

TATIANE STEIN

**AVALIAÇÃO DE DESCRITORES DE TEXTURA  
PARA SEGMENTAÇÃO DE IMAGENS**

Dissertação apresentada como requisito parcial à obtenção do grau de Mestre. Programa de Pós-Graduação em Informática, Setor de Ciências Exatas, Universidade Federal do Paraná.  
Orientador: Prof. Dr. Hélio Pedrini

CURITIBA

2005

## AGRADECIMENTOS

Primeiramente agradeço a Deus pela vida e por permitir minha chegada até aqui.

Agradeço aos meus pais, Sônia e Erni, meus primeiros “orientadores” nesta vida. Amo muito vocês.

Ao meu orientador Prof. Dr. Hélio Pedrini, tenho muito a agradecer. Obrigada pelos ensinamentos, pelo constante incentivo e dedicação. Hoje o considero um amigo e serei sempre grata pela confiança e consideração que teve por mim.

Aos meus irmãos Cristiane e Leandro, meu lindo sobrinho Mateus, Frederico e a todos os meus queridos amigos, obrigada pela paciência, compreensão e pela presença constante, apesar das distâncias e da falta de tempo.

Aos meus avós, em especial ao meu avô Germano que faleceu durante o desenvolvimento deste trabalho. Alguém com quem aprendi grandes lições.

Aos meus parceiros neste trabalho Gustavo e Herton, pois sem eles tudo teria sido mais difícil. Obrigada pela paciência e pela ajuda.

Desejo expressar aqui os meus sinceros agradecimentos a todos, sem exceções, que estiveram ao meu lado durante essa caminhada e, de alguma forma, contribuíram para que mais um degrau em minha vida fosse alcançado. Pessoas como vocês significam muito mais do que as palavras podem dizer, os sentimentos conseguem expressar ou o pensamento imaginar.

Nós conseguimos, vencemos juntos essa batalha. Vocês merecem cada pedacinho desse momento maravilhoso da minha vida que hoje dedico a vocês, com todo o meu carinho e eterna gratidão.

## SUMÁRIO

<b>GLOSSÁRIO</b>	<b>iv</b>
<b>LISTAS DE FIGURAS</b>	<b>viii</b>
<b>LISTA DE TABELAS</b>	<b>ix</b>
<b>RESUMO</b>	<b>x</b>
<b>1 INTRODUÇÃO</b>	<b>1</b>
1.1 Textura . . . . .	2
1.2 Áreas de Aplicação . . . . .	4
1.3 Objetivos e Contribuições . . . . .	5
1.4 Organização do Trabalho . . . . .	6
<b>2 TÓPICOS RELACIONADOS</b>	<b>7</b>
2.1 Vetor de Características . . . . .	7
2.2 Análise de Componentes Principais . . . . .	8
2.3 Abordagens para Análise de Textura . . . . .	14
2.3.1 Abordagem Estatística . . . . .	14
2.3.2 Abordagem Estrutural . . . . .	17
2.3.3 Abordagem Espectral . . . . .	19
2.4 Matrizes de Co-ocorrência . . . . .	23
2.4.1 Definição da Matriz de Co-ocorrência . . . . .	24
2.4.2 Construção da Matriz de Co-ocorrência . . . . .	25
2.4.3 Propriedades da Matriz de Co-ocorrência . . . . .	28
2.5 Classificadores de Imagens . . . . .	29
2.5.1 Redes Neurais Artificiais . . . . .	29
2.5.2 Classificador de Distância Mínima . . . . .	35

<b>3</b>	<b>MÉTODO PROPOSTO</b>	<b>37</b>
3.1	Percurso das Imagens . . . . .	39
3.2	Características de Textura . . . . .	42
3.3	Matrizes de Co-ocorrência 3D . . . . .	48
3.4	Normalização das Medidas . . . . .	50
<b>4</b>	<b>RESULTADOS EXPERIMENTAIS</b>	<b>51</b>
4.1	Experimentos Realizados . . . . .	51
4.1.1	Experimento 1 . . . . .	51
4.1.2	Experimento 2 . . . . .	52
4.1.3	Experimento 3 . . . . .	55
4.1.4	Experimento 4 . . . . .	56
4.1.5	Experimento 5 . . . . .	57
4.1.6	Experimento 6 . . . . .	58
4.1.7	Experimento 7 . . . . .	60
4.1.8	Experimento 8 . . . . .	62
4.1.9	Experimento 9 . . . . .	62
4.1.10	Experimento 10 . . . . .	65
4.1.11	Experimento 11 . . . . .	67
4.1.12	Experimento 12 . . . . .	68
4.1.13	Experimento 13 . . . . .	69
4.2	Análise dos Resultados . . . . .	70
<b>5</b>	<b>CONCLUSÕES E TRABALHOS FUTUROS</b>	<b>73</b>
	<b>REFERÊNCIAS BIBLIOGRÁFICAS</b>	<b>75</b>
	<b>APÊNDICE A FORMATO DAS IMAGENS</b>	<b>82</b>
A.1	Formato PGM . . . . .	82
A.2	Formato PPM . . . . .	83

## GLOSSÁRIO

DISPLAY	<i>Image Magick</i>
GIMP	<i>GNU Image Manipulation Program</i>
INNS	<i>International Conference on Neural Networks Society</i>
JavaNNS	<i>Java Neural Network Simulator</i>
JPEG	<i>Joint Photographers Expert Group</i>
PCA	<i>Principal Component Analysis</i>
PGM	<i>Portable Graymap File Format</i>
PPM	<i>Portable Pixmap File Format</i>
RNA	<i>Rede Neural Artificial (Artificial Neural Network)</i>
XV	<i>Image Viewer</i>

## LISTA DE FIGURAS

1.1	Exemplos de Textura - Álbum de Brodatz [7]. . . . .	4
2.1	Regra da cadeia. . . . .	19
2.2	(a) Imagem mostrando uma textura periódica; (b) espectro; (c) gráfico de $S(r)$ ; (d) gráfico de $S(\theta)$ . . . . .	21
2.3	(a) Imagem contendo um tipo diferente de textura periódica; (b) gráfico de $S(\theta)$ . . . . .	22
2.4	Ângulos para o cálculo na matriz de co-ocorrência. . . . .	24
2.5	Imagem com níveis de cinza de 0 a 2. . . . .	25
2.6	Neurônio artificial. . . . .	30
2.7	Camadas de uma rede neural artificial. . . . .	32
2.8	Classificador da distância mínima. . . . .	35
3.1	Etapas do método proposto. . . . .	37
3.2	Algoritmo para percorrer a imagem de amostra na forma seqüencial. . . . .	40
3.3	Varredura da imagem de maneira seqüencial. . . . .	41
3.4	Algoritmo para percorrer a imagem de amostra na forma cascata. . . . .	41
3.5	Varredura da imagem em cascata. . . . .	42
3.6	Relações entre as bandas de cores, onde os números iguais indicam as variações de tons consideradas, sempre passando pelo pixel central da banda H. . . . .	48
4.1	Experimento 1. (a) Imagem original e (b) amostras utilizadas. . . . .	52
4.2	Resultados da segmentação da imagem letra. Parâmetros comuns: Características - identificadas pela PCA (segundo momento angular, contraste, desvio padrão, média e correlação), classificador Redes Neurais, Normalização, Seqüencial. (a) janela 3x3, (b) janela 5x5, (c) janela 7x7 e (d) janela 15x15. . . . .	53
4.3	Experimento 2. (a) Imagem original e (b) amostras utilizadas. . . . .	53

4.4	Resultados da segmentação da imagem beans. Parâmetros comuns: Características - identificadas pela PCA (segundo momento angular, contraste, desvio padrão, média e correlação) e classificador Redes Neurais. (a) janela 3x3 e seqüencial, (b) janela 3x3 e cascata, (c) janela 3x3 e aleatório, (d) janela 15x15 e seqüencial e (e) janela 3x3 e seqüencial. Em (e) não foi utilizada a Normalização. . . . .	54
4.5	Resultados da classificação da imagem beans corrompida com ruído impulsivo sal-e-pimenta ((a) imagem original e (b) resultado da classificação) e com ruído Gaussiano ((c) imagem original e (d) resultado da classificação). Parâmetros comuns: Características - identificadas pela PCA (segundo momento angular, contraste, desvio padrão, média e correlação), classificador Redes Neurais, janela 3x3, seqüencial, com utilização da Normalização. . . . .	55
4.6	Experimento 4. (a) Imagem original e (b) amostras utilizadas. . . . .	56
4.7	Resultados da segmentação da imagem mosaico 1. Parâmetros comuns: classificador Redes Neurais, cascata e normalização. Em (a), (b) e (c) - características identificadas pela PCA (segundo momento angular, contraste, desvio padrão, média e correlação). Em (d) as características foram: tendência de agrupamento, inércia, curtose, probabilidade máxima e simetria. (a) janela 5x5 e 30 iterações, (b) janela 15x15 e 30 iterações, (c) janela 15x15 percorrendo toda a imagem e (d) janela 15x15 e 30 iterações. . . . .	57
4.8	Experimento 5. (a) Imagem original e (b) amostras utilizadas. . . . .	58
4.9	Resultados da classificação da imagem mosaico 2. Parâmetros comuns: seqüencial e normalização. De (a) até (d) - características identificadas pela PCA (segundo momento angular, contraste, desvio padrão, média e correlação). Em (e) as características foram: tendência de agrupamento, inércia, curtose, probabilidade máxima e simetria. (a) e (e) janela 15x15 e redes neurais, (b) janela 5x5 e redes neurais, (c) janela 3x3 e redes neurais e (d) janela 15x15 e distância mínima. . . . .	59
4.10	Experimento 6. (a) Imagem original e (b) amostras utilizadas. . . . .	60

4.11	Resultados da segmentação da imagem Kodiak. Parâmetros comuns: classificador redes neurais e normalização. De (a) até (e) - características identificadas pela PCA (segundo momento angular, contraste, desvio padrão, média e correlação). Em (f) as características foram: tendência de agrupamento, inércia, curtose, probabilidade máxima e simetria. (a) e (f) janela 3x3 e seqüencial, (b) janela 3x3 e aleatório, (c) janela 3x3 e cascata com 30 iterações, (d) janela 3x3 e cascata percorrendo toda a imagem e (e) janela 15x15 e seqüencial. . . . .	61
4.12	Experimento 7. (a) Imagem original e (b) amostras utilizadas. . . . .	62
4.13	Resultados da segmentação da imagem Chesapeake. Parâmetros comuns: características identificadas pela PCA (segundo momento angular, contraste, desvio padrão, média e correlação), seqüencial e normalização. (a) janela 3x3 e redes neurais, (b) janela 3x3 e distância mínima, (c) janela 5x5 e redes neurais e (d) janela 15x15 e redes neurais. . . . .	63
4.14	Experimento 8. (a) Imagem original e (b) amostras utilizadas. . . . .	64
4.15	Resultados da segmentação da imagem Moreno Glacier. Parâmetros comuns: características identificadas pela PCA (segundo momento angular, contraste, desvio padrão, média e correlação), seqüencial e normalização. (a) janela 3x3 e redes neurais, (b) janela 3x3 e distância mínima e (c) janela 15x15 e redes neurais. . . . .	64
4.16	Experimento 9. (a) Imagem original e (b) amostras utilizadas . . . . .	65
4.17	Resultados da segmentação da imagem Politécnico. Parâmetros comuns: características identificadas pela PCA (segundo momento angular, contraste, desvio padrão, média e correlação) e normalização. De (a) até (e) redes neurais, em (f) distância mínima. (a) janela 15x15 e seqüencial, (b) janela 15x15 e cascata, (c) janela 15x15 e aleatório, (d) janela 3x3 e seqüencial, (e) janela 3x3 e cascata e (f) janela 3x3 e seqüencial. . . . .	66
4.18	Experimento 10. (a) Imagem original e (b) amostras utilizadas. . . . .	67

4.19	Resultados da segmentação da imagem Satélite1. Características propostas por Dacheng <i>et al.</i> [14] (segundo momento angular, contraste, correlação e entropia), seqüencial, janela 3x3, classificador redes neurais e normalização. . .	67
4.20	Experimento 11. (a) Imagem original e (b) amostras utilizadas. . . . .	68
4.21	Resultado da segmentação da imagem Minnesota. Características propostas por Dacheng <i>et al.</i> [14] (segundo momento angular, contraste, correlação e entropia), seqüencial, janela 3x3, classificador redes neurais e normalização. . .	68
4.22	Experimento 12. (a) Imagem original e (b) amostras utilizadas. . . . .	69
4.23	Resultados da segmentação da imagem Satélite2. Características propostas por Dacheng <i>et al.</i> [14] (segundo momento angular, contraste, correlação e entropia), seqüencial, janela 3x3, classificador redes neurais e normalização. . .	69
4.24	Experimento 13. (a) Imagem original e (b) amostras utilizadas. . . . .	70
4.25	Resultado da segmentação da imagem Patagônia. Características propostas por Dacheng <i>et al.</i> [14] (segundo momento angular, contraste, correlação e entropia), seqüencial, janela 3x3, classificador redes neurais e normalização. . .	70
A.1	Exemplo de arquivo PGM. . . . .	83
A.2	Exemplo de imagem no formato PGM. . . . .	83
A.3	Exemplo de arquivo PPM. . . . .	84
A.4	Exemplo de imagem no formato PPM. . . . .	84

## LISTA DE TABELAS

2.1	Conjunto de características. . . . .	10
2.2	Subtração da média de cada amostra. . . . .	10
2.3	Conjunto de novos componentes. . . . .	13
2.4	Resultado da classificação pela distância mínima. . . . .	36

## RESUMO

A análise ou interpretação de imagens é uma das principais tarefas envolvidas em um sistema de visão computacional. Seu objetivo é obter informações suficientes para distinguir diferentes objetos de interesse. Normalmente, o processo de análise é baseado nos níveis de cinza ou cores, características de forma e textura dos objetos presentes nas imagens.

Diversas aplicações utilizam técnicas para reconhecimento de objetos em imagens, como a análise de fotografias aéreas ou de imagens de satélite, identificação de faces, recuperação de objetos através de seu conteúdo em uma base de imagens ou vídeo, automação industrial e diagnósticos médicos.

A segmentação de imagens é uma etapa complexa e bastante importante para a análise automática de imagens, consistindo no particionamento da imagem em um conjunto de regiões ou objetos homogêneos. Para que o processo de segmentação seja efetuado de maneira satisfatória é necessário que as propriedades de cada região da imagem sejam descritas de maneira precisa.

Este trabalho descreve um método para segmentação de imagens baseado na análise de textura. A textura é definida por um conjunto de medidas locais descrevendo as variações espaciais de intensidade de níveis de cinza ou cor. Tais medidas, calculadas utilizando-se matrizes de co-ocorrência, são submetidas a um classificador para realizar a rotulação dos objetos presentes na imagem.

Diversos aspectos que influenciam o processo de segmentação são considerados, tais como a escolha das amostras de textura, a presença de ruído e a redução da dimensionalidade do conjunto de características. Um conjunto de descritores, aplicado sobre diversas imagens sintéticas e reais, é utilizado para demonstrar a eficácia do método desenvolvido.

## CAPÍTULO 1

### INTRODUÇÃO

O objetivo da análise ou interpretação de imagens é obter informações suficientes para distinguir diferentes objetos de interesse. Este processo é normalmente baseado na forma, textura, níveis de cinza ou cores dos objetos presentes nas imagens.

Uma dificuldade inerente ao processo de análise de imagens é seu caráter multidisciplinar, em que diversos domínios de conhecimento são comumente necessários para solucionar satisfatoriamente o problema, tais como geometria computacional, visualização científica, psicofísica, estatística, teoria da informação e muitos outros.

Quando é necessária a existência de informações sobre as classes para as quais uma amostra será atribuída, este tipo de classificação é denominado como *classificação supervisionada*, na qual parâmetros devem ser obtidos em uma etapa anterior à execução do algoritmo de classificação. Os referidos parâmetros são obtidos a partir de amostras previamente identificadas, compondo um conjunto denominado *conjunto de treinamento*. Quando não se dispõe de parâmetros ou informações sobre as classes existentes na imagem, a classificação é denominada *não supervisionada*. Neste tipo de classificação, as amostras (não identificadas previamente) que possuem atributos similares são agrupadas em uma mesma classe, enquanto aquelas com atributos distintos são separadas.

Como a interpretação dos dados contidos em imagens digitais é uma atividade complexa, um processo intermediário de *segmentação* é necessário para particionar o conjunto de dados de entrada em regiões homogêneas, de modo a produzir estruturas de mais alto nível. Estas estruturas correspondem a objetos ou partes de objetos que auxiliarão o processo de interpretação das imagens. Portanto, um processo de segmentação que identifique corretamente a localização, a topologia e a forma dos objetos é um requisito de fundamental importância para que as informações resultantes de um sistema de análise de imagens sejam confiáveis.

Processar uma imagem de modo a segmentá-la em um número de objetos é uma tarefa

difícil e extremamente dependente da correta extração de características dos objetos, especialmente em imagens ruidosas. Ruídos podem levar os métodos de segmentação a distorcer as formas dos objetos, comprometendo seu reconhecimento, tal que regiões distintas poderiam ser incorretamente identificadas como uma única região ou, por outro lado, uma região homogênea poderia ser dividida em regiões menores.

A textura contém informações sobre a distribuição espacial das variações de tonalidade de um objeto, podendo ser definida por meio de um conjunto de medidas locais descrevendo as variações espaciais de intensidade ou cor.

A necessidade do processo de análise de imagens vem aumentando significativamente ao longo dos anos, uma vez que a inspeção visual humana nem sempre é precisa e depende de fatores externos, principalmente quando se trata de processos repetitivos.

As próximas seções apresentam a definição de textura, áreas de aplicação da análise de imagens baseada em textura, os objetivos e contribuições deste trabalho, bem como a organização do texto.

## 1.1 Textura

Apresentar uma definição universal para textura não é uma tarefa simples. Muitas definições diferentes têm sido propostas ao longo dos últimos anos. Coggins [12] agrupou diversas dessas definições de texturas encontradas na literatura sobre visão computacional. Alguns exemplos são apresentados a seguir:

1. "Pode-se considerar uma textura como uma região macroscópica. Sua estrutura é atribuída a padrões repetitivos em que os elementos ou primitivas são arranjados de acordo com uma determinada regra."
2. "Uma textura é compreendida como uma primitiva que é repetida sobre uma região que é maior em comparação ao tamanho da primitiva."
3. "A textura é utilizada principalmente em pré-processamento de informação visual, em especial em classificação. Ainda não há uma definição de textura que seja completamente

aceita. Isto mostra que é preciso um modelo mais rico, mais desenvolvido, o qual servirá de representação para muitos níveis diferentes de abstração. Esses níveis utilizarão descritores de borda e de orientação para a superfície, talvez descritores volumétricos. Dadas essas estruturas multi-níveis, parece evidente a necessidade de uma definição para descritores de texturas."

4. "Regiões de textura são interpretadas de maneiras diferentes devido a determinados parâmetros como distância e grau de atenção visual."

Por meio desse conjunto de definições, compreende-se que o conceito de textura depende fortemente da aplicação em questão. Apesar da falta de uma definição universal sobre textura, a maior parte dos pesquisadores concorda em dois aspectos:

1. Existe uma variação entre pixels vizinhos, ou seja, uma não homogeneidade dos níveis de cinza ou tons de cor.
2. Textura é uma propriedade homogênea em alguma escala espacial maior que a definição da imagem.

Levando-se em conta esses dois aspectos, Haralick *et al.* [22] definiram textura a partir de medidas de uniformidade, densidade, aspereza, regularidade e intensidade, dentre outras características da imagem. Para esses autores, textura é um conceito bidimensional, onde uma dimensão caracteriza as propriedades de tonalidades dos pixels e a outra dimensão caracteriza o relacionamento espacial entre esses pixels.

A textura é uma informação valiosa, que pode ser associada à interpretação de padrões relacionados às áreas de aplicação. Ela fornece informações sobre a distribuição espacial das variações de tonalidade de um objeto ou da repetição tonal de alguns grupos de objetos não identificáveis individualmente. Textura é o produto da forma, tamanho, sombra e tonalidade de cada objeto, permitindo a impressão visual de rugosidade ou suavidade de determinada superfície.

A figura 1.1 apresenta imagens com exemplos de texturas.

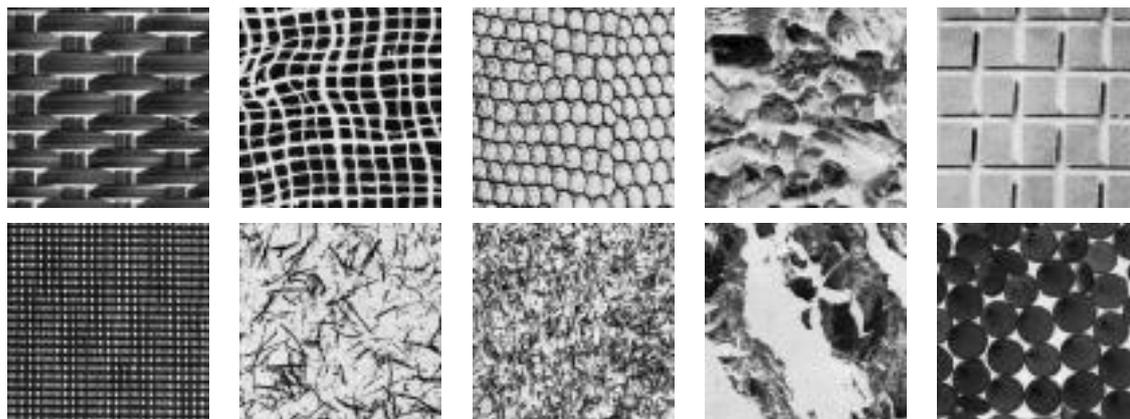


Figura 1.1: Exemplos de Textura - Álbum de Brodatz [7].

## 1.2 Áreas de Aplicação

Vários domínios de conhecimento podem ser beneficiados pela análise de imagens baseada no reconhecimento de texturas sintéticas ou naturais, como será descrito a seguir [17].

- a) *Sensoriamento Remoto*: Fotografias aéreas devem ser avaliadas e catalogadas com fidelidade e eficiência para a obtenção rápida de mapas cartográficos e estudos geográficos, tais como análise do solo, mapeamento de áreas florestais e áreas urbanas [8, 13, 35].
- b) *Medicina*: Diagnósticos médicos podem ser auxiliados por meio de imagens capturadas através de ressonância magnética, tomografia computadorizada e ultrassonografia. A análise e interpretação dessas imagens auxilia, por exemplo, a busca por células atingidas pelo câncer e por anormalidades congênitas ou a contagem de células sanguíneas [39, 65].
- c) *Microscopia*: Análise de imagens capturadas por meio de microscópios ópticos ou eletrônicos, em áreas que variam desde a biologia até a metalurgia, identificando e classificando células ou partículas [23, 33].
- d) *Identificação de faces*: Reconhecimento de pessoas em cenas naturais através de seqüências de vídeo, possibilitando distinguir indivíduos rapidamente. Áreas de aplicação incluem a identificação de pessoas em fichas criminais, criação de retrato falado, monitoramento em sistemas de segurança e de controle de acesso [6, 9].

- e) *Recuperação de imagens em vídeo*: A utilização de dados multimídia, especialmente relacionados a imagens, tem sido considerada para identificar e recuperar dados em bases gráficas, eliminando determinadas restrições de bases textuais [38, 44].
- f) *Controle de qualidade*: A aplicação de processamento de imagens na área de controle de qualidade é muito vasta. Praticamente qualquer processo industrial que necessite de algum monitoramento óptico ou visual pode ser automatizado com vantagens. Aplicações comuns incluem a verificação de falhas em placas de circuito impresso e separação de peças em uma linha de montagem [21, 62, 67].

### 1.3 Objetivos e Contribuições

Este trabalho propõe um método para segmentação de imagens monocromáticas e coloridas utilizando a análise de textura. A textura é descrita por um conjunto de medidas calculadas a partir de matrizes de co-ocorrência.

Dentre os objetivos gerais a serem alcançados pelo trabalho, podem-se citar a escolha das amostras de textura para testes; a extração e seleção das medidas mais representativas dos objetos ou regiões das imagens, armazenando-as em um vetor de características; a aplicação e avaliação da metodologia proposta em imagens sintéticas e reais.

O processo de análise de imagens requer a investigação de diversos parâmetros que afetam seu desempenho, tais como o número de amostras de entrada, valores de limiar, tamanho da janela de vizinhança entre pixels, dimensão do vetor de características e muitos outros. Essa avaliação deve seguir critérios rigorosos para permitir a redução dos custos computacionais e a geração de resultados mais precisos, possibilitando sua aplicação em uma maior variedade de imagens.

Dentre os principais objetivos específicos a serem alcançados, os seguintes tópicos serão analisados e discutidos:

- *Imagens de entrada*: verificação da influência da qualidade da imagem de entrada e de suas amostras de textura no processo de segmentação.

- *Descritores de textura*: proposição e avaliação de medidas para caracterização de texturas.
- *Ruído*: realização de testes em imagens com ruído, analisando sua influência no processo de segmentação.
- *Imagens coloridas*: proposição e avaliação de métricas para segmentação de imagens coloridas utilizando matrizes de co-ocorrência [42, 45].
- *Dimensão do vetor de características*: identificação e remoção de descritores que sejam redundantes ao processo de segmentação, reduzindo a dimensionalidade do vetor de características.
- *Janela de treinamento e janela de classificação*: identificação da influência do tamanho destas janelas no resultado da segmentação da imagem.

