

JOÃO PAULO RIBEIRO DO NASCIMENTO

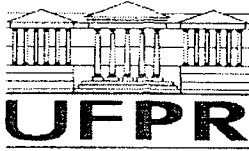
**ANÁLISE E CLASSIFICAÇÃO DE IMAGENS
BASEADAS EM CARACTERÍSTICAS DE TEXTURA
UTILIZANDO MATRIZES DE CO-OCORRÊNCIA**

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

2003



Ministério da Educação
Universidade Federal do Paraná
Mestrado em Informática

PARECER

Nós, abaixo assinados, membros da Banca Examinadora da defesa de Dissertação de Mestrado em Informática, do aluno João Paulo Ribeiro do Nascimento, avaliamos o trabalho intitulado, "*Análise e Classificação de Imagens Baseadas em Características de Textura Utilizando Matrizes de Co-ocorrência*", cuja defesa foi realizada no dia 14 de abril de 2003, às quatorze horas, no Auditório da Informática da Universidade Federal do Paraná. Após a avaliação, decidimos pela aprovação do candidato.

Curitiba, 14 de abril de 2003.

Prof. Dr. Hélio Pedrini
DINF/UFPR (Orientador)

Prof. Dr. Carlos Hitoshi Morimoto
IME/USP

Prof. Dr. André Luiz Pires Guedes
DINF/UFPR

AGRADECIMENTOS

Primeiramente agradeço a Deus pela vida e por permitir minha chegada até aqui. Tantos foram os obstáculos, mas essa força superior e indescritível sempre me deu ânimo para continuar lutando de cabeça erguida.

Agradeço aos meus pais, Sonia e Honório, meus primeiros “orientadores” nesta vida. Mãe, você é uma pessoa maravilhosa que ilumina a minha vida com a sua existência. Em quantos momentos desanimei e você com seu otimismo, sua dedicação e amor de mãe me ensinou que a arte de viver é acordar a cada dia, levantar a cada tombo e sorrir. Pai, obrigado a você também, o destino um dia nos separou, mas sempre recebi sua luz, sua força e sua presença durante toda essa caminhada. Vocês moram no meu coração.

Ao meu orientador Prof. Dr. Hélio Pedrini, tenho muito a agradecer. Obrigado pelos ensinamentos, pela objetividade e pela coerência com que conduziu este trabalho. Seu constante incentivo, dedicação e competência sempre indicaram a direção correta a ser tomada nos momentos de dificuldade. Obrigado ainda pela amizade, confiança e consideração que teve por mim.

À minha irmã Juliana, aos meus avós Adolfina e Nelson, a meu tio Paulo, à minha namorada Betânia e amigos, obrigado pela paciência, compreensão e pela presença constante, apesar das distâncias e da falta de tempo.

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

LISTAS DE FIGURAS	vi
LISTA DE TABELAS	vii
RESUMO	viii
ABSTRACT	ix
1 INTRODUÇÃO	1
1.1 MOTIVAÇÃO	1
1.2 PROBLEMÁTICA	2
1.3 OBJETIVOS	3
1.4 ORGANIZAÇÃO	4
2 TEXTURA	5
2.1 O QUE É TEXTURA?	5
2.2 ABORDAGENS ESTATÍSTICAS	6
2.3 ABORDAGENS ESTRUTURAIS	8
2.4 ABORDAGENS ESPECTRAIS	10
2.5 OUTRAS ABORDAGENS	12
2.6 ANÁLISE DE TEXTURA	14
2.7 TEXTURA × PADRÃO	20
3 TÉCNICAS DE CLASSIFICAÇÃO	22
3.1 METODOLOGIAS DE RECONHECIMENTO DE PADRÕES	23
3.1.1 Casamento de Padrões	23
3.1.2 Abordagem Estatística	24
3.1.3 Redes Neurais	26
3.1.3.1 Histórico	28

3.1.3.2	Redes neurais artificiais	29
3.1.3.3	Aprendizado supervisionado para reconhecer padrões	36
3.1.3.4	Princípios básicos das redes de retropropagação	37
3.1.4	Lógica Nebulosa (Fuzzy)	38
3.2	CLASSIFICADORES	40
3.2.1	Distância Mínima	40
3.2.2	K-vizinhos mais Próximos	42
3.2.3	Máxima Verossimilhança	43
3.2.4	Intervalos	45
4	MÉTODO DE CLASSIFICAÇÃO DE IMAGENS BASEADO EM CARAC- TERÍSTICAS TEXTURAIIS	48
4.1	MATRIZ DE CO-OCORRÊNCIA DE NÍVEIS DE CINZA	50
4.1.1	Ordem das Medidas Estatísticas	51
4.1.2	Relacionamento Espacial entre Dois Pixels	52
4.1.3	Separação entre Dois Pixels	53
4.1.4	Propriedades da Matriz de Co-ocorrência	54
4.1.5	Expressando a MC como uma Probabilidade	55
4.1.6	Características de Textura	56
4.2	REDUÇÃO DE DIMENSIONALIDADE	61
4.3	EXTRAÇÃO DE CARACTERÍSTICAS	63
4.4	SELEÇÃO DE CARACTERÍSTICAS	64
5	RESULTADOS EXPERIMENTAIS	66
5.1	CLASSIFICAÇÃO 1	74
5.2	CLASSIFICAÇÃO 2	81
5.3	CLASSIFICAÇÃO 3	84
5.4	CLASSIFICAÇÃO 4	86
5.5	CLASSIFICAÇÃO 5	90
5.6	ANÁLISE DOS RESULTADOS	93

	iv
6 CONCLUSÕES	96
REFERÊNCIAS BIBLIOGRÁFICAS	98
APÊNDICE A FORMATO DAS IMAGENS	104

LISTA DE FIGURAS

2.1	Exemplos de texturas do álbum Brodatz.	7
2.2	Exemplos de primitivas de textura.	9
2.3	Exemplo de formação de um padrão de textura estrutural.	9
2.4	(a) Imagem mostrando uma textura periódica; (b) espectro; (c) gráfico de $S(r)$; (d) gráfico de $S(\theta)$	12
2.5	(a) Imagem contendo um tipo diferente de textura periódica; (b) gráfico de $S(\theta)$	13
3.1	Exemplo de padrão em um espaço tridimensional.	24
3.2	Exemplo de espaço de características.	24
3.3	Esquema dos constituintes de uma célula neural.	27
3.4	Neurônio artificial simples com n entradas e uma saída.	32
3.5	Função de ativação (F_a) e função de transferência (F_t).	33
3.6	Funções de transferência.	34
3.7	Neurônio completo.	35
3.8	Rede neural.	35
3.9	Exemplo de limites de decisão - Distância Mínima.	41
3.10	Exemplo de limites de decisão através do método dos K -vizinhos mais próximos.	43
3.11	Limite de aceitação de uma classificação.	45
3.12	Exemplo de decisão entre dois intervalos de classes.	47
4.1	Esquema do sistema proposto para classificação de imagens.	49
4.2	Ângulos para o cálculo na matriz de co-ocorrência.	52
4.3	Imagem 4×4 com níveis de cinza de 0 a 3.	53
5.1	Amostras de texturas extraídas do álbum Brodatz.	67
5.2	Gráficos comparativos para cada característica de textura extraída das imagens.	68
5.3	Gráficos comparativos para cada característica de textura extraída das imagens.	69

5.4	Imagem do <i>Valley of the Kings</i> , região do Rio Nilo.	72
5.5	Amostras de solo da região do Rio Nilo.	72
5.6	Amostras de vegetação da região do Rio Nilo.	72
5.7	Classificação - Letra.	75
5.8	Rede Neural - Letra.	80
5.9	Classificação - Satélite.	82
5.10	Rede Neural - Satélite.	83
5.11	Classificação - Mosaico I.	84
5.12	Rede Neural - Mosaico I.	85
5.13	Classificação - Mosaico II.	88
5.14	Rede Neural - Mosaico II.	89
5.15	Classificação - Grãos de Feijão.	91
5.16	Rede Neural - Grãos de Feijão.	92
5.17	Efeito do tamanho (em pixels) da janela na classificação de texturas.	95
A.1	Exemplo de arquivo PGM.	105
A.2	Exemplo de imagem no formato PGM.	105

LISTA DE TABELAS

2.1	Padrão 1 - Soma dos valores igual a 240.	21
2.2	Padrão 2 - Soma dos valores igual a 240.	21
3.1	Resultado da classificação - Distância Mínima.	42
4.1	Matriz de combinações possíveis para níveis de cinza.	53
5.1	Cálculo obtido para a imagem D9.PGM nas direções 0°, 45°, 90° e 135°. . . .	70
5.2	Cálculo obtido para a imagem D29.PGM nas direções 0°, 45°, 90° e 135°. . . .	70
5.3	Cálculo obtido para a imagem D37.PGM nas direções 0°, 45°, 90° e 135°. . . .	71
5.4	Cálculo obtido para a imagem D38.PGM nas direções 0°, 45°, 90° e 135°. . . .	71
5.5	Média das características calculadas a partir das matrizes de co-ocorrência das amostras de solo e vegetação recortadas da região do Rio Nilo.	73
5.6	Distância Mínima - Letra.	74
5.7	Intervalos - Letra.	76
5.8	Distância Mínima - Satélite.	81
5.9	Intervalos - Satélite.	81
5.10	Distância Mínima - Mosaico I.	85
5.11	Intervalos - Mosaico I.	85
5.12	Distância Mínima - Mosaico II.	86
5.13	Intervalos - Mosaico II	87
5.14	Distância Mínima - Grãos de Feijão.	90
5.15	Intervalos - Grãos de Feijão.	90

RESUMO

Os métodos de classificação de imagens baseados em características texturais têm sido amplamente utilizados no meio científico e industrial, possuindo aplicações na medicina, microscopia, sensoriamento remoto, controle de qualidade, recuperação de dados em bases gráficas, dentre outras. A característica de textura é uma fonte importante de informações para o processo de análise e interpretação de imagens. Inicialmente, este trabalho apresenta conceitos sobre textura e uma revisão bibliográfica sobre os principais estudos envolvendo técnicas para análise de texturas, principalmente estatísticas de segunda ordem. São avaliados diversos métodos de classificação supervisionada de imagens baseada em informação de textura, tais como o classificador de distância mínima, k -vizinhos mais próximos, máxima verossimilhança, redes neurais e um novo classificador baseado em intervalos de decisão. Neste trabalho, a textura é definida através de um conjunto de medidas locais descrevendo as variações espaciais de intensidade ou cor. Tais medidas são calculadas utilizando-se matrizes de co-ocorrência. São abordados também detalhes relevantes sobre redução de dimensionalidade, extração e seleção de características que influenciam a precisão e desempenho do classificador. Resultados experimentais utilizando uma grande variedade de imagens reais demonstram a aplicabilidade dos métodos.

ABSTRACT

Image classification methods based on textural characteristics have been greatly used in scientific and industrial area, with applications in medicine, microscopy, remote sensing, control of quality, retrieval of information in graphic databases, among others. The characteristic of texture is an important source of information for the process of image analysis and interpretation. Initially, this work presents concepts of texture and a bibliographic review about the main studies related to techniques for texture analysis, especially statistics of second order. Several image supervised classification methods based on textural information are evaluated, such as minimum distance classifier, nearest neighbor classifier, maximum likelihood classifier, neural networks and a new classifier based on decision intervals. In this work, texture is defined through a set of local features indicating the spatial distribution of intensity or color. Such features are calculated by using co-occurrence matrices. Relevant details on dimensionality reduction, feature extraction and selection that affect the precision and performance of the classifier are also described. Experimental results using a great variety of real images demonstrate the applicability of the methods.

CAPÍTULO 1

INTRODUÇÃO

1.1 MOTIVAÇÃO

Reconhecimento de objetos é uma das tarefas mais importantes em aplicações envolvendo um sistema de visão computacional. O objetivo é obter uma descrição que contém informação suficiente para distinguir entre diferentes objetos de interesse. Normalmente, o processo de reconhecimento é baseado em níveis de cinza ou cores, características de forma e textura dos objetos.

Diversas áreas de aplicação utilizam técnicas de classificação para reconhecimento de objetos, alguns exemplos são listados a seguir:

- **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.
- **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.
- **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.
- **Identificação de impressões digitais:** Através de processo adequado de armazenamento e classificação, pode-se identificar uma impressão digital utilizando um banco de dados associado à uma base de dados de imagens.

- **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, monitoração em sistemas de segurança e de controle de acesso.
- **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.
- **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 alguma monitoração óptica 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.

A textura é uma informação valiosa que pode ser associada à interpretação de padrões relacionados às áreas descritas acima e em inúmeras outras. 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 um descritor da forma, tamanho, sombra e tonalidade de cada objeto, permitindo a impressão visual de rugosidade ou suavidade de determinada superfície.

1.2 PROBLEMÁTICA

A análise de imagens tem sido utilizada em várias linhas de pesquisa para o reconhecimento de texturas sintéticas e naturais. Entretanto, o número de estudos sobre a avaliação de desempenho dos vários métodos permanece reduzido. Estudos existentes comparam poucos métodos utilizando textura ou utilizam um pequeno número de amostras para tentar chegar a conclusões significativas [44].

Por outro lado, a necessidade de automação da inspeção visual em diversas tarefas tem aumentado significativamente. O desempenho da inspeção humana é geralmente variável e

inadequado para processos que necessitam de precisão e velocidade de identificação. Em muitos casos, o mesmo tipo de textura deve ser analisado repetidamente para se detectar anomalias, requerendo uma alta concentração do observador. Vários estudos demonstram a incapacidade do ser humano para executar tarefas monótonas e repetitivas. Além disso, existem ambientes inapropriados onde a presença humana pode ser perigosa. Tais fatores demonstram a necessidade e também os benefícios da utilização de sistemas automáticos. Dentre eles, pode-se citar o aumento de produtividade, melhoria na qualidade de produtos ou serviços prestados, ausência de riscos pessoais, alta velocidade de inspeção e redução de custos [52].

O atributo textura é considerado uma fonte importante de dados usado na identificação de objetos ou regiões de interesse em uma imagem, contribuindo para uma melhoria na exatidão do processo de classificação [2]. Entretanto, técnicas de segmentação e classificação baseadas em textura necessitam da definição de um conjunto de medidas que a represente de forma efetiva [5].

1.3 OBJETIVOS

Este trabalho visa investigar a relevância de informação textural presente em imagens digitais, levando-se em conta que textura é um dos aspectos mais importantes em análise de imagens.

Para isso, algumas abordagens para classificação de imagens através de características de textura foram estudadas e avaliadas, principalmente, os métodos estatísticos e as redes neurais.

No presente trabalho, a textura é definida através de um conjunto de medidas locais descrevendo as variações espaciais de intensidade de cinza ou cor. Tais medidas são calculadas utilizando-se matrizes de co-ocorrência, com o objetivo de alcançar uma classificação mais eficiente.

O método para a classificação de imagens baseado em características texturais definido envolve etapas que vão desde a aquisição de dados e a extração de atributos até a determinação da classe a qual pertence cada pixel de uma imagem. Todos os processos dentro dessas etapas requerem a escolha de técnicas adequadas para garantir um bom desempenho do classificador.

Os parâmetros necessários dentro de um sistema de classificação são muito diversificados.

Entre eles, pode-se citar o tamanho da janela de classificação, a extração e seleção dos atributos mais representativos das classes e outros.

A metodologia proposta é avaliada e testada através de experimentos utilizando-se amostras de dados reais e sintéticos.

1.4 ORGANIZAÇÃO

A seguir é demonstrada a organização do presente texto. O capítulo 2 apresenta o conceito de textura e as principais abordagens utilizadas para descrevê-la, bem como uma revisão bibliográfica dos principais estudos realizados em análise de textura.

O capítulo 3 sumariza as principais técnicas de classificação utilizadas em análise de imagens, abordando aspectos relevantes das metodologias de reconhecimento de padrões e dos classificadores mais comuns, tais como redes neurais, classificador de distância mínima, k -vizinhos mais próximos e máxima verossimilhança. Um novo classificador baseado em intervalos de decisão é proposto.

A metodologia proposta de classificação de imagens baseada em características texturais utilizando matrizes de co-ocorrência de níveis de cinza é descrita no capítulo 4.

Resultados experimentais envolvendo a análise e classificação de imagens com informação textural e a análise dos resultados são apresentados no capítulo 5.

Alguns comentários finais e propostas para trabalhos futuros são tratados no capítulo 6.

CAPÍTULO 2

TEXTURA

2.1 O QUE É TEXTURA?

Textura é uma das características mais importantes para classificar e reconhecer objetos e cenas, e pode ser caracterizada por variações locais em valores de pixels que se repetem de maneira regular ou aleatória ao longo do objeto ou imagem. Vários métodos para classificação de imagens utilizando características de textura têm sido propostos na literatura [3, 16, 24, 25, 30, 36, 41, 42, 44, 47, 52, 54, 56]. Entretanto, não há um método genérico ou uma abordagem formal que seja útil para uma grande variedade de imagens [1].

Haralick, Shanmugam e Dinstein [14] definiram textura como uniformidade, densidade, aspereza, regularidade, intensidade, entre outras características da imagem. No trabalho citado, os autores caracterizam textura como um conceito bidimensional, onde uma dimensão contém as propriedades primitivas da tonalidade e a outra corresponde aos relacionamentos espaciais entre elas. Eles indicaram que os conceitos de tonalidade e textura não são independentes, tal que em algumas imagens a tonalidade é dominante e, em outras, a textura é dominante.

