoport, A. (1951). Connectivity of random nets. *Bulletin of Mathematical Biology*, 13(2):107–117. Citado nas páginas 2 e 10.
- Stanley, H. E. and Buldyrev, S. V. (2001). Statistical physics - the salesman and the tourist. *Nature (London)*, 413(6854):373–374. Citado nas páginas 20 e 71.
- Stivanello, M. E. and Gomes, P. C. R. (2006). Inspeção visual industrial automatizada por análise de forma com descritores de fourier e redes neurais artificiais. *XV Seminário de Computação*. Citado na página 1.
- Su, J. and Zhang, H. (2006). Full bayesian network classifiers. In *ICML '06: Proceedings of the 23rd international conference on Machine learning*, pages 897–904, New York, NY, USA. ACM. Citado na página 60.
- Terçariol, C. A. and Martinez, A. S. (2005). Analytical results for the statistical distribution related to a memoryless deterministic walk: dimensionality effect and mean-field models. *Physica Rev. E*, 72. Citado nas páginas 19 e 20.

- Tisse, C., Martin, L., Torres, L., and Robert, M. (2002). Person identification technique using human iris recognition. In *Proc. Vision Interface*, pages 294–299. Citado na página 1.
- Watts, D. J. (1999). Networks, dynamics, and the small-world phenomenon. *The American Journal of Sociology*, 105(2):493–527. Citado nas páginas 2 e 11.
- Watts, D. J. and Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature*, 393(6684):440–442. Citado nas páginas 2, 10, e 11.
- West, G. B., Brown, J. H., and Enquist, B. J. (1999). A general model for the structure and allometry of plant vascular systems. *Nature*, 400:664–667. Citado na página 8.
- Weszka, J., Dyer, C., and Rosenfeld, A. (1976). A comparative study of texture measures for terrain classification. *IEEE Trans SMC*, 6:269–285. Citado na página 28.
- Zhan, F. B. and Noon, C. E. (1998). Shortest path algorithms: An evaluation using real road networks. *Transportation Science*, 32(1):65–73. Citado na página 8.
- Zhang, H. (2004). The optimality of naive bayes. In *FLAIRS Conference*. Citado nas páginas 39 e 60.