## 1.4 Organização do Trabalho

Este capítulo apresentou o conceito de textura, algumas áreas de aplicação de análise de imagens, bem como os principais objetivos e contribuições deste trabalho.

O capítulo 2 apresenta uma revisão dos principais fundamentos e abordagens relacionados à análise de imagens utilizando texturas. Os conceitos de vetor de características e de matriz de co-ocorrência são introduzidos. A análise de componentes principais é apresentada como abordagem a ser utilizada na redução da dimensionalidade do vetor de características. Além disso, os classificadores baseados nas redes neurais e na distância mínima são descritos.

O capítulo 3 descreve o método proposto para segmentação de imagens a partir de informação textural. Cada etapa do método é descrito, desde a preparação das amostras de imagens até a determinação da classe à qual pertence cada pixel da imagem de teste.

Resultados experimentais utilizando uma grande variedade de imagens sintéticas e reais são descritos no capítulo 4.

Finalmente, o capítulo 5 apresenta a conclusão do trabalho e propostas para trabalhos futuros.

## CAPÍTULO 2

### TÓPICOS RELACIONADOS

Este capítulo descreve os principais conceitos e abordagens relacionados ao tema desenvolvido neste trabalho. O primeiro tópico a ser abordado é o vetor de características, o que é e sua influência na segmentação. Em seguida, é apresentado o método de Análise de Componentes Principais que é utilizado na identificação das características mais representativas a serem utilizadas.

Dentre as principais técnicas para análise de textura, são estudadas as abordagens estatística, estrutural e espectral. As matrizes de co-ocorrência, um dos temas centrais deste trabalho, são apresentadas e discutidas. Finalmente, os classificadores baseados nas redes neurais e na distância mínima são apresentados.

#### 2.1 Vetor de Características

O processo de análise de imagens normalmente utiliza uma representação mais sucinta da imagem ou de seus componentes (objetos), conhecida como *vetor de características*, o qual armazena os atributos mais representativos das regiões da imagem.

O número de atributos ou características determina a dimensão do vetor de características e, normalmente, depende da área de aplicação e das propriedades que se deseja discriminar. De maneira geral, busca-se reduzir a dimensionalidade do vetor de características por meio da seleção dos atributos que melhor descrevem as propriedades dos componentes da imagem.

Os principais atributos extraídos das imagens para auxiliar o processo de análise são a *cor*, a *textura* e a *forma* dos objetos [11]. Algumas técnicas para extração de medidas de cada um desses atributos são discutidas a seguir.

As cores presentes em uma imagem desempenham um papel significativo no processo de análise realizado tanto pelos seres humanos quanto pelos computadores. Sabe-se que o olho humano pode discernir milhares de tons e intensidades de cores, comparado com ape-

nas algumas dezenas de níveis de cinza interpretados pelo computador. A abordagem mais comum para extração de características baseadas em distribuição de cores utiliza histogramas de cor [11, 34]. Os histogramas de cor são invariantes quanto à translação e rotação das imagens, sendo que, com a normalização dos histogramas, pode-se obter também a invariância com respeito à escala. Entretanto, os histogramas de cor não indicam a localização espacial dos pixels na imagem.

A textura fornece informações sobre a distribuição espacial das variações de tonalidade de um objeto ou de alguns grupos de objetos não identificáveis individualmente, permitindo o aspecto visual de rugosidade ou suavidade de determinada superfície. Embora o sistema visual humano apresente relativa facilidade no reconhecimento e descrição de textura, definir um conjunto de descritores de textura não é uma tarefa simples. Tal dificuldade é refletida pela grande quantidade de definições e métodos de análise de textura encontrados na literatura [4, 5, 24, 51, 56, 60, 63].

A forma dos objetos é outro atributo importante na análise de imagens. Algumas abordagens para obter descritores eficientes em termos morfológicos dos objetos são os momentos invariantes, os descritores de Fourier, a transformada de Hough, a morfologia matemática, as aproximações por curvas paramétricas e os descritores topológicos [19, 27, 59].

## 2.2 Análise de Componentes Principais

Conforme mencionado na seção anterior, o termo *dimensionalidade* é atribuído ao número de atributos ou características da representação dos objetos presentes na imagem.

Um dos grandes problemas inerentes ao processo de identificação de regiões ou objetos de interesse em imagens é a alta dimensionalidade do vetor de características, o que compromete o desempenho e a precisão do classificador [26].

A redução de dimensionalidade consiste em restringir o conjunto de características, a fim de identificar um número mínimo que seja suficiente para permitir a discriminação dos objetos da imagem. Isto é possível porque nem todas as características podem ser relevantes para a tomada de decisão.

Para efetuar a redução de dimensionalidade, há basicamente duas abordagens: a extração

e a seleção de características. Os métodos de extração criam novas características a partir de transformações ou combinações do conjunto de características original. Os métodos de seleção visam determinar, segundo critérios pré-estabelecidos, o melhor subconjunto de características capazes de discriminar os objetos.

A *análise de componentes principais (PCA - Principal Component Analysis)* é um método estatístico multivariado muito utilizado na área de reconhecimento de padrões, devido à sua simplicidade e velocidade com que é executado. A PCA pode ser vista como uma transformação linear de  $m$  variáveis originais em  $m$  variáveis novas. O conjunto de novas variáveis é ordenado conforme a relevância das características. Por esse motivo, pode-se dizer que a PCA é um método tanto de extração quanto de seleção de características [47].

Segundo Kendall [29], a PCA pode ser utilizada em diferentes enfoques, dentre os quais se destacam:

- a) *Simplificação do Conjunto de Dados*: o objetivo é encontrar uma maneira simplificada de representar o universo de estudo. Isto pode ocorrer pela transformação (combinação linear ou não linear) de um conjunto de variáveis independentes em outro conjunto independente ou em um conjunto de menor dimensão.
- b) *Classificação*: permite analisar se indivíduos estão dispersos aleatoriamente no espaço, podendo também agrupar variáveis.
- c) *Análise de Interdependência*: o objetivo é estudar se um conjunto de variáveis (uma ou mais) é dependente com relação às restantes. Isto pode ser realizado por meio da análise de regressão múltipla ou pela análise de correlação canônica.
- d) *Formulação e Prova de Hipóteses*: a partir de um conjunto de dados é possível encontrar modelos que permitam formular hipóteses em função de parâmetros estimados. A prova desse novo modelo requer uma recompilação dos dados a fim de garantir a necessária independência e validade das conclusões.

Para ilustrar o uso da PCA, a seguir é apresentado um exemplo, adaptado de Smith [57]. Seja o espaço bidimensional formado por duas características, cada uma com dez amostras, conforme a tabela 2.1.

<b>x</b>	2,5	0,5	2,2	1,9	3,1	2,3	2,0	1,0	1,5	1,1
<b>y</b>	2,4	0,7	2,9	2,2	3,0	2,7	1,6	1,1	1,6	0,9

Tabela 2.1: Conjunto de características.

A média aritmética, dada pela equação 2.1, é calculada para cada dimensão, resultando para este exemplo nos valores:  $\bar{X} = 1,81$  e  $\bar{Y} = 1,91$ .

$$\bar{X} = \frac{\sum_{i=1}^N X_i}{N} \quad (2.1)$$

Subtraindo de cada amostra da tabela 2.1 o valor da média, obtém-se o conjunto de dados apresentado na tabela 2.2.

<b>x</b>	0,69	-1,31	0,39	0,09	1,29	0,49	0,19	-0,81	-0,31	-0,71
<b>y</b>	0,49	-1,21	0,99	0,29	1,09	0,79	-0,31	-0,81	-0,31	-1,01

Tabela 2.2: Subtração da média de cada amostra.

A partir dos valores da tabela 2.2, constrói-se a matriz de covariância. Para a construção desta, utilizam-se os conceitos de variância, desvio padrão e covariância, revistos a seguir.

A *variância* de uma variável aleatória é uma medida da sua dispersão estatística, indicando quão longe, em geral, os seus valores se encontram do valor esperado, ou seja, da média. A unidade de variância é o quadrado da unidade de observação. Por exemplo, a variância de um conjunto de alturas medidas em centímetros será dada em centímetros quadrados. Este fato é inconveniente e levou muitos estatísticos a usar a raiz quadrada da variância, conhecida como o *desvio padrão* [57].

As equações para o desvio padrão e a variância são mostradas nas equações 2.2 e 2.3, respectivamente.

$$\sigma = \sqrt{\frac{\sum_{i=1}^N (X_i - \bar{X})^2}{N - 1}} \quad (2.2)$$

$$\sigma^2 = \frac{\sum_{i=1}^N (X_i - \bar{X})^2}{N - 1} \quad (2.3)$$

O desvio padrão e a variância operam apenas em uma dimensão, de modo que só é possível calcular estas medidas para cada um dos conjuntos de dados, independentemente das demais dimensões. Contudo, é útil ter uma medida semelhante para encontrar a variação da média com respeito às outras dimensões [57].

A *covariância* é mostrada na equação 2.4. A covariância de uma dimensão com ela mesma resulta na variância.

$$\text{cov}(X, Y) = \frac{\sum_{i=1}^N (X_i - \bar{X})(Y_i - \bar{Y})}{N - 1} \quad (2.4)$$

A covariância é sempre medida entre duas dimensões, assim, para calcular a covariância de três dimensões, por exemplo,  $(x, y, z)$ , é necessário calcular a covariância entre as dimensões  $x$  e  $y$ , entre as dimensões  $x$  e  $z$  e entre as dimensões  $y$  e  $z$ .

A definição da *matriz de covariância* de dimensão  $n$  é dada por:

$$C_{N \times N} = (C_{i,j}, C_{i,j} = \text{cov}(\text{Dim}_i, \text{Dim}_j)) \quad (2.5)$$

Dessa forma, para um conjunto de dados tridimensionais  $x$ ,  $y$  e  $z$ , a matriz de covariância é dada por:

$$C = \begin{pmatrix} \text{cov}(x, x) & \text{cov}(x, y) & \text{cov}(x, z) \\ \text{cov}(y, x) & \text{cov}(y, y) & \text{cov}(y, z) \\ \text{cov}(z, x) & \text{cov}(z, y) & \text{cov}(z, z) \end{pmatrix}$$

Para o exemplo apresentado, a matriz de covariância é:

$$\text{cov} = \begin{pmatrix} 0,6165555556 & 0,6154444444 \\ 0,6154444444 & 0,7165555556 \end{pmatrix}$$

A Análise de Componentes Principais baseia-se na transformação linear das observações originais. Esta transformação linear é conhecida na Álgebra Vetorial como generalização de

*autovetores e autovalores*, ou também vetores e valores característicos.

Os autovalores representam o comprimento dos eixos das componentes principais de uma imagem e são medidos em unidade de variância. Associado a cada autovalor existe um vetor de módulo unitário chamado autovetor. Os autovetores representam as direções dos eixos das componentes principais.

Se, para algum escalar  $m$  e para algum vetor  $v$ ,  $T(v) = m.v$  diz-se que  $v$  é um autovetor de  $T$  com relação ao autovalor  $m$  [32].

Os autovalores são encontrados pela solução das equações 2.6 e 2.7:

$$|S - lI|t_i = 0 \quad (2.6)$$

onde  $l$  é um escalar e

$$u_i = \frac{t_i}{\sqrt{t_i^T t_i}} \quad (2.7)$$

para  $i = 1, 2, 3, \dots, p$ , tal que  $t$  são os autovalores,  $S$  é a matriz de covariância e  $I$  é a matriz identidade. Portanto,

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad lI = \begin{pmatrix} l & 0 \\ 0 & l \end{pmatrix}$$

Assumindo para este exemplo  $i = 1$ , obtém-se:

$$|S - lI| = \begin{pmatrix} 0,616555556 - l & 0,615444444 \\ 0,615444444 & 0,716555556 - l \end{pmatrix} = 0$$

Resolvendo o determinante, obtém-se:

$$l_1 = 0,049083399 \text{ e } l_2 = 1,284027713$$

$$|S - lI| = \begin{bmatrix} 0,616555556 - 0,049083399 & 0,615444444 \\ 0,615444444 & 0,716555556 - 0,049083399 \end{bmatrix} = \begin{bmatrix} t_{11} \\ t_{21} \end{bmatrix} = 0$$

Sendo  $t_{11}=1$ , então  $t_{21}= 0,922052611$ , resultando

$$u_1 = \begin{pmatrix} -0,735178655 \\ 0,677873399 \end{pmatrix}$$

De maneira análoga, sendo  $t_{21}=1$ ,

$$u_2 = \begin{pmatrix} -0,677873399 \\ -0,735178656 \end{pmatrix}$$

Com os autovetores, cria-se a matriz  $U$  a seguir:

$$U = \begin{bmatrix} -0,735178655 & | & -0,677873399 \\ 0,677873399 & | & -0,735178656 \end{bmatrix}$$

Para obter o novo conjunto de dados, utiliza-se a equação 2.8.

$$Z_i = [u_i^T [X - \bar{X}]] \quad (2.8)$$

Para o exemplo:

$$Z_1 = \begin{bmatrix} -0,677873399 & | & -0,735178656 \\ -0,735178656 & | & -0,677873399 \end{bmatrix} \begin{bmatrix} 2,5 - 1,8 \\ 2,4 - 1,9 \end{bmatrix}$$

$$Z_1 = \begin{bmatrix} -0,842100707 \\ -0,17568836 \end{bmatrix}$$

Calculando a fórmula para todas amostras da tabela 2.1, o conjunto de novos componentes resultantes é dado pela tabela 2.3.

<b>x</b>	-0,842	1,777	-0,992	-0,274	-1,675	-0,912	0,099	1,144	0,438	1,223
<b>y</b>	-0,175	0,142	0,384	0,130	-0,209	0,175	-0,349	0,464	0,017	-0,162

Tabela 2.3: Conjunto de novos componentes.

A partir do novo conjunto de dados obtido e dos autovetores, são selecionados os valores mais relevantes dependendo da aplicação [57]. A Análise de Componentes Principais será utilizada neste trabalho como o método para identificar as características mais relevantes na classificação, reduzindo a dimensionalidade do vetor de características. Após o cálculo da PCA, foram selecionadas as cinco primeiras características na ordem crescente dos autovalores.

## 2.3 Abordagens para Análise de Textura

A análise de textura visa estudar as propriedades da superfície de um objeto, compreender como os seres humanos discriminam diferentes texturas e modelar algoritmos que tenham a capacidade de realizar essa tarefa.

As principais técnicas utilizadas para descrever textura em imagens podem ser classificadas em abordagens *estatística*, *estrutural* e *espectral*. As próximas seções descrevem essas três abordagens em mais detalhes, com ênfase na abordagem estatística, a qual foi utilizada no método desenvolvido neste trabalho.

### 2.3.1 Abordagem Estatística

Na abordagem estatística, a textura é descrita por um conjunto de medidas locais extraídas da amostra. Medidas estatísticas típicas incluem a entropia, correlação, contraste e variância.

Ao invés de tentar compreender explicitamente a estrutura hierárquica da textura, os métodos estatísticos descrevem a textura por meio de propriedades não determinísticas que definem as distribuições e relacionamentos entre os níveis de cinza ou tons de cor de uma imagem.

Métodos baseados em estatísticas de segunda ordem têm demonstrado alto poder de discriminação em diversos tipos de imagens. Uma técnica para análise de textura, proposta por Haralick *et al.* [22], utiliza matrizes de co-ocorrência para representar a distribuição de probabilidade da dependência espacial da intensidade dos pixels.

As matrizes de co-ocorrência são construídas assumindo-se que cada pixel, exceto os externos, possui 8 vizinhos (direções horizontal, vertical e diagonal). Um conjunto de quatorze

medidas de características, extraídas a partir de cada uma dessas matrizes, é definido para descrever as propriedades contidas nas texturas, tais como uniformidade, densidade, aspereza, regularidade e intensidade.

Uma descrição sobre a importância do uso de características que são invariantes sobre transformações monotônicas de níveis de cinza também é apresentada em [22]. Duas fotografias contendo a mesma textura podem ter, por exemplo, níveis de cinza claros e escuros, respectivamente. A maioria dos seres humanos facilmente identificaria que a textura contida nas imagens é a mesma. Porém, para que uma máquina seja capaz de fazer tal identificação, cada imagem deve ser adequadamente quantizada e suas características calculadas. O segundo momento angular, a entropia, a soma da entropia, a diferença da entropia, as medidas de informação de correlação e o coeficiente máximo de correlação apresentam essa propriedade da invariância e permitem a obtenção de resultados mais eficazes de classificação.

Shaban e Dikshit [54] demonstram a importância da textura na segmentação de imagens e discutem a variação de determinados parâmetros em seus experimentos, alguns descritos a seguir. As características de textura são extraídas por meio de uma janela de tamanho arbitrário que percorre as amostras de textura. Há um tamanho ótimo de janela para cada uma das características extraídas. Segundo seus experimentos, resultados indicam que, para o segundo momento angular, contraste e entropia, o tamanho ótimo de janela é 7, para homogeneidade é 5 e para variância e correlação é 9.

Qualquer aumento no tamanho da janela que não seja o ótimo provavelmente tornará o processo computacionalmente mais custoso, sem prover qualquer melhora na precisão, podendo também levar a interpretações incorretas pelo fato de englobar mais do que uma classe de textura. Com relação à quantização dos níveis de cinza, os valores 16, 32, 64 e 128 produziram resultados estatisticamente similares [54]. O uso da média das características de textura, calculadas para as quatro direções (horizontal, vertical e duas diagonais), normalmente gera resultados melhores do que quando apenas uma orientação é utilizada.

Muhamad e Deravi [40] apresentam as matrizes de co-ocorrência na forma de imagens para observar as diferenças ou similaridades na representação das características de textura. Cada elemento da matriz de co-ocorrência é considerado como um pixel, onde seu nível de

cinza é o valor numérico do elemento correspondente. Algumas amostras são extraídas de classes de textura e as matrizes de co-ocorrência calculadas. Quatro casos são computados, utilizando os seguintes valores de quantização dos níveis de cinza:  $256 \times 256$ ,  $64 \times 64$ ,  $16 \times 16$  e  $4 \times 4$ . Com a diminuição do tamanho da matriz, há também diminuição na quantidade de informação, tornando-se mais difícil distinguir entre algumas matrizes do mesmo grupo. Entretanto, nota-se que mesmo com  $4 \times 4$  pixels, a maioria das matrizes preserva informação suficiente para deixá-las distinguíveis das demais.

Um método interessante para melhorar a qualidade das características das matrizes de co-ocorrência foi proposto por Walker *et al.* [64]. Eles classificaram as características propostas por Haralick e seus colaboradores [22] através do cálculo e ponderação de cada valor do elemento da matriz ou sua localização espacial. Por exemplo, medidas de energia e entropia foram calculadas com base no valor dos elementos, e o segundo momento angular, energia, entropia, correlação e variância foram calculados com base na localização espacial. Os autores propuseram que é melhor anular os elementos da matriz que possuem pouca habilidade de discriminação.

Então, com base no cálculo da distância de Bhattacharya, encontrou-se quais elementos eram mais discriminatórios. Uma matriz de discriminação contendo esses pesos pode ser multiplicada pela matriz original para produzir uma melhor representação dos valores mais representativos. Com esses valores, pode-se calcular as medidas de características ponderando a soma de elementos. Baseado nos novos cálculos, os autores encontraram que, em 6 das 8 medidas, foram obtidos desempenhos melhores nas tarefas de classificação com valores sempre acima de 70% de precisão.

Weszka *et al.* [66] propõem uma aplicação de análise de textura baseada em matrizes de co-ocorrência para inspeção de materiais. A idéia do estudo foi utilizar métodos de textura para avaliar a qualidade da superfície de um material para distinguir entre amostras pobres e aceitáveis. Para isso, doze imagens digitais de superfícies de materiais foram consideradas. Para cada imagem, uma pequena janela de pixels foi utilizada para calcular as características de textura. As medidas incluíram o espectro de Fourier e as matrizes de co-ocorrência.

Para as matrizes de co-ocorrência, quatro medidas foram calculadas em quatro direções

principais, sendo o segundo momento angular, entropia, contraste e correlação. Ainda, para cada característica, a média, o desvio padrão, o máximo, o mínimo e a média dessas três estatísticas foram calculadas e usadas como características. O objetivo do estudo era encontrar quais dos conjuntos de características obtinha a melhor classificação que se assemelhava à discriminação de um ser humano. Os autores observaram que os melhores resultados foram obtidos para as características de entropia, desvio padrão da entropia e máxima correlação. Eles também concluíram que a decisão a respeito da qualidade da textura de uma superfície encontrada nem sempre corresponde com a propriedade visual óbvia do material.

Estatísticas de terceira ordem ou superior também têm sido aplicadas em análise de textura. O estudo realizado por Murino *et al.* [41] propõe uma técnica para classificação de texturas que é baseada no uso de estatísticas de ordem mais alta (*higher order statistics*) para definir medidas discriminatórias de textura. Combinações de momentos de alta ordem são usadas como características. Os autores analisam o desempenho dessas características em imagens com diferentes proporções de ruído. Eles verificam que uma classificação quase perfeita é possível para ruído zero e um comportamento robusto é mostrado em direção ao ruído aditivo.

Outra abordagem para análise de texturas é baseada em campos aleatórios de Markov (MRF, *Markov random fields*) [2, 30, 31, 52, 53], onde uma textura é considerada como uma amostra extraída de um processo estocástico definido por um conjunto de parâmetros que, servindo como modelo para textura, resumem suas características.

Chen e Huang [10] citam que, dentre as principais vantagens de se utilizar um MRF, está o fato de que seus parâmetros capturam a direcionalidade e a aspereza da textura, além de não necessitar de uma etapa subsequente de seleção ou nova extração de características.

### 2.3.2 Abordagem Estrutural

A abordagem estrutural utiliza a idéia de que texturas são compostas de primitivas dispostas de forma aproximadamente regular e repetitiva, de acordo com regras bem definidas [19, 43]. Como exemplo, pode-se citar a descrição da textura baseada em linhas paralelas regularmente espaçadas.

Para entender melhor essa abordagem, pode-se utilizar uma regra do tipo  $S \rightarrow aSA$ .

Neste tipo de regra, as letras maiúsculas  $A$  e  $S$  são variáveis e as letras minúsculas  $a$ ,  $b$  e  $c$  são símbolos terminais. Uma variável é escolhida como símbolo inicial para começar a construção da cadeia de caracteres.

Supondo que além da regra acima, existam ainda as seguintes regras:

$$\begin{array}{ll} (1) S \rightarrow aSA & (4) A \rightarrow cA \\ (2) S \rightarrow bA & (5) A \rightarrow c \\ (3) S \rightarrow a & (6) A \rightarrow b \end{array}$$

Um exemplo utilizando essas regras é agora apresentado. Seja  $S$  o símbolo inicial. Desta forma, pelas regras anteriores, o sinal  $\rightarrow$  indica que  $S$  pode ser substituído por  $aSa$  (caso seja utilizada a primeira regra),  $bA$  (caso seja utilizada segunda regra) ou  $a$  (caso a terceira regra seja utilizada), ou seja, toda vez que existir uma variável em uma regra, esta poderá ser substituída pelo conteúdo localizado do lado direito do símbolo  $\rightarrow$ . Caso uma variável possua mais de uma regra, como ocorre neste exemplo, qualquer uma das regras poderá ser utilizada na substituição.

Com essas regras, podem-se gerar vários tipos de cadeias. Uma delas poderia ser, por exemplo,  $aabbcccb$ , que foi obtida pelos seguintes passos:

$$\begin{array}{ll} S \text{ (regra inicial)} & aaScAc b \text{ (regra 6)} \\ aSA \text{ (regra 1)} & aaScccb \text{ (regra 5)} \\ aScA \text{ (regra 4)} & aabAcccb \text{ (regra 2)} \\ aaSAcA \text{ (regra 1)} & aabbcccb \text{ (regra 6) - Seqüência final} \\ aaScAcA \text{ (regra 4)} & \end{array}$$

Vale observar que, dependendo das regras, uma cadeia de símbolos terminais poderá ser gerada por diferentes combinações de regras. Além disso, uma regra poderá ser do tipo  $S \rightarrow \emptyset$ , que indica que a variável será eliminada da cadeia sem ser substituída por um símbolo terminal.

Para construir uma textura utilizando a regra da cadeia, pode-se definir uma ação para

cada símbolo terminal da cadeia gerada como, por exemplo, *a* círculo à direita, *b* círculo abaixo e *c* círculo à esquerda. A primitiva de textura do exemplo construído é dada pela figura 2.1.

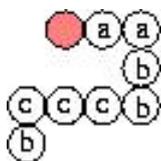


Figura 2.1: Regra da cadeia.

Como observado no exemplo, é importante ressaltar que, dependendo das regras e de sua combinação, estruturas irregulares podem ser geradas. Uma primitiva de textura simples pode ser usada na formação de padrões complexos de textura através de algumas regras que limitem o número de arranjos possíveis da primitiva.

A vantagem da abordagem estrutural é que ela provê uma boa descrição simbólica da imagem. Entretanto, essa característica é normalmente mais útil em tarefas de síntese do que em análise de textura.

### 2.3.3 Abordagem Espectral

A abordagem espectral baseia-se em propriedades do espectro de Fourier [19], sendo principalmente utilizada na detecção de periodicidade global em uma imagem através da identificação de picos de alta energia no espectro.

O espectro é útil para este propósito por possuir três características: (a) picos proeminentes no espectro fornecem a direção dos padrões de textura; (b) a posição dos picos no plano da frequência fornece o período espacial fundamental dos padrões e (c) a eliminação de quaisquer componentes periódicos através da filtragem deixa os elementos não-periódicos na imagem, que podem ser descritos por técnicas estatísticas.