As três abordagens principais usadas em classificação de imagens para a descrição de texturas são a estatística, a estrutural e a espectral, conforme descritas a seguir:

- *Abordagens Estatísticas:* a textura é definida por um conjunto de medidas locais extraídas do padrão. Medidas estatísticas comuns incluem entropia, correlação, contraste e variância;
- *Abordagens Estruturais:* utilizam a idéia de que texturas são compostas de primitivas dispostas de forma aproximadamente regular e repetitiva, de acordo com regras bem definidas. Como exemplo, pode-se citar a descrição da textura baseada em linhas paralelas regularmente espaçadas;

- *Abordagens Espectrais*: baseiam-se em propriedades do espectro de Fourier, sendo principalmente utilizadas na detecção de periodicidade global em uma imagem através da identificação de picos de alta energia no espectro.

Outros trabalhos que definem texturas de modo semelhante foram desenvolvidos. Um deles, como aponta Aksoy em [1], é o de Rosenfeld e Troy, que também define textura como um arranjo repetitivo de padrões sobre uma área e tenta medir a aspereza da textura observando fatores como as dependências de níveis de cinza, auto-correlação, etc.

A quantificação do conteúdo de textura é uma abordagem importante para a descrição de regiões [13]. Esses padrões homogêneos ou arranjos espaciais de pixels não podem ser descritos somente por intensidades locais ou cores. A figura 2.1 ilustra alguns exemplos de imagens em níveis de cinza obtidos da base de texturas Brodatz [6].

As texturas mostradas na figura 2.1 pertencem às classes: D1 (fio de cobre), D56 (junco), D95 (tijolo), D20 (tela), D9 (grama), D92 (pele de porco), D110 (papel com barbantes), D4 (cortiça prensada), D87 (lã), D57 (papel amassado), D111 (bolhas plásticas), D66 (grãos), D11 (corda), D103 (tecido grosseiro), D29 (areia), D15 (palha).

2.2 ABORDAGENS ESTATÍSTICAS

Os métodos estatísticos descrevem a textura de uma maneira adequada para o reconhecimento estatístico de padrões. Como resultado, cada textura é descrita por um vetor de características que representa um ponto em um espaço multi-dimensional de características. O objetivo é encontrar uma regra de decisão probabilística associando a textura à alguma classe específica [48].

As abordagens estatísticas não tentam compreender explicitamente a estrutura hierárquica da textura. Elas representam a textura indiretamente por propriedades não determinísticas que definem as distribuições e relacionamentos entre os níveis de cinza de uma imagem.

Métodos baseados em estatísticas de segunda ordem têm demonstrado alto poder de discriminação em diversos tipos de imagens. O mais popular deles é o de análise de textura utilizando matrizes de co-ocorrência, definido em [14]. Tal método é baseado na ocorrência

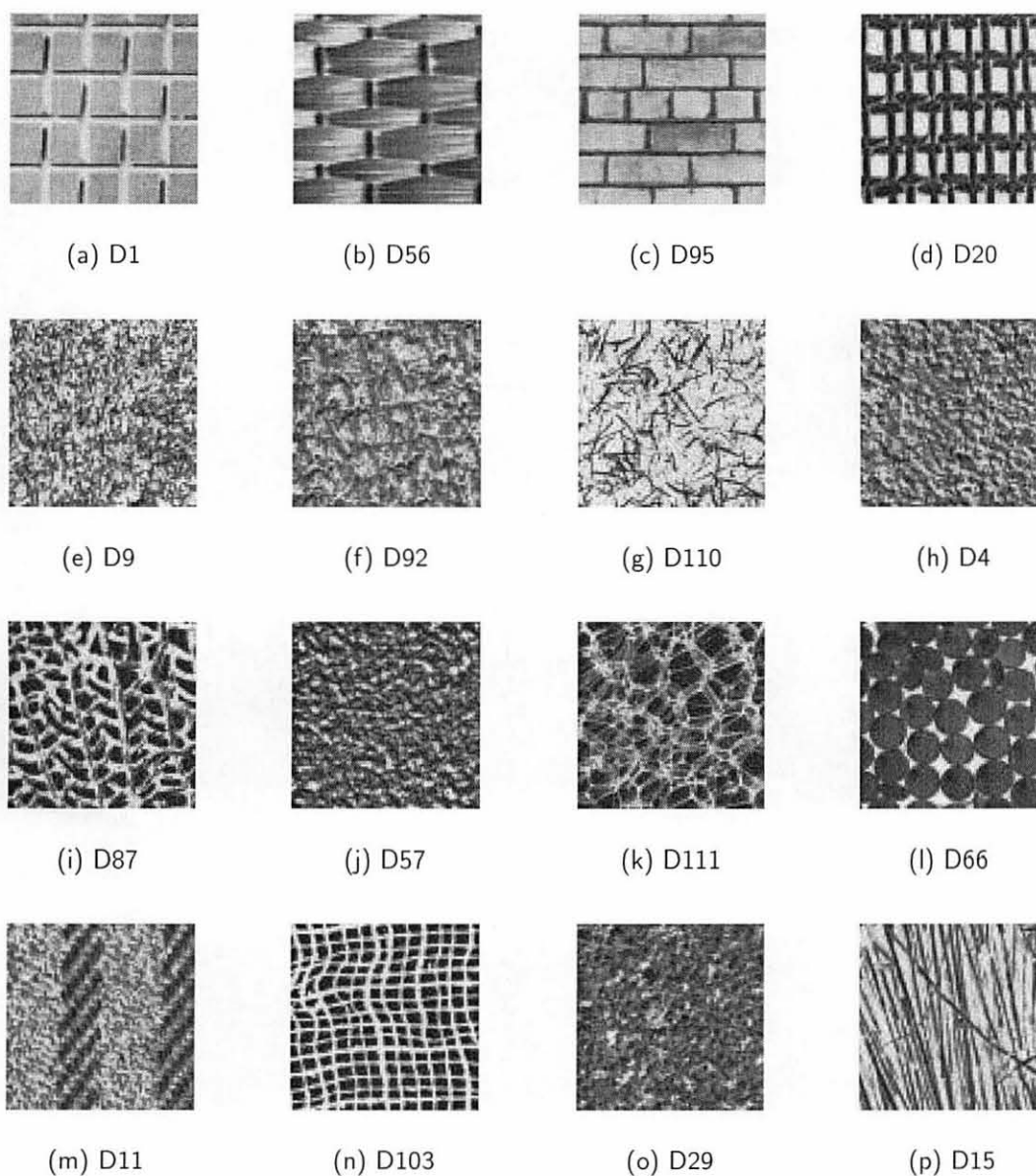


Figura 2.1: Exemplos de texturas do álbum Brodatz.

repetida da configuração de alguns níveis de cinza na textura, variando rapidamente com a distância em texturas finas e lentamente em texturas ásperas. O método de descrição de textura utilizando matrizes de co-ocorrência consiste em construir tais matrizes em determinadas direções e distâncias de uma imagem, e computar vetores de características para representar a textura presente na imagem. Por ser esse o objeto de estudo do presente trabalho, maiores detalhes serão descritos no capítulo 4 .

2.3 ABORDAGENS ESTRUTURAIS

As abordagens estruturais são consideradas apropriadas quando a primitiva de textura é grande o suficiente para ser individualmente segmentada e descrita. A primitiva de textura é a estrutura geométrica básica que forma a textura. Em texturas muito finas de imagens digitais, pode ser o próprio pixel.

Nesse método, a textura é definida por sub-padrões, chamados primitivas, que ocorrem repetidamente de acordo com regras bem definidas, dentro de um padrão [33].

Se for possível extrair de uma imagem um conjunto razoável de primitivas de textura, a descrição da textura pode ser feita a partir de estatísticas de propriedades dessas primitivas. Exemplos incluem a média ou desvio padrão dos níveis de cinza, área, perímetro, orientação, etc.

Uma idéia mais simples é extrair blocos de homogeneidade máxima e descrever a imagem em termos da estatística dos tamanhos dos blocos. Esta descrição pode ser hierárquica, ou seja, as primitivas podem ser compostas em sub-primitivas. A textura, nesse caso, é descrita através de gramáticas estocásticas, nas quais as regras a serem aplicadas são escolhidas com base em uma função densidade de probabilidade [39].

A relação espacial entre as primitivas de textura pode ser aleatória, ou um certo número de primitivas pode ser mutuamente dependente. As figuras 2.2(a) e (b) mostram que o mesmo número e o mesmo tipo de primitivas não necessariamente resultam na mesma textura. As figuras 2.2(a) e (c) mostram que o mesmo relacionamento espacial de primitivas não garante que as texturas sejam iguais e, portanto, não é suficiente para descrever texturas [48].

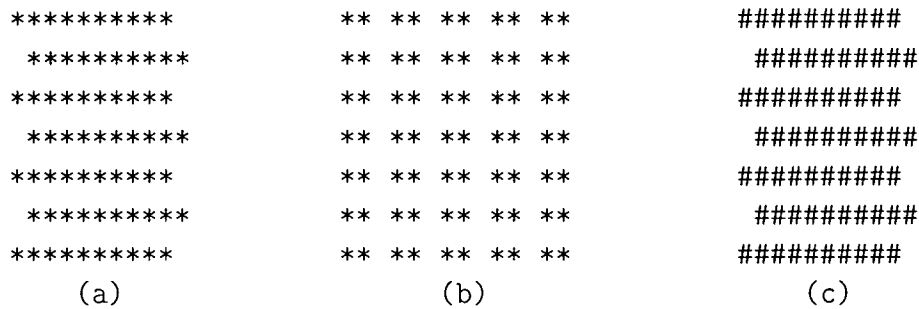


Figura 2.2: Exemplos de primitivas de textura.

Quando a textura é descrita através de uma abordagem estrutural, a idéia básica é a de que uma primitiva de textura simples possa ser usada na formação de padrões complexos de textura através de regras que limitem o número de arranjos possíveis [13].

Supondo que a primitiva de textura seja um círculo e o significado de cada primitiva seja a = círculo à direita, b = círculo abaixo e c = círculo à esquerda, um conjunto de regras pode ser definido. Considerando que temos uma regra da forma $S > aS$, que indica que o símbolo S pode ser reescrito como aS e, aplicando-se três vezes essa regra seria formada a cadeia $aaaS$. Acrescentando-se novas regras a esse esquema: $S - aS$, $S - bA$, $A - cA$, $A - c$, $A - bS$, $S - a$, forma-se um conjunto de regras a partir do qual, é possível gerar a cadeia $aaabccbaa$ que corresponde a uma matriz 3×3 de círculos, conforme mostra a figura 2.3. Maiores padrões de textura podem ser criados facilmente da mesma maneira, gerando estruturas de formas bem variadas.

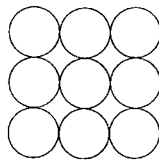


Figura 2.3: Exemplo de formação de um padrão de textura estrutural.

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

2.4 ABORDAGENS ESPECTRAIS

Os métodos no domínio espacial referem-se ao próprio plano da imagem, sendo baseados na manipulação direta dos pixels. A transformada de Fourier transforma imagens do domínio espacial para o domínio de frequência, levando-se em conta que certas operações são executadas mais facilmente e com maior rapidez no domínio de frequência.

O espectro de Fourier é idealmente adaptado para a descrição da orientação de padrões periódicos ou quase periódicos em uma imagem. Esses padrões globais de textura, embora facilmente distinguíveis como concentrações de agrupamentos de alta-energia no espectro, são geralmente difíceis de se detectar com métodos espaciais devido à natureza local dessas técnicas [13].

Para descrever texturas através dessa abordagem utilizam-se três características: (1) picos proeminentes no espectro fornecem a direção dos padrões de textura; (2) a posição dos picos no plano de frequência fornece o período espacial fundamental dos padrões; e (3) a eliminação de quaisquer componentes periódicos através de filtragem mantém os elementos não periódicos na imagem, que podem ser descritos por técnicas estatísticas.

O espectro de uma imagem real é simétrico em torno da origem, de maneira que apenas metade do plano frequencial necessita ser considerado. Portanto, para propósito de análise, cada padrão periódico é associado a apenas um pico no espectro, ao invés de dois.

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 θ são variáveis nesse sistema de coordenadas. Para cada direção θ , $S(r, \theta)$ 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 θ 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

$$S(r) = \sum_{\theta=0}^{\pi} S_{\theta}(r) \quad (2.1)$$

$$S(\theta) = \sum_{r=1}^R S_r(\theta) \quad (2.2)$$

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 2.1 e 2.2 constituem um par de valores $[S(r), S(\theta)]$ para cada par de coordenadas (r, θ) . 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 2.4 ilustra o uso das equações 2.1 e 2.2 para a descrição global de textura. A figura 2.4(a) mostra uma imagem com uma textura periódica, enquanto a figura 2.4(b) apresenta o espectro correspondente. As figuras 2.4(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° , correspondendo claramente à periodicidade no conteúdo de textura da imagem.

Para ilustrar como o gráfico de $S(\theta)$ poderia ser usado para diferenciar dois padrões diferentes de textura, a figura 2.5(a) mostra outra imagem cujo padrão de textura está predominantemente nas orientações horizontal e vertical. A figura 2.5(b) mostra o gráfico de $S(\theta)$ para o espectro dessa imagem. Esse gráfico mostra picos em intervalos de 90° . 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 [33].

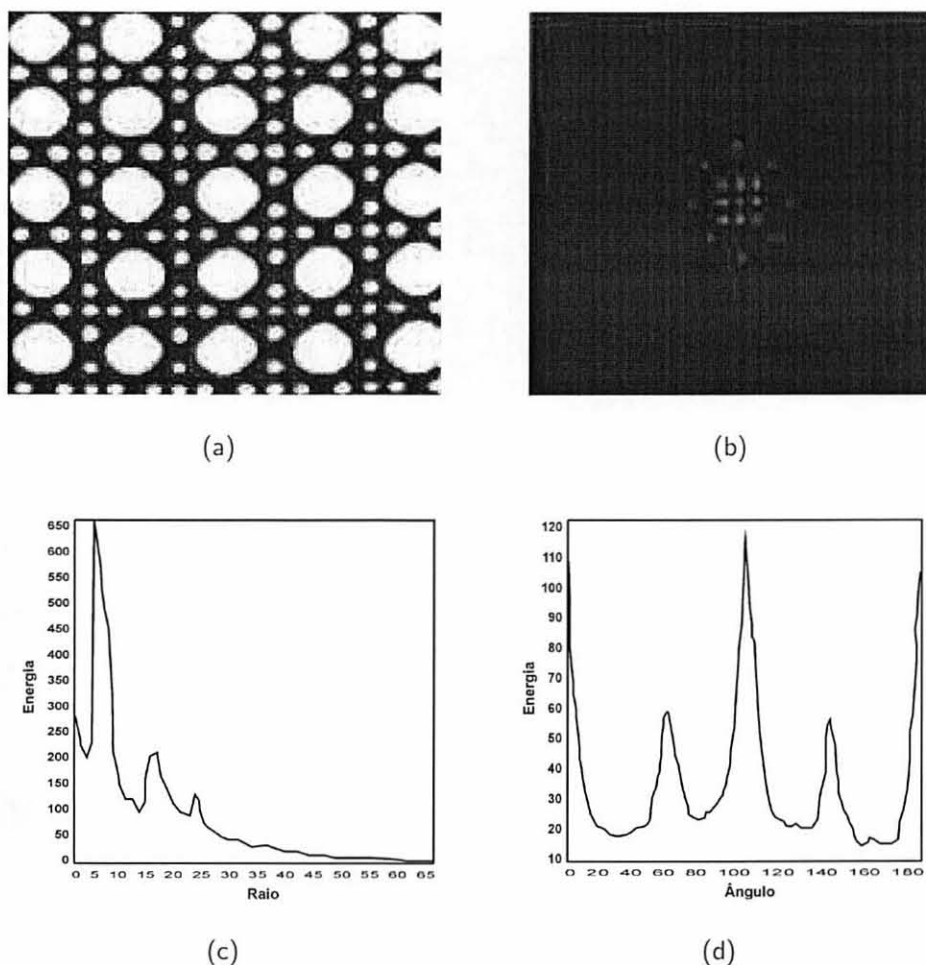


Figura 2.4: (a) Imagem mostrando uma textura periódica; (b) espectro; (c) gráfico de $S(r)$; (d) gráfico de $S(\theta)$.

2.5 OUTRAS ABORDAGENS

Existem ainda diversas outras abordagens propostas para descrever textura. A seguir serão citadas algumas delas, explicando brevemente seus princípios.

Como a textura é altamente dependente da escala, pode-se diminuir seu grau de sensibilidade, descrevendo-a em múltiplas resoluções. Uma escala apropriada pode ser escolhida para se obter a máxima discriminação da textura. As *transformadas de Gabor* e *wavelets* são adequadas para a caracterização de textura multi-escalar [48]. 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

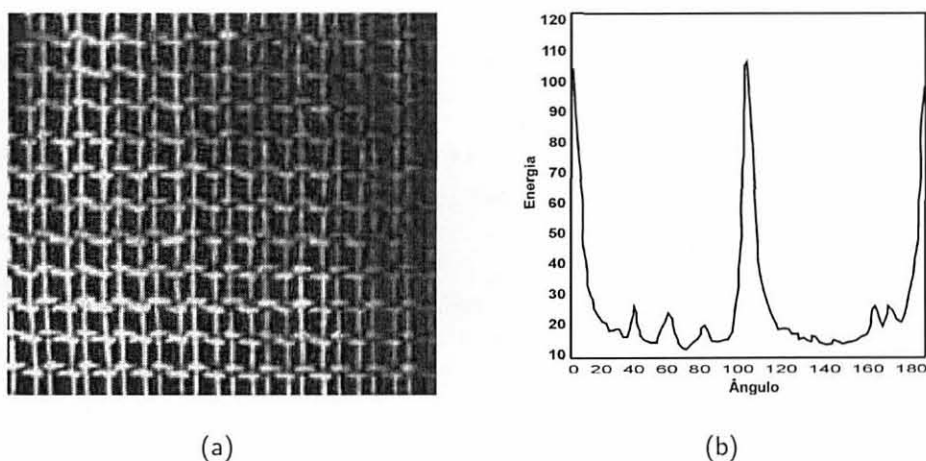


Figura 2.5: (a) Imagem contendo um tipo diferente de textura periódica; (b) gráfico de $S(\theta)$.

é limitada na prática porque não existem filtros de resolução simples que possam localizar uma estrutura espacial em texturas naturais. Comparados com a *transformada 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 mais adequada para a análise de textura em uma aplicação específica. Isso torna a transformada *wavelet* apropriada para segmentação de textura [26].

Existem também métodos híbridos que combinam mais de uma abordagem, tornando a descrição de textura parcialmente pertencente a uma ou outra técnica. Um exemplo chamado *descrição de textura multi-nível* é baseado na definição de primitivas e na descrição espacial de relações entre as primitivas. O método considera propriedades de intensidades de cinza e estruturais, consistindo em vários passos. Primeiramente, primitivas de textura devem ser extraídas, descritas e classificadas. Como resultado desse estágio de processamento, um classificador aprende a classificar primitivas de textura. Texturas conhecidas são apresentadas ao sistema de reconhecimento em uma segunda etapa de aprendizado. Primitivas de textura são extraídas da imagem e o classificador de primeiro nível reconhece suas classes. A partir das primitivas de textura reconhecidas, relacionamentos espaciais entre classes de primitivas são avaliados para cada textura do conjunto de treinamento. Relacionamentos espaciais entre primitivas de textura são descritos por meio de um vetor de características usado para ajustar um

classificador de segundo nível. Se o classificador de segundo nível é selecionado, o processo de aprendizado do segundo nível está pronto, e texturas desconhecidas podem ser apresentadas ao sistema de reconhecimento. As primitivas são classificadas pelo classificador de primeiro nível, propriedades de primitivas espaciais são calculadas e o classificador de segundo nível associa a textura à uma das classes de textura [48].

2.6 ANÁLISE DE TEXTURA

A análise de textura em uma imagem é extremamente importante. Seu objetivo é 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. Antes de analisar a textura de uma imagem deve-se considerar a seguinte questão: existe algum tipo de textura na imagem?

Karu *et al* [21] respondeu essa questão em seu trabalho determinando quando uma dada imagem possui textura para análise. A maioria das definições de textura trata-a como uma medida de aspereza em uma imagem. Imagens que não possuem textura podem ser classificadas como imagens com ruído ou imagens contendo objetos distinguíveis muito grandes para caracterizarem textura. Então, para uma imagem conter textura, ela deveria estar entre esses dois extremos.

Em [23], os autores propuseram o uso de características geométricas (nebulosas - *fuzzy*) para classificação por textura. Em sua abordagem, primeiro detecta-se na imagem um conjunto de valores 2D em extremos locais. Eles são usados como sementes de regiões e crescem enquanto elas não se encontram com outras regiões. As regiões resultantes são chamadas de regiões de influência. Um número de características é calculado a partir do conjunto incluindo área *fuzzy*, perímetro, solidez, altura e largura formando a classificação de textura. Em um total de 16 imagens extraídas do álbum Brodatz de tamanho 128×128 pixels, foi obtida a classificação correta de 90% dos pixels envolvidos.

Em [8] e [9], foi proposto o uso de pilhas binárias para a caracterização de textura. O método de características geométricas estatísticas (*Statistical Geometric Features - SGF*) é baseado na geração de L diferentes imagens binárias a partir da imagem original que possui um

total de L níveis de cinza, usando a cada etapa um limiar um pouco maior para segmentação. Uma coleção das imagens resultantes é chamada de pilha binária correspondendo à imagem original.

Para cada imagem binária, todos os valores 1 de pixels são agrupados dentro de regiões conectadas. O mesmo é feito para todos os valores 0 de pixels. Para cada uma das regiões conectadas, medidas de irregularidade ou circularidade, definidas de maneira diferente, são calculadas. Para cada imagem binária, as quatro características seguintes são computadas: o número de regiões 1 conectadas, o número de regiões 0 conectadas e as duas medidas de irregularidade para essas regiões conectadas. Para cada uma dessas quatro características, derivam-se outras quatro seguintes: valor máximo, valor médio, média da amostra e desvio padrão da amostra, alcançando um total de 16 características. Tais características são utilizadas para classificação da textura.

Estudos comparam tal abordagem com matrizes de características estatísticas (*Statistical Feature Matrix - SFM*) e características de Fourier utilizando o conjunto completo de 112 imagens do álbum Brodatz. Os resultados são gerados usando o classificador dos k -vizinhos mais próximos. O desempenho de reconhecimento encontrado foi: SGF (85.6%), SFM (72.8%) e descritores Fourier (32.7%).

Haralick e seus colaboradores propuseram em [14] uma das mais importantes técnicas para classificação de imagens utilizando textura. Esse estudo está relacionado com a tarefa de desenvolver um conjunto de características para classificação e categorização de dados pictóricos. A textura é escolhida como a mais importante característica para representar imagens. Textura contém informações importantes sobre o arranjo estrutural das unidades e seus relacionamentos. Essa informação discriminatória pode ser usada para classificar imagens.

Os autores apresentam um procedimento geral para extrair propriedades de textura de blocos de imagens. Essas características são calculadas no domínio espacial e assumiu-se que a informação de textura em uma imagem I está contida na média do relacionamento espacial que os níveis de cinza na imagem têm uns com os outros.

Um conjunto de matrizes de distribuição de probabilidade da dependência espacial dos níveis de cinza é calculado e um conjunto de quatorze características de textura, que podem ser

extraídas de cada uma dessas matrizes, é proposto. As matrizes são construídas assumindo-se que cada pixel, exceto os externos, possui 8 vizinhos (horizontalmente, verticalmente e diagonalmente em 45 graus). Também assumiu-se que a informação do contexto da textura está adequadamente especificada pela matriz de frequências relativas. Algumas dessas medidas de características de textura, obtidas das matrizes, relatam específicas características texturais da imagem, tais como a homogeneidade, contraste e a presença de uma estrutura organizada dentro da imagem.

Outras medidas caracterizam a complexidade das transições dos níveis de cinza que ocorrem na imagem. A utilidade das características de textura para categorizar imagens foi testada em três conjuntos de imagens.

No primeiro caso, fotografias microscópicas de pedras de areia, importantes na indústria petrolífera, foram categorizadas. O conjunto de dados era constituído de 243 imagens divididas em 5 classes. Tal conjunto foi dividido em dados de treinamento e dados de teste. Um conjunto de 8 variáveis, englobando a média e a variância de características de textura obtidas das matrizes, foi utilizado para a classificação. O resultado obteve 89% de precisão.

No segundo caso, um conjunto de fotos aéreas foi classificado usando a regra de decisão *mini-max*. O conjunto de dados consistia em 170 imagens divididas em 8 categorias. Quatro dependências espaciais de níveis de cinza e 11 características de textura foram usadas na classificação que resultou em 82.3% de classificação correta. No terceiro caso, um conjunto de imagens de satélite foi classificado. As 624 amostras de tamanhos iguais foram divididas em conjuntos de treinamento e de teste. A média e a variância de 4 características de textura e 8 características espectrais foram usadas para a classificação. O resultado foi de 83.5% de classificação correta em oposição a 74-77% de classificação correta obtida usando somente características espectrais.

Essa combinação de características espectrais e texturais mostra que ocorre uma melhoria significativa na precisão de classificação quando as características de textura são usadas como entrada para o classificador. Desde a publicação desse trabalho, as características sugeridas têm sido utilizadas na maioria dos estudos relacionados à textura. A importância teórica e visual dessas medidas é apresentada por Baraldi e Parmigianni [4].

Uma descrição sobre a importância do uso de características que são invariantes sobre transformações monotônicas de níveis de cinza também é apresentada em [14]. Isso ocorre devido ao fato de que a textura é independente da intensidade de cinza. 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 as 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.

As seguintes observações extraídas de [42] mostram a importância da textura na classificação de imagens:

- A precisão de classificação combinando textura com abordagens estruturais e espectrais é estatisticamente melhor do que utilizar tais abordagens independentemente, especialmente para tamanhos de janela maiores que 3 pixels. Esse tamanho de janela de 3 pixels é muito pequeno e altamente variável para expressar as propriedades características de cada classe em consideração.
- Existe um tamanho ótimo de janela para cada uma das características extraídas. 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 melhoria na precisão, podendo também levar a interpretações incorretas pelo fato de englobar mais do que uma classe de textura.
- No tamanho ótimo de janela, a maioria das características fornece precisão de classificação estatisticamente similar.
- Para os diferentes níveis de quantização radiométricos, 16, 32, 64 e 128, foram obtidos resultados de classificação estatisticamente similares.

- Tratando-se de imagens espectrais, o uso de diferentes bandas espectrais para o cálculo de características de textura não produz resultados estatisticamente diferentes de classificação.
- O uso da média de características de textura determina resultados estatisticamente melhores do que quando uma simples orientação é usada.
- Precisoões de classificação similares foram obtidas para métricas de distância de valor 1 até 4. Esta observação se deve ao fato de que para distâncias maiores, as características de textura são altamente correlacionadas.

Em [31], as matrizes de co-ocorrência foram apresentadas na forma de imagens para observar as diferenças ou similaridades na representação das características de textura. Cada elemento da matriz de co-ocorrência foi considerado como um pixel, onde seu nível de cinza era o valor numérico do elemento correspondente. Algumas amostras foram extraídas de classes de textura e as matrizes de co-ocorrência calculadas. Foram computados quatro casos, utilizando os seguintes números de níveis de cinza: 256×256 , 64×64 , 16×16 e 4×4 . Observou-se que cada matriz possui um padrão distinto, embora algumas delas mostraram uma certa similaridade. Com a diminuição do tamanho da matriz, houve também diminuição na quantidade de informação e, tornou-se mais difícil distinguir entre algumas matrizes do mesmo grupo. Entretanto, nota-se que mesmo com 4×4 , a maioria das matrizes preservou informação suficiente para deixá-las distinguíveis das demais.

Um método interessante para melhorar a qualidade das características das matrizes de co-ocorrência foi proposto em [53]. Os autores classificaram as características propostas por Haralick e seus colaboradores [14] 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 importantes. 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.

Uma metodologia proposta em [2] consiste na caracterização das classes texturais presentes na imagem por meio das frequências espaciais mais significativas, conforme identificadas nas amostras de cada classe. Cada frequência 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 em geral ocasionará resultados pobres no processo de classificação, restando os atributos de textura para se obter uma classificação com maior exatidão.

Segundo [35], a textura contém mais informações do que as cores. Nesse trabalho foram comparadas as matrizes de co-ocorrência e os mapas auto-organizáveis através de um classificador que obteve melhores resultados em imagens coloridas utilizando as matrizes de co-ocorrência.

Em [55], foi proposta uma aplicação de análise de textura baseada em matrizes de co-ocorrência para inspeção de materiais. A idéia do estudo era utilizar métodos de textura que são usados para avaliar a qualidade da superfície de um material para distinguir entre amostras pobres e aceitáveis. Para isso, considerou-se 12 imagens digitais de superfícies de materiais. 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 3 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 à propriedade visual óbvia do material.

Estatísticas de terceira ordem ou superior também têm sido aplicadas em análise de textura. O estudo em [32] propõe uma nova técnica para classificação de texturas que é baseada no uso de estatísticas de ordem mais alta (*Higher Order Statistics - HOS*) para definir características discriminatórias. Os autores definem funções *HOS* e suas propriedades para a classificação 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.

A maior vantagem de utilizar atributos de textura é obviamente sua facilidade. Entretanto, eles podem não caracterizar completamente a informação de textura. Julesz [19] encontrou através de seus famosos experimentos em percepção humana visual de textura que, para uma grande classe de texturas, “nenhuma textura pode ser discriminada se ela não está de acordo com sua estatística de segunda ordem”. Segundo ele, um sistema não pode processar globalmente estatísticas de terceira ou mais altas ordens. Mesmo se contra-exemplos forem encontrados para essa hipótese, a importância de estatísticas de segunda ordem é comprovada. Então, o melhor método estatístico usado em análise de textura é baseado nas distribuições de probabilidade da união de pares de pixels [26].

2.7 TEXTURA × PADRÃO

A textura geralmente é confundida ou utilizada como um sinônimo de padrão. Neste trabalho, o termo *padrão* é usado para descrever alguma regularidade espacial dentro de uma área de interesse. Por outro lado, *textura* é uma medida da variabilidade espacial, aleatória ou não

aleatória, dentro de uma área de interesse. Muitas medidas de textura podem não distinguir padrões de outras texturas, visto que a técnica não é espacial dentro da escala de medida.

As tabelas 2.1 e 2.2 mostram como algumas medidas de textura podem não distinguir um padrão. Se os números são simplesmente somados em cada tabela, a soma de ambos é igual a 240, sendo esse número independente da ordem em que os valores ocorrem na matriz. Isso ocorre com medidas de textura utilizando estatísticas de primeira ordem.

Medidas de textura que usam estatísticas de ordem mais alta, como a matriz de co-ocorrência, são explicitamente espaciais e a ordem dos valores torna-se significativa. Maiores detalhes sobre medidas de segunda ordem serão citados nos capítulos seguintes, por serem objetos de estudo deste trabalho.

5	5	10	5	5
5	20	10	20	5
10	10	20	10	10
5	20	10	20	5
5	5	10	5	5

Tabela 2.1: Padrão 1 - Soma dos valores igual a 240.

20	5	10	5	5
5	20	10	5	20
10	5	5	10	5
5	20	20	10	5
10	5	10	10	5

Tabela 2.2: Padrão 2 - Soma dos valores igual a 240.

CAPÍTULO 3

TÉCNICAS DE CLASSIFICAÇÃO

Classificação é o processo de extração de informação em imagens para reconhecer padrões ou objetos homogêneos que são utilizados para mapear áreas que correspondem a temas de interesse [49].

Quando existem regiões da imagem em que se dispõe de informações que permitem a identificação de uma classe de interesse, a classificação é chamada *supervisionada*. Neste caso, o padrão de entrada é identificado como um membro de uma classe pré-definida pelos padrões de treinamento, que são rotulados com suas classes.

Para um treinamento supervisionado, deve-se identificar na imagem uma área representativa de cada classe. É importante que a área de treinamento seja uma amostra homogênea da respectiva classe, mas ao mesmo tempo, deve incluir toda a variabilidade dos níveis de cinza. Recomenda-se adquirir mais de uma área de treinamento, utilizando o maior número de informações disponíveis, como trabalhos de campo, mapas, fotografias, etc. O número de pixels de treinamento necessários para a precisão do reconhecimento de uma classe aumenta com o aumento da variabilidade entre as classes.

Quando o padrão é associado a uma classe que é aprendida com base na similaridade entre os padrões de treinamento, ou seja, são utilizados algoritmos para reconhecer as classes presentes na imagem, a classificação é chamada *não-supervisionada*. Ao definir áreas para o treinamento não-supervisionado, não é necessário se preocupar com a homogeneidade das classes. As áreas escolhidas devem ser heterogêneas para assegurar que todas as possíveis classes e suas variabilidades sejam incluídas. Os pixels dentro de uma área de treinamento são submetidos a um algoritmo de agrupamento (*clustering*) que determina o grupo dos dados em um arranjo espacial de dimensão igual ao número de bandas presentes. Esse algoritmo assume que cada grupo (*cluster*) representa a distribuição de probabilidade de uma classe.

3.1 METODOLOGIAS DE RECONHECIMENTO DE PADRÕES

As técnicas de reconhecimento de padrões são inúmeras e podem ser aplicadas em vários domínios, porém, ainda não se chegou a uma abordagem individual ótima e específica para uma determinada aplicação. Portanto, métodos e abordagens múltiplas devem ser utilizadas, combinando várias modalidades de sensores e métodos de classificação. O projeto de um sistema de reconhecimento de padrões essencialmente envolve três etapas: aquisição de dados e pré-processamento, representação dos dados e tomada de decisões. Assim, geralmente o maior desafio encontra-se na escolha de técnicas para realizar essas três etapas.

As abordagens mais conhecidas para o reconhecimento de padrões são: casamento de padrões, abordagem estatística, redes neurais e lógica nebulosa (*fuzzy*). A seguir é feita uma descrição a respeito de cada uma delas, explicando seus principais fundamentos.

3.1.1 Casamento de Padrões

Casamento de padrões (*Template Matching*) é uma das abordagens de reconhecimento mais antigas e mais simples. De maneira geral, dispõe-se de um modelo (*template*) que é um protótipo do padrão a ser reconhecido. No caso de imagens, o modelo é uma forma 2D. Para classificar, é realizado o casamento do modelo com a imagem que deve conter o padrão a ser reconhecido, levando em consideração variações como translação, rotação e mudanças na escala. A medida de similaridade, normalmente relacionada à idéia de correlação, deve ser otimizada com base no conjunto de treinamento.

A maior desvantagem dessa abordagem está em seu custo computacional. No caso do uso de *templates* rígidos, ocorrem falhas ao reconhecer padrões distorcidos ou com grande variação intra-classe. Os modelos de *templates* deformáveis podem ser usados quando as deformações não podem ser modeladas diretamente.

Maiores detalhes a respeito dessa abordagem podem ser encontrados em [51].