O espectro de Fourier é idealmente adaptado para a descrição da orientação de padrões periódicos ou quase periódicos em uma imagem. Os padrões globais de textura, embora facilmente distinguíveis como concentrações de agrupamentos de alta energia no espectro, são geralmente difíceis de se detectar em métodos espaciais devido à natureza local dessas técnicas [19].

Geralmente é feita uma simplificação na detecção e interpretação das características do espectro citadas anteriormente, expressando-se o espectro em coordenadas polares, levando a uma função  $S(r, \theta)$ , sendo que  $S$  é uma função de espectro, e  $r$  e  $\theta$  são variáveis nesse sistema de coordenadas. Para cada direção  $\theta$ ,  $S(r, \theta)$  pode ser considerado uma função  $S_\theta(r)$ . De maneira similar, para cada  $r$ ,  $S_r(\theta)$  é uma função unidimensional. A análise de  $S_\theta(r)$  para um valor fixo de  $\theta$  fornece o comportamento do espectro (como a presença de picos) ao longo de uma direção radial a partir da origem, enquanto a análise de  $S_r(\theta)$  para um valor fixo de  $r$  leva ao comportamento ao longo de uma circunferência centrada na origem.

Uma descrição global é obtida através das funções

$$S(r) = \sum_{\theta=0}^{\pi} S_\theta(r) \quad (2.9)$$

$$S(\theta) = \sum_{r=1}^R S_r(\theta) \quad (2.10)$$

em que  $R$  é o raio de uma circunferência centrada na origem. Para um espectro de  $N \times N$ ,  $R$  é tipicamente escolhido como  $N/2$ .

Os resultados das equações 2.9 e 2.10 constituem um par de valores  $[S(r), S(\theta)]$  para cada par de coordenadas  $(r, \theta)$ . Variando-se essas coordenadas, pode-se gerar funções unidimensionais  $S(r)$  e  $S(\theta)$  que constituem descrições de energia espectral da textura para uma imagem ou região em questão. Além disso, os descritores das próprias funções podem ser calculados para uma caracterização quantitativa. Descritores tipicamente usados para esse propósito são a posição do valor mais alto, a média e a variância da amplitude e as variações de eixo, e a distância entre a média e o maior valor da função.

A figura 2.2 ilustra o uso das equações 2.9 e 2.10 para a descrição global de textura. A figura 2.2(a) mostra uma imagem com uma textura periódica, enquanto a figura 2.2(b) apresenta o espectro correspondente. As figuras 2.2(c) e (d) mostram os gráficos de  $S(r)$  e  $S(\theta)$ , respectivamente. O gráfico de  $S(r)$  possui uma estrutura típica com um conteúdo de alta energia perto da origem e progressivamente menor para as altas frequências. O gráfico de  $S(\theta)$  mostra picos proeminentes em intervalos de  $45^\circ$ , correspondendo claramente à periodicidade

no conteúdo de textura da imagem.

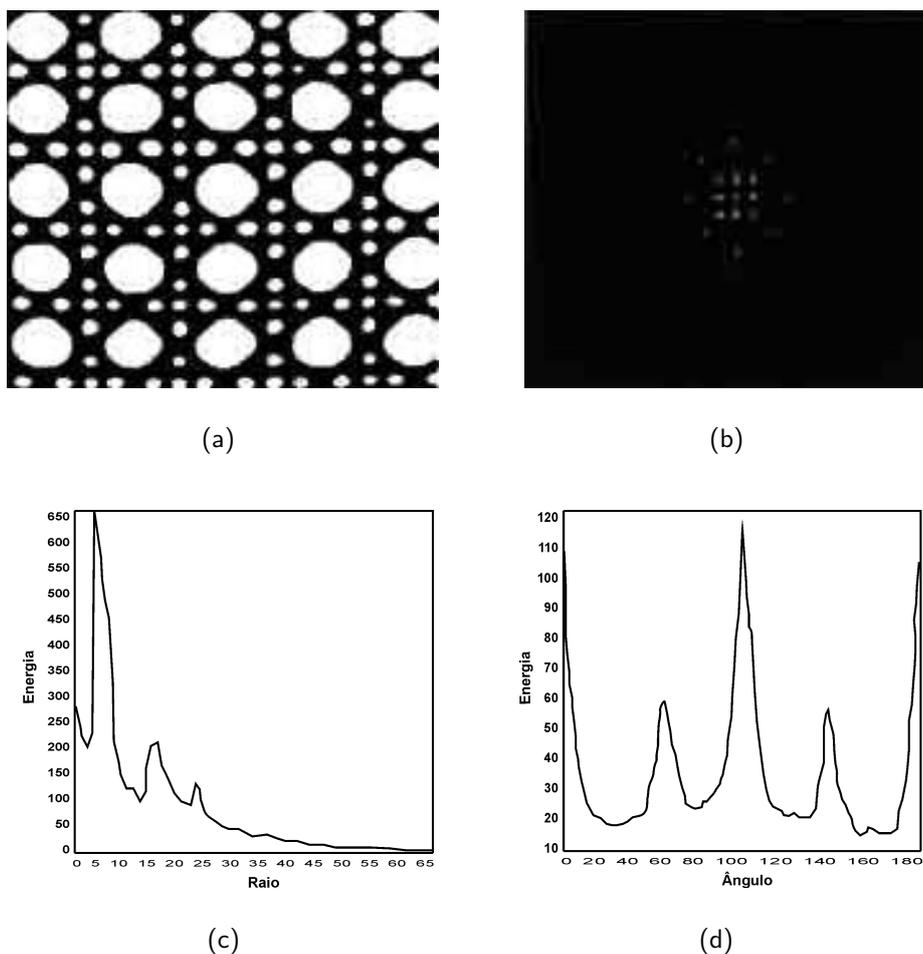


Figura 2.2: (a) Imagem mostrando uma textura periódica; (b) espectro; (c) gráfico de  $S(r)$ ; (d) gráfico de  $S(\theta)$ .

Para ilustrar como o gráfico de  $S(\theta)$  poderia ser usado para diferenciar dois padrões diferentes de textura, a figura 2.3(a) mostra outra imagem cujo padrão de textura está predominante nas orientações horizontal e vertical. A figura 2.3(b) mostra o gráfico de  $S(\theta)$  para o espectro dessa imagem. Esse gráfico mostra picos em intervalos de  $90^\circ$ . A discriminação entre os dois padrões de textura poderia ser feita através da análise dos gráficos de  $S(\theta)$  correspondentes. Maiores detalhes a respeito dessa abordagem podem ser encontrados em [43].

Modelos fractais [28], transformadas *wavelets* [16, 34, 51] e filtros de Gabor [24] também apresentam-se como técnicas eficientes para descrever textura, especialmente em análise multi-resolução de imagens.

Como a textura é altamente dependente da escala, pode-se diminuir seu grau de sensibili-

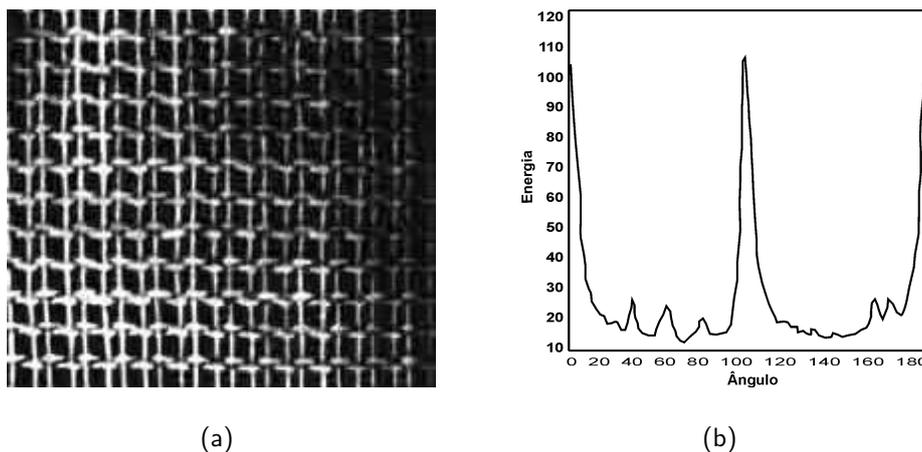


Figura 2.3: (a) Imagem contendo um tipo diferente de textura periódica; (b) gráfico de  $S(\theta)$ .

dade, descrevendo-a em múltiplas resoluções. Uma escala apropriada pode ser escolhida para se obter a máxima discriminação da textura. Os filtros de Gabor e as transformadas *wavelets* são adequados para a caracterização de textura multi-escalar [59]. Ambas abordagens representam uma imagem em um espaço cujo sistema de coordenadas possui uma interpretação que está altamente relacionada com as características de uma textura, tal como freqüência ou tamanho.

Os *filtros de Gabor* permitem uma melhor localização espacial, entretanto, sua utilidade é limitada na prática porque não existem filtros de resolução simples que possam localizar uma estrutura espacial em texturas naturais. Comparado com o filtro de Gabor, as características da transformada *wavelet* possuem várias vantagens: variando a resolução espacial, permite-se representar texturas em uma escala mais adequada; existe uma ampla faixa de escolhas para a função *wavelet*, tornando essa abordagem adequada para a análise de textura em uma aplicação específica. Isso torna a transformada *wavelet* apropriada para segmentação de textura [36].

Uma metodologia, proposta por Angelo e Haertel [3], consiste na caracterização das classes texturais presentes na imagem por meio das freqüências espaciais mais significativas, conforme identificadas nas amostras de cada classe. Cada freqüência espacial dá origem a um filtro de Gabor. Os filtros assim gerados, quando aplicados na imagem, originam um número correspondente de imagens filtradas, denominadas de bandas texturais. Esses dados multivariados

podem então ser utilizados em um classificador, gerando uma imagem classificada com base em atributos de textura.

A aplicação dos filtros de Gabor na classificação supervisionada de imagens é especialmente útil quando as classes envolvidas apresentam comportamento espectral muito semelhante, mas características de textura distintas. Nesse caso, a utilização de atributos espectrais no processo de classificação ocasionará, em geral, resultados pobres no processo de classificação, restando os atributos de textura para se obter uma classificação com maior exatidão.

## 2.4 Matrizes de Co-ocorrência

Conforme mencionado nas primeiras seções deste capítulo, a abordagem estatística utilizada para descrever texturas utiliza apenas os valores dos níveis de cinza ou tons de cor para calcular os descritores de textura. Entretanto, apenas o cálculo de tais descritores não é suficiente para o reconhecimento de texturas, pois, em uma imagem monocromática, texturas diferentes podem conter exatamente os mesmos valores de níveis de cinza ou, ainda, uma mesma textura pode conter valores de tons de cinza diferentes por apresentar um contraste maior ou menor.

Neste trabalho, *matrizes de co-ocorrência* são utilizadas para considerar o relacionamento espacial entre os pixels de uma imagem de acordo com seus níveis de cinza ou tons de cor, ou seja, uma análise espacial da variação de tons de cinza ou de cor na imagem de acordo com alguns fatores que serão explicados ainda neste capítulo.

Para se entender o conceito de matrizes de co-ocorrência, algumas definições são apresentadas a seguir. Nesta seção, considera-se que as imagens são monocromáticas. O capítulo 3 estenderá esse conceito para imagens coloridas.

1. *Vizinhos Imediatos*: seja  $p$  um pixel da imagem. Os vizinhos imediatos deste pixel são aqueles que estão conectados a  $p$  nas direções horizontal, vertical e diagonal. Desta forma, os vizinhos imediatos de  $p = (x, y)$  são:

- na horizontal:  $(x - 1, y)$  e  $(x + 1, y)$
- na vertical:  $(x, y + 1)$  e  $(x, y - 1)$

- na diagonal:  $(x - 1, y + 1), (x + 1, y + 1), (x + 1, y - 1), (x - 1, y - 1)$

Esses vizinhos imediatos podem ser considerados com relação à direção, conforme ilustrado na figura 2.4.

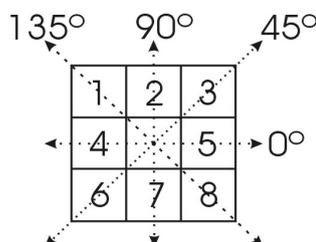


Figura 2.4: Ângulos para o cálculo na matriz de co-ocorrência.

2. *Combinação de Intensidades*: como já foi mencionado anteriormente, a matriz de co-ocorrência reflete o relacionamento espacial dos pixels de acordo com as intensidades dos pixels da imagem. É preciso então, antes da construção da matriz, verificar todas as combinações possíveis de transição entre as intensidades.

### 2.4.1 Definição da Matriz de Co-ocorrência

Uma matriz de co-ocorrência pode ser definida como  $P(i, j, d, q)$  onde:  $i$  e  $j$  são pixels vizinhos de acordo com a distância  $d$ . Esta distância entre os pixels  $i$  e  $j$  é analisada de acordo com uma direção  $q$ . Os quatro valores possíveis para  $q$  são apresentados na figura 2.4.

Desta forma, seja  $A$  uma matriz (obrigatoriamente quadrada) de co-ocorrência. O elemento  $a_{ij}$  representa o número de vezes que ocorreu uma transição do nível de cinza  $Z_i$  para  $Z_j$  considerando a distância  $d$  entre os pixels vizinhos  $i$  e  $j$  na direção  $q$ .

As freqüências das funções de ângulo e distância de uma matriz de co-ocorrência podem

ser representadas formalmente como [59]:

$$\begin{aligned}
P(i, j, d, 0^\circ) &= |\{(k, l), (m, n)] \in D : k - m = 0, |l - n| = d, f(k, l) = i, f(m, n) = j\}| \\
P(i, j, d, 45^\circ) &= |\{(k, l), (m, n)] \in D : (k - m = d, l - n = -d) \text{ ou} \\
&\quad (k - m = -d, l - n = d), f(k, l) = i, f(m, n) = j\}| \\
P(i, j, d, 90^\circ) &= |\{(k, l), (m, n)] \in D : |k - m| = d, l - n = 0, f(k, l) = i, f(m, n) = j\}| \\
P(i, j, d, 135^\circ) &= |\{(k, l), (m, n)] \in D : (k - m = d, l - n = d) \text{ ou} \\
&\quad (k - m = -d, l - n = -d), f(k, l) = i, f(m, n) = j\}|
\end{aligned}
\tag{2.11}$$

onde  $k, l, m$  e  $n$  são índices da matriz de co-ocorrência.

## 2.4.2 Construção da Matriz de Co-ocorrência

A partir da definição de matriz de co-ocorrência e os conceitos de vizinhos imediatos e combinação de níveis de cinza, a construção da matriz de co-ocorrência torna-se bem simples.

Para iniciar a construção, é preciso ter uma imagem com  $Z_n$  níveis de cinza. Para este exemplo será utilizada a imagem ilustrada na figura 2.5.

0	0	0	1	2
1	1	0	1	1
2	2	1	0	0
1	1	0	2	0
0	0	1	0	1

Figura 2.5: Imagem com níveis de cinza de 0 a 2.

Como se pode observar, essa imagem possui apenas 3 níveis de cinza:  $Z_0 = 0$ ,  $Z_1 = 1$  e  $Z_2 = 2$ .

O próximo passo é construir a matriz de co-ocorrência, que será obviamente uma matriz quadrada, pois conterá o mesmo número de linhas e de colunas que equivalem ao número total de níveis de cinza. A combinação possível de níveis de cinza que forma a matriz é ilustrada a seguir:

	<b>0</b>	<b>1</b>	<b>2</b>
<b>0</b>			
<b>1</b>			
<b>2</b>			

Para esse exemplo, a distância será definida como  $d = 1$ . Para cada imagem sempre haverá quatro matrizes de co-ocorrência, uma para cada direção, ou seja, uma matriz para  $q = 0^\circ$ ,  $q = 45^\circ$ ,  $q = 90^\circ$  e  $q = 135^\circ$ . Sendo a direção  $q = 0^\circ$ , para cada pixel analisado, deve-se que verificar os pixels imediatamente à direita e também à esquerda, pois a distância é 1 e a direção é  $0^\circ$ .

Cada posição da matriz de co-ocorrência conterà a variação dos níveis de cinza referentes aos parâmetros informados no parágrafo anterior, considerando os índices da matriz. Então, por exemplo, a posição  $(0, 0)$  da matriz de co-ocorrência conterà a quantidade de vezes que houve variação no nível de cinza 0 para o nível de cinza 0 na horizontal, ou seja, à direita ou à esquerda. Nesse exemplo, o valor da posição  $(0, 0)$  será igual a 8, pois, como pode ser visto na imagem, a situação descrita ocorre como mostrado a seguir:

$$\begin{array}{ll}
 I_{(0,0)} \rightarrow I_{(0,1)} & I_{(2,3)} \rightarrow I_{(2,4)} \\
 I_{(0,1)} \rightarrow I_{(0,0)} & I_{(2,4)} \rightarrow I_{(2,3)} \\
 I_{(0,1)} \rightarrow I_{(0,2)} & I_{(4,0)} \rightarrow I_{(4,1)} \\
 I_{(0,2)} \rightarrow I_{(0,1)} & I_{(4,1)} \rightarrow I_{(4,0)}
 \end{array} \quad (2.12)$$

Se  $d = 2$ , ocorreriam  $M_{(0,0)} = 6$  transições, como apresentado a seguir:

$$\begin{array}{ll}
 I_{(0,0)} \rightarrow I_{(0,2)} & I_{(3,4)} \rightarrow I_{(3,2)} \\
 I_{(0,2)} \rightarrow I_{(0,0)} & I_{(4,1)} \rightarrow I_{(4,3)} \\
 I_{(3,2)} \rightarrow I_{(3,4)} & I_{(4,3)} \rightarrow I_{(4,1)}
 \end{array} \quad (2.13)$$

Caso  $d = 2$ , apenas as transições mostradas nas expressões 2.13 seriam consideradas, não

sendo válidas as transições mostradas nas expressões 2.12. Para  $d = 2$ , os pixels analisados seriam:

- na horizontal:  $(x - 2, y)$  e  $(x + 2, y)$
- na vertical:  $(x, y + 2)$  e  $(x, y - 2)$
- na diagonal:  $(x - 2, y + 2)$ ,  $(x + 2, y + 2)$ ,  $(x - 2, y - 2)$ ,  $(x + 2, y - 2)$

É muito importante observar que os valores da diagonal principal sempre serão números pares, isto porque a diagonal principal contém a análise das transições de níveis de cinza iguais. Desta forma, sendo  $Z_0 = 0$ ,  $Z_1 = 1$  e  $Z_2 = 2$ , uma transição  $Z_0 \rightarrow Z_0$  será analisada nos dois sentidos para uma mesma direção, ou seja,  $Z_0 \rightarrow Z_0$  e  $Z_0 \leftarrow Z_0$ . Já uma transição  $Z_0 \rightarrow Z_1$  é diferente de uma transição  $Z_1 \rightarrow Z_0$ .

A matriz resultante é:

	<b>0</b>	<b>1</b>	<b>2</b>
<b>0</b>	8	8	2
<b>1</b>	8	6	2
<b>2</b>	2	2	2

As matrizes de direção  $45^\circ$ ,  $90^\circ$  e  $135^\circ$  são, respectivamente:

	<b>0</b>	<b>1</b>	<b>2</b>
<b>0</b>	6	5	2
<b>1</b>	5	4	4
<b>2</b>	2	4	0

	<b>0</b>	<b>1</b>	<b>2</b>
<b>0</b>	4	9	2
<b>1</b>	9	2	5
<b>2</b>	2	5	0

	<b>0</b>	<b>1</b>	<b>2</b>
<b>0</b>	8	4	1
<b>1</b>	4	6	4
<b>2</b>	1	4	0

Para utilizar as matrizes de co-ocorrência no cálculo dos descritores apresentados no capítulo 3, é preciso realizar a normalização de todas elas, obtendo assim quatro novas matrizes normalizadas.

Seja  $M$  uma matriz quadrada. A probabilidade  $P_{i,j}$  indica a probabilidade de um certo valor aparecer na matriz, sendo então calculado pela equação 2.14.

$$P_{i,j} = \frac{M_{i,j}}{\sum_{i,j=0}^{N-1} (M_{i,j})} \quad (2.14)$$

onde

- $i$  é o número da linha e  $j$  o número da coluna
- $M$  é o conteúdo da célula  $(i, j)$
- $P_{i,j}$  é a probabilidade da célula  $(i, j)$
- $N$  é o número de linhas ou colunas, uma vez que  $M$  é uma matriz quadrada

Desta forma, a matriz normalizada é obtida pela divisão de cada posição da matriz pelo total de seus valores não nulos. A partir das matrizes obtidas neste exemplo, a matriz de direção  $q = 0^\circ$  normalizada é dada a seguir:

	<b>0</b>	<b>1</b>	<b>2</b>
<b>0</b>	0,2	0,2	0,05
<b>1</b>	0,2	0,15	0,05
<b>2</b>	0,05	0,05	0,05

### 2.4.3 Propriedades da Matriz de Co-ocorrência

Algumas propriedades importantes da matriz de co-ocorrência são [20]:

1. Toda matriz de co-ocorrência é uma matriz quadrada.
2. O número de linhas e de colunas é igual ao nível de quantização da imagem.

Essas duas primeiras propriedades ficam evidentes através da construção de uma matriz de co-ocorrência. Como visto anteriormente, para construir uma matriz de co-ocorrência

é preciso inicialmente identificar o número de níveis de cinza presentes na imagem. A matriz de co-ocorrência irá conter um número de linhas e de colunas de acordo com o número de níveis de cinza da imagem. Desta forma, se a imagem possui  $k$  níveis de cinza, sua matriz de co-ocorrência possui  $k$  linhas e  $k$  colunas, sendo portanto uma matriz  $k \times k$ , ou seja, uma matriz quadrada.

3. A matriz de co-ocorrência é simétrica ao redor da diagonal.

## 2.5 Classificadores de Imagens

As próximas seções descrevem os dois classificadores utilizados neste trabalho para rotular as classes presentes nas imagens, as redes neurais artificiais e o classificador da distância mínima.

### 2.5.1 Redes Neurais Artificiais

As *redes neurais artificiais* (RNA) são técnicas computacionais que apresentam um modelo matemático inspirado na estrutura neuronal de organismos inteligentes e que adquirem conhecimento através da experiência.

As primeiras informações a respeito da neuro-computação surgiram em 1943, em artigos de McCulloch e Pitts [37], os quais sugeriam a construção de uma máquina inspirada no cérebro humano.

Máquinas baseadas nessas idéias foram sendo construídas e, devido a um entusiasmo exagerado, muitos livros e artigos foram publicados com previsões pouco confiáveis para a época sobre uma máquina que fosse tão poderosa quanto o cérebro humano. Nos anos 80, pesquisadores publicaram propostas para exploração de redes neurais, sugerindo aplicações, fazendo com que esse ramo da Inteligência Artificial passasse a ser utilizado com sucesso em diversas aplicações, dentre elas a análise e processamento de sinais, controle de processos, robótica, classificação de dados, reconhecimento de padrões em linhas de montagem, filtros para atenuação de ruídos eletrônicos, análise de imagens, análise de voz, avaliação de crédito, análise de aroma e odor e análise de reconhecimento de padrões acústicos.

Para representar RNA, utiliza-se um grafo direcionado onde os nós são denominados *elementos de processamento*, ou seja, representam os neurônios. A transmissão de sinais de um neurônio para outro é representada por meio das arestas desse grafo.

## Neurônio Artificial de McCulloch-Pitts

Como mencionado na seção 2.5.1, no ano de 1943, McCulloch e Pitts [37] desenvolveram um modelo computacional que possuía elementos abstraídos das propriedades fisiológicas de neurônios biológicos e de suas conexões, denominado neurônio de McCulloch-Pitts.

O neurônio de McCulloch-Pitts é um dispositivo binário, ou seja, possui apenas um tipo de saída: pulso ou não pulso. Porém, possui diversas entradas, simulando os diversos estímulos recebidos por um neurônio biológico. Para cada entrada de um neurônio artificial é associado um valor, denominado *peso*, que define um caráter excitatório ou inibitório para o impulso [1].

O resultado de um estímulo é determinado pelo cálculo da soma ponderada das entradas com seus respectivos pesos. Se a soma for maior ou igual a um limiar definido, o neurônio torna-se ativo, ou seja, um pulso é transmitido; caso contrário o neurônio permanece inativo. A representação de um neurônio artificial de McCulloch-Pitts é apresentada na figura 2.6.

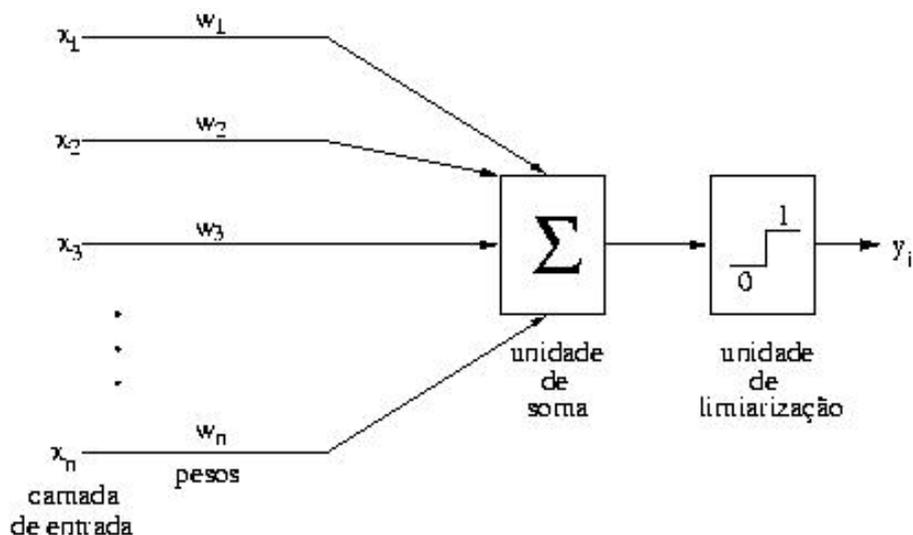


Figura 2.6: Neurônio artificial.

Dessa forma, tem-se o seguinte [1]:

1. Estímulos  $x = [x_1, x_2, \dots, x_n]$  chegam ao neurônio;

2. Cada estímulo é multiplicado pelo seu respectivo peso sináptico pré-definido, podendo inibi-lo ( $w < 0$ ) ou excitá-lo ( $w > 0$ ), onde  $w = [w_1, w_2, \dots, w_n]$ ;
3. É feita a integração sináptica, ou seja, a soma ponderada dos estímulos que produzem um nível de atividade ( $y$ );
4. Se este nível de atividade atingir ou exceder um certo limiar (*threshold*), a unidade produz uma determinada resposta de saída ( $f$ ).

Esse comportamento pode ser descrito pela expressão:

$$y = \sum_{i=1}^n w_i x_i \quad (2.15)$$

em que  $y$  é o nível de atividade interna produzida pelos estímulos,  $x_i$  é o conjunto de entradas e  $w_i$  são os pesos do neurônio.

## Redes Perceptrons

O *perceptron* é uma rede de múltiplos neurônios para discriminação linear, o qual foi criado por Rosenblatt [50] no final da década de 1950 inspirado em um modelo de retina, dando continuidade ao trabalho de McCulloch e Pitts [37].

Na implementação das funções Booleanas de duas variáveis utilizando o neurônio de McCulloch-Pitts, podem-se determinar pesos sinápticos e o limiar das funções Booleanas. Esses neurônios poderiam realizar uma operação ou-exclusivo (XOR). Entretanto, seriam incapazes de aprendê-la. Para problemas com um grande número de variáveis, a implementação McCulloch-Pitts torna-se inviável; dessa forma, Rosenblatt [50] propôs um modelo matemático capaz de aprender através de exemplos, ou seja, os pesos de suas conexões são ajustados de acordo com os padrões apresentados à rede e às suas respostas esperadas.

As redes perceptrons são organizadas em camadas. Um exemplo é apresentado na figura 2.7.

- Camada de entrada: os padrões são apresentados à rede.

- Camada intermediária ou oculta: onde é realizada a maior parte do processamento, por meio das conexões ponderadas. São os neurônios da camada interna que tornam possível a resolução de problemas lineares não separáveis.
- Camada de saída: o resultado final é apresentado.

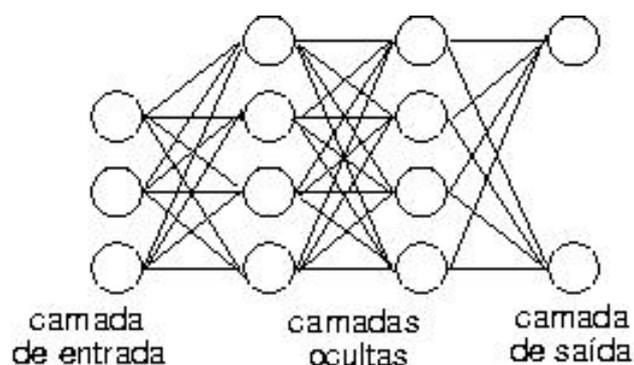


Figura 2.7: Camadas de uma rede neural artificial.

## Backpropagation

*Backpropagation* ou retropropagação é um algoritmo que se baseia no modelo de aprendizado supervisionado, retropropagando os erros da camada de saída para o treinamento dos pesos das camadas intermediárias [15].

Durante o treinamento com o algoritmo *backpropagation*, a rede opera em uma seqüência de dois passos. Primeiro, um padrão é apresentado à camada de entrada da rede. A atividade resultante flui através da rede, camada por camada, até que a resposta seja produzida pela camada de saída. No segundo passo, a saída obtida é comparada à saída desejada para esse padrão particular. Se esta não estiver correta, o erro é calculado. O erro é propagado a partir da camada de saída até a camada de entrada, e os pesos das conexões das unidades das camadas internas vão sendo modificados conforme o erro é retropropagado [15].

O treinamento das redes neurais com *backpropagation* pode demandar muitos passos no conjunto de treinamento, resultando um tempo de treinamento consideravelmente longo. Se for encontrado um mínimo local, o erro para o conjunto de treinamento pára de diminuir e estaciona em um valor maior que o aceitável.

## Aprendizado em uma Rede Neural

A propriedade mais importante das redes neurais é a habilidade de aprender com seu ambiente e com isso melhorar seu desempenho. Existem dois tipos de aprendizado:

- **Aprendizado supervisionado:** utiliza um agente externo que indica à rede um comportamento bom ou ruim de acordo com o padrão de entrada.
- **Aprendizado não-supervisionado:** não utiliza um agente externo indicando a resposta desejada para os padrões de entrada. Utiliza-se, entretanto, de exemplos de problemas semelhantes para que a rede responda de maneira semelhante.

Para realizar o aprendizado em uma rede neural, antes de mais nada, é preciso selecionar os dados que serão utilizados neste processo. Os dados coletados são separados em duas categorias: dados de treinamento, que, como o próprio nome diz, serão dados utilizados para treinar a rede e dados de testes.

Posteriormente, seguem-se algumas etapas [61]:

### 1. Configuração da Rede

É preciso selecionar um paradigma neural apropriado à aplicação e determinar a topologia de rede a ser utilizada (número de camadas, número de neurônios em cada camada). Depois disso, determinam-se os parâmetros do algoritmo de treinamento e a função de ativação. Esses dois últimos itens possuem um grande impacto no desempenho da rede. Uma configuração adequada requer grande experiência do projetista.

### 2. Treinamento

É nesta fase que a rede neural, através da modificação dos pesos das conexões, irá realizar a aprendizagem.

O treinamento deve ser interrompido quando a rede apresentar uma boa capacidade de generalização e quando a taxa de erro for suficientemente pequena, ou seja, menor que o erro admissível. Desta forma, o ideal é encontrar um ponto ótimo de parada com erro mínimo e capacidade de generalização máxima. Nesta fase é importante considerar

alguns aspectos como a inicialização da rede, o modo de treinamento e o tempo do treinamento.

### 3. Teste

Nesta fase, o conjunto de teste é inserido na rede neural para verificar se os dados de treinamento fizeram com que a rede aprendesse corretamente o que lhe foi apresentado.

O desempenho da rede medido nesta fase é uma boa indicação de seu desempenho real.

## Vantagens da Utilização de Redes Neurais Artificiais

Algumas vantagens da utilização de redes neurais artificiais são citadas a seguir [18]:

- *habilidade de tratar sistemas não-lineares*: relevante nas aplicações em identificação de sistemas dinâmicos e classificação de padrões;
- *tolerância a falhas*: o conhecimento é distribuído pela RNA, mais que em uma simples localização de memória. Uma parte das conexões pode estar inoperante, sem mudanças significativas no desempenho de toda a RNA;
- *adaptabilidade*: capacidade da RNA em se auto-ajustar. Os aspectos de aprendizado, auto-organização, generalização e treinamento estão intrinsecamente ligados a esta característica;
- *aprendizado*: uma RNA pode modificar seu comportamento em resposta ao ambiente. Quando é apresentado um conjunto de entradas, as RNAs se ajustam para gerar as respostas apropriadas;
- *generalização*: consiste na RNA mapear entradas similares em saídas similares;
- *treinamento*: é a forma pela qual a RNA aprende;
- *processamento paralelo*: as RNAs são estruturalmente paralelas. A seqüência de processamento das RNs é realizada em paralelo e simultaneamente;
- *abstração*: muitas RNAs são capazes de abstrair a essência de um conjunto de entradas.

## 2.5.2 Classificador de Distância Mínima

Dada uma amostra desconhecida  $x$ , representada pelo vetor de características  $v$ , o classificador de distância mínima visa encontrar a amostra  $p_i$ , pertencente ao conjunto de treinamento  $P = \{p_1, \dots, p_N\}$ , que apresenta menor distância de  $x$ , atribuindo essa amostra à mesma classe de  $p_i$ .

Com o intuito de reduzir a taxa de erro de classificação, pode-se utilizar os  $k$  vizinhos mais próximos ao invés de apenas um. Dessa maneira, atribui-se a amostra  $x$  à classe que apresenta maior número de vizinhos mais próximos dentre os  $k$  selecionados a partir do conjunto de treinamento.

A figura 2.8 ilustra a classificação de quatro pixels, representados por  $p_1, p_2, p_3$  e  $p_4$ , em três classes  $A, B$  e  $C$  utilizando o classificador da distância mínima.

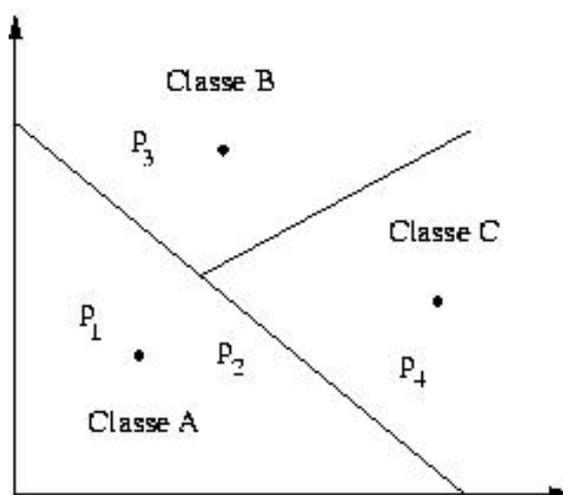


Figura 2.8: Classificador da distância mínima.

Normalmente, utiliza-se a distância Euclidiana para calcular a proximidade entre a amostra e as classes. O classificador compara a distância Euclidiana da amostra com a média de cada agrupamento, tal que a amostra é incorporada ao agrupamento que apresenta a menor distância Euclidiana. A média é recalculada levando-se em conta a nova amostra. Esse procedimento é repetido até que toda a imagem seja classificada.

Caso a amostra esteja distante de todas as classes com relação a um certo limiar, ela será classificada como desconhecida ou uma nova classe será formada. Este método é matematicamente simples e computacionalmente eficiente, tanto na fase de treinamento quanto na

fase de teste.

Os círculos pretos representam o centróide (média) de cada classe. O resultado para este exemplo é dado na tabela 2.4 e a classe escolhida é determinada pela menor distância entre o pixel e o centro da classe.

Pixel	$p_1$	$p_2$	$p_3$	$p_4$
Classe	A	A	B	C

Tabela 2.4: Resultado da classificação pela distância mínima.

## CAPÍTULO 3

### MÉTODO PROPOSTO

O desenvolvimento de uma metodologia para análise eficiente de imagens é uma tarefa complexa e, embora diversas técnicas tenham sido propostas, não há um método genérico que apresente bom desempenho em diferentes domínios de aplicação.

Este capítulo descreve o método desenvolvido para segmentação de imagens utilizando descritores estatísticos de textura. Os descritores, que formam o vetor de características, são extraídos por meio de matrizes de co-ocorrência calculadas a partir de imagens monocromáticas e coloridas. Os descritores de textura são submetidos a um classificador supervisionado para gerar a imagem classificada. Um diagrama ilustrando as principais etapas do método proposto é mostrado na figura 3.1.

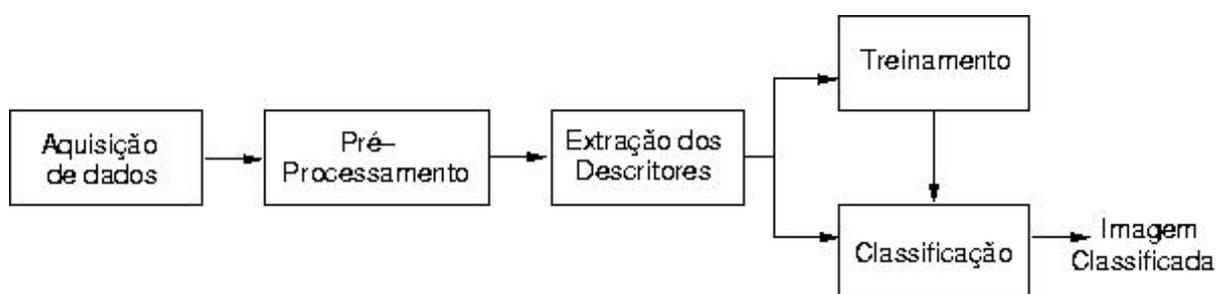


Figura 3.1: Etapas do método proposto.

Cada uma dessas etapas é descrita em mais detalhes a seguir:

- *aquisição de dados*: seleção das imagens a serem segmentadas e identificação de quantas e quais classes serão consideradas no processo de classificação.
- *pré-processamento*: preparação das imagens em formato adequado (de acordo com o Apêndice A) e definição das amostras de textura de acordo com as classes identificadas na etapa anterior. Uma análise visual é realizada para identificar as regiões da imagem que melhor representam as classes. Em seguida, são recortadas porções (de mesmo tamanho) da imagem que serão salvas como as imagens de amostras de texturas.

- *extração dos descritores*: utilizando as imagens selecionadas na etapa de aquisição de dados (imagem original) e pré-processamento (amostras de textura), os cálculos das matrizes de co-ocorrência são realizados para geração dos arquivos de treinamento e classificação. Por meio de um programa desenvolvido neste trabalho, diversos parâmetros são considerados e avaliados como medidas de características, modo de normalização das medidas, tamanho das janelas de treinamento e classificação, número de iterações para percorrer a imagem de amostra, modo como a imagem é percorrida, tipo de imagem (monocromática ou colorida), tipo do classificador (rede neural ou de distância mínima). Apesar da importância da seleção das características, não há regras ou procedimentos definitivos para esta tarefa em uma aplicação particular, principalmente quando o número de características disponíveis for grande [17]. Desta forma, visando principalmente à redução da dimensionalidade do vetor de características, utilizou-se o método de Análise de Componentes Principais (já descrito no capítulo 2) para identificar quais características seriam utilizadas na segmentação.

O número de características a serem utilizadas (informação que não é fornecida pela PCA) foi determinado por experimentos, que identificou cinco características por estratégia simples de tentativa e erro. Dessa forma, definiu-se um conjunto de cinco características como sendo o número mínimo aceitável para o sucesso da segmentação. Como será apresentado no capítulo 4, foram realizados testes com outros conjuntos de características para comparação dos resultados.

O tamanho das janelas de treinamento e classificação, o modo como a imagem é percorrida e o tipo do classificador foram parâmetros amplamente variados durante os testes (sem nenhum método de seleção específico), de forma a permitir uma análise da influência desses parâmetros no processo de segmentação. Padronizou-se em trinta o número de iterações em todos os experimentos. Esse número foi identificado nos primeiros testes realizados.

Os dados são coletados das imagens de maneiras diferentes para geração dos conjuntos de treinamento e classificação. Para gerar o conjunto de treinamento, as imagens de

amostras de textura são percorridas de acordo com o número de vezes indicado pelo número de iterações, conforme o tipo de percurso (seqüencial, cascata ou aleatório) escolhida para percorrer esta imagem. Para gerar o conjunto de classificação, a imagem original é percorrida pixel-a-pixel utilizando sempre a maneira seqüencial.

- **treinamento:** nessa etapa, o conjunto de treinamento, gerado na etapa anterior, é armazenado em um arquivo adequado ao tipo de classificador. Cada amostra de textura possui seu conjunto respectivo de descritores. Nos testes realizados, utilizaram-se dois tipos de classificador, um baseado em redes neurais artificiais e outro baseado no cálculo de distância mínima. O simulador de redes neurais utilizados nos experimentos foi o JavaSNNS [58]. O número de nodos da camada de entrada é igual ao número de características selecionadas. A camada de saída deve conter o número de classes desejado, enquanto que a(s) camada(s) oculta(s) pode(m) ter um número variado de nodos, conforme o tipo de imagem a ser classificada. O tipo de rede utilizado em todos os experimentos foi o Backpropagation (já explicado na seção 2.5.1).
- **classificação:** com base nos descritores de textura extraídos e nas informações resultantes do aprendizado utilizando os dados do conjunto de treinamento, a classificação dos pixels da imagem original é finalmente realizada. Nessa etapa, cada pixel da imagem original é agrupado em sua classe correspondente, representada por uma cor específica. Os mesmos classificadores descritos na etapa de treinamento foram utilizados para rotular cada pixel da imagem de teste.

### 3.1 Percurso das Imagens

Como descrito na seção anterior, para gerar o arquivo de treinamento, as imagens de amostra devem ser percorridas para que sejam realizados os cálculos das matrizes de co-ocorrência.

Três maneiras diferentes foram propostas para percorrer as imagens de amostra de textura. Os pixels selecionados em cada uma das iterações será utilizado para posicionar o pixel central da janela de treinamento na imagem de amostra, coletando os dados para a realização dos cálculos.

Para explicar esses três tipos de percurso, a seguir são apresentados os algoritmos referentes a cada um deles. Nesses algoritmos são utilizadas as seguintes variáveis:  $l$  linha da imagem da amostra,  $c$  coluna da imagem da amostra,  $c_{max}$  é o último pixel que pode ser o pixel central na janela ( $c_{max} = \text{número de colunas da imagem da amostra} - (\text{tamanho da janela} + 1)$ ),  $i$  número de iterações selecionadas pelo usuário,  $j$  tamanho da janela de treinamento,  $c_{anterior}$  última coluna percorrida na imagem e  $contador$  utilizado para incrementar o número de iterações.

a) seqüencial

O algoritmo seqüencial é apresentado na figura 3.2.

```

l = 1
c = 1
contador = 0
Enquanto (contador <= i)
    Enquanto(c <= c_max)
        c = c + 1
    Fim_Enquanto
    l = l + 1
    c = 1
    contador = contador + 1
Fim_Enquanto

```

Figura 3.2: Algoritmo para percorrer a imagem de amostra na forma seqüencial.

Este é ilustrado pela figura 3.3. Em (a) pode ser vista a primeira iteração e em (b) a segunda. A imagem (c) mostra a janela da primeira iteração depois que toda a primeira linha foi percorrida, onde  $k \leq i$ . Vale ressaltar que, se  $i < c_{max}$ , a segunda linha da imagem não será percorrida.

Para criação do arquivo de classificação, independente da maneira de percorrer a imagem de amostra na etapa de treinamento, a imagem sempre é percorrida da maneira seqüencial. Isto porque os cálculos realizados para geração do conjunto de classificação são efetuados com os valores de cada um dos pixels da imagem. Dessa forma, se a imagem original tiver largura  $M$  e altura  $N$  pixels, o arquivo de classificação terá  $M \times N$  valores. A única diferença para o algoritmo apresentado na figura 3.2 é a condição de

término que ao invés de utilizar o número de iterações, verifica o tamanho da imagem original como condição de parada.

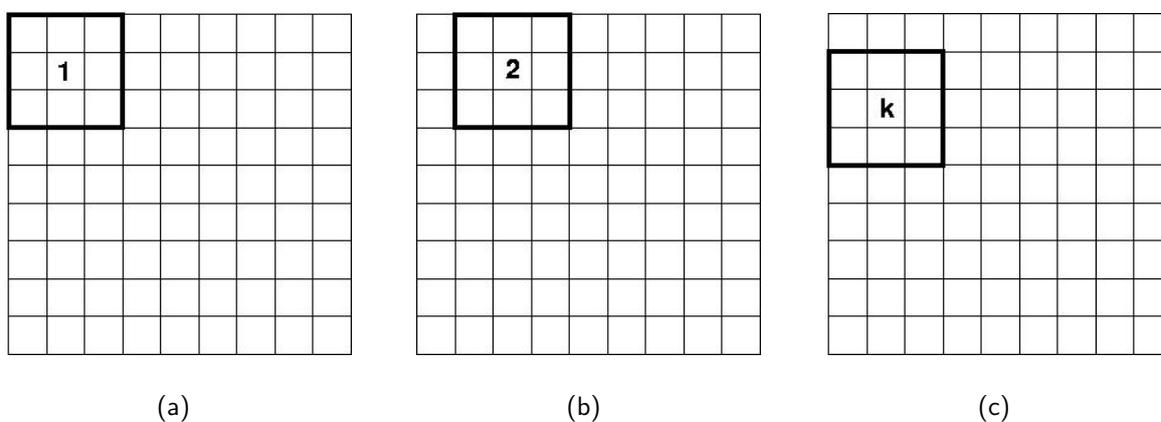


Figura 3.3: Varredura da imagem de maneira seqüencial.

b) cascata

```

j = definido pelo usuário
l = 1
c = 1
contador = 0
Enquanto (contador <= i)
    c_anterior = c
    Enquanto(c <= c_max)
        c = c + j
    Fim_Enquanto
    l = l + j
    c = (c_anterior + 1) mod j
    contador = contador + 1
Fim_Enquanto

```

Figura 3.4: Algoritmo para percorrer a imagem de amostra na forma cascata.

A figura 3.5 mostra os passos do algoritmo em cascata que é apresentado na figura 3.4. Em (a) pode ser vista a primeira iteração e em (b) a segunda. A imagem (c) mostra a janela da primeira iteração depois que toda a primeira linha foi percorrida, onde  $k \leq i$ . É importante lembrar que o algoritmo apresentado na figura 3.4 considera que um número máximo de iterações foi escolhido, ou seja, não está sendo considerado o fato da imagem da amostra de textura ser completamente percorrida na geração do arquivo de treinamento. Entretanto, basta trocar a condição de término de número de iterações

para as dimensões máximas da imagem que o algoritmo funcionaria de maneira análoga.

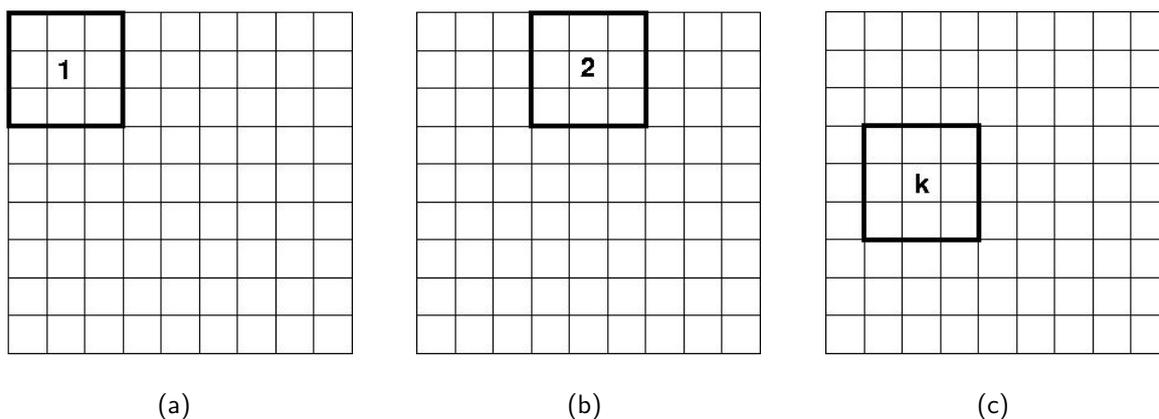


Figura 3.5: Varredura da imagem em cascata.

c) aleatório: os valores de  $l$  e  $c$  são escolhidos de maneira aleatória enquanto  $contador \leq i$ .

## 3.2 Características de Textura

Com o objetivo de extrair a informação contida nas matrizes de co-ocorrência de níveis de cinza, foram utilizadas diversas medidas estatísticas para avaliar as características texturais como homogeneidade, contraste, organização estrutural, complexidade e a natureza das transições das intensidades dos pixels [55]. As primeiras quinze medidas são aplicadas às imagens monocromáticas, enquanto as últimas quatro são aplicadas às imagens coloridas. Essas medidas foram baseadas nos trabalhos de Haralick *et al.* [22] e Dacheng *et al.* [14].

### • Segundo Momento Angular e Energia

Tanto o segundo momento angular (SMA) quanto a energia usam  $P_{i,j}$  como probabilidades. Valores elevados de SMA ou energia ocorrem quando a imagem está muito ordenada.

$$SMA = \sum_{i,j=0}^{N-1} P_{i,j}^2 \quad (3.1)$$

A raiz quadrada do SMA é também utilizada como medida de textura, e é chamada de energia.

$$Energia = \sqrt{SMA} \quad (3.2)$$

- **Contraste**

É uma estimativa das variações locais dos níveis de cinza entre pares de pixels. Essa medida é também chamada de soma do quadrado da variância.

Medidas relacionadas ao contraste usam pesos relacionados com a distância da diagonal da matriz de co-ocorrência (MC). Valores na diagonal da (MC) representam pouco contraste, aumentando quando a distância da diagonal aumenta.

$$\text{Contraste} = \sum_{i,j=0}^{N-1} P_{i,j}(i-j)^2 \quad (3.3)$$

Para elementos localizados na diagonal, ou seja, quando  $i$  e  $j$  são iguais, estes valores representam pixels totalmente iguais aos seus vizinhos, tendo um peso 0. Se  $i$  e  $j$  diferem em 1, existe uma pequena semelhança, e o peso é 1. Se  $i$  e  $j$  diferem em 2, o contraste aumenta e o peso é 4. Os pesos continuam a aumentar com o aumento do valor  $(i-j)$ .