3.1.2 Abordagem Estatística

Na abordagem estatística, cada padrão é representado em termos de características (*features*) ou atributos, sendo visto como um ponto em um espaço d -dimensional (Figura 3.1).

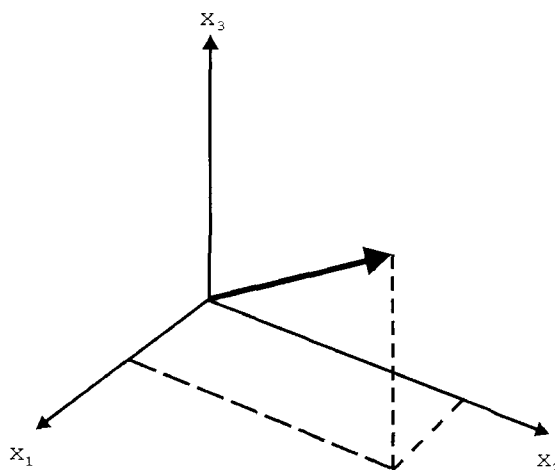


Figura 3.1: Exemplo de padrão em um espaço tridimensional.

O objetivo é fazer com que esse espaço seja gerado de modo que suas características permitam que vetores de padrões pertencentes a categorias diferentes fiquem dispostos de maneira a simplificar o processo de classificação (Figura 3.2). Por exemplo, pode ser desejável que os vetores ocupem regiões compactas e disjuntas em um espaço de características de d dimensões. A efetividade do espaço de características é determinada pelo grau com que os padrões de classes diferentes podem ser separados.

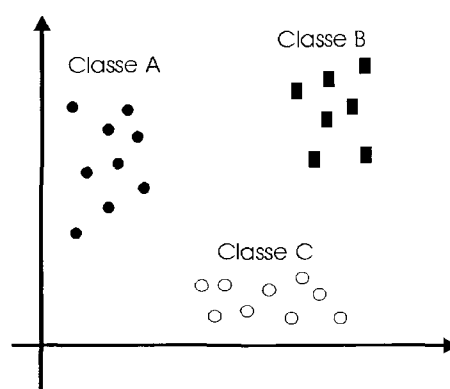


Figura 3.2: Exemplo de espaço de características.

Assume-se que cada padrão é uma observação obtida aleatoriamente segundo uma certa probabilidade condicional a uma determinada classe. Dado um conjunto de padrões de treina-

mento contendo elementos de todas as classes, o objetivo é estabelecer fronteiras de decisão no espaço de características que separem os padrões de classes diferentes.

Os padrões são tratados como vetores aleatórios pois um padrão desconhecido pode ser o representante de uma classe conhecida que sofreu alterações aleatórias devido a ruídos oriundos do método de aquisição (sensores), da influência de outros fatores externos ou mesmo dos mecanismos de extração de características intrínsecos ao sistema de reconhecimento.

Na teoria de decisão estatística, a fronteira de decisão é determinada para cada classe, que pode ser especificada (aprendizado supervisionado) ou aprendida (aprendizado não-supervisionado). Normalmente é aplicado o aprendizado supervisionado, isto é, as classes das imagens de treinamento possuem um rótulo que as identifica.

Para efetuar a classificação, também pode-se usar uma abordagem baseada em análise de funções discriminantes: primeiramente, uma forma paramétrica da fronteira de decisão é especificada (por exemplo, linear ou quadrática); posteriormente, a melhor fronteira de decisão da forma especificada é encontrada com base na classificação dos padrões de treinamento. Tais fronteiras podem ser construídas usando, por exemplo, o critério do erro médio quadrático.

A dificuldade no estabelecimento de funções de probabilidade adequadas para as variáveis envolvidas nos processos de classificação estatísticos é em geral um fator de descontentamento no desenvolvimento desses sistemas. A grande quantidade de dados necessários, o levantamento das estatísticas sobre os dados obtidos e, em certas ocasiões, as informações insuficientes, são responsáveis pela não obtenção de um bom desempenho de classificação.

O paradigma pode ser resumido da seguinte forma:

- Cada padrão é representado por um vetor de características

$$\mathbf{x} = (x_1, x_2, \dots, x_n) \quad (3.1)$$

- Um dado padrão deve ser classificado em apenas uma classe com base no seu vetor de características;
- Assume-se que o vetor de características possui uma *assinatura* típica da sua classe;

- Conceitos da teoria estatística de decisão são usados para estabelecer fronteiras de decisão entre as classes.

Em [18], foi realizado um breve levantamento dentre os artigos publicados na revista *IEEE Transactions on Pattern Analysis and Machine Intelligence* e descobriu-se que, desde 1979, 350 artigos enquadravam-se na área de reconhecimento de padrões, dos quais 300 relacionavam-se à abordagem estatística. Um dos principais fatores que tornam os métodos estatísticos tão utilizados é a maneira como as imagens são geradas e armazenadas. Geralmente, os sinais registrados por sensores sofrem influência de uma grande quantidade de fatores aleatórios como, por exemplo, a atmosfera quando o sensor é um satélite.

Existem várias técnicas estatísticas de reconhecimento de padrões, cada qual com seu nicho de aplicações [5, 33, 7].

3.1.3 Redes Neurais

As Redes Neurais fundamentam-se em estudos sobre a estrutura do cérebro humano, constituindo um sistema de circuitos que tentam simular sua forma inteligente de processar informação, inclusive seu comportamento, aprendendo, errando e fazendo novas descobertas.

Estima-se que a quantidade de neurônios existentes no cérebro humano esteja na casa dos bilhões, constituindo o mais fascinante processador baseado em carbono existente. Todas as tarefas do organismo como aquisição, armazenamento, representação, recuperação e análise de dados sensoriais muito complexos estão relacionados ao funcionamento dessas pequenas células. Os neurônios estão conectados uns aos outros por meio de sinapses, e juntos formam uma grande rede massivamente paralela chamada Rede Neural.

Alguns estudos da neurofisiologia consideram que a riqueza computacional do cérebro humano está associada ao grande número de neurônios, interconectados por uma rede complexa de sinapses.

As sinapses transmitem estímulos, disparados em velocidades medidas em milissegundos, através de diferentes concentrações de sódio e potássio, e o resultado disto pode ser estendido por todo o corpo humano. Essa grande rede proporciona uma enorme capacidade de processamento e armazenamento de informação. A sofisticada manipulação de dados no cérebro

permite construir imagens mentais de pessoas, lugares e eventos baseando-se em fragmentos de informação, como sons ou cheiros familiares, ou simplesmente ao lembrar de uma determinada data de calendário [50].

Os principais componentes dos neurônios são (figura 3.3):

- Os dendritos, cuja função é receber os estímulos transmitidos pelos outros neurônios;
- O corpo de neurônio, também chamado de soma, que é responsável por coletar e combinar informações vindas de outros neurônios;
- O axônio, que é constituído de uma fibra tubular que pode alcançar até alguns metros, e é responsável por transmitir os estímulos para outras células.

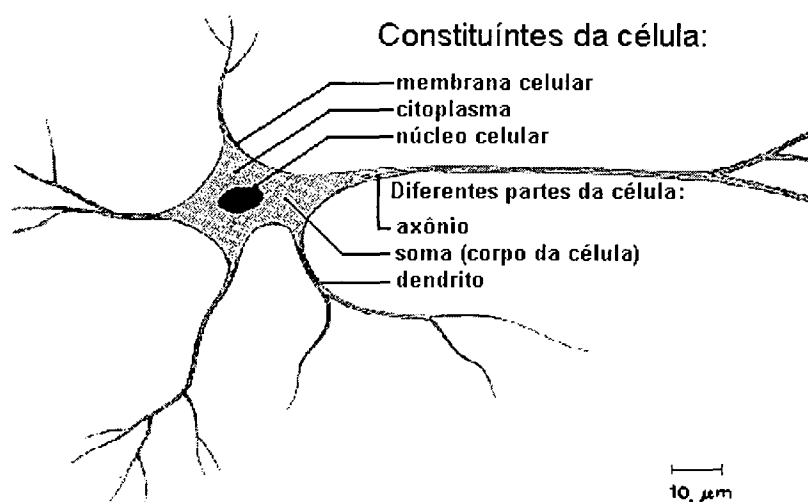


Figura 3.3: Esquema dos constituintes de uma célula neural.

O sistema nervoso é formado por um conjunto extremamente complexo de neurônios. Nos neurônios, a comunicação é realizada através de impulsos. Quando um impulso é recebido, o neurônio o processa e, passado um limite de ação, dispara um segundo impulso que produz uma substância neurotransmissora, o qual flui do corpo celular para o axônio, que por sua vez pode ou não estar conectado a um dendrito de outra célula. O neurônio que transmite o pulso pode controlar a frequência de pulsos aumentando ou diminuindo a polaridade na membrana pós-sináptica. Eles têm um papel essencial na determinação do funcionamento, comportamento e do raciocínio do ser humano.

Ao contrário das redes neurais artificiais, redes neurais naturais não transmitem sinais negativos; sua ativação é medida pela frequência com que emite pulsos, frequência esta de pulsos contínuos e positivos. As redes naturais não são uniformes como as redes artificiais, e apresentam uniformidade apenas em alguns pontos do organismo. Seus pulsos não são síncronos ou assíncronos, devido ao fato de não serem contínuos, o que a difere de redes artificiais. Uma grande rede neural artificial pode ter centenas ou milhares de unidades de processamento, enquanto que o cérebro de um mamífero pode ter muitos bilhões de neurônios.

3.1.3.1 Histórico

O surgimento das redes neurais artificiais teve seu marco inicial na década de 40 pelo neurofisiologista McCulloch e pelo matemático Walter Pitts da Universidade de Illinois [27], os quais publicaram um trabalho sobre os neurônios formais fazendo uma analogia entre células nervosas vivas e o processo eletrônico. O trabalho consistia em um modelo de resistores variáveis e amplificadores representando conexões sinápticas de um neurônio.

Em 1949, houve um avanço com a publicação do livro *The Organization of Behavior*, do psicólogo Donald Hebb [15] que descobriu a base de aprendizado nas redes neurais quando explicou o que ocorre em nível celular, durante o processo de aprendizagem no cérebro. A lei de aprendizagem de Hebb diz que, se um neurônio A é repetidamente estimulado por outro neurônio B , ao mesmo tempo que ele está ativo, ele ficará mais sensível aos estímulos de B , e a conexão sináptica de B para A será mais eficiente. Desse modo, B achará mais fácil estimular A para produzir uma saída.

Em 1954, Minsky [28] abordou um problema relacionado com o aprendizado em sua tese de doutorado em redes neurais. Durante meados dos anos 1950 e início dos anos 1960, uma classe das chamadas máquinas que aprendem, criada por Rosenblatt [38], causou grande interesse entre pesquisadores da teoria de reconhecimento de padrões. A razão do grande interesse nessas máquinas, chamadas *perceptrons*, foi o desenvolvimento de provas matemáticas mostrando que os perceptrons, quando treinados com conjuntos de treinamento linearmente separáveis, convergiriam para uma solução em um número finito de passos iterativos. A solução tomaria a forma de coeficientes de hiperplanos capazes de separar corretamente as classes representadas

pelos padrões no conjunto de treinamento.

Minsky e Papert [29], em 1969, utilizaram conceitos matemáticos para demonstrar a limitação das máquinas do tipo *perceptron* de única camada. Essa visão foi mantida até meados da década de 80 quando ressurgiram os interesses na área.

O desenvolvimento de um novo método de treinamento para *perceptrons* multicamadas em 1986, por Rumelhart, Hinton e Williams [40], chamado de aprendizado por retropropagação (*backpropagation*), causou um grande impacto científico. O algoritmo tem sido usado com sucesso em vários problemas de interesse prático. Esse sucesso estabeleceu as máquinas do tipo *perceptron* com múltiplas camadas como um dos principais modelos de redes neurais atualmente em uso. Muitos outros trabalhos seguiram a esse fazendo com que as redes neurais representem hoje uma área de pesquisa multidisciplinar.

3.1.3.2 Redes neurais artificiais

Inúmeras definições ou interpretações tentam demonstrar o que realmente representa uma rede neural artificial (RNA), visto que sua grande complexidade não permite uma única definição.

Pode-se dizer que são técnicas computacionais para solucionar problemas de inteligência artificial que apresentam um modelo inspirado na estrutura neural de organismos inteligentes e que adquirem conhecimento através da experiência.

Uma rede neural artificial é um modelo de grafo orientado disposto em uma rede em que os nós representam neurônios artificiais e as arestas orientadas são as conexões entre as entradas e saídas dos neurônios. Redes neurais podem ser vistas como máquinas massivamente paralelas com muitos processadores simples e muitas interconexões, possuindo algumas restrições e definições próprias.

Os nós desse grafo são chamados elementos de processamento. Suas arestas são conexões que funcionam como caminhos de condução instantânea de sinais em uma única direção, de forma que seus elementos de processamento podem receber qualquer número de conexões de entrada. Essas estruturas podem possuir memória local e também possuir qualquer número de conexões de saída, desde que os sinais nessas conexões sejam os mesmos. Portanto, esses elementos têm na verdade uma única conexão de saída, que pode dividir-se em cópias para

formar múltiplas conexões, sendo que todos carregam o mesmo sinal.

Então, a única entrada permitida para a função de transferência (que cada elemento de processamento possui) são os valores armazenados na memória local do elemento de processamento e os valores atuais dos sinais de entrada nas conexões recebidas pelo elemento de processamento. Os únicos valores de saída permitidos a partir da função de transferência são valores armazenados na memória local do elemento de processamento, e o sinal de saída do mesmo.

A função de transferência pode operar continuamente ou esporadicamente. No segundo caso, deve existir uma entrada chamada de ativação que estimula a função de transferência com o sinal de entrada corrente e com valores da memória local, e produzir um sinal de saída atualizado (ocasionalmente alterando valores da memória). No primeiro caso, os elementos estão sempre ativados, e a ativação chega através de uma conexão de um elemento de processamento que também é parte da rede.

Sinais de entrada para uma rede neural a partir de fora da rede chegam através de conexões que se originam do mundo externo. Saídas da rede para o mundo externo são conexões que deixam a rede.

De forma geral, a operação de uma célula da rede se resume em:

- Sinais são apresentados à entrada;
- Cada sinal é multiplicado por um peso que indica sua influência na saída da unidade;
- É feita a soma ponderada dos sinais que produz um nível de atividade;
- Se este nível excede um limiar, a unidade produz uma saída.

As redes neurais podem ser hábeis a aprender relações entre entradas e saídas complexas e não lineares. Dentre outras características, inclui-se a possibilidade de usar procedimentos seqüenciais de treinamento, e o poder de adaptação aos dados. Há várias famílias de redes neurais [13, 35]. Dentre elas, uma das mais comuns em reconhecimento de padrões é a rede retroalimentada (*feed-forward network*), que inclui *perceptron* multicamadas e redes funções de bases radiais (*Radial-Basis Function - RBF*). Essas redes são organizadas em camadas

possuindo conexões unidirecionais entre uma camada e outra. Outra classe popular de redes é a dos mapas auto-organizáveis (*self-organizing maps - SOM*), ou redes de Kohonen, as quais são usadas principalmente para agrupamento de dados (*clustering*) e em mapeamento de características. O processo de aprendizado envolve atualização da arquitetura da rede e dos pesos das conexões, de forma que a rede pode efetuar tarefas específicas de classificação e análise de agrupamentos.

A popularidade de redes neurais cresceu devido ao fato de que, aparentemente, elas têm uma baixa dependência a um domínio específico, de forma que o mesmo tipo de rede pode ser aplicado a vários problemas, o que não ocorre com as abordagens baseadas em modelos ou em regras. Outro aspecto importante na popularidade é o fato de existirem muitos algoritmos eficientes de aprendizado com redes neurais.

As redes neurais disponibilizam um novo conjunto de algoritmos não lineares para extração de características (usando camadas escondidas) e classificação (por exemplo, *multi-layer perceptrons*). Além disso, algoritmos de extração de características podem ser mapeados em arquiteturas de redes neurais para implementação eficiente em *hardware*. Apesar de possuir princípios diferentes, a maioria dos modelos de redes neurais populares é implicitamente equivalente ou similar a métodos clássicos de reconhecimento estatístico de padrões. Entretanto, redes neurais oferecem várias vantagens, como abordagens unificadas para extração e classificação, e procedimentos flexíveis para encontrar boas soluções não lineares.

O neurônio artificial

O neurônio artificial possui muitos nomes diferentes na bibliografia atual. Entretanto, todos esses nomes estão intencionados a expor o mesmo elemento em questão, porém, com o título alterado conforme o autor. Entre os sinônimos mais conhecidos, encontram-se o elemento de processamento e o nodo. De qualquer forma, esses neurônios possuem, na sua estrutura e funcionamento, uma similaridade muito grande com os conhecidos neurônios biológicos.

Sinais de entrada e saída

O neurônio possui, a princípio, um ou mais sinais de entrada e um sinal de saída. Sua

semelhança com um neurônio biológico é muito grande, pois ambos os neurônios possuem disparos de saída podendo receber muitas entradas, conforme mostrado na figura 3.4.

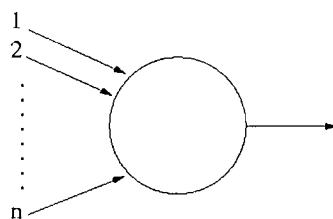


Figura 3.4: Neurônio artificial simples com n entradas e uma saída.

As entradas de um neurônio artificial podem ser comparadas exatamente como estímulos para o neurônio natural. Todos esses estímulos são trazidos até o neurônio simultaneamente, ou seja, se um neurônio possuir 5 entradas, os sinais das 5 entradas chegarão até o núcleo de processamento ao mesmo tempo, ou seja, paralelamente.