- **Média**

A média na MC difere da média dos valores dos pixels na imagem. Na MC, essa medida não é simplesmente a média de todos os valores dos pixels originais na imagem. Ela é expressa em termos da matriz de co-ocorrência dos níveis de cinza. O valor do pixel é calculado não por sua frequência de ocorrência própria (como em uma equação da média regular), mas por sua frequência de ocorrência em combinação com um certo valor do pixel vizinho.

$$\mu_i = \sum_{i,j=0}^{N-1} i(P_{i,j}) \quad (3.4)$$

$$\mu_j = \sum_{i,j=0}^{N-1} j(P_{i,j}) \quad (3.5)$$

A equação 3.4 calcula a média baseada nos pixels de referência  $\mu_i$ . É possível também calcular a média usando-se os pixels vizinhos  $\mu_j$ , como na equação 3.5. Para a MC simétrica, onde cada pixel na imagem é contado uma vez como referência e outra como vizinho, os dois valores são idênticos.

- **Variância e Desvio Padrão**

A variância quando utilizada na textura executa a mesma tarefa que a variância utilizada na estatística descritiva. Ela baseia-se na dispersão dos dados numéricos em torno de um valor médio de valores dentro da MC. Entretanto, a variância da MC trata especificamente das combinações do pixel de referência e do seu vizinho. Então, isso não é o mesmo que a simples variância de níveis de cinza na imagem original.

$$\sigma_i^2 = \sum_{i,j=0}^{N-1} P_{i,j} (i - \mu_i)^2 \quad (3.6)$$

$$\sigma_j^2 = \sum_{i,j=0}^{N-1} P_{i,j} (j - \mu_j)^2 \quad (3.7)$$

A variância calculada usando  $i$  e  $j$  dá o mesmo resultado, pois a MC é simétrica. O desvio padrão é obtido tomando-se a raiz quadrada da variância, ou seja:

$$s_i = \sqrt{\sigma_i^2} \quad (3.8)$$

$$s_j = \sqrt{\sigma_j^2} \quad (3.9)$$

Algumas propriedades do desvio padrão, que resultam imediatamente da definição, são: o desvio padrão é sempre não negativo e será tanto maior quanto maior variabilidade houver entre os dados; se  $s = 0$ , então não existe variabilidade, isto é, os dados são todos iguais.

- **Correlação**

O coeficiente de correlação utilizado em estatística é uma medida do grau de associação linear (negativa ou positiva) entre duas variáveis quantitativas. A correlação em textura mede a dependência linear de pixels em relação a sua vizinhança.

$$\text{Correlação} = \sum_{i,j=0}^{N-1} P_{i,j} \left[ \frac{(i - \mu_i)(j - \mu_j)}{(s_i)(s_j)} \right] \quad (3.10)$$

- **Homogeneidade**

Os valores das probabilidades da homogeneidade são inversamente proporcionais as probabilidades do contraste, com pesos diminuindo quanto mais distantes da diagonal. Dissimilaridade e contraste resultam em grandes valores para imagens compostas principalmente por níveis de cinza diferentes (por exemplo, brancos e pretos), enquanto que a homogeneidade resulta em grandes valores para níveis de cinza similares. Se os valores diminuem distantes da diagonal, o resultado será maior para imagens com pouco contraste.

$$\text{Homogeneidade} = \sum_{i,j=0}^{N-1} \frac{P_{i,j}}{1 + (i - j)^2} \quad (3.11)$$

- **Soma da Média**

$$\text{Soma da média} = \sum_{i=2}^{2N} iP_{x+y}(i) \quad (3.12)$$

onde

$$P_{x+y}(i) = \sum_{j,k;j+k=i} P(j, k) \quad (3.13)$$

- **Soma da Variância**

$$\text{Soma da variância} = \sum_{i=2}^{2N} \left( i + \sum_{j=2}^{2N} P_{x+y}(j) \log(P_{x+y}(j)) \right)^2 P_{x+y}(i) \quad (3.14)$$

- **Diferença da Variância**

$$\text{Diferença da variância} = \text{variância de } P_{x-y} \quad (3.15)$$

- **Entropia**

Como a função logarítmica não é definida para zero e valores negativos, assume-se que  $0 \cdot \log(0) = 0$ .

$$\text{Entropia} = - \sum_{i,j=0}^{N-1} \log(P_{i,j}) P_{i,j} \quad (3.16)$$

- **Soma da Entropia**

$$\text{Soma da entropia} = - \sum_{i=2}^{2N} \log(P_{x+y}(i)) P_{x+y}(i) \quad (3.17)$$

- **Diferença da Entropia**

$$\text{Diferença da entropia} = - \sum_{i=0}^{N-1} \log(P_{x-y}(i)) P_{x-y}(i) \quad (3.18)$$

- **Informação de Medida de Correlação 1**

$$\text{Medida de correlação 1} = \frac{HXY - HXY1}{\max(HX, HY)} \quad (3.19)$$

onde  $HX$  e  $HY$  são entropias de  $P_x$  e  $P_y$ , e

$$HXY = - \sum_{i,j} \log(P(i, j)) P(i, j)$$

$$HXY1 = - \sum_{i,j} \log(P_x(i) P_y(j)) P(i, j)$$

- **Informação de Medida de Correlação 2**

$$\text{Medida de correlação 2} = \sqrt{1 - e^{-2(HXY2 - HXY)}} \quad (3.20)$$

onde

$$HXY2 = - \sum_{i,j} P_x(i) p_y(j) \log(P_x(i) P_y(j)) \quad (3.21)$$

- **Dissimilaridade**

Utilizando-se das medidas relacionadas com a dissimilaridade, ao invés dos pesos aumentarem na proporção (0,1,4,9,...) com a distância da diagonal, eles aumentam linearmente (0,1,2,3,...).

$$\text{Dissimilaridade} = \sum_{i,j=0}^{N-1} P_{i,j} |i - j| \quad (3.22)$$

- **Tendência de Agrupamento**

Mede o agrupamento dos pixels que possuem valores de níveis de cinza similares.

$$\text{Tendência de Agrupamento} = \sum_{i,j=0}^{N-1} (i + j - 2\mu)^k P(i, j) \quad (3.23)$$

- **Inércia**

A inércia é uma medida de variação local dos níveis de cinza.

$$\text{Inércia} = \frac{1}{\sigma^3} \sum_{i=0}^{N-1} (i - j)^2 P(i, j) \quad (3.24)$$

- **Curtose**

A curtose mede a desvio da distribuição de intensidade com relação à distribuição normal.

A curtose da distribuição normal é igual a 3.

$$\text{Curtose} = \frac{1}{\sigma^4} \sum_{i=0}^{N-1} (i - \mu)^4 P(i) - 3 \quad (3.25)$$

- **Probabilidade Máxima**

Esta medida resulta no par de pixels que é mais predominante na imagem.

$$\text{Probabilidade Máxima} = \max P_{i,j} \quad (3.26)$$

- **Simetria**

A simetria caracteriza o grau de dispersão das intensidades ao redor da média. Se o valor de simetria é negativo, os dados estão mais concentrados à esquerda da média, enquanto que, se o valor é positivo, os dados estão mais concentrados à direita da média.

$$\text{Simetria} = \frac{1}{\sigma^3} \sum_{i=0}^{N-1} (i - \mu)^3 P(i) \quad (3.27)$$

### 3.3 Matrizes de Co-ocorrência 3D

Dacheng *et al.* [14] estenderam o conceito de matriz de co-ocorrência para o espaço tridimensional, tal que elas possam ser utilizadas para descrever texturas coloridas que estejam no espaço cromático HSI, quantizado, respectivamente, em 8, 4 e 4 tons para as bandas H, S e I.

A idéia central do método desenvolvido pelos autores é relacionar tanto os pixels vizinhos, como efetuado pela abordagem tradicional de matriz de co-ocorrência, quanto extrair o relacionamento existente entre as bandas do espaço HSI e, com isso, gerar matrizes de co-ocorrência tridimensionais que resumem essas transições.

A figura 3.6 ilustra como são consideradas as variações de intensidades em cada banda de cor.

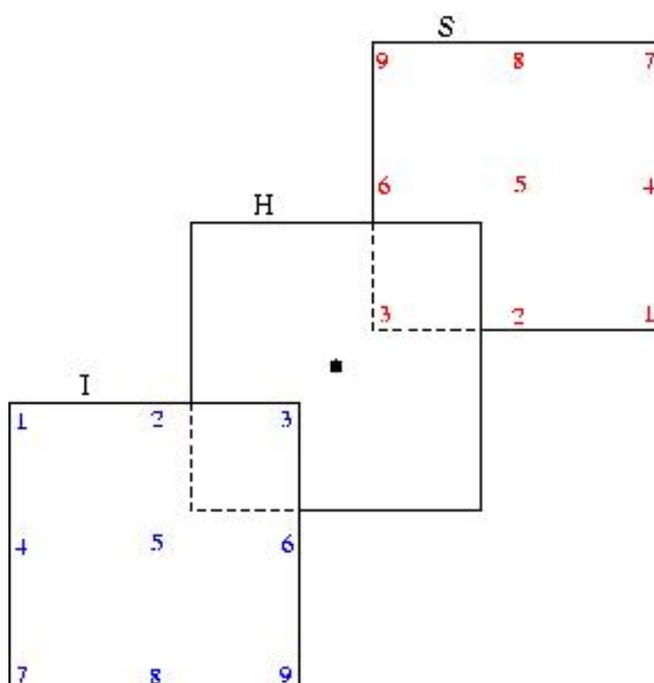


Figura 3.6: Relações entre as bandas de cores, onde os números iguais indicam as variações de tons consideradas, sempre passando pelo pixel central da banda H.

Conforme mostrado na figura 3.6, nove direções são consideradas para descrever as relações inter-canais e inter-bandas e, com isso, são criadas nove matrizes, uma para extrair as características de cada relacionamento. Dada a quantização apresentada anteriormente, essas matrizes apresentam dimensões  $8 \times 4 \times 4$  e o elemento  $(h, s, i)$  contém o número de transições que apresentam valor  $h$ ,  $s$  e  $i$ , nas bandas H, S e I, enquanto  $p_{h,s,i}$  denota a probabilidade de

ocorrer tal transição.

Dentre as características a serem extraídas dessas matrizes, estão quatro medidas calculadas para matriz de co-ocorrência bidimensional, entretanto, estendidas para três dimensões, sendo elas o segundo momento angular, contraste, correlação e a entropia, onde  $G_h$ ,  $G_s$ ,  $G_i$  denotam o valor do tom máximo em cada uma das bandas do espaço cromático HSI. Dessa maneira, a partir das nove matrizes são obtidas 36 características.

- **Segundo Momento Angular**

$$\text{Segundo Momento Angular} = \sum_{h=0}^{G_h} \sum_{s=0}^{G_s} \sum_{i=0}^{G_i} P_{h,s,i}^2 \quad (3.28)$$

- **Contraste**

$$\begin{aligned} \text{Contraste} = & \frac{1}{G_s} \sum_{k=0}^{G_s} \sum_{n=0}^{\max(G_h, G_i)} n^2 \sum_{|i-j|=n} P_{j,k,i} + \\ & \frac{1}{G_h} \sum_{j=0}^{G_h} \sum_{n=0}^{\max(G_s, G_i)} n^2 \sum_{|i-k|=n} P_{j,k,i} + \\ & \frac{1}{G_i} \sum_{i=0}^{G_i} \sum_{n=0}^{\max(G_s, G_h)} n^2 \sum_{|k-j|=n} P_{j,k,i} \end{aligned} \quad (3.29)$$

- **Correlação**

$$\text{Correlação} = \sum_{h=0}^{G_h} \sum_{s=0}^{G_s} \sum_{i=0}^{G_i} \frac{h s i p_{h,s,i} - \mu_h \mu_s \mu_i}{\sigma_h \sigma_s \sigma_i} \quad (3.30)$$

- **Entropia**

$$\text{Entropia} = - \sum_{h=0}^{G_h} \sum_{s=0}^{G_s} \sum_{i=0}^{G_i} \log(P_{h,s,i}) P_{h,s,i} \quad (3.31)$$

Idéia semelhante à matriz de co-ocorrência utilizada para descrever imagens coloridas proposta por Dacheng *et al.* é abordada no trabalho de Palm [46]. Em seu trabalho, Palm utiliza

o espaço cromático LUV e o relacionamento entre as bandas de cores é efetuado apenas duas a duas, ao invés de considerar as três bandas de maneira simultânea.

### 3.4 Normalização das Medidas

Um aspecto importante da segmentação, identificado por meio dos experimentos, é a necessidade de normalização das medidas extraídas a partir das matrizes de co-ocorrência. Como os intervalos de valores das características são, em geral, bastante distintos, uma característica não deveria predominar sobre as outras devido a essa variabilidade relativa.

Conforme a função de ativação utilizada na rede neural, ou seja, dependendo de como a rede combina as variáveis de entrada para gerar os valores que serão utilizados nas camadas subseqüentes (camadas ocultas e a camada de saída), o uso de intervalos diferentes para as características pode produzir resultados pouco satisfatórios.

A idéia inicial para evitar esse problema foi normalizar as características de modo que todas elas ficassem no intervalo [0..1]. Entretanto, embora essa normalização tenha gerado resultados melhores com relação às medidas não normalizadas, observou-se que algumas medidas (por exemplo, o contraste) perderam sua capacidade em discriminar certas texturas.

A forma escolhida para normalizar as características foi subtrair cada valor por uma medida de localização e dividir por um fator de escala [25]. Considerando o vetor de características como sendo formado por valores que seguem uma distribuição Gaussiana, a normalização foi realizada subtraindo-se cada valor pela média e dividindo-se pelo desvio padrão, ou seja

$$S_i = \frac{x_i - \mu}{\sigma} \quad (3.32)$$

onde  $\mu$  e  $\sigma$  são a média e o desvio padrão, respectivamente, das medidas para uma determinada característica.

## CAPÍTULO 4

### RESULTADOS EXPERIMENTAIS

Com a finalidade de avaliar a metodologia proposta, foram desenvolvidos experimentos utilizando um conjunto de imagens sintéticas e reais.

A plataforma de desenvolvimento utilizada neste trabalho foi um computador com processador AMD Athlon XP 2400+, 512 MBytes RAM, sistema operacional Linux versão 2.4.27. Os programas foram desenvolvidos em linguagem de programação C++.

Dentre as ferramentas auxiliares utilizadas estão o GIMP (*Gnu Image Manipulation Program*), versão 2.2 (*Spencer Kimball & Peter Mattis*); XV (*Image Viewer*), versão 3.10a (*John Bradley - 1994*); DISPLAY (*Image Magick*), versão 5.4.4 (*Image Magick Studio LLC - 2002*). O simulador de redes neurais utilizado foi o JavaNNS (*Java Neural Network Simulator*), versão 1.1 (*Igor Fischer, Favian Hennecke, Christian Bannes - 2002*). Para todos os experimentos foi utilizado o tipo de rede Backpropagation.

As próximas seções descrevem um conjunto de experimentos realizados e uma discussão dos resultados obtidos.

#### 4.1 Experimentos Realizados

Um grande conjunto de imagens sintéticas e reais foi utilizado nos experimentos para avaliar o método desenvolvido. Algumas imagens escolhidas apresentam detalhes finos e várias reentrâncias, tornando complexo o processo de segmentação. A seguir são apresentados treze experimentos realizados.

##### 4.1.1 Experimento 1

A figura 4.1(a) mostra uma imagem contendo um objeto simples. Em 4.1(b), são apresentadas as amostras que indicam, respectivamente, a textura do objeto e o fundo da imagem. As

dimensões das imagens são  $250 \times 250$  pixels para a imagem original e  $32 \times 32$  pixels para as amostras.

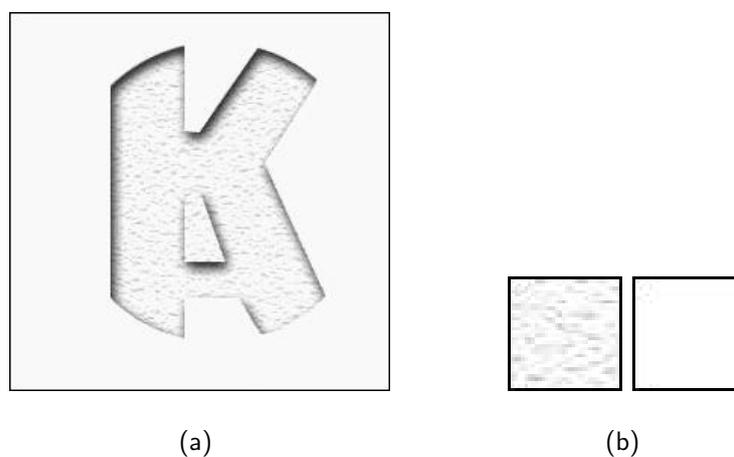


Figura 4.1: Experimento 1. (a) Imagem original e (b) amostras utilizadas.

Para todos os testes realizados com esta imagem foram utilizadas as características: segundo momento angular, contraste, desvio padrão, média e correlação, identificadas pelo método de Análise de Componentes Principais (PCA). A figura 4.2 apresenta os resultados. A diferença entre as imagens (a) até (d) reflete a influência do tamanho da janela, sendo utilizados, respectivamente,  $3 \times 3$ ,  $5 \times 5$ ,  $7 \times 7$  e  $15 \times 15$ . Todas as imagens foram percorridas de maneira seqüencial (algoritmo 3.2), utilizando o classificador SNNS, normalização de características e 30 iterações.

### 4.1.2 Experimento 2

A imagem apresentada na figura 4.3(a) contém um conjunto de grãos de feijão. Em 4.3(b) são apresentadas as amostras que indicam respectivamente os grãos de feijão e o fundo da imagem. As dimensões das imagens são  $179 \times 200$  pixels para imagem original e  $32 \times 32$  pixels para as amostras.

Para todos os testes realizados com esta imagem foram utilizadas 30 iterações e as características: segundo momento angular, contraste, desvio padrão, média e correlação, selecionadas por meio da Análise de Componentes Principais (PCA). A figura 4.4 apresenta os resultados. Em (a), (b) e (c) foi utilizada a janela  $3 \times 3$  e as imagens foram percorridas de maneira seqüencial, cascata e aleatório, respectivamente. O resultado mostrado em (d) foi

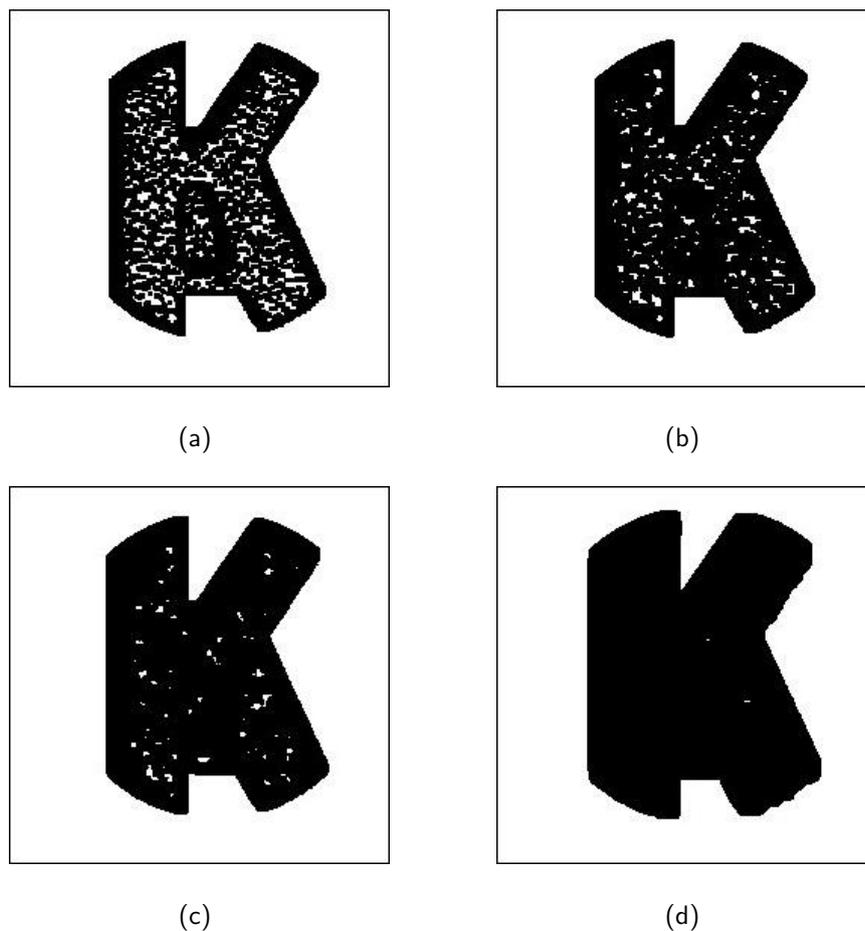


Figura 4.2: Resultados da segmentação da imagem letra. Parâmetros comuns: Características - identificadas pela PCA (segundo momento angular, contraste, desvio padrão, média e correlação), classificador Redes Neurais, Normalização, Seqüencial. (a) janela 3x3, (b) janela 5x5, (c) janela 7x7 e (d) janela 15x15.

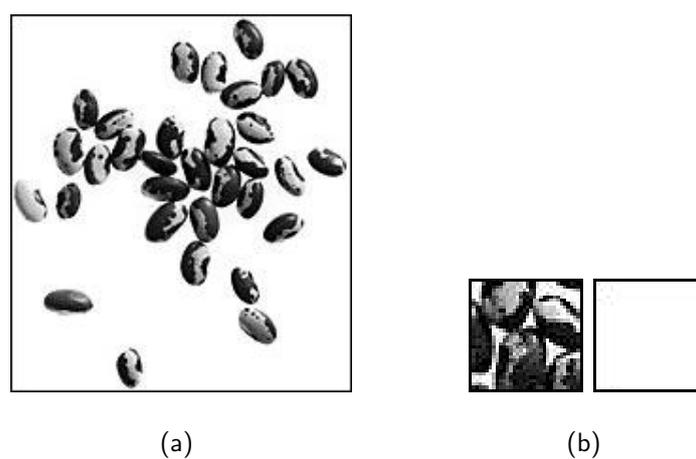


Figura 4.3: Experimento 2. (a) Imagem original e (b) amostras utilizadas.

obtido com janela  $15 \times 15$  de maneira seqüencial. O classificador SNNS e a normalização das medidas foram utilizadas nas imagens de (a) até (d). Em (e), apesar do classificador utilizado

também ter sido o SNNS, a normalização das medidas não foi utilizada. Os demais parâmetros de (e) foram os mesmos utilizados em (a).

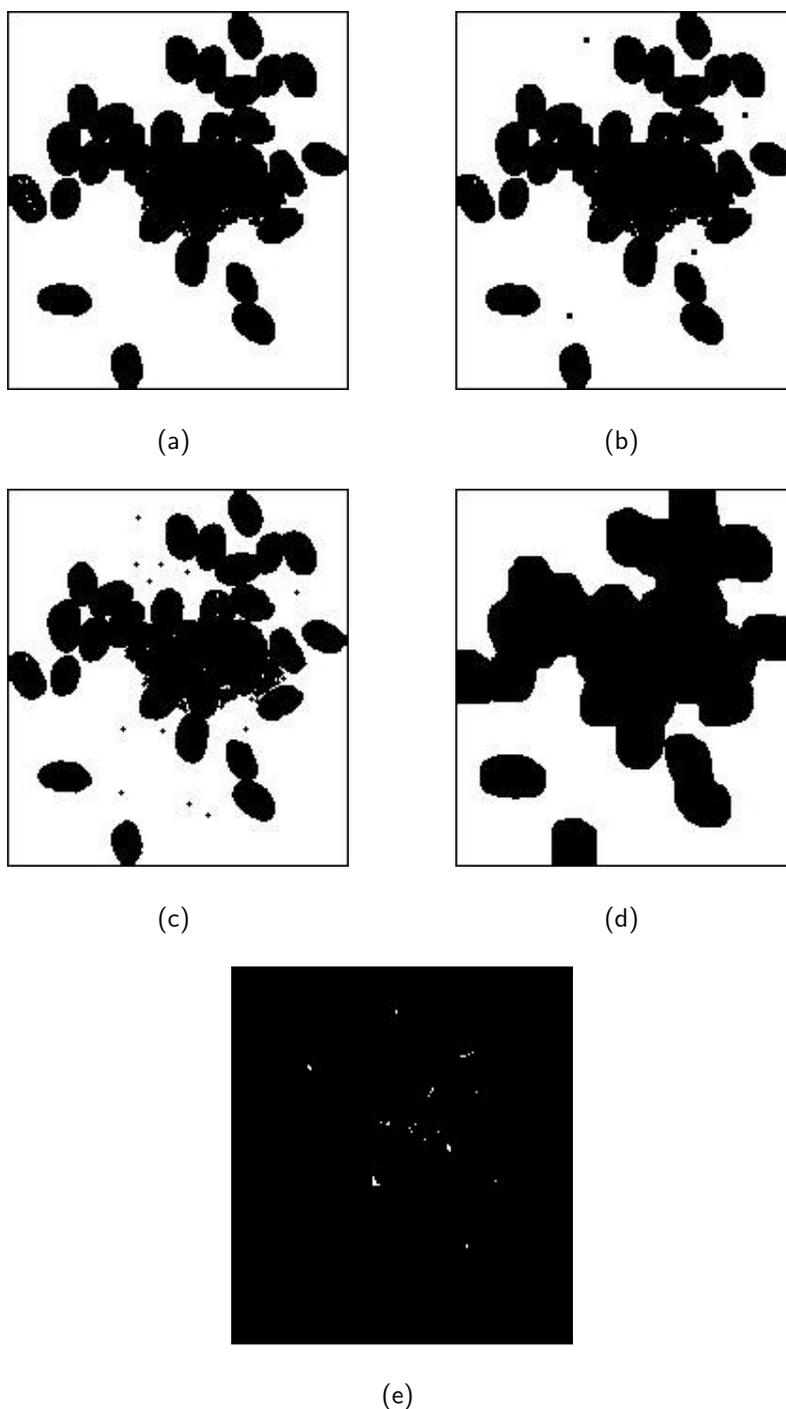


Figura 4.4: Resultados da segmentação da imagem beans. Parâmetros comuns: Características - identificadas pela PCA (segundo momento angular, contraste, desvio padrão, média e correlação) e classificador Redes Neurais. (a) janela 3x3 e seqüencial, (b) janela 3x3 e cascata, (c) janela 3x3 e aleatório, (d) janela 15x15 e seqüencial e (e) janela 3x3 e seqüencial. Em (e) não foi utilizada a Normalização.