A simulação de um ambiente paralelo em computadores é possível, e é desta forma que ocorre o processamento para as redes neurais. O modelo matemático simula o paralelismo da rede neural através de um algoritmo.

Pesos

Os pesos são atributos muito importantes do neurônio. Pode-se compará-los com os dendritos realizando as suas sinapses em outros neurônios. Graças a essa comparação, os pesos são chamados *pesos sinápticos*. Os pesos, denotados por w (peso - *weight*), são valores que representam o grau de importância que determinada entrada possui em relação aquele determinado neurônio. Ou seja, o peso varia em função da intensidade do sinal de entrada e, dessa forma, o peso muda o seu valor representativo para a rede. Isso significa que, quando uma entrada é bastante estimulada, acaba estimulando também o peso correspondente a sua conexão. Um peso, quando é bastante estimulado, automaticamente terá cada vez mais influência no resultado do sinal de saída.

Os pesos podem ser vistos, matematicamente, como um vetor de valores (w_1, w_2, \dots, w_n) . Havendo mais de um neurônio na rede, pode-se então ter uma coleção de vetores, ou seja, uma matriz de pesos, onde cada vetor corresponde a um neurônio. Quando as entradas

(x_1, x_2, \dots, x_n) são apresentadas para o neurônio, elas são multiplicadas pelos pesos, e a soma desses resultados é, então, o sinal de excitação do neurônio. As entradas multiplicadas pelos pesos recebem, depois dessa operação, o nome de entradas ponderadas.

A função do neurônio é, depois de acumulado o valor somado dos produtos ocorridos entre as entradas e os pesos, comparar esse valor com um limiar (um valor pré-definido) e, atingindo-o, o valor é então passado adiante através da saída. Esse processo é chamado de função de transferência. Caso contrário, se o valor não atinge o limiar, o sinal não é transferido adiante. Em ambos os casos, com sinal ou sem sinal, a resposta é significativa, pois afetará diretamente a resposta final da rede ou os neurônios da próxima camada. A lógica neural expõe que a intensidade dos sinais de entrada dispara ou não o sinal do neurônio, fazendo com que este estimule o neurônio seguinte.

Função de ativação e função de transferência

A função de ativação antecede a função de transferência, e tem por atribuição repassar o sinal para a saída do neurônio. A função de ativação é uma função de ordem interna, cuja atribuição é fazer acontecer um nível de ativação dentro do próprio neurônio, ou seja, é uma decisão tomada pelo neurônio sobre o que fazer com o valor resultante do somatório das entradas ponderadas. Essa decisão terá um efeito restrito ao próprio neurônio (figura 3.5).

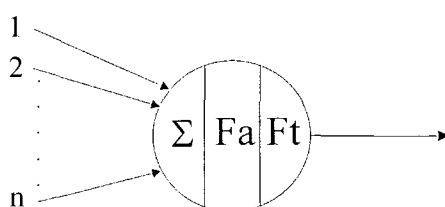


Figura 3.5: Função de ativação (Fa) e função de transferência (Ft).

Em modelos simples de redes neurais, a função de ativação pode ser a própria função de soma das entradas ponderadas do neurônio. Em modelos mais complexos, a função de ativação possui um processamento atribuído. Esse processamento pode usar, por exemplo, o valor prévio de saída como uma entrada para o próprio neurônio (uma auto-excitação). Após o valor ter sido processado pela função de ativação, este é então passado para a função de

transferência que produzirá o valor de saída do neurônio.

A função de transferência pode ter muitas formas e métodos, podendo ser simples ou complexa. A função de transferência também é conhecida como limiar lógico (*threshold*). Essa função é quem define e quem envia para fora do neurônio, o valor passado pela função de ativação. As funções de transferência mais conhecidas são apresentadas na figura 3.6.

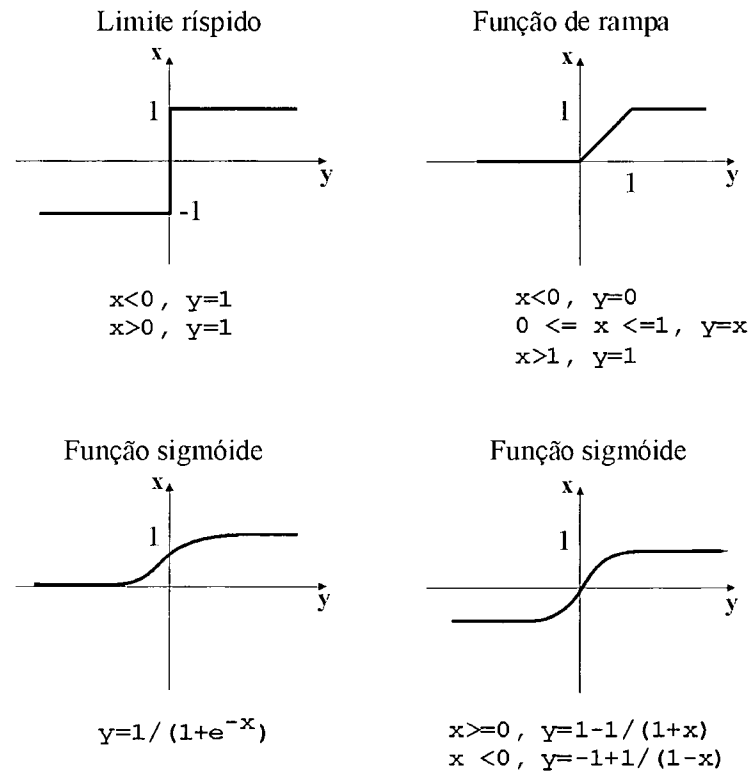


Figura 3.6: Funções de transferência.

Às vezes, a saída do neurônio é conduzida para outros neurônios ao invés de ser conduzida como resposta final. Existem também redes que conduzem as saídas dos neurônios para ambos os sentidos, como uma retroalimentação, ou seja, a saída do neurônio afeta ele próprio (auto-excitação).

As funções de rampa e limite possuem decisões ríspidas, principalmente para valores extremos. Essas funções refletem a saída dentro de uma faixa (por exemplo, entre 0 e 1). Quando é ultrapassado um certo limite, a função dispara o valor 0, ou 1, ou -1, dependendo da função e da forma que será utilizada.

As funções sigmóides são utilizadas normalmente para decisões onde as saídas limites são disparadas quando existe uma saturação muito alta do valor de ativação. Isso significa que

essa função produziria um 0 de saída somente quando o valor passado da função de ativação tivesse um grande valor negativo. e produziria o valor de saída 1 quando a ativação tivesse um grande valor positivo. Essa função faz a transição entre os extremos de forma suave.

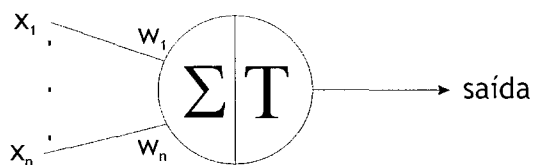


Figura 3.7: Neurônio completo.

Após definidos todos os aspectos, características, funções e valores do neurônio artificial, pode-se considerá-lo como completo (figura 3.7), restando apenas comentar sobre as estruturas das redes de neurônios, chamadas redes neurais. O neurônio, a partir desse momento pode assumir seu símbolo universal, demonstrando em seu interior as duas peças chaves para seu perfeito funcionamento: a função de soma (Σ) e a função de transferência (T). A função de ativação será considerada igual à função soma.

A rede neural

Combinando diversos neurônios pode-se formar o que é chamado de rede de neurônios ou rede neural. As entradas podem ser conectadas em muitos neurônios com vários pesos, resultando assim em uma série de saídas (figura 3.8). Essas conexões, em comparação com o sistema biológico, representam o contato dos dendritos com outro neurônio formando a sinapse.

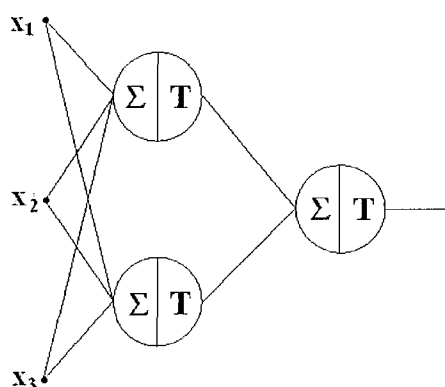


Figura 3.8: Rede neural.

A camada de entrada tem como função armazenar a informação de entrada para ser passada para a camada seguinte de neurônios. Essa camada pode ser chamada também de neurônios de entrada ou camada de distribuição.

As redes neurais podem possuir também camadas escondidas, também chamadas de intermediárias ou ocultas. Essas camadas intermediárias se situam entre a camada de entrada e a camada de saída da rede neural. Não existe uma regra que defina o número de camadas escondidas.

Essas camadas são compostas por neurônios que possuem exatamente a mesma estrutura que os neurônios da camada de saída. Sua única diferença é não ter contato com o mundo externo e, assim, os sinais são passados para outros neurônios obedecendo as regras de transferência de cada neurônio. O número de neurônios da camada escondida é livre, não obedecendo nenhuma regra específica.

3.1.3.3 Aprendizado supervisionado para reconhecer padrões

O aprendizado supervisionado das RNAs caracteriza-se pela disponibilidade de um conjunto de treinamento composto por dados de entrada previamente classificados. Ou seja, estímulos de entrada e respostas desejadas a estes estímulos.

O processo de ajuste dos pesos se dá da seguinte forma: os estímulos de entrada disponíveis no conjunto de treinamento são apresentados à rede, que calcula uma resposta, utilizando como parâmetros os valores de pesos atuais, ou seja, sem ajuste.

A partir daí, faz-se uma comparação entre a resposta oferecida pela rede atual e a resposta desejada por aqueles estímulos. Com base na similaridade entre as duas respostas, os pesos são ajustados. Esse procedimento ocorre até que os pesos possibilitem à rede classificar o mais corretamente possível as entradas de treinamento e, conseqüentemente, as entradas apresentadas durante a fase de teste. Isso inclui o caso especial de aprendizado por reforço, onde a única informação disponível é se cada saída está certa ou errada, e não qual é a resposta correta.

Essas regras de ajuste supervisionado geralmente baseiam-se na determinação do erro, calculado como a diferença entre a resposta que a rede oferece utilizando os pesos atuais e a

resposta que deveria oferecer ou resposta desejada, disponível no conjunto de treinamento.

Um fator importante no processo de aprendizado é definir se a rede será treinada por padrões ou por grupos. No primeiro modo, os pesos são ajustados à medida em que os padrões vão sendo apresentados à rede. Assim, em cada apresentação do conjunto de treinamento, tantas alterações de pesos ocorrerão quantos forem os elementos deste conjunto. No modo por grupos, os pesos normalmente são ajustados após todos os elementos do conjunto de treinamento serem apresentados à rede, ou seja, apenas um cálculo ocorre ao final de cada época (intervalo no qual todo o conjunto de treinamento é apresentado à RNA). No aprendizado por grupos, há a possibilidade de ocorrer a apresentação de grupos de padrões à rede; dessa forma, em uma época ocorrerão tantos ajustes de pesos quantos forem os grupos apresentados.

Há dois modelos de redes neurais utilizados na prática como classificadores passíveis de serem gerados através de aprendizado supervisionado. Ambos os modelos baseiam-se nos *perceptrons feed-forward*, variando o número de camadas e a função de ativação e, por conseguinte, a regra de aprendizado: (a) as redes de retropropagação (*backpropagation*) e (b) as redes de base radial. Por serem as mais utilizadas e, do ponto de vista prático, mais importantes, maiores detalhes sobre as redes de retropropagação, também chamadas redes-BP, serão descritos a seguir. As redes de base radial, conhecidas também por redes-RBF podem ser usadas, em teoria, para representar os mesmos tipos de problemas que uma rede de retropropagação equivalente, sendo consideradas em alguns casos mais eficientes durante o treinamento. Por outro lado, a compreensão de seu algoritmo de aprendizado envolve aspectos matemáticos mais complexos. Como são utilizadas para resolver o mesmo tipo de problemas onde a retropropagação encontra aplicação, sem vantagens consideráveis na qualidade do resultado final, elas não serão tratadas aqui.

3.1.3.4 Princípios básicos das redes de retropropagação

A idéia básica de toda rede neural *feed-forward* com aprendizado supervisionado é que, durante o seu treinamento, é aplicada à sua camada de entrada o padrão que se deseja aprender e propaga-se a ativação gerada por este padrão na camada de neurônios de entrada, camada a

camada, até que seja gerada uma ativação nos neurônios da camada de saída.

Cada neurônio i de uma camada está tipicamente ligado a todos os neurônios j da camada anterior e recebe o sinal h de todos estes, cada qual ponderado pelo peso w_{ij} da conexão correspondente. Cada neurônio possui uma função de ativação $g()$, cuja variável independente é a entrada h do neurônio. Todo neurônio i possui um sinal de saída, que será denominado V_i se o sinal for dirigido para outra camada da rede (neurônios de entrada e internos) ou O_i se o neurônio for de saída. A relação típica entre o sinal de saída e a ativação $g(h)$ do neurônio é a identidade.

A separação da ativação e do sinal de saída em duas variáveis permite o sincronismo de toda a rede (inerentemente paralela) quando se simula a mesma em uma máquina de von Neumann (seqüencial). Os valores de saída dos neurônios de uma camada somente são atualizados após o cálculo de todas as ativações desta camada. Isto é especialmente importante em redes BP recorrentes e outros modelos de rede, onde neurônios possuem ligações com neurônios da própria camada.

3.1.4 Lógica Nebulosa (Fuzzy)

A lógica nebulosa foi criada com inspiração no comportamento humano, que se baseia na interpretação do mundo sem precisão e na descrição deste por atributos lingüísticos.

Inicialmente, define-se um conjunto não-nebuloso (*crisp*) dividindo os indivíduos de determinado universo em dois grupos: membros (indivíduos que certamente pertencem a determinado conjunto) e não-membros (indivíduos que com certeza não fazem parte do conjunto). Essa teoria de conjuntos não-nebulosos é muito utilizada, porém, há muitas categorias que não exibem tal característica binária (verdadeiro ou falso), por exemplo classes de pessoas altas, carros caros, números muito maiores que 1. Ao invés disso, os seus limites parecem vagos, e a transição de membro para não membro é gradual ao invés de abrupta [22].

A partir de conjuntos não-nebulosos (*crisp*), a criação de seus respectivos conjuntos nebulosos (*fuzzificação*) é feita através da determinação do grau de pertinência de cada um dos elementos aos conjuntos.

Nessa etapa são verificadas a análise do ambiente, pois é de extrema importância conhecer

o problema a ser manipulado; como o especialista trabalha e como ele entende sua realidade; qual a linguagem usada pelo especialista para tratar os problemas; quais as decisões e quando elas são tomadas. A partir da aquisição desse conhecimento, pode-se saber em qual contexto o sistema baseado em lógica nebulosa será aplicado.

Como geralmente os dados de entrada são valores precisos, resultados de medidas ou observações, torna-se necessário efetuar o mapeamento desses dados precisos para os conjuntos nebulosos de entrada. Dessa forma, são definidos os conjuntos nebulosos, escolhendo-se o tipo de conjunto a ser adotado, as variáveis lingüísticas e as funções de pertinência.

As funções de pertinência são baseadas na relação entre cada elemento e o(s) suporte(s) dos conjuntos. Essa relação pode ser, por exemplo, inversamente proporcional à distância entre os elementos e os suportes. Um suporte pode ser determinado, por exemplo, pelo centróide do conjunto de treinamento.

A classificação com lógica nebulosa normalmente é feita através de regras nebulosas e métodos de *defuzzificação*. As regras nebulosas normalmente são regras de inferência que utilizam delimitadores dos valores de pertinência de cada padrão aos conjuntos. Elas se baseiam no conhecimento adquirido do problema a ser solucionado. Tais regras são expressas na forma de proposições, tais como:

SE x é alto e/ou y é baixo ENTÃO z é médio

onde

x e y : são variáveis de entrada (valores conhecidos);

z : é uma variável de saída (um valor a ser calculado);

alto: é uma variável lingüística com sua respectiva função de pertinência definida em x ;

baixo: é uma variável lingüística com sua respectiva função de pertinência definida em y ;

médio: é uma variável lingüística com sua respectiva função de pertinência definida em z ;

e/ou: são os operadores lógicos *AND* e *OR* que representam as operações *fuzzy* de interseção e união, respectivamente.

A parte da regra entre o SE e o ENTÃO é a premissa (condição) ou antecedente da regra, e a parte posterior ao ENTÃO é a sua conclusão (resultado ou ação). O conjunto *fuzzy* final é gerado através da execução das proposições *fuzzy* correlacionadas. O processo de correlação

se baseia na aplicação dos operadores *AND* e *OR*, definidos nas regras, os quais conectam as proposições.

Em um sistema *fuzzy*, a ordem de execução das regras não é importante, exceto quando uma regra depende dos resultados de execução de outras regras.

O processo de *defuzzificação* consiste na utilização dos resultados das regras para a criação de conjuntos *crisp* que contenham informações a respeito da classificação dos elementos. Nessa etapa, procura-se obter a melhor representação para o conjunto de saída *fuzzy*, onde o valor da variável lingüística de saída inferida pelas regras será traduzida para um valor discreto. Para selecionar o método apropriado de *defuzzificação*, pode-se utilizar um enfoque baseado no centro de área ou nos valores máximos que ocorrem a partir da função de pertinência resultante.

Maiores detalhes a respeito de conjuntos nebulosos, lógica *fuzzy* e aplicações podem ser encontrados em [11, 20, 45].

3.2 CLASSIFICADORES

3.2.1 Distância Mínima

O classificador de distância mínima é bastante eficiente em termos de esforço computacional, tanto na fase de treinamento quanto na fase de teste. Essa característica deve-se à simplicidade de seu algoritmo.

A fase de treinamento consiste na determinação dos protótipos, no mínimo um para cada classe. Os protótipos são vetores no espaço de características que normalmente são criados a partir de informações obtidas do conjunto de treinamento ou da distribuição probabilística das classes. Um exemplo bastante comum de protótipo utilizado é a média (baricentro) do conjunto de treinamento das classes.

Na fase de classificação, cada padrão é classificado de acordo com o protótipo mais próximo. Normalmente, utiliza-se a distância Euclidiana para calcular a proximidade entre os padrões e os protótipos. Cada pixel será incorporado a um agrupamento, através da análise

da medida de similaridade de distância Euclidiana, que é dada por

$$d(p, m) = \sqrt{(p^2 - m^2)} \quad (3.2)$$

onde p = medida do pixel que está sendo testado e m = média de um agrupamento.

O classificador compara a distância Euclidiana da medida do pixel à média de cada agrupamento. O pixel será incorporado ao agrupamento que apresenta a menor distância Euclidiana. Este procedimento é repetido até que toda a imagem seja classificada.

Caso o pixel esteja distante de todas as classes mais do que um certo limiar, ele será classificado como desconhecido. Assim, para cada padrão de teste, é necessário realizar apenas $C - 1$ comparações, sendo C o número de classes existentes.

Este método é matematicamente simples e computacionalmente eficiente, mas possui algumas limitações. A mais importante é a sua insensibilidade aos diferentes graus de variância dos dados espectrais. Tendo em vista esse problema, o método não é muito utilizado em aplicações onde classes espectrais estão muito próximas uma das outras e possuem grande variância. A desvantagem dessa abordagem é a qualidade dos resultados, pois os protótipos muitas vezes não contêm informações suficientes sobre a forma da distribuição das classes.

A figura 3.9 ilustra a classificação de 4 pixels representados pelas letras p_1, p_2, p_3 e p_4 dentro de 3 classes utilizando a técnica da distância mínima. Os pontos pretos representam o centro de cada classe. O resultado para este exemplo é dado na tabela 3.1 e a classe escolhida é determinada pela menor distância entre o pixel e o centro da classe.

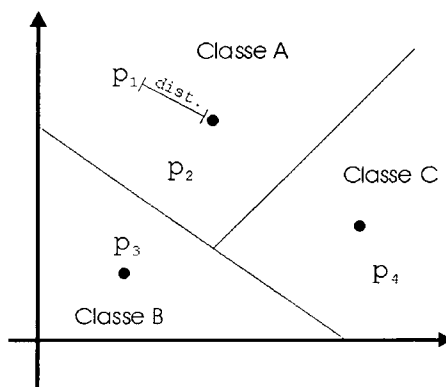


Figura 3.9: Exemplo de limites de decisão - Distância Mínima.

Pixel	Classe
p_1	A
p_2	A
p_3	B
p_4	C

Tabela 3.1: Resultado da classificação - Distância Mínima.

3.2.2 K-vizinhos mais Próximos

Trata-se de um método de classificação que não possui processamento na fase de treinamento, pois não é necessário estimar as distribuições de probabilidades das classes. Entretanto, é necessário um grande número de padrões de treinamento (padrões cuja classe é conhecida *a priori*), pois pode-se dizer que as tarefas de estimativa e de classificação são unidas em uma única tarefa.

O classificador dos K -vizinhos mais próximos (*K-Nearest Neighbor - KNN*) é um classificador que cria fronteiras de decisão complexas. Dado um padrão de teste C (desconhecido), sua classificação é realizada da seguinte maneira: inicialmente, calcula-se a distância entre C e todos os padrões de treinamento; verifica-se a quais classes pertencem os K padrões mais próximos; a classificação é feita associando-se o padrão de teste à classe que for mais freqüente entre os K padrões mais próximos de C .

A principal vantagem desse método é que ele cria uma superfície de decisão que se adapta à forma de distribuição dos dados de treinamento de forma detalhada, possibilitando a obtenção de boas taxas de acerto quando o conjunto de treinamento é grande ou representa todas as variações possíveis dos dados, principalmente no caso de $K = 1$. Valores maiores de K não promovem bons resultados quando a distribuição dos conjuntos possui poucas sobreposições, e quando, por exemplo, um padrão de treinamento de uma classe A adjacente por vários da classe B pode ser tratado como uma exceção, e os padrões de teste que ficarem na região podem ser classificados como elementos da classe B .

A principal desvantagem está na complexidade do teste, visto que, para cada padrão de teste, além de ser necessário realizar P medições de distância, sendo P o número de padrões de treinamento, é necessário ordenar os resultados.

A figura 3.10 ilustra um exemplo de classificação através dos K -vizinhos mais próximos. É calculada a menor distância entre o padrão desconhecido localizado ao centro e os K padrões mais próximos a ele. A classe atribuída ao padrão desconhecido é a que tiver maior ocorrência entre os K vizinhos encontrados. Para o exemplo em questão é calculada a distância entre o padrão e três vizinhos mais próximos a ele. Verifica-se que dois dos vizinhos pertencem a classe B, a qual é então atribuída ao padrão.

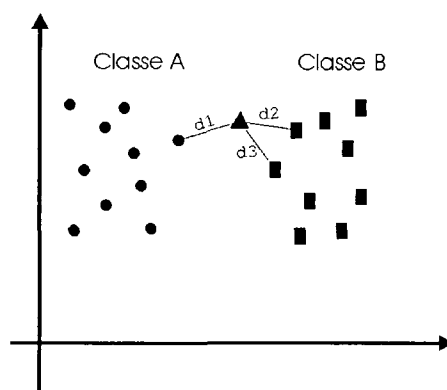


Figura 3.10: Exemplo de limites de decisão através do método dos K -vizinhos mais próximos.

3.2.3 Máxima Verossimilhança

A classificação por máxima verossimilhança é um tipo de classificação supervisionada baseada no princípio de que a classificação errada de um pixel particular não possui mais significado do que a classificação errada de qualquer outro pixel na imagem. Ela considera a ponderação das distâncias entre médias dos níveis digitais das classes, utilizando parâmetros estatísticos [49].

Para que a classificação por máxima verossimilhança seja precisa o suficiente, é necessário um número razoavelmente elevado de amostras, para cada conjunto de treinamento. Os conjuntos de treinamento definem o diagrama de dispersão das classes e suas distribuições de probabilidade, considerando a distribuição de probabilidade normal para cada classe de treinamento.

Esse algoritmo é considerado paramétrico, pois utiliza a média e covariância das amostras de treinamento, e assume que as classes possuem distribuição normal. Com esses parâmetros, pode ser calculada a probabilidade de um pixel desconhecido pertencer a uma ou outra classe.

Funções de densidade de probabilidade são utilizadas para classificar um pixel desconhecido em uma determinada classe. Depois de avaliada a probabilidade para cada classe, o pixel é assinalado à classe cuja probabilidade tenha sido a mais alta, ou a nenhuma delas no caso dos valores de probabilidade estarem abaixo de um determinado limite estipulado pelo especialista.

A eficácia da Máxima Verossimilhança depende principalmente de uma precisão razoável da estimativa do vetor médio \mathbf{m} e da matriz de covariância Σ de toda classe espectral, que depende da quantidade de pixels incluídos nas amostras de treinamento. Sendo \mathbf{x} o vetor correspondente a um pixel nas N classes envolvidas, o vetor médio dos pixels pertencentes a uma classe é dado por [37]

$$\mathbf{m} = E(\mathbf{x}) = \frac{1}{K} \sum_{i=1}^K \mathbf{x}_i \quad (3.3)$$

onde K é o número de pixels na classe e $E(\mathbf{x})$ a esperança de \mathbf{x} , uma notação estatística para estimar a média de \mathbf{x} . Já a matriz de covariância será dada por

$$\Sigma = \frac{1}{K-1} \sum_{i=1}^K (\mathbf{x}_i - \mathbf{m})(\mathbf{x}_i - \mathbf{m})^t \quad (3.4)$$

Para exemplificar, considere duas classes, A e B , com distribuições de probabilidade distintas. Essas distribuições de probabilidade representam a probabilidade de um pixel pertencer a uma ou outra classe, dependendo da posição do pixel em relação a esta distribuição.

Observa-se uma região onde as duas curvas se sobrepõem, indicando que um determinado pixel tem igual probabilidade de pertencer às duas classes. Nessa situação, estabelece-se um critério de decisão a partir da definição de limiares. Os limites de classificação são definidos a partir de pontos de mesma probabilidade de classificação de uma e de outra classe.

A figura 3.11 apresenta o limite de aceitação de uma classificação, no ponto onde as duas distribuições se cruzam. Dessa forma, um pixel localizado na região sombreada, apesar de pertencer à classe B , será classificado como classe A , pelo limite de aceitação estabelecido.

O limiar de aceitação indica a percentagem de pixels da distribuição de probabilidade de uma classe que será classificada como pertencente a esta classe. Um limite de 99%, por exemplo, engloba 99% dos pixels, sendo que 1% serão ignorados (os de menor probabilidade), compensando a possibilidade de alguns pixels terem sido incorretamente introduzidos nesta

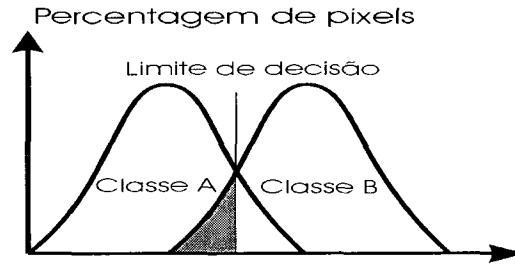


Figura 3.11: Limite de aceitação de uma classificação.

classe durante o treinamento, ou estarem no limite entre duas classes. Um limiar de 100% resultará em uma imagem classificada sem rejeição, ou seja, todos os pixels serão classificados.

Para diminuir a confusão entre as classes, ou seja, reduzir a sobreposição entre as distribuições de probabilidades das classes, aconselha-se a aquisição de amostras significativas de alvos distintos e a avaliação da matriz de classificação das amostras.

A matriz de classificação apresenta a distribuição de porcentagem de pixels classificados correta e erroneamente. Uma matriz de classificação ideal deve apresentar os valores da diagonal principal próximos a 100%, indicando que não houve confusão entre as classes. Contudo, essa é uma situação difícil em imagens com alvos de características espectrais semelhantes.

O método da Máxima Verossimilhança deve ser aplicado quando se conhece bem a imagem a ser classificada, para que seja possível definir classes representativas. O resultado é melhor quanto maior o número de pixels em uma amostra de treinamento para implementá-los na matriz de covariância. Se o tamanho das amostras de treinamento para as classes é limitado, recomenda-se um método de classificação mais simples e rápido, que não utilize matriz de covariância, por exemplo, o método da distância mínima.

3.2.4 Intervalos

O método de classificação supervisionada baseado em intervalos foi definido visando descrever estatisticamente as classes através de vetores de intervalos numéricos.

A cada nova amostra incluída no processo de treinamento, suas medidas estatísticas são calculadas e os limites inferiores e superiores da classe em questão são verificados. Cada medida é tratada separadamente. Se o valor apresentado for menor que o limite inferior, ele

passa a ser o limite inferior. Por outro lado, se o valor for maior que o limite superior, ele passa a ser o limite superior. Se o valor estiver apenas entre um intervalo já definido, o intervalo permanece inalterado.

Após todas amostras terem sido apresentadas, tem-se uma configuração da distribuição das classes.

A classificação é feita de acordo com o intervalo em que se enquadra a medida do pixel desconhecido. Existem regiões onde os intervalos de classes se sobrepõem, indicando que um determinado pixel pode pertencer a mais do que uma classe. Nesse caso, estabelece-se um critério de decisão para definir somente uma classe onde o pixel será classificado.

O critério de decisão é baseado no seguinte cálculo de distâncias entre uma medida para o pixel desconhecido e cada intervalo dessa mesma medida para as diferentes classes

$$\max \left[\min \left(\frac{|x - L_i^a|}{|L_s^a - L_i^a|}, \frac{|x - L_s^a|}{|L_s^a - L_i^a|} \right), \min \left(\frac{|x - L_i^b|}{|L_s^b - L_i^b|}, \frac{|x - L_s^b|}{|L_s^b - L_i^b|} \right) \right] \quad (3.5)$$

onde x é o valor da medida do pixel a ser classificado; a e b são dois intervalos de duas classes distintas; L_i é o limite inferior da classe; L_s é o limite superior da classe.

A figura 3.12 ilustra um exemplo onde determinada medida estatística de um pixel a ser reconhecido ficou localizada entre dois intervalos, cada um pertencente a uma classe distinta (A e B). O critério de decisão encontra primeiramente um valor mínimo para cada classe e, posteriormente, o valor máximo dentre os mínimos, o qual indica a classe B na qual será classificado o pixel. Tal cálculo busca uma maior proximidade à média do intervalo da classe.

A aquisição de amostras significativas para cada classe é um fator que deve ser levado em conta para evitar confusão entre as classes, reduzindo a sobreposição de dois ou mais intervalos.

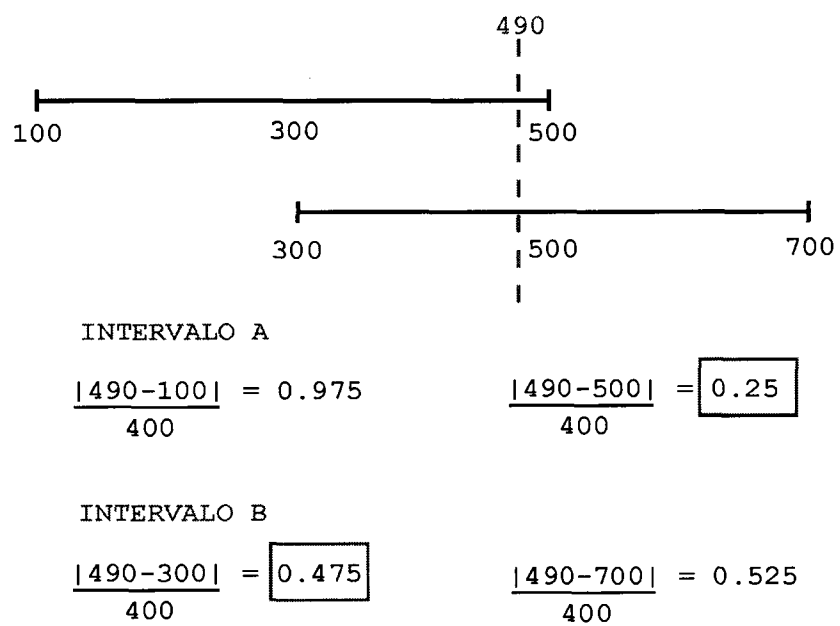


Figura 3.12: Exemplo de decisão entre dois intervalos de classes.

CAPÍTULO 4

MÉTODO DE CLASSIFICAÇÃO DE IMAGENS BASEADO EM CARACTERÍSTICAS TEXTURAIS

Este capítulo apresenta a estrutura do sistema proposto para classificação supervisionada de imagens utilizando características de textura. Tal sistema, ilustrado na figura 4.1, é composto pelos seguintes módulos:

- *aquisição de dados*: prepara as imagens e amostras de textura;
- *pré-processamento*: determina a quantidade e o tamanho das amostras;
- *extração de atributos*: cria um vetor de características com dados extraídos das imagens e amostras de textura;
- *seleção de atributos*: analisa o conjunto de características, eliminando as mais redundantes;
- *medição das características*: responsável por calcular a proximidade entre padrões de treinamento e de classificação;
- *classificação*: consiste em um classificador que toma decisões baseando-se no aprendizado obtido a partir de um conjunto de treinamento, o qual deve conter exemplos de padrões das classes mais representativas existentes no sistema.

O sistema de classificação é avaliado através de um conjunto de testes, preferencialmente composto por padrões de textura de todas as classes, mas que não estão necessariamente no conjunto de treinamento.

Para a execução do algoritmo implementado para classificação supervisionada de imagens através de textura utilizando matrizes de co-ocorrência [34], seguem-se os seguintes passos básicos:

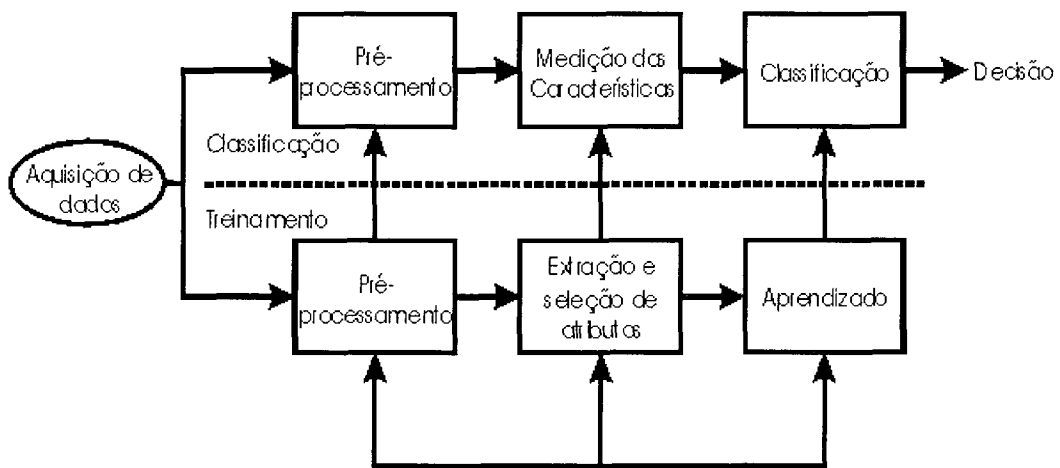


Figura 4.1: Esquema do sistema proposto para classificação de imagens.