### 4.1.3 Experimento 3

Para este experimento, a imagem original 4.3(a) foi corrompida por dois tipos de ruído: impulsivo (sal-e-pimenta) e Gaussiano. A figura 4.5 apresenta em (a) e (c) a imagem corrompida com ruído sal-e-pimenta e Gaussiano, respectivamente e, em (b) e (d) o resultado da segmentação. Para ambos os casos, foram utilizadas as características segundo momento angular, contraste, desvio padrão, média e correlação (identificadas pela PCA), janela  $3 \times 3$  pixels, a imagem foi percorrida pelo algoritmo seqüencial, 30 iterações e a normalização de características foi utilizada.

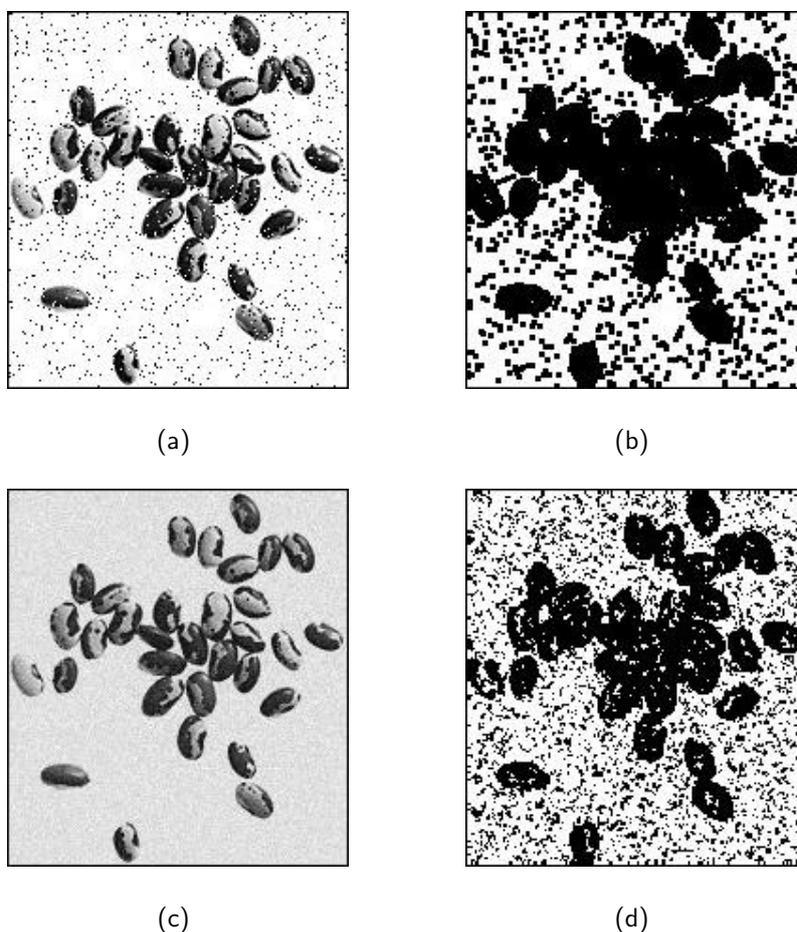


Figura 4.5: Resultados da classificação da imagem beans corrompida com ruído impulsivo sal-e-pimenta ((a) imagem original e (b) resultado da classificação) e com ruído Gaussiano ((c) imagem original e (d) resultado da classificação). Parâmetros comuns: Características - identificadas pela PCA (segundo momento angular, contraste, desvio padrão, média e correlação), classificador Redes Neurais, janela  $3 \times 3$ , seqüencial, com utilização da Normalização.

#### 4.1.4 Experimento 4

A imagem apresentada na figura 4.6(a) contém a união de duas texturas do álbum Brodatz [7], nomeadas D6 (fio de alumínio entrelaçado) e D57 (papel amassado), respectivamente. O tamanho original desta imagem é de  $512 \times 256$  pixels. Para todos os testes referentes a esta imagem foram coletadas duas amostras, sendo uma de cada textura. Estas são apresentadas na figura 4.6(b) e possuem  $64 \times 64$  pixels cada uma. O classificador utilizado foi o SNNS. O algoritmo utilizado para percorrer a imagem foi o cascata.

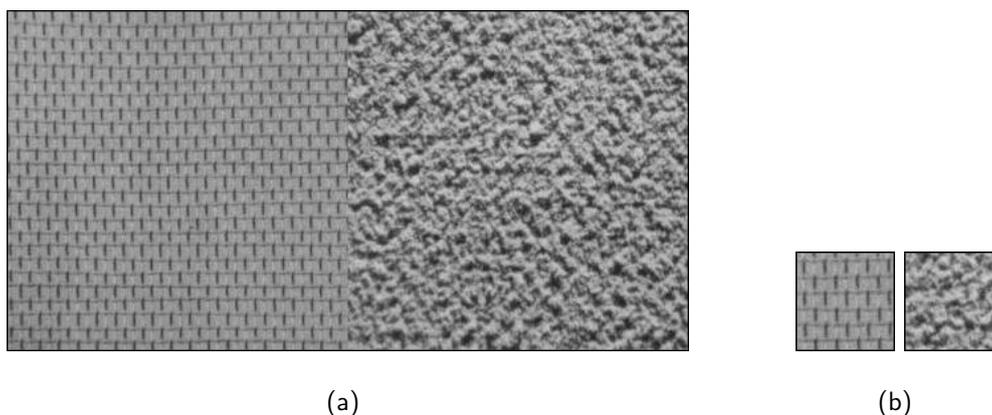


Figura 4.6: Experimento 4. (a) Imagem original e (b) amostras utilizadas.

Os resultados da segmentação dessa imagem, apresentados na figura 4.7, ilustram novamente a influência do tamanho da janela e as diferenças de se percorrer a imagem pelo algoritmo cascata utilizando um número limitado de iterações e percorrer em cascata utilizando toda a imagem. Em (a), foi utilizada uma janela  $5 \times 5$  pixels e 30 iterações. Em (b) e (c), a janela utilizada foi de  $15 \times 15$  pixels, com a diferença que em (b) utilizaram-se 30 iterações para coletar dados para o treinamento e em (c) utilizou-se a imagem toda para a coleta de dados para o treinamento.

Para melhorar o resultado em (c), foi utilizada a opção de normalização. Vale ressaltar que testes realizados com janela  $3 \times 3$  pixels não segmentaram a imagem, independente da variação de outros parâmetros. Para as imagens segmentadas de (a) a (c) foram utilizadas as seguintes características: segundo momento angular, contraste, desvio padrão, média e correlação, identificadas pela PCA. Para a segmentação da imagem (d), as características utilizadas foram: tendência de agrupamento, inércia, curtose, probabilidade máxima e simetria.

Em (d), os demais parâmetros foram os mesmos utilizados para (c).

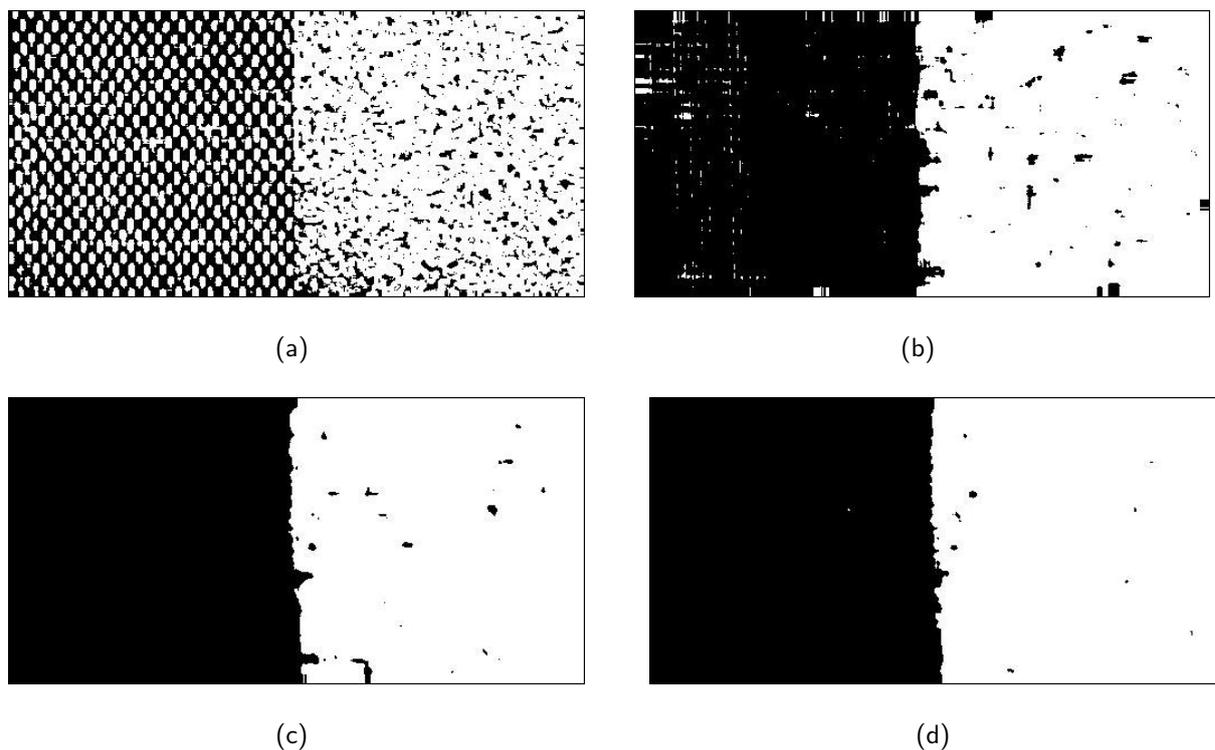


Figura 4.7: Resultados da segmentação da imagem mosaico 1. Parâmetros comuns: classificador Redes Neurais, cascata e normalização. Em (a), (b) e (c) - características identificadas pela PCA (segundo momento angular, contraste, desvio padrão, média e correlação). Em (d) as características foram: tendência de agrupamento, inércia, curtose, probabilidade máxima e simetria. (a) janela 5x5 e 30 iterações, (b) janela 15x15 e 30 iterações, (c) janela 15x15 percorrendo toda a imagem e (d) janela 15x15 e 30 iterações.

### 4.1.5 Experimento 5

A imagem de teste mostrada na figura 4.8(a) contém um mosaico com cinco texturas do álbum Brodatz [7]. Mosaicos são muito utilizados para testes de classificadores. A figura 4.8(b) apresenta as cinco amostras de textura desta imagem. Como se pode observar, algumas texturas são bastante parecidas o que aumenta a complexidade da classificação. A imagem original possui tamanho  $128 \times 128$  e as amostras  $32 \times 32$  pixels.

Na figura 4.9 são apresentadas as imagens de resultado da segmentação da imagem 4.8(a) com diversas variações de parâmetros. De (a) a (d) foram utilizadas as características indicadas pela PCA: segundo momento angular, contraste, desvio padrão, média e correlação. Em (e) utilizaram-se as características: tendência de agrupamento, inércia, curtose, probabilidade

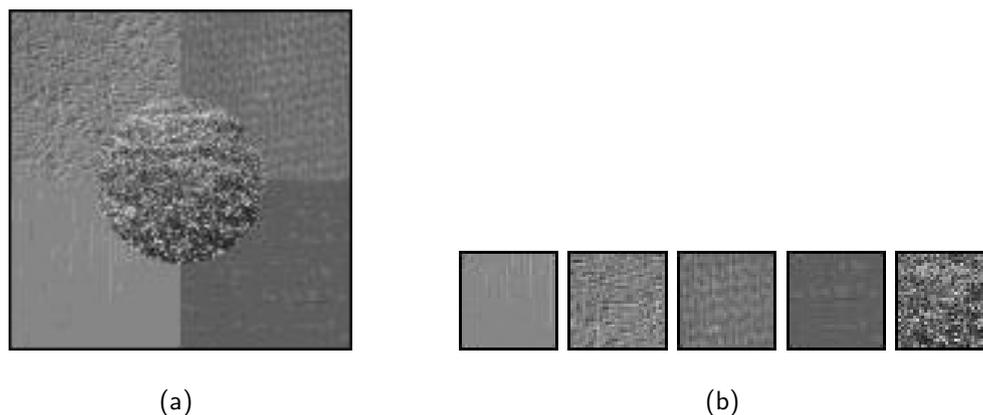


Figura 4.8: Experimento 5. (a) Imagem original e (b) amostras utilizadas.

máxima e simetria. Em (a), (d) e (e), a janela foi de  $15 \times 15$  pixels, em (b)  $5 \times 5$  pixels e em (c) utilizou-se janela de  $3 \times 3$  pixels. O classificador utilizado foi o SNNS, exceto em (d) que utilizou o classificador de distância mínima. Em todos os testes foi utilizada a normalização das características e 30 iterações.

#### 4.1.6 Experimento 6

A imagem de teste mostrada na figura 4.10(a), cujo tamanho é de  $432 \times 432$  pixels, corresponde à região chamada Kodiak Island, Alaska, Estados Unidos, capturada pelo satélite LANDSAT.

Na mesma figura, em (b), são ilustradas as duas amostras utilizadas em todos os testes com essa imagem. As amostras representam, respectivamente, água e continente, ambas com tamanho de  $55 \times 55$  pixels.

A figura 4.11 apresenta os resultados da segmentação da imagem 4.10. Para os testes de (a) até (e) foram utilizadas as características: segundo momento angular, contraste, desvio padrão, média e correlação, selecionadas por meio da Análise de Componentes Principais (PCA). Em (f), as características utilizadas foram a tendência de agrupamento, inércia, curtose, probabilidade máxima e simetria. O classificador SNNS e a normalização das medidas foram utilizados em todos os testes realizados.

Em (a) e (f) foi utilizada uma janela de treinamento  $3 \times 3$  pixels e a imagem foi percorrida de maneira seqüencial utilizando 30 iterações. O percurso aleatório foi utilizado na varredura das imagens em (b), também com janela  $3 \times 3$  pixels, 30 iterações. As imagens (c) e (d) foram

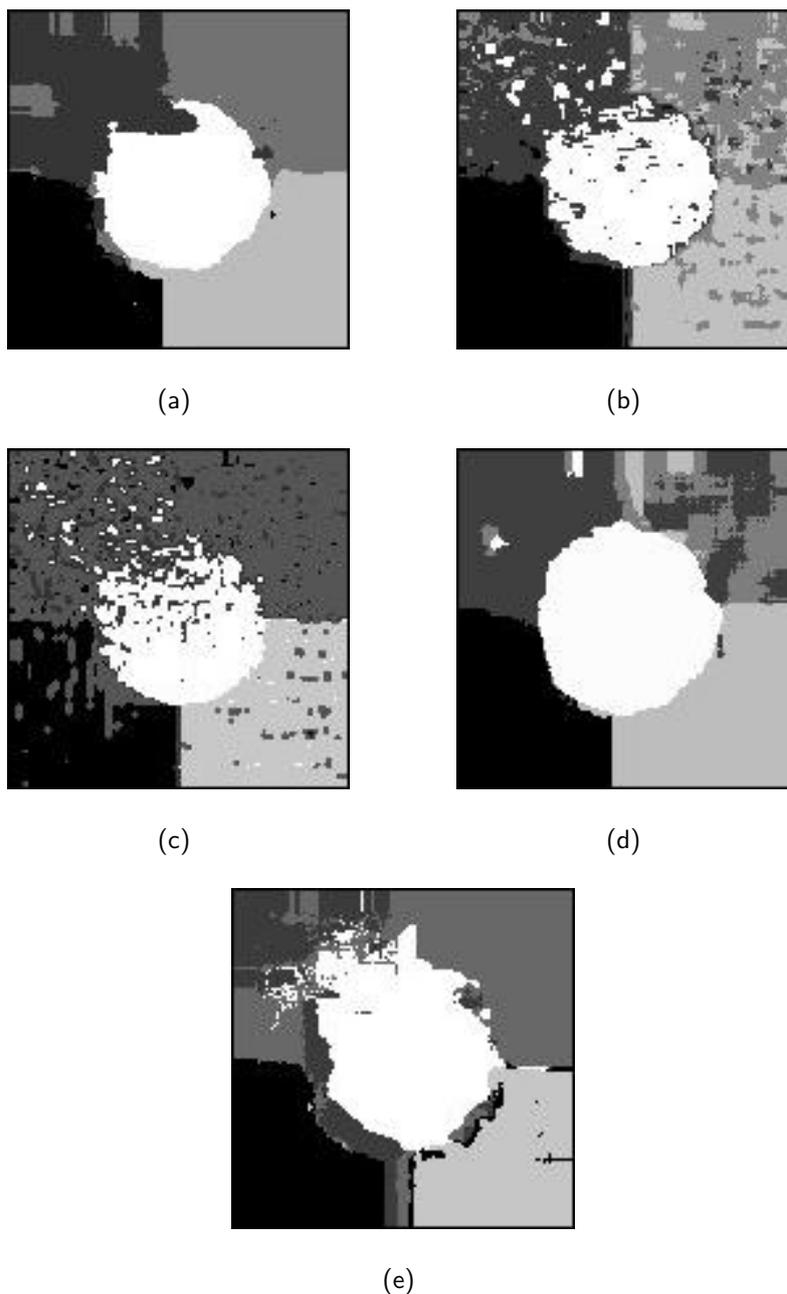


Figura 4.9: Resultados da classificação da imagem *mosaico 2*. Parâmetros comuns: seqüencial e normalização. De (a) até (d) - características identificadas pela PCA (segundo momento angular, contraste, desvio padrão, média e correlação). Em (e) as características foram: tendência de agrupamento, inércia, curtose, probabilidade máxima e simetria. (a) e (e) janela 15x15 e redes neurais, (b) janela 5x5 e redes neurais, (c) janela 3x3 e redes neurais e (d) janela 15x15 e distância mínima.

percorridas pelo algoritmo cascata mostrado na figura 3.4, ambas com janela  $3 \times 3$  pixels. A primeira mostra a imagem segmentada pelas amostras coletadas durante as 30 iterações, a segunda apresenta o resultado da imagem inteira sendo utilizada no treinamento. Em (e), pode-se identificar a alteração no resultado devido ao tamanho da janela. Para esta imagem,

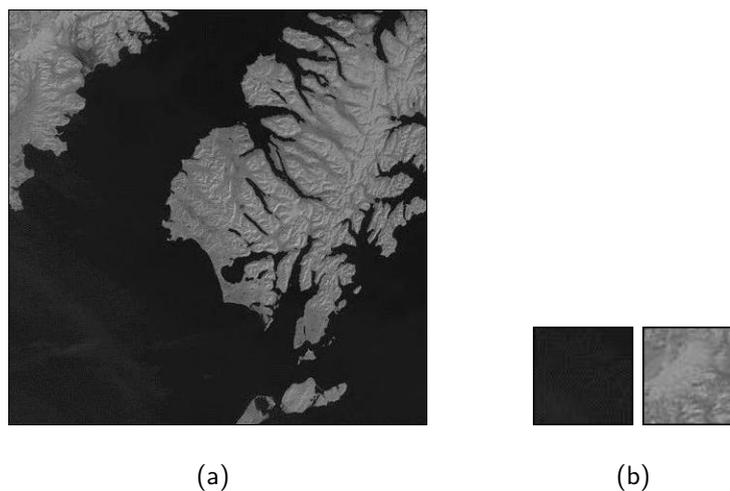


Figura 4.10: Experimento 6. (a) Imagem original e (b) amostras utilizadas.

os parâmetros utilizados foram o percurso seqüencial, 30 iterações e de janela  $15 \times 15$  pixels.

#### 4.1.7 Experimento 7

A figura 4.12(a) mostra uma imagem da região de Chesapeake Bay, Anápolis, MA, Estados Unidos, seu tamanho de  $320 \times 320$  pixels. Em (b) são apresentadas as amostras recortadas da imagem original (ambas com  $35 \times 35$  pixels) que foram utilizadas na segmentação. A primeira amostra representa a classe continente e a segunda representa a classe oceano. Para todos os testes foram utilizadas as características: segundo momento angular, contraste, desvio padrão, média e correlação, identificadas pela PCA. O classificador utilizado foi o SNNS.

A figura 4.13 apresenta o resultado dos testes realizados tendo como imagem de entrada a figura 4.12(a). Em (a) e (b) foi utilizada a janela  $3 \times 3$  pixels. Para (c) e (d) as janelas foram respectivamente  $5 \times 5$  pixels e  $15 \times 15$  pixels. Todos os resultados foram obtidos utilizando 30 iterações e as características sugeridas pela PCA (segundo momento angular, contraste, desvio padrão, média e correlação), foram percorridas de maneira seqüencial e utilizaram a normalização. Em (a), (c) e (d) o classificador utilizado foi o SNNS em (b) utilizou-se o classificador de distância mínima.

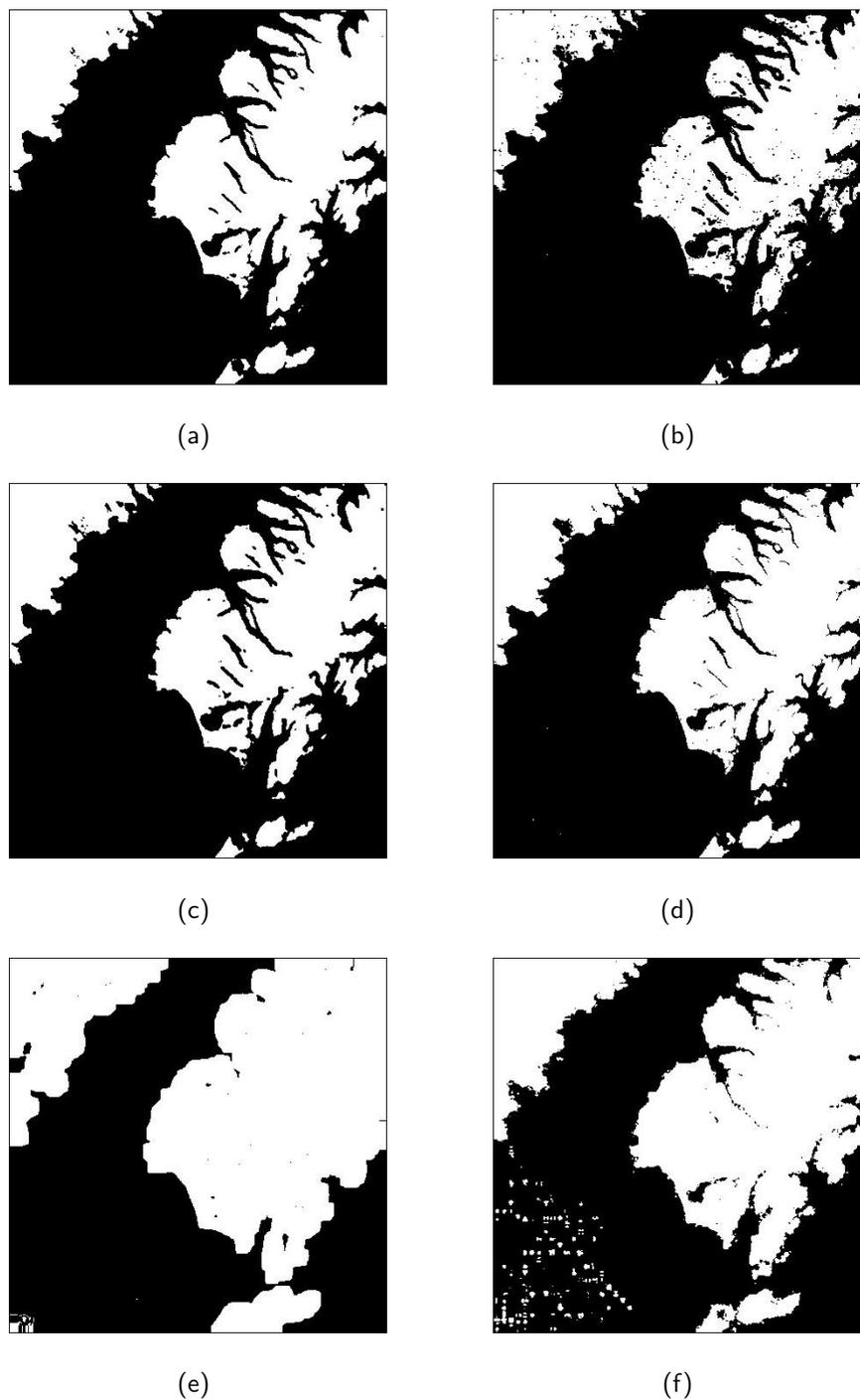


Figura 4.11: Resultados da segmentação da imagem Kodiak. Parâmetros comuns: classificador redes neurais e normalização. De (a) até (e) - características identificadas pela PCA (segundo momento angular, contraste, desvio padrão, média e correlação). Em (f) as características foram: tendência de agrupamento, inércia, curtose, probabilidade máxima e simetria. (a) e (f) janela 3x3 e seqüencial, (b) janela 3x3 e aleatório, (c) janela 3x3 e cascata com 30 iterações, (d) janela 3x3 e cascata percorrendo toda a imagem e (e) janela 15x15 e seqüencial.