1. Fase de treinamento

- (a) para cada classe de texturas, selecionar áreas representativas de treinamento, as quais serão chamadas de amostras de textura;
- (b) as amostras de textura devem ser apresentadas uma a uma ao sistema ou pode-se utilizar um processo automático, onde a partir de uma amostra grande podem ser recortadas $n \times n$ janelas menores de textura. No segundo caso, o número de amostras e o tamanho da janela de treinamento deve ser informado;
- (c) para cada nova amostra, deve-se indicar a qual classe ela pertence, seguindo os conceitos de aprendizado supervisionado. É calculada sua matriz de co-ocorrência e um vetor de m características.
- (d) os dados de treinamento são armazenados em um arquivo de acordo com o tipo de treinamento adequado a cada um dos classificadores, ou seja, distância mínima, máxima verossimilhança, intervalos e redes neurais;
- (e) para utilizar o classificador da distância mínima, o treinamento é feito calculando-se a média ponderada entre a entrada já existente no vetor e a nova entrada;
- (f) para o classificador da máxima verossimilhança, utiliza-se o conceito de diagramas de dispersão e distribuições de probabilidade para efetuar o treinamento. Para pon-

tos com a mesma probabilidade de classificação são definidos limites de aceitação.

O pacote SPRING [49] foi utilizado para a determinação das classes;

- (g) para o classificador baseado em intervalos, limites inferiores e superiores são estabelecidos para cada classe, baseando-se na verificação de cada entrada e nos limites já calculados;
- (h) para a utilização da técnica de redes neurais foi utilizada uma ferramenta poderosa para a execução dos testes propostos, evitando a necessidade de implementar as redes. O SNNS [46] é um dos melhores simuladores de redes neurais existentes e possui duas versões, SNNS Padrão e JavaNNS. A versão padrão é fornecida em código-fonte e pode ser compilada para qualquer plataforma Unix/Linux. Ela possui uma série de utilitários que permitem a integração das redes neurais treinadas com o sistema em programas aplicativos. O JavaNNS é uma nova implementação com finalidade exclusiva de ensino. O simulador está também disponível para plataforma Windows, porém oferece apenas o ambiente de treinamento e teste interativo de redes neurais, não oferecendo nenhuma ferramenta de integração das redes em programas aplicativos. Nesse caso, é gerado um arquivo de treinamento para ser utilizado no simulador de redes neurais SNNS.

2. fase de classificação

- para cada pixel da imagem:

- (a) calcular as k matrizes de co-ocorrência M_k em uma vizinhança $n \times n$ do pixel corrente, sendo k o número total de pixels na imagem;
- (b) identificar a classe do objeto a qual pertence o pixel, conforme parâmetros da classe e do classificador desejado.

4.1 MATRIZ DE CO-OCORRÊNCIA DE NÍVEIS DE CINZA

A primitiva mais simples que pode ser definida em uma imagem digital em níveis de cinza é um pixel, que tem como propriedade seu nível de cinza. Conseqüentemente, a distribuição

dos níveis de cinza dos pixels pode ser descrita por estatísticas de primeira ordem, como média, variância, desvio padrão, inclinação (*skewness*) ou estatísticas de segunda ordem, como a probabilidade de dois pixels terem um determinado nível de cinza ocorrendo com um relacionamento espacial particular. Essa informação pode ser resumida em matrizes de co-ocorrência (MC) bidimensionais, calculadas para diferentes distâncias e orientações. Para texturas grosseiras, a distribuição varia lentamente com a distância, e para texturas finas a distribuição varia rapidamente com a distância [1].

Os primeiros trabalhos na área de discriminação de texturas usando dependências espaciais de níveis de cinza foram realizados no começo dos anos 70. Em estudos preliminares, os pesquisadores observaram que matrizes de dependência espacial de níveis de cinza obtiveram ótimos resultados na discriminação de imagens com texturas relativamente homogêneas. Embora existam alguns contra-exemplos, já foi comprovado que a discriminação de texturas por um ser humano faz uso de estatísticas de segunda ordem [1].

A MC extrai informações a partir de uma imagem I verificando o relacionamento espacial entre os níveis de cinza existentes nesta imagem [14]. A idéia principal da MC é caracterizar texturas em uma imagem através de um conjunto de estatísticas para as ocorrências de cada nível de cinza em pixels diferentes ao longo de diferentes direções. O termo característica é usado na classificação de textura para descrever um conjunto de estatísticas extraídas de uma matriz de co-ocorrência, caracterizando a textura, tais como energia, entropia e contraste [35].

4.1.1 Ordem das Medidas Estatísticas

A ordem das medidas estatísticas define de que maneira deve ser considerada a disposição espacial entre os pixels existentes na imagem, para que eles sejam adequadamente relacionados.

As medidas podem ser definidas como:

- *primeira ordem*: são as medidas estatísticas de textura calculadas a partir dos valores originais da imagem (por exemplo a variância), e não consideram o relacionamento entre os pixels.
- *segunda Ordem*: consideram o relacionamento entre grupos de dois pixels na imagem

original.

- *terceira e mais altas ordens*: consideram o relacionamento entre três ou mais pixels. Devido ao tempo de cálculo e a dificuldade de interpretação, elas não são geralmente utilizadas na prática.

4.1.2 Relacionamento Espacial entre Dois Pixels

Os vizinhos imediatos de qualquer pixel podem estar em quatro possíveis direções: 0° , 45° , 90° e 135° , como mostra a figura 4.2.

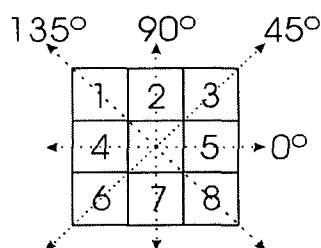


Figura 4.2: Ângulos para o cálculo na matriz de co-ocorrência.

A MC de textura considera a relação entre dois pixels por vez, um chamado de pixel referência e o outro de pixel vizinho. O pixel vizinho escolhido pode, por exemplo, estar a leste (direita) de cada pixel referência. Isto pode ser expresso como uma relação $(1,0)$, ou seja, 1 pixel na direção x , 0 pixels na direção y . Cada pixel dentro da imagem torna-se o pixel referência, iniciando no canto superior esquerdo e procedendo até o inferior direito. Os pixels situados na margem direita não possuem vizinhos à direita, portanto, não são utilizados para esta contagem.

Co-ocorrência, na sua forma geral, pode ser especificada por uma matriz de frequências relativas $P(i, j; d, \theta)$, na qual dois elementos de textura vizinhos, separados por uma distância d em uma orientação θ ocorrem na imagem, um com propriedade i e o outro com propriedade j . Instanciando essa definição para co-ocorrência de níveis de cinza, os elementos de textura são pixels e as propriedades são os níveis de cinza. Por exemplo, para um relacionamento com um ângulo $\theta = 0^\circ$, $P(i, j; d, \theta)$ é a média aritmética da probabilidade de uma transição, da

esquerda para a direita, do nível de cinza i para o nível de cinza j , com uma distância d . Essas matrizes podem ser normalizadas através da divisão de cada entrada da matriz pelo número de pixels vizinhos usados para calcular aquela matriz [1].

4.1.3 Separação entre Dois Pixels

Os exemplos citados a seguir usam sempre um pixel de referência e seu vizinho direto. Se a imagem é grande o suficiente, é perfeitamente possível utilizar uma distância maior, não havendo diferença no método de cálculo. Combinações de níveis de cinza que são possíveis para a imagem de teste e sua posição na matriz são mostradas na tabela 4.1.

valor do pixel referência	valor do pixel vizinho			
	0	1	2	3
0	0,0	0,1	0,2	0,3
1	1,0	1,1	1,2	1,3
2	2,0	2,1	2,2	2,3
3	3,0	3,1	3,2	3,3

Tabela 4.1: Matriz de combinações possíveis para níveis de cinza.

A interpretação da matriz anterior é feita da seguinte maneira: a posição superior esquerda da matriz deve ser preenchida com o número de vezes em que a combinação (0,0) ocorre, ou seja, quantas vezes a imagem contém um pixel com nível de cinza 0 (pixel vizinho) ao lado de outro pixel com nível de cinza 0 (pixel de referência).

Um exemplo de cálculo de matriz de co-ocorrência é mostrado a seguir, levando-se em consideração a imagem apresentada na figura 4.3, a qual possui 4×4 pixels com níveis de cinza no intervalo [0..3].

0	0	1	1
0	0	1	1
0	2	2	2
2	2	3	3

Figura 4.3: Imagem 4×4 com níveis de cinza de 0 a 3.

$$P(i, j; 1, 0^\circ) = \begin{pmatrix} 4 & 2 & 1 & 0 \\ 2 & 4 & 0 & 0 \\ 1 & 0 & 6 & 1 \\ 0 & 0 & 1 & 2 \end{pmatrix} \quad P(i, j; 1, 45^\circ) = \begin{pmatrix} 4 & 1 & 0 & 0 \\ 1 & 2 & 2 & 0 \\ 0 & 2 & 4 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

$$P(i, j; 1, 90^\circ) = \begin{pmatrix} 6 & 0 & 2 & 0 \\ 0 & 4 & 2 & 0 \\ 2 & 2 & 2 & 2 \\ 0 & 0 & 2 & 0 \end{pmatrix} \quad P(i, j; 1, 135^\circ) = \begin{pmatrix} 2 & 1 & 3 & 0 \\ 1 & 2 & 1 & 0 \\ 3 & 1 & 0 & 2 \\ 0 & 0 & 2 & 0 \end{pmatrix}$$

4.1.4 Propriedades da Matriz de Co-ocorrência

As matrizes de co-ocorrência apresentam as seguintes propriedades importantes:

- **Quadrada**

A MC sempre apresenta o número de linhas igual ao número de colunas.

- **O número de linhas e colunas é igual ao nível de quantização da imagem**

A imagem do exemplo ilustrado na figura 4.3 possui quatro níveis de cinza (0, 1, 2 e 3). Entretanto, cada nível de cinza poderia assumir 256 valores, podendo produzir uma matriz quadrada de 256×256 , com 65.536 posições.

- **É simétrica ao redor da diagonal**

A MC é uma matriz simétrica, ou seja, possui os mesmos valores para os elementos situados em lados opostos da diagonal principal. Por exemplo, o valor do elemento da posição (3,2) é igual ao valor da posição (2,3). Os cálculos de textura são melhor executados em matrizes simétricas.

4.1.5 Expressando a MC como uma Probabilidade

É possível transformar a MC em uma tabela de probabilidades. Isto resulta em apenas uma aproximação, já que uma probabilidade real requer valores contínuos, enquanto que os níveis de cinza são valores inteiros ou discretos.

Esse processo é chamado de *normalização* da matriz e envolve uma divisão pela soma total dos valores

$$P_{i,j} = \frac{V_{i,j}}{\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} V_{i,j}} \quad (4.1)$$

onde:

- i é o índice da linha e j é o índice da coluna;
- V é o valor do elemento i, j na MC;
- $P_{i,j}$ é a probabilidade para o elemento i, j ;
- N é o número de linhas ou colunas.

Um exemplo de normalização de uma matriz de co-ocorrência é ilustrado a seguir:

Imagem de teste:

0	0	1	1
0	0	1	1
0	2	2	2
2	2	3	3

Matriz de combinações possíveis:

0,0	0,1	0,2	0,3
1,0	1,1	1,2	1,3
2,0	2,1	2,2	2,3
3,0	3,1	3,2	3,3

MC (direção sul):

3	0	2	0
0	2	2	0
0	0	1	2
0	0	0	0

MC transposta (direção norte):

3	0	0	0
0	2	0	0
2	2	1	0
0	0	2	0

MC simétrica (direção vertical):

6	0	2	0
0	4	2	0
2	2	2	2
0	0	2	0

Dividindo a MC simétrica pela soma dos elementos, obtém-se a MC normalizada:

0.250	0	0.083	0
0	0.166	0.083	0
0.083	0.083	0.083	0.083
0	0	0.083	0

A MC normalizada representa todas as probabilidades possíveis de ocorrência de relacionamentos entre os pixels da imagem original em uma determinada direção (ângulo). Essa informação é aplicada no cálculo de medidas estatísticas, cuja finalidade é caracterizar a textura particular presente na imagem e auxiliar o futuro processo de classificação.

4.1.6 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 dos níveis de cinza. Tais medidas, apesar de antigas, são a base para a maioria dos trabalhos realizados nessa área [43]. As medidas estatísticas utilizam as seguintes definições:

L indica o número de níveis de cinza da imagem.

$p(i, j)$ é a entrada (i, j) da matriz de co-ocorrência normalizada, $P(i, j)/R$, sendo R uma constante de normalização.

$p_x(i)$ é a entrada i da matriz de probabilidades obtida pela soma das linhas de $p(i, j) = \sum_{j=1}^L P(i, j)$.

$$p_y(j) = \sum_{i=1}^L p(i, j).$$

$$p_{x+y}(k) = \sum_{i=1}^L \sum_{j=1}^L p(i, j), \quad i + j = k, \quad k = 2, 3, \dots, 2L.$$

$$p_{x-y}(k) = \sum_{i=1}^L \sum_{j=1}^L p(i, j), \quad |i - j| = k, \quad k = 0, 1, \dots, L - 1.$$

- **Segundo Momento Angular e Energia**

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

$$SMA = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} p(i, j)^2 \quad (4.2)$$

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

$$Energia = \sqrt{SMA} \quad (4.3)$$

- **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 a contraste usam pesos relacionados com a distância da diagonal da MC. Valores na diagonal da MC representam pouco contraste, aumentando quando a distância da diagonal aumenta.

$$\text{Contraste} = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} p(i, j)(i - j)^2 \quad (4.4)$$

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.

$$\mu_i = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} i(p(i, j)) \quad (4.5)$$

$$\mu_j = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} j(p(i, j)) \quad (4.6)$$

A equação 4.5 calcula a média baseada nos pixels de referência μ_i . É possível também calcular a média usando-se os pixels vizinhos μ_j , como na equação 4.6. 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.

$$\sigma_i^2 = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} p(i, j) (i - \mu_i)^2 \quad (4.7)$$

$$\sigma_j^2 = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} p(i, j) (j - \mu_j)^2 \quad (4.8)$$

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:

$$s_i = \sqrt{\sigma_i^2} \quad (4.9)$$

$$s_j = \sqrt{\sigma_j^2} \quad (4.10)$$

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 à sua vizinhança.

$$\text{Correlação} = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} p(i, j) \left[\frac{(i - \mu_i)(j - \mu_j)}{(s_i)(s_j)} \right] \quad (4.11)$$

- **Homogeneidade**

Os valores dos pesos da homogeneidade são inversamente proporcionais aos pesos 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 pesos diminuem distantes da diagonal, o resultado será maior para imagens com pouco contraste.

$$\text{Homogeneidade} = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} \frac{p(i, j)}{1 + (i - j)^2} \quad (4.12)$$

- **Soma da Média**

$$\text{Soma da média} = \sum_{i=2}^{2L} i p_{x+y}(i) \quad (4.13)$$

- **Soma da Variância**

$$\text{Soma da variância} = \sum_{i=2}^{2L} \left(i + \sum_{j=2}^{2L} p_{x+y}(j) \log(p_{x+y}(j)) \right)^2 p_{x+y}(i) \quad (4.14)$$

- **Diferença da Variância**

$$\text{Diferença da variância} = \text{variância de } p_{x-y} \quad (4.15)$$

- **Entropia**

Entropia é normalmente denominada como uma medida de primeiro grau, mas deveria propriamente ser uma medida de zero grau. Como a função logarítmica não é definida para zero e valores negativos, assume-se que $0 \cdot \log(0) = 0$.

$$\text{Entropia} = - \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} \log(p(i, j)) p(i, j) \quad (4.16)$$

- **Soma da Entropia**

$$\text{Soma da entropia} = - \sum_{i=2}^{2L} \log(p_{x+y}(i)) p_{x+y}(i) \quad (4.17)$$

- **Diferença da Entropia**

$$\text{Diferença da entropia} = - \sum_{i=0}^{L-1} \log(p_{x-y}(i)) p_{x-y}(i) \quad (4.18)$$

- **Informação de Medidas de Correlação 1**

$$\text{Medidas de correlação 1} = \frac{HXY - HXY1}{\max(HX, HY)} \quad (4.19)$$

onde HX e HY são entropias de p_x e p_y , e

$$HXY = - \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} \log(p(i, j)) p(i, j)$$

$$HXY1 = - \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} \log(p_x(i) p_y(j)) p(i, j)$$

- **Informação de Medidas de Correlação 2**

$$\text{Medidas de correlação 2} = \sqrt{1 - e^{-2(HXY2 - HXY)}} \quad (4.20)$$

onde

$$HXY2 = - \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} p_x(i) p_y(j) \log(p_x(i) p_y(j)) \quad (4.21)$$

- **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,...).

$$\text{Dissimilaridade} = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} p(i, j) |i - j| \quad (4.22)$$

4.2 REDUÇÃO DE DIMENSIONALIDADE

O termo dimensionalidade é atribuído ao número de características de uma representação de padrões, ou seja, a dimensão do espaço de características (N). As duas principais razões para

que a dimensionalidade seja a menor possível são: custo de medição e precisão do classificador. Quando o espaço de características contém somente as características mais representativas, o classificador será mais rápido e ocupará menos memória [18]. Além disso quando o conjunto de exemplos de treinamento não é muito grande, o problema da dimensionalidade pode ser evitado usando-se um espaço de características pequeno. Isso também propicia a obtenção de menores taxas de erro de classificação.