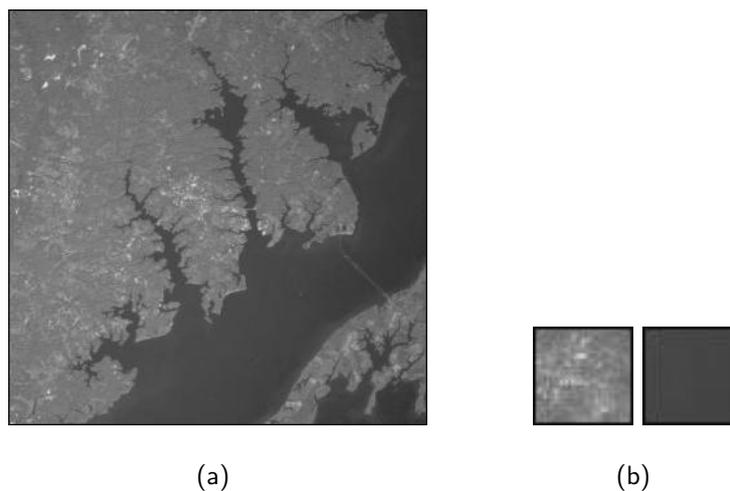


Figura 4.12: Experimento 7. (a) Imagem original e (b) amostras utilizadas.

#### 4.1.8 Experimento 8

Na figura 4.14(a) é apresentada uma imagem da região de Moreno Glacier, Argentina, sendo seu tamanho de  $340 \times 340$  pixels. Em (b), são apresentadas as amostras recortadas da imagem original, todas de tamanho  $30 \times 30$  pixels. As amostras indicam, respectivamente, vegetação, oceano e neve. Para todos os testes foram utilizadas as características: segundo momento angular, contraste, desvio padrão, média e correlação, identificadas pela PCA.

Para todos os testes as imagens foram percorridas de maneira seqüencial (com 30 iterações), utilizou-se a normalização e as características identificadas pela PCA (segundo momento angular, contraste, desvio padrão, média e correlação) foram novamente utilizadas. A figura 4.15 ilustra os resultados obtidos pela segmentação. Em (a) e (b) utilizou-se janela de tamanho  $3 \times 3$  pixels e em (c) janela de tamanho  $15 \times 15$  pixels. Em (a) e (c) o classificador utilizado foi o SNNS, em (b) o de distância mínima.

#### 4.1.9 Experimento 9

A figura 4.16(a) apresenta o recorte de uma imagem do Centro Politécnico da Universidade Federal do Paraná (Curitiba-PR), cedida pela empresa Esteio Engenharia e Aerolevantamentos S.A. Esse recorte possui tamanho de  $512 \times 504$  pixels. Em (b) são apresentadas as amostras recortadas da imagem original (ambas com  $60 \times 60$ ) que foram utilizadas na segmentação.

A primeira amostra representa a classe de áreas construídas (prédios, ruas, calçadas, etc)

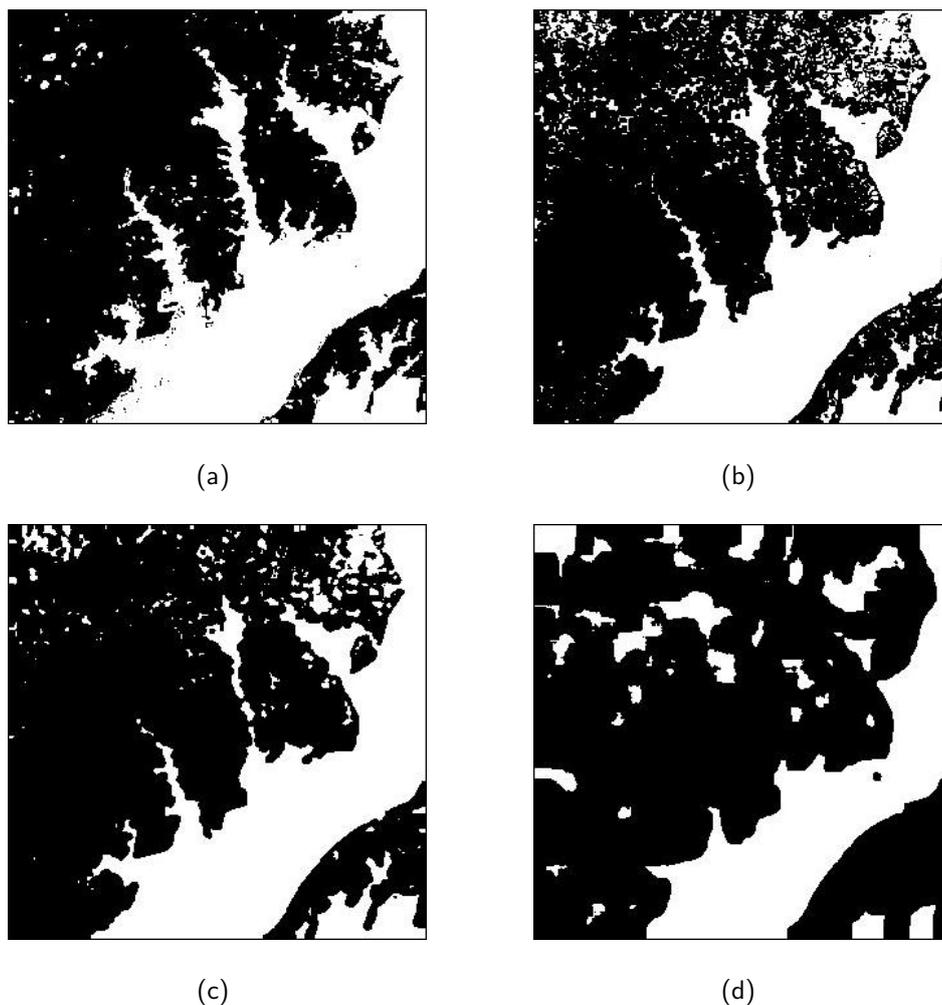


Figura 4.13: Resultados da segmentação da imagem Chesapeake. Parâmetros comuns: características identificadas pela PCA (segundo momento angular, contraste, desvio padrão, média e correlação), seqüencial e normalização. (a) janela 3x3 e redes neurais, (b) janela 3x3 e distância mínima, (c) janela 5x5 e redes neurais e (d) janela 15x15 e redes neurais.

e a segunda representa a classe de vegetação. Para todos os testes foram utilizadas as características: segundo momento angular, contraste, desvio padrão, média e correlação, identificadas pela PCA. O classificador utilizado foi o SNNS.

Na figura 4.17 são apresentados os resultados da segmentação da imagem 4.16. A imagem original é extremamente complexa e, dessa forma, as amostras recortadas podem não representar as classes como deveriam. Em (a), (b) e (c), utilizou-se janela 15x15 pixels, embora a maneira de percorrer a imagem seja diferente ((a) seqüencial, (b) cascata e (c) aleatório, todos com 30 iterações), pode-se observar que os resultados não foram satisfatórios. Em (d), (e) e (f), as janelas utilizadas são de tamanho 3x3 pixels, percorrendo a imagem com 30

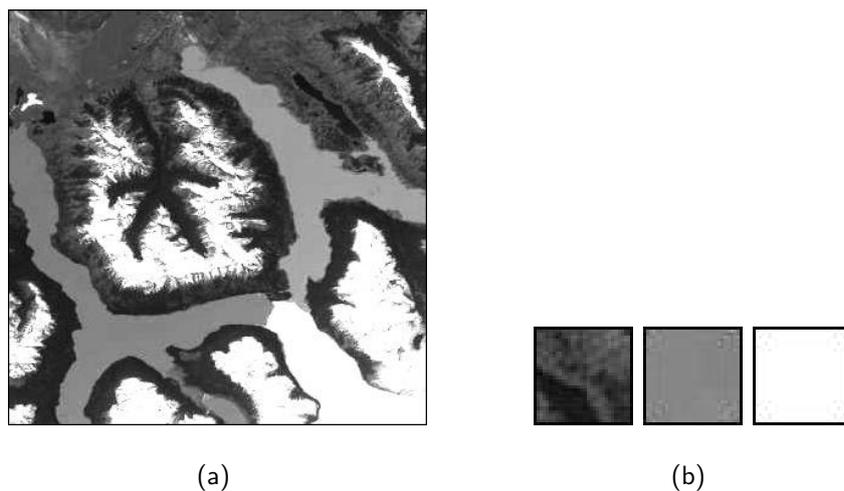


Figura 4.14: Experimento 8. (a) Imagem original e (b) amostras utilizadas.

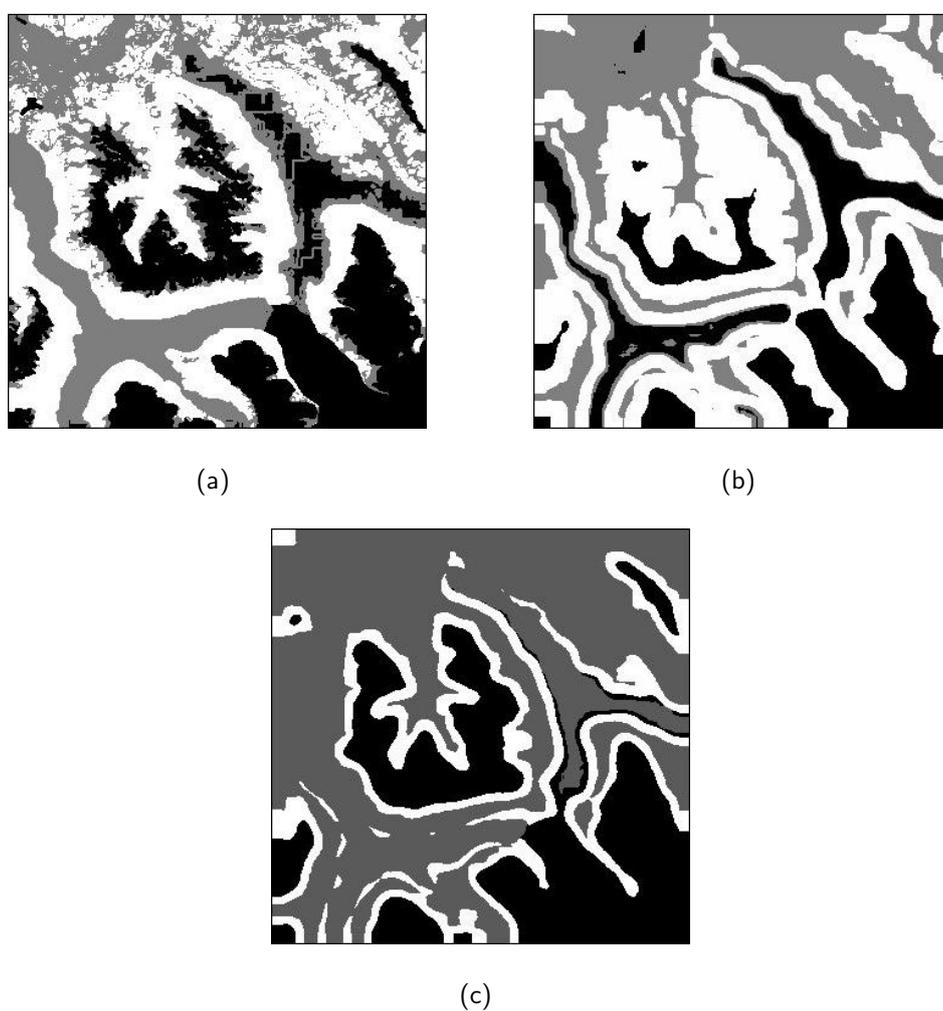


Figura 4.15: Resultados da segmentação da imagem Moreno Glacier. Parâmetros comuns: características identificadas pela PCA (segundo momento angular, contraste, desvio padrão, média e correlação), seqüencial e normalização. (a) janela 3x3 e redes neurais, (b) janela 3x3 e distância mínima e (c) janela 15x15 e redes neurais.

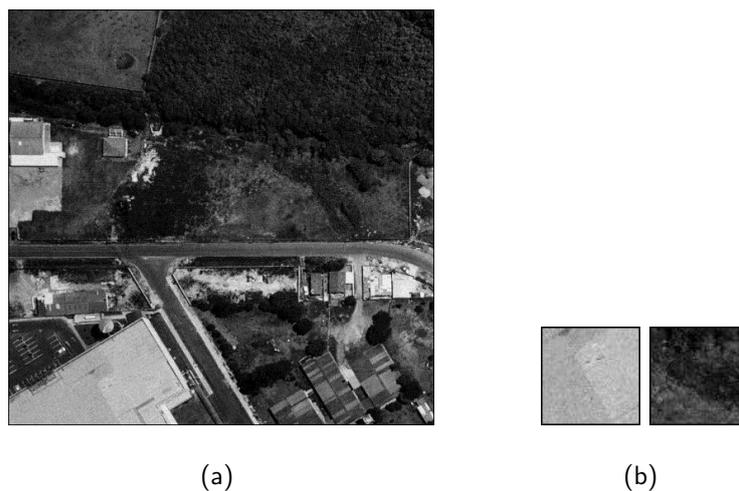


Figura 4.16: Experimento 9. (a) Imagem original e (b) amostras utilizadas

iterações e utilizando a normalização das características. Em (d), a imagem é percorrida de forma seqüencial, em (e) percorrida em cascata, ambas utilizando o classificador SNNS. Em (f), a imagem também é percorrida utilizando o tipo seqüencial, entretanto, o classificador utilizado é o de distância mínima.

#### 4.1.10 Experimento 10

A imagem colorida apresentada na figura 4.18(a) é um recorte de uma imagem de satélite. Em 4.18(b) são apresentadas as amostras que indicam, respectivamente, continente e oceano. As dimensões das imagens são  $861 \times 1377$  pixels para a imagem original e  $100 \times 100$  pixels para as amostras.

Para os testes realizados com imagens de entrada colorida, os parâmetros selecionados foram sempre os mesmos. As características utilizadas foram segundo momento angular, contraste, correlação e entropia (propostas por Dacheng *et al.* [14] para matrizes de co-ocorrência 3D), a maneira de percorrer a imagem foi seqüencial, utilizando janela  $3 \times 3$  pixels, 30 iterações e normalização das medidas. O resultado da segmentação é apresentado na figura 4.19.

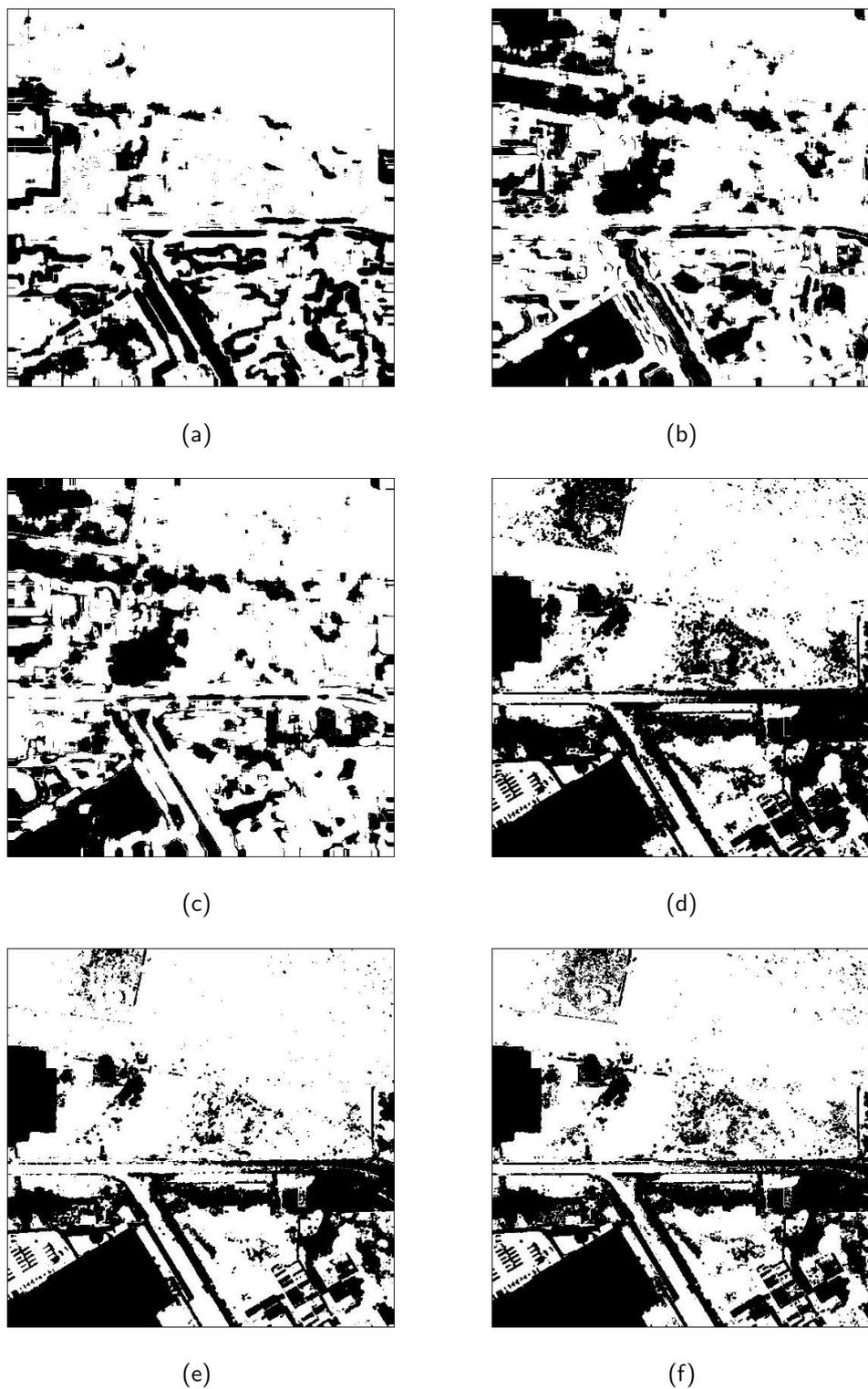


Figura 4.17: Resultados da segmentação da imagem Politécnico. Parâmetros comuns: características identificadas pela PCA (segundo momento angular, contraste, desvio padrão, média e correlação) e normalização. De (a) até (e) redes neurais, em (f) distância mínima. (a) janela 15x15 e seqüencial, (b) janela 15x15 e cascata, (c) janela 15x15 e aleatório, (d) janela 3x3 e seqüencial, (e) janela 3x3 e cascata e (f) janela 3x3 e seqüencial.

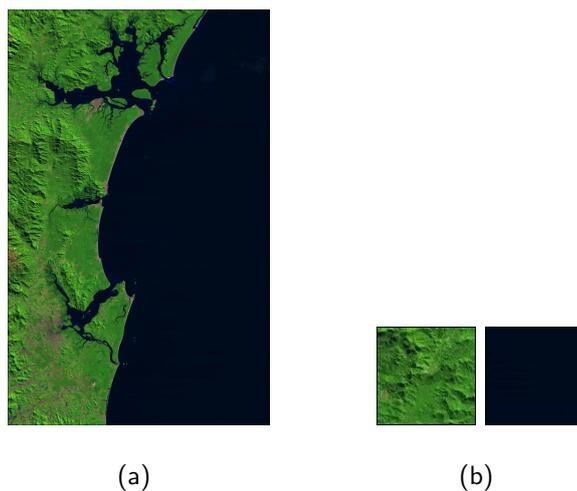


Figura 4.18: Experimento 10. (a) Imagem original e (b) amostras utilizadas.



Figura 4.19: Resultados da segmentação da imagem Satélite1. Características propostas por Dacheng *et al.* [14] (segundo momento angular, contraste, correlação e entropia), seqüencial, janela 3x3, classificador redes neurais e normalização.

#### 4.1.11 Experimento 11

A imagem apresentada na figura 4.20(a) é um recorte de uma imagem de satélite da região de Minnesota, Estados Unidos. Em 4.20(b) são apresentadas as amostras que indicam, respectivamente, continente e oceano. As dimensões das imagens são  $368 \times 400$  pixels para imagem original e  $50 \times 50$  pixels para as amostras.

Os parâmetros utilizados neste experimento foram os mesmos do experimento 10. O resultado é mostrado na figura 4.21.

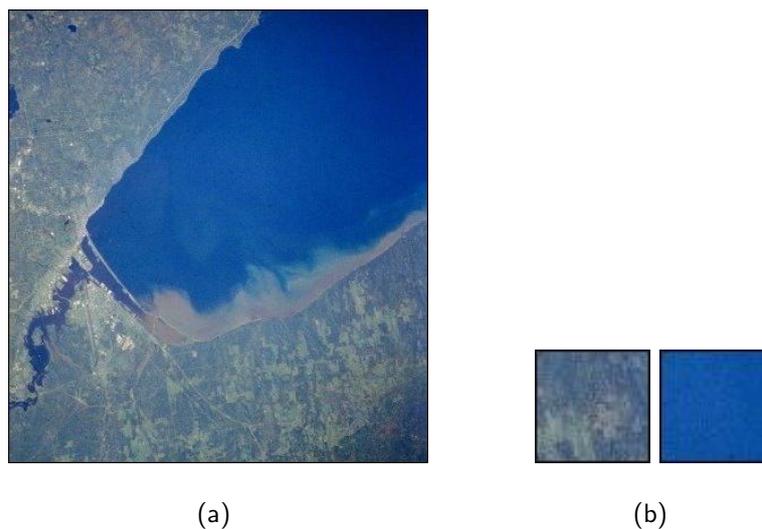


Figura 4.20: Experimento 11. (a) Imagem original e (b) amostras utilizadas.



Figura 4.21: Resultado da segmentação da imagem Minnesota. Características propostas por Dacheng *et al.* [14] (segundo momento angular, contraste, correlação e entropia), seqüencial, janela 3x3, classificador redes neurais e normalização.

#### 4.1.12 Experimento 12

A imagem apresentada na figura 4.22(a) é um recorte de uma imagem de satélite. Em 4.22(b) são apresentadas as amostras que indicam respectivamente continente e oceano. As dimensões das imagens são  $150 \times 204$  pixels para imagem original e  $35 \times 35$  pixels para as amostras.

Os parâmetros utilizados neste experimento foram os mesmos do experimento 10. O resultado é mostrado na figura 4.23.

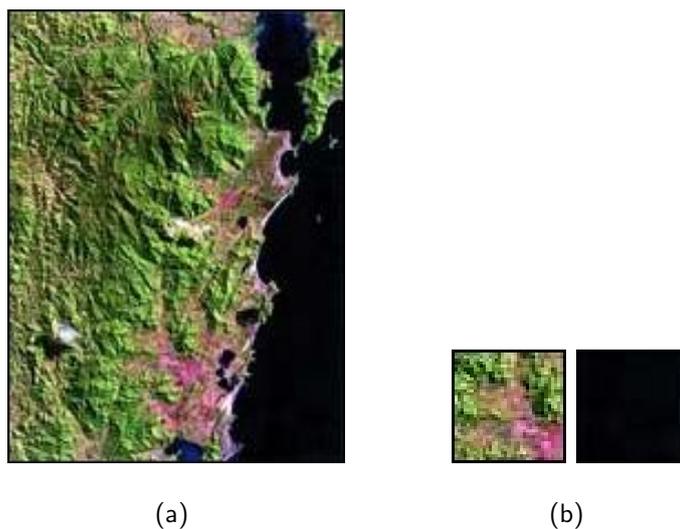


Figura 4.22: Experimento 12. (a) Imagem original e (b) amostras utilizadas.



Figura 4.23: Resultados da segmentação da imagem Satélite2. Características propostas por Dacheng *et al.* [14] (segundo momento angular, contraste, correlação e entropia), seqüencial, janela 3x3, classificador redes neurais e normalização.

### 4.1.13 Experimento 13

A imagem apresentada na figura 4.24(a) é um recorte de uma imagem de satélite da região da Patagônia, Argentina. Em 4.24(b) são apresentadas as amostras que indicam, respectivamente, continente e oceano. As dimensões das imagens são  $400 \times 400$  pixels para a imagem original e  $50 \times 50$  pixels para as amostras.

Os parâmetros utilizados neste experimento foram os mesmos do experimento 10. O resultado é mostrado na figura 4.25.

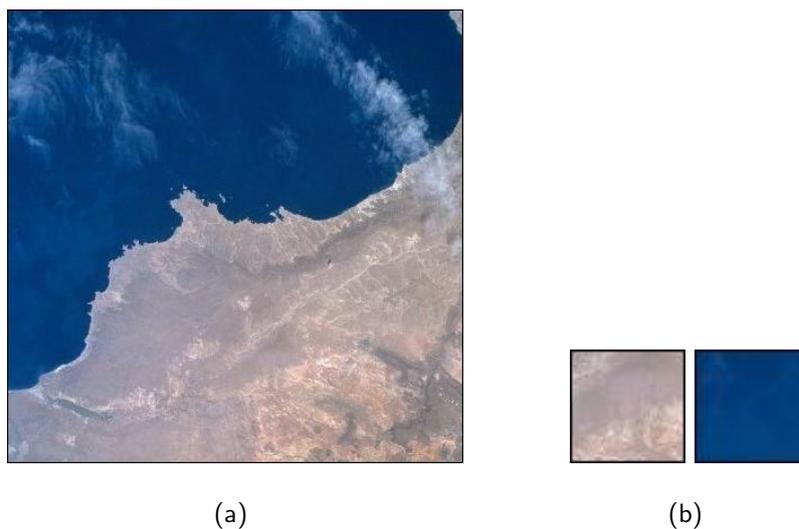


Figura 4.24: Experimento 13. (a) Imagem original e (b) amostras utilizadas.

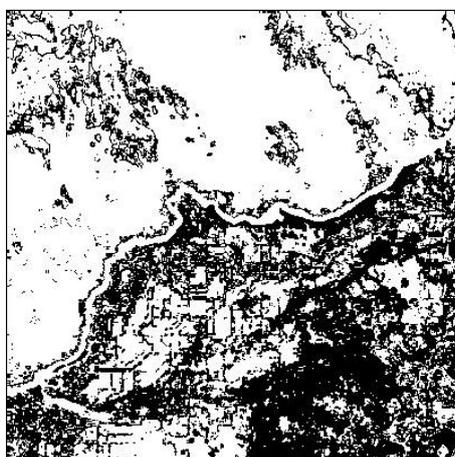


Figura 4.25: Resultado da segmentação da imagem Patagônia. Características propostas por Dacheng *et al.* [14] (segundo momento angular, contraste, correlação e entropia), seqüencial, janela 3x3, classificador redes neurais e normalização.

## 4.2 Análise dos Resultados

Com base nos resultados apresentados na seção 4.1 deste capítulo, pode-se observar que existe uma extensa possibilidade de variação de parâmetros, tais como medidas de características utilizadas, tamanho das janelas de treinamento e classificação, tipo do classificador, normalização das medidas e número de iterações.

A seleção das características mostrou-se um fator importante na segmentação. Embora o custo computacional seja grande, foram realizados testes com todas as características disponíveis no programa desenvolvido e, nesses casos, a segmentação não foi realizada com

sucesso. Buscou-se, então, determinar qual o número mínimo e quais seriam as características que segmentariam as imagens adequadamente. Por método simples de tentativa e erro, verificou-se, dentre diversos conjuntos diferentes de características, que o número mínimo aceitável era de cinco características. Para simplificar a escolha de quais características seriam utilizadas utilizou-se a PCA. Com algumas variações de ordenação, as características identificadas pelo método foram sempre as mesmas: segundo momento angular, contraste, desvio padrão, média e correlação.

Também foram realizados testes com o conjunto de características formado pela tendência de agrupamento, inércia, curtose, probabilidade máxima e simetria. Esse conjunto de característica produziu resultados satisfatórios tendo em vista que possuem um vetor de características com dimensionalidade reduzida, o que diminui bastante o custo computacional.

Outro fator identificado como crucial no resultado da segmentação é o tamanho da janela de treinamento e classificação. Pela análise dos resultados, verificou-se que o tamanho da janela está diretamente relacionado ao nível de detalhe dos objetos da imagem. Para imagens com objetos pequenos ou reentrâncias, como por exemplo as imagens de terrenos, janelas menores ( $3 \times 3$  ou  $5 \times 5$  pixels) produzem melhores resultados. Para imagens com objetos grandes ou poucos detalhes, como as imagens de mosaico, janelas maiores (por exemplo,  $13 \times 13$  ou  $15 \times 15$  pixels), resultam em melhores segmentações. Caso esta regra não seja seguida, imagens podem perder detalhes ou não serem segmentadas como deveriam.

A normalização das medidas se mostrou um parâmetro crucial para a segmentação, principalmente nas imagens coloridas. Vários testes foram realizados e, quando a normalização não era utilizada, a imagem não era segmentada corretamente. Nesses casos, as imagens resultantes são parecidas com a apresentada na figura 4.4 (e).

A escolha da amostra que caracteriza a classe também é um fator importante no processo de classificação. Ela deve ser escolhida de forma que a textura seja completamente descrita, ou seja, a amostra não pode ser muito pequena, pois desta forma pode não conter todos os relacionamentos de intensidades que a caracterizam. Entretanto, a amostra também não pode ser muito grande a ponto de conter informações redundantes, extrapolando as características pertencentes à classe em questão.

Quanto às maneiras diferentes pelas quais a imagem de amostra pode ser percorrida durante o treinamento, verificou-se que o tipo aleatório é o mais ineficiente. Os tipos seqüencial e cascata se mostraram praticamente equivalentes, desde que para o tipo cascata utilize-se a opção para percorrer a imagem inteira, caso contrário, o tipo seqüencial apresenta melhores resultados conforme a inspeção visual. É importante observar que a maneira como a imagem é percorrida, diferente do que se pensou inicialmente, não é um parâmetro que cause grandes mudanças no resultado da segmentação. Certamente as características selecionadas, o tamanho das janelas e a escolha correta das amostras possuem maior influência nos resultados.

Observou-se que o método é sensível à presença de ruído. Dessa forma, verificando os resultados dos testes realizados, recomenda-se que imagens com ruído sejam previamente tratadas por métodos específicos de eliminação de ruído para somente então serem submetidas à segmentação.

Analisando os resultados quanto aos classificadores utilizados, pode-se concluir que o SNNS é mais eficiente que o classificador de distância mínima. Porém, existe muito a ser pesquisado sobre a influência da rede neural na segmentação. Descobriu-se também que valores muito discrepantes como entrada da rede neural podem produzir resultados indesejados e chegou-se à conclusão de que a normalização melhoraria os resultados, como de fato aconteceu.

## CAPÍTULO 5

### CONCLUSÕES E TRABALHOS FUTUROS

Este trabalho apresentou um método de segmentação de imagens monocromáticas e coloridas baseado em um conjunto de características de textura extraídas a partir de matrizes de co-ocorrência.

Diversos parâmetros foram analisados para verificar sua influência no processo de segmentação. Observou-se que o tamanho da janela utilizado nos algoritmos para analisar a textura influenciou o resultado da segmentação. A redução de dimensionalidade, ou seja, a diminuição do número de características selecionadas da imagem contribuiu significativamente para um melhor desempenho do classificador em termos de memória utilizada e tempo computacional.

A escolha da amostra que caracteriza a classe também é um fator importante na segmentação. Ela deve ser escolhida de forma que contenha todo o padrão da textura e não extrapole as características pertencentes à classe em questão.

Observou-se também que o método é sensível à presença de ruído sendo necessário submeter as imagens a um método de redução de ruídos antes de serem segmentadas.

Os experimentos demonstraram que a abordagem é bastante promissora, podendo ser aplicada a uma grande variedade de imagens.

A partir da investigação realizada, novas propostas surgem como trabalhos futuros. Novas medidas estatísticas devem ser pesquisadas para caracterizar mais eficientemente a informação de textura, aumentando a precisão do classificador. Outra proposta é a utilização de conjuntos nebulosos (*fuzzy*) no processo de classificação, já que nem sempre é possível definir fronteiras rígidas entre as diferentes classes.

A influência do classificador de rede neural na segmentação de imagens também pode ser melhor estudada a fim de identificar como a alteração de parâmetros do classificador pode afetar os resultados. Outro item a ser pesquisado são os tipos de seleção dos componen-

tes principais gerados pela PCA. Verificou-se neste trabalho que existe uma relação entre o tamanho da janela e o tipo de objetos presentes na imagem, tal que um estudo futuro pode associar o tamanho da janela não apenas aos objetos, mas também às características utilizadas. Também como trabalho futuro pode-se estudar o motivo pelo qual o algoritmo de percurso cascata não foi mais eficiente que o modo seqüencial na varredura da imagem durante a etapa de extração dos descritores.

## REFERÊNCIAS BIBLIOGRÁFICAS

- [1] Redes neurais artificiais, Acesso em 2004. <http://www.geocities.com/CapeCanaveral/Runway/4303/index2.html>.
- [2] ANDREY, P. E TARROUX, P. Unsupervised Segmentation of Markov Random Field Modeled Textured Images Using Selectionist Relaxation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 20, 3 (Março 1998), 252–262.
- [3] ANGELO, N. P. E HAERTEL, V. Aplicação dos filtros de Gabor na classificação supervisionada de imagens digitais. In *I Congresso Brasileiro de Computação* (Porto Alegre, RS, 2001).
- [4] ARIVAZHAGAN, S. E GANESAN, L. Texture Classification Using Wavelet Transform. *Pattern Recognition Letters* 24, 9–10 (Junho 2003), 1513–1521.
- [5] AYALA, W. E DOMINGO, J. Spatial Size Distributions: Applications to Shape and Texture Analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 23, 12 (Dezembro 2001), 1430–1442.
- [6] BEYMER, D. J. Face recognition under varying pose. No. AIM-1461, p. 14. <http://citeseer.ist.psu.edu/beymer94face.html>.
- [7] BRODATZ, P. Brodatz textures, Acesso em 2003. <http://www.ux.his.no/~tranden/brodatz.html>.
- [8] CENTENO, J. A. S. *Sensoriamento Remoto e Processamento de Imagens Digitais*. Editora da UFPR, Brasil, 2003.
- [9] CHEN, C., GAGAUDAKIS, G. E ROSIN, P. Similarity-based image browsing, 2000. <http://www.pages.drexel.edu/>.
- [10] CHEN, C.-C. E HUANG, C.-L. Markov Random Fields for Texture Classification. *Pattern Recognition Letters* 14, 11 (1993), 907–914.

- [11] CHENG, Y.-C. E CHEN, S.-Y. Image Classification Using Color, Texture and Regions. *Image and Vision Computing* 21, 9 (Setembro 2003), 759–776.
- [12] COGGINS, J. M. A framework for texture analysis based on spatial filtering. Dissertação de Mestrado, Computer Science Department, Michigan State University, East Lansing, Michigan, 1982.
- [13] COOPER, G. R. J. The textural analysis of gravity data using co occurrence-matrices. *Computers Geosciences* 30 (2004), 107–115.
- [14] DACHENG, T., XUELONG, L., YUAN, Y., NENGHAI, Y., ZHENGKAI, L. E XIAO-OU, T. A Set of Novel Textural Features Based on 3D Co-occurrence Matrix for Content-based Image Retrieval. In *Proceedings of the Fifth International Conference on Information Fusion* (2002), vol. 2, pp. 1403–1407.
- [15] DE CARVALHO, A. C. P. L. F. Backpropagation, Acesso em 2004. <http://www.icmc.usp.br/~andre/research/neural/MLP.htm>.
- [16] DE WOUWER, G. V., SCHEUNDERS, P. E D. VAN DYCK, D. Statistical Texture Characterization from Discrete Wavelet Representations. *IEEE Transactions on Image Processing* 8, 4 (Abril 1999), 592–598.
- [17] DO NASCIMENTO, J. P. R. Análise e classificação de imagens baseadas em características de textura utilizando matrizes de co-ocorrência. Dissertação de Mestrado, Universidade Federal do Paraná, Departamento de Informática, 2003.
- [18] DOS SANTOS COELHO, L. Introdução às redes neurais artificiais - Notas de aula da disciplina de Controle Avançado do Curso de Engenharia Mecatrônica (PUC-PR), Acesso em 2004. <http://www.produtronica.pucpr.br/leandro>.
- [19] GONZALEZ, R. C. E WOODS, R. E. *Processamento de Imagens Digitais, tradução do original Digital Image Processing*. Edgard Blücher, São Paulo, 2000.
- [20] HALL-BEYER, M. GLCM texture: A tutorial, Acesso em 2003. [http://www.ucalgary.ca/~mhallbey/texture/texture\\\_tutorial.html](http://www.ucalgary.ca/~mhallbey/texture/texture\_tutorial.html).

- [21] HANDMANN, U. E KALINKE, T. Fusion of texture and contour based methods for object recognition. *IEEE Conference on Intelligent Transportation Systems* (1997), 876–881.
- [22] HARALICK, R. M., SHANMUGAM, K. E DINSTEIN, I. Textural features for image classification. *IEEE Transactions on Systems, Man and Cybernetics* 3, 6 (Novembro 1973), 610–621.
- [23] ICHI MURATA, S., HERMAN, P. E LAKOWICZ, J. R. Texture Analysis of Fluorescence Lifetime Images of AT- and GC-rich Regions in Nuclei. *Journal of Histochemistry and Cytochemistry* 49 (2001), 1443–1452.
- [24] IDRISSE, M. E ACHEROY, M. Texture Classification Using Gabor Filters. *Pattern Recognition Letters* 23, 9 (Julho 2002), 1095–1102.
- [25] IGLEWICZ, B. Robust scale estimators and confidence intervals for location. In *Understanding Robust and Exploratory Data Analysis*, D. C. Hoaglin, F. Mosteller, and J. Tukey, Eds. John Wiley & Sons, Inc., 1983.
- [26] JAIN, A. K., DUIN, R. P. W. E MAO, J. Statistical pattern recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22, 1 (Janeiro 2000), 4–37.
- [27] JAIN, R., KASTURI, R. E SCHUNCK, B. G. *Machine Vision*. McGraw-Hill, Inc., New York, NY, Estados Unidos, 1995.
- [28] KASPARIS, T., CHARALAMPIDIS, D., GEORGIPOULOS, M. E ROLLAND, J. Segmentation of Textured Images Based on Fractals and Image Filtering. *Pattern Recognition* 34, 10 (Outubro 2001), 1963–1973.
- [29] KENDALL, M. *Multivariate Analysis*. Griffin, Londres, 1980.
- [30] KERVRANN, C. E HEITZ, F. A Markov Random Field Model-Based Approach to Unsupervised Texture Segmentation Using Local and Global Spatial. *IEEE Transactions on Image Processing* 4, 6 (Junho 1995), 856–862.

- [31] KIM, J. H., YUN, I. D. E LEE, S. U. Unsupervised Segmentation of Textured Image Using Markov Random Field in Random Spatial Interaction. In *Proceedings 1998 International Conference on Image Processing* (Outubro 1998), vol. 3, pp. 756–760.
- [32] KOLMAN, B. *Introdução à Álgebra Linear - Com Aplicações, tradução do original Introductory Linear Algebra with Applications*. Prentice-Hall do Brasil, Rio de Janeiro, 1998.
- [33] LEHMANN, T., BREDNO, J. E SPITZER, K. Texture-adaptive active contour models, 2001.
- [34] LIAPIS, S., SIFAKIS, E. E TZIRITAS, G. Colour and Texture Segmentation Using Wavelet Frame Analysis, Deterministic Relaxation, and Fast Marching Algorithms. *Journal of Visual Communication and Image Representation* 15, 1 (2004), 1–26.
- [35] MARTINS, M. P., GUIMARÃES, L. N. F. E FONSECA, L. M. G. Classificador de texturas por redes neurais. *Anais do II Congresso Brasileiro de Computação 1* (2002), 1–12.
- [36] MATERKA, A. E STRZELECKI, M. Texture analysis methods - a review, 1998.
- [37] MCCULLOCH, W. S. E PITTS, W. H. A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics* 5 (1943), 115–133.
- [38] MITCHELL, O. R. E CARLTON, S. G. Image segmentation using a local extrema texture measure. *Pattern Recognition* 10 (1978), 205–210.
- [39] MUDIGONDA, N. R., RANGAYYAN, R. M. E DESAUTELS, J. E. L. Gradient and texture analysis for the classification of mammographic masses. *IEEE Transaction on Medical Imaging* 19 (2000), 1032–1034.
- [40] MUHAMAD, A. K. E DERAVI, F. Neural network texture classifiers using direct input co-occurrence matrices. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (1993), 117–120.

- [41] MURINO, V., OTTONELLO, C. E PAGNAN, S. Noisy texture classification: a higher order statistics approach. *Pattern Recognition* 34 (1998), 383–393.
- [42] MÄENPÄÄ, T. E PIETIKÄINEN, M. Classification with color and texture: jointly or separately? *Pattern Recognition* 37 (2004), 1629–1640.
- [43] NADLER, M. E SMITH, E. P. *Pattern Recognition Engineering*. John Wiley - Sons, 1993.
- [44] OZYILDIZ, E., KRAHNSTÖVER, N. E SHARMA, R. Adaptive texture and color segmentation for tracking moving objects. *Pattern Recognition* 35 (2002), 2013–2029.
- [45] PALM, C. Color texture classification by integrative co-occurrence matrices. *Pattern Recognition* 37 (2004), 965–976.
- [46] PALM, C. Color Texture Classification by Integrative Co-occurrence Matrices. *Pattern Recognition* 37, 5 (Maio 2004), 965–976.
- [47] PLA, L. E. *Análisis Multivariado: Método de Componentes Principales*. Departamento de Producción Vegetal. Area de Ciencias del Agro y del Mar, Universidad Nacional Experimental Francisco de Miranda. Coro, Falcón, Venezuela, 1986.
- [48] POSKANZER, J. *Manual PGM - Portable Pixmap File Format*, 1991.
- [49] POSKANZER, J. *Manual PPM - Portable Graymap File Format*, 1991.
- [50] ROSENBLATT, F. The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review* 65, 6 (1958), 386–408.
- [51] RUIZ, L. A., FDEZ-SARRÍA, A. E RECIO, J. A. Texture Feature Extraction for Classification of Remote Sensing Data Using Wavelet Decomposition: A Comparative Study. In *Proceedings International Society for Photogrammetry and Remote Sensing Congress* (Istambul, Turquia, Julho 2004).

- [52] SCHWARTZ, W. R. Segmentação de Imagens Baseada em Dependência Espacial Utilizando Campo Aleatório de Markov Associado com Características de Texturas. Dissertação de Mestrado, Universidade Federal do Paraná, 2005.
- [53] SCHWARTZ, W. R. E PEDRINI, H. Segmentação de Imagens de Terrenos Baseada na Associação de Características de Texturas com Dependência Espacial Modelada por Campo Aleatório de Markov. In *XII Simpósio Brasileiro de Sensoriamento Remoto* (Goiânia-GO, Abril 2005), pp. 4311–4318.
- [54] SHABAN, M. A. E DIKSHIT, O. Textural classification of high resolution digital satellite imagery. *IEEE Transactions on Computers* (1998), 2590–2592.
- [55] SHARMA, M. Performance evaluation of image segmentation and texture extraction methods in scene analysis. Dissertação de Mestrado, University of Exeter, Reino Unido, 2001.
- [56] SINGH, M. E SINGH, S. Spatial Texture Analysis: A Comparative Study. In *International Conference on Pattern Recognition* (2002), vol. 1, pp. 676–679.
- [57] SMITH, L. I. A tutorial on principal component analysis, 2002. [http://www.cs.otago.ac.nz/cosc453/student\\_tutorials/principal\\_components.pdf](http://www.cs.otago.ac.nz/cosc453/student_tutorials/principal_components.pdf).
- [58] SNNS. *Stuttgart Neural Network Simulator*. University of Stuttgart, Acesso em 2003. <http://www-ra.informatik.uni-tuebingen.de/SNNS/>.
- [59] SONKA, M., HLAVAC, V. E BOYLE, R. *Image Processing, Analysis and Machine Vision*. Brooks and Cole Publishing, 1998.
- [60] SUN, J., GU, D., ZHANG, S. E CHEN, Y. Hidden Markov Bayesian Texture Segmentation Using Complex Wavelet Transform. In *IEE Proceedings on Vision, Image and Signal Processing* (Junho 2004), vol. 151, pp. 215–223.
- [61] TATIBANA, C. Y. E KAETSU, D. Y. Redes neurais artificiais, Acesso em 2004. <http://www.din.uem.br/ia/neurais/neural>.

- [62] TOBIAS, O. J., SEARA, R., SOARES, F. A. P. E BERMUDEZ, J. C. M. Automatic visual inspection using the co-occurrence approach. *IEEE Midwest Symposium on Circuits and Systems 38* (1995), 154–157.
- [63] TSENG, D.-C. E LAI, C.-C. A Genetic Algorithm for MRF-based Segmentation of Multi-spectral Textured Images. *Pattern Recognition Letters 20*, 14 (Dezembro 1999), 1499–1510.
- [64] WALKER, R., JACKWAY, P. E LONGSTAFF, I. Improving co-occurrence matrix feature discrimination. In *Proc. of DICTA95 - 3rd International Conference on Digital Image Computing: Techniques and Applications* (1995), pp. 643–648.
- [65] WALKER, R. F. E JACKWAY, P. T. Statistical geometric features-extensions for cytological texture analysis. *Pattern Recognition. Proceedings of the 13th International Conference. 2* (1996), 790–794.
- [66] WESZKA, J. S., DYER, C. R. E ROSENFELD, A. A Comparative Study of Texture Measures for Terrain Classification. *IEEE Transactions on Systems, Man, and Cybernetics SMC-6*, 4 (Abril 1976), 269–285.
- [67] WESZKA, J. S. E ROSENFELD, A. An application of texture analysis to materials inspection. *Pattern Recognition 8* (1976), 195–200.

## APÊNDICE A

### FORMATO DAS IMAGENS

#### A.1 Formato PGM

O formato PGM (*Portable Gray Map*) é um formato de armazenamento de imagens em tons de cinza. Este formato apresenta as seguintes características [49]:

- Pode ser escrito no padrão ASCII (*American Standard Code for Information Interchange*) sendo assim, portátil em diversas arquiteturas.
- Possui um identificador de tipo de formato sendo P2 (ASCII) e P5 (Binário).
- Os valores dos pixels da imagem variam de 0 (preto) a 255 (branco), sendo o intervalo entre estes valores a variação de níveis de cinza.
- A leitura da imagem deve ser realizada linha por linha, de cima para baixo, da esquerda para direita.

Uma imagem exemplo com dimensões de  $24 \times 7$  pixels e nível máximo de tons de cinza por pixel igual a 12 é mostrada na figura A.1. A primeira linha do arquivo contém informações essenciais para composição da imagem. Os dois primeiros caracteres indicam o tipo do formato P2 para ASCII ou P5 para binário, o próximo carácter é um número que indica a largura da imagem em pixels, em seguida outro número é necessário pra indicar a altura da imagem em pixels e por fim, o último número da primeira linha do arquivo expressa o nível de cinza máximo presente na imagem. É importante ressaltar que existe um espaço em branco entre cada item exigido para a primeira linha. O resultado é apresentado na figura A.2 que teve sua escala alterada para ser melhor visualizada.

```

P2 24 7 12
0 0 1 1 2 2 3 3 4 4 5 5 6 6 7 7 8 8 9 10 10 11 11
0 0 1 1 2 2 3 3 4 4 5 5 6 6 7 7 8 8 9 10 10 11 11
0 0 1 1 2 2 3 3 4 4 5 5 6 6 7 7 8 8 9 10 10 11 11
0 0 1 1 2 2 3 3 4 4 5 5 6 6 7 7 8 8 9 10 10 11 11
0 0 1 1 2 2 3 3 4 4 5 5 6 6 7 7 8 8 9 10 10 11 11
0 0 1 1 2 2 3 3 4 4 5 5 6 6 7 7 8 8 9 10 10 11 11
0 0 1 1 2 2 3 3 4 4 5 5 6 6 7 7 8 8 9 10 10 11 11

```

Figura A.1: Exemplo de arquivo PGM.

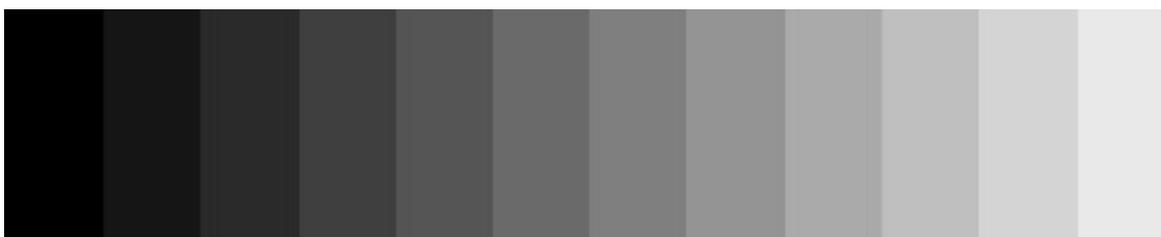


Figura A.2: Exemplo de imagem no formato PGM.

## A.2 Formato PPM

Assim como no formato PGM, o formato *Portable Pixmap* (PPM) é composto de um cabeçalho e das informações de cores para cada pixel. Neste cabeçalho, P3 indica o tipo da imagem, e os valores indicam respectivamente a altura, a largura e o valor máximo de intensidade para cada componente de cor. Cada informação é escrita em números decimais no código ASCII. Como a imagem é colorida, existe um valor decimal para cada componente de cor e para cada pixel. Assim, o número de valores para os pixels é igual a *altura x largura x 3*. Cada pixel consiste em três valores seguidos, um para o vermelho, outro para o verde e outro para o azul, nesta ordem [48]. A figura A.3 mostra o exemplo da imagem apresentada na figura A.4. Esta possui dimensão  $4 \times 4$  pixels e 15 intensidades para cada componente de cor. A imagem resultante teve sua escala alterada para ser melhor visualizada.

```
P3 4 4 15
15 0 0 15 0 0 15 0 0 15 0 0
0 15 0 0 15 0 0 15 0 0 15 0
0 0 15 0 0 15 0 0 15 0 0 15
15 15 0 0 15 15 15 0 15 0 0 0
```

Figura A.3: Exemplo de arquivo PPM.

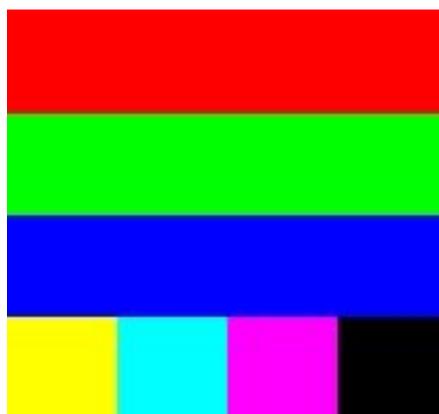


Figura A.4: Exemplo de imagem no formato PPM.