Em visão computacional, a necessidade de redução de dimensionalidade é acentuada, pois a dimensionalidade de imagens é muito grande. O espaço de imagens possui características que podem ser eliminadas para efetuar o reconhecimento de objetos. Uma imagem de largura w e altura h (em pixels) pode ser vista como um padrão no espaço de imagens, o qual possui dimensionalidade $N = h \times w$. Esse pode ser um valor muitíssimo elevado em imagens obtidas por *scanners* ou câmeras. Além disso, qualquer alteração em translação, rotação ou escala dos objetos contidos nessa imagem fará com que ocorram grandes erros de classificação. Por isso, é necessária a utilização de algoritmos de redução de dimensionalidade que propiciem a obtenção de representações dos padrões (obtidos das imagens) de forma robusta a essas alterações.

Além da necessidade de utilizar a menor dimensionalidade possível, há outro fator analisado que diz ser possível fazer dois padrões arbitrários ficarem similares se esses forem codificados com um número suficientemente grande de características similares. Isso enfatiza a necessidade de uma escolha cuidadosa de características.

Para efetuar redução de dimensionalidade, existem basicamente duas abordagens: extração de características e seleção de características. Em linhas gerais, os algoritmos de extração criam novas características a partir de transformações ou combinações do conjunto original de características. Já os algoritmos de seleção escolhem, segundo determinado critério, o melhor subconjunto do conjunto original de características.

Freqüentemente, a extração de características precede a seleção, de forma que inicialmente, é feita a extração de características a partir dos dados de entrada, seguida por um algoritmo de seleção de características que elimina os atributos mais irrelevantes segundo um determinado critério, reduzindo a dimensionalidade.

A escolha entre seleção e extração de características depende do domínio de aplicação e do conjunto específico de dados de treinamento disponíveis. Em geral, a seleção de características reduz o custo de medição de dados, e as características selecionadas mantêm sua interpretação física original, mantendo as propriedades que possuíam quando foram criadas. Já as características transformadas geradas por extração podem prover uma habilidade de discriminação melhor que o subconjunto das características originais. Entretanto, as novas características (combinações lineares ou não lineares das características originais) podem não possuir um significado físico.

É importante lembrar que, se a redução de dimensionalidade for excessiva, o classificador pode ter seu poder de discriminação reduzido. Por isso, é importante analisar a variação do comportamento do classificador com o número de características, de forma que seja possível estimar a dimensionalidade ideal para determinado classificador e conjunto de dados. A seguir, encontram-se maiores detalhes sobre a extração e a seleção de características.

4.3 EXTRAÇÃO DE CARACTERÍSTICAS

O método de extração de características abordado neste trabalho utiliza a matriz de co-ocorrência. Um método de extração de características cria um novo espaço a partir de transformações ou combinações das características do espaço original. Formalmente, dado um espaço de características I de dimensão N , um método de extração de características H é uma função $H : I \rightarrow F$, em que F possui dimensão m . Assim, dado um padrão x em um espaço de características I , temos

$$H(x) = y \quad (4.23)$$

tal que $y(y \in F)$ é a nova representação do padrão no espaço F .

Normalmente, $m < N$, mas nem sempre a redução de dimensionalidade é promovida diretamente pelos métodos de extração de características. Em geral, eles criam um novo espaço de características em que a determinação dos vetores mais representativos de sua base é muito simples. Por exemplo, pode-se nos piores casos, criar um espaço de características com $m = N$. Entretanto, geralmente basta selecionar os primeiros vetores da base criada

para reduzir a dimensionalidade de forma eficiente.

4.4 SELEÇÃO DE CARACTERÍSTICAS

Métodos automáticos de seleção de características são importantes em muitas situações em que se tem disponível um conjunto grande de características e deseja-se selecionar um subconjunto adequado. Além de ser uma forma de redução de dimensionalidade, uma aplicação importante é a fusão de dados procedentes de múltiplas modalidades de sensores ou de múltiplos modelos de dados. A importância da redução de dimensionalidade foi demonstrada na seção 4.2.

A seleção automática de características é uma técnica de otimização que, dado um conjunto de características, tenta selecionar um subconjunto de tamanho m , onde ($m < N$), que maximiza uma função critério $f(x)$. Um exemplo simples é definir $f(x) = 1 - E$, sendo E a taxa ou probabilidade de erro de um classificador.

É desejável que a função critério seja maior quanto menor for a redundância entre as características e quanto maior a facilidade de discriminar padrões de classes diferentes. Dessa forma, o algoritmo de seleção de características poderá reduzir a dimensionalidade de forma que ocorra a menor queda possível no poder de distinção das classes por um classificador no espaço de características. Uma consequência da aplicação de um bom algoritmo de seleção de atributos é a redução do número necessário de amostras de treinamento para se obter bons resultados com um classificador, ou seja, a redução do problema da dimensionalidade.

Além da escolha da função critério, também é importante determinar a dimensionalidade apropriada do espaço de características reduzido. Uma forma simples de resolver esse problema é efetuar a seleção de características para vários valores de m . Conforme foi mencionado em [18], os autores defendem que, em problemas práticos, sendo $|T|$ o tamanho do conjunto de treinamento, o problema da dimensionalidade pode ser evitado se forem usadas menos que $|T|/10$ características.

Apesar da importância de seleção de atributos, não há regras ou procedimentos definitivos para essa tarefa em cada aplicação particular, principalmente quando o número de características disponíveis for grande. Por esse motivo, um grande conjunto de algoritmos de seleção de atributos tem sido proposto. Em [17], foi proposta uma taxonomia sobre esse

tópico.

Para se obter o desempenho máximo de um classificador, é necessário investigar qual é a dimensionalidade ideal para um determinado problema de reconhecimento de padrões. Para isso, pode ser aplicada uma estratégia simples de tentativa e erro em relação à dimensionalidade, usando um método de redução de dimensionalidade (incluindo extração e seleção de características) até que o ponto de máximo desempenho de um classificador seja atingido. Nessa estratégia, são realizados testes de redução de dimensionalidade para a obtenção de sub-espacos de características de vários tamanhos diferentes até que seja obtida a dimensionalidade que minimiza o erro da classificação [10].

CAPÍTULO 5

RESULTADOS EXPERIMENTAIS

Com a finalidade de testar a metodologia proposta, foram desenvolvidos experimentos utilizando um conjunto de imagens contendo informação textural. O formato de imagem utilizado para a realização dos testes foi do tipo PGM (*Portable Gray Map*). Para maiores detalhes sobre esse formato, ver Apêndice A.

O algoritmo descrito no capítulo 4 foi desenvolvido para calcular um conjunto de características de textura em imagens do tipo PGM. As características de textura calculadas são listadas a seguir, conforme definidas na seção 4.1.6:

- (1) Segundo Momento Angular
- (2) Contraste
- (3) Correlação
- (4) Variância
- (5) Homogeneidade
- (6) Soma da Média
- (7) Soma da Variância
- (8) Soma da Entropia
- (9) Entropia
- (10) Diferença da Variância
- (11) Diferença da Entropia
- (12) Informação de Medidas de Correlação 1
- (13) Informação de Medidas de Correlação 2

Os recursos computacionais utilizados no desenvolvimento da ferramenta consistem em um computador com processador AMD Athlon K7-700 MHz e 128 Mbytes de memória RAM, sistema operacional Conectiva Linux 6, linguagem C, o pacote SPRING [49] e o simulador de

Redes Neurais SNNS [46].

No primeiro experimento, quatro matrizes de co-ocorrência foram construídas para cada amostra de textura, de tamanho 75×75 pixels, mostrada na figura 5.1, nas direções 0° , 45° , 90° e 135° , com uma distância de pixel unitária ($d = 1$). As treze características listadas acima foram computadas para as quatro matrizes, bem como para a média delas, resultando em um total de 65 características de textura para cada imagem baseada em estatísticas de co-ocorrência em níveis de cinza. A distância $d = 1$ foi escolhida para assegurar um bom desempenho na classificação. Características de textura calculadas para distâncias maiores ($d = 3, 5$ e 7) não produziram uma melhor classificação, além de aumentarem o esforço computacional.

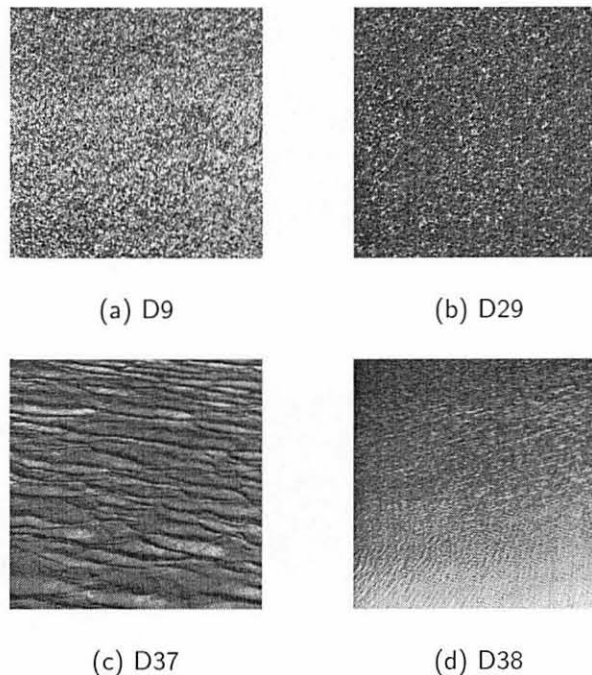


Figura 5.1: Amostras de texturas extraídas do álbum Brodatz.

As figuras 5.2 e 5.3 comparam os valores da média de cada característica de textura para as imagens D9.PGM (grama), D29.PGM (areia), D37.PGM (ondas) e D38.PGM (água). A partir dos gráficos gerados observa-se que o contraste, por exemplo, encontrou valores disjuntos para as 4 classes, fator que auxilia o processo de classificação. Por outro lado, a diferença da entropia obteve valores muito próximos, podendo ser descartada na etapa de seleção de características.

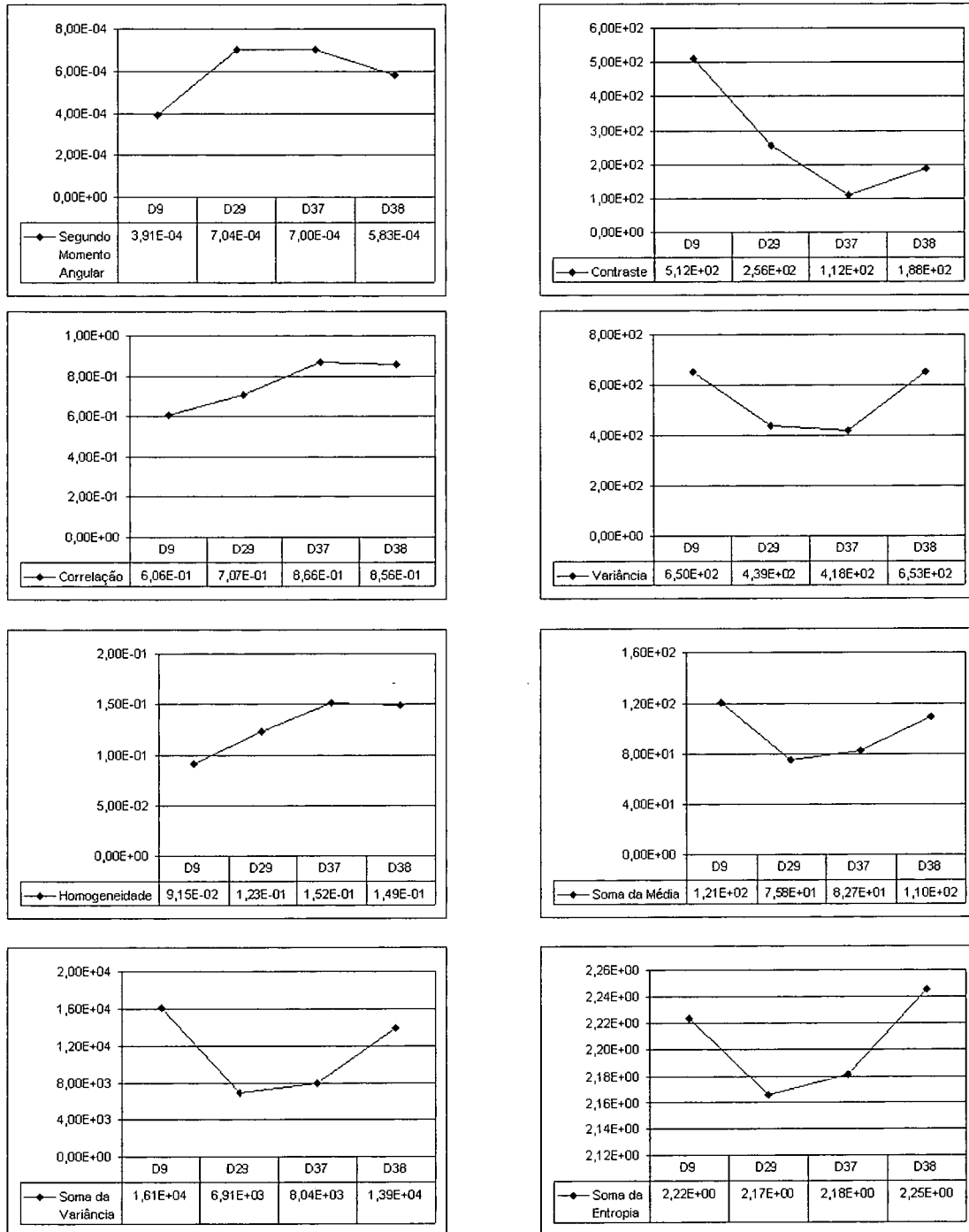


Figura 5.2: Gráficos comparativos para cada característica de textura extraída das imagens.

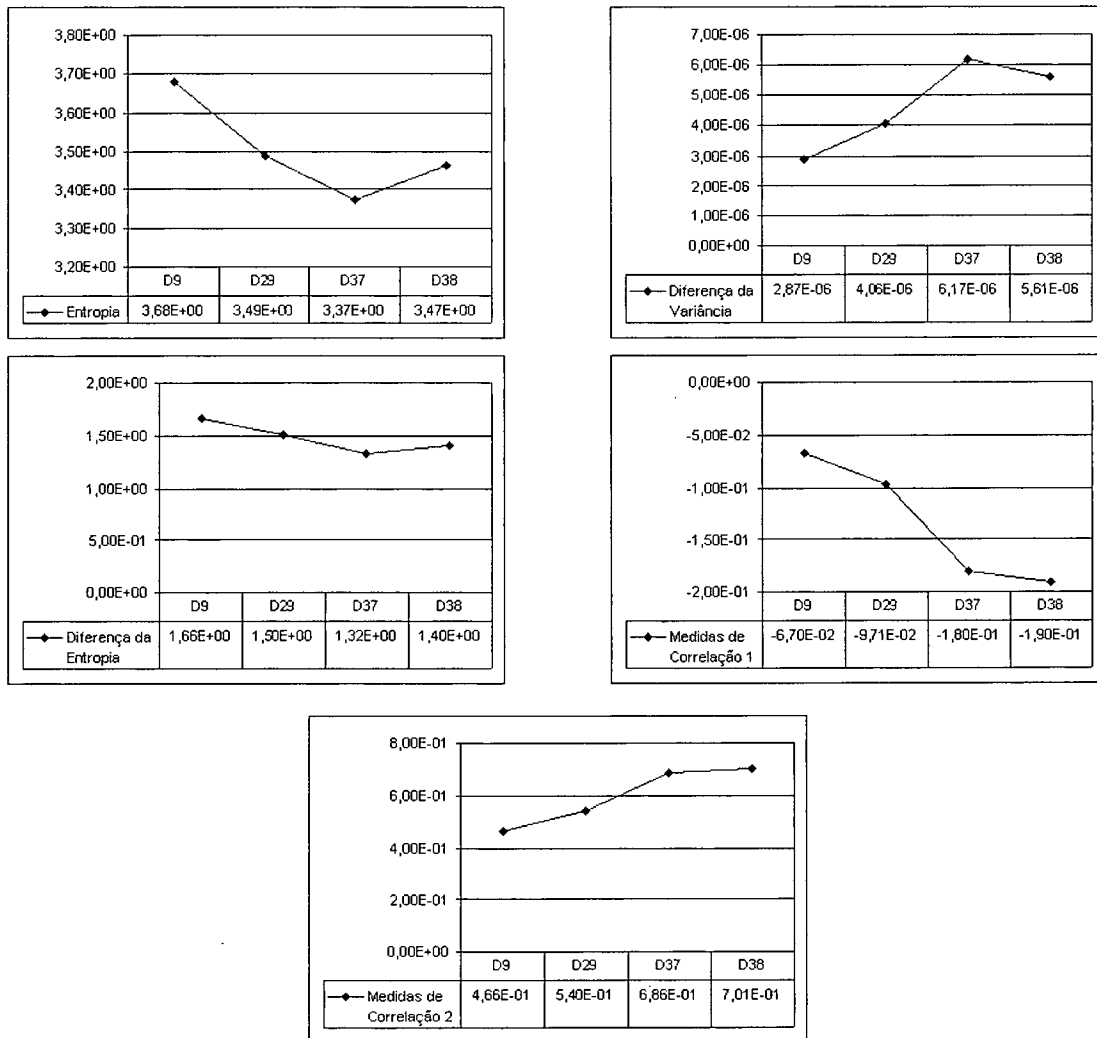


Figura 5.3: Gráficos comparativos para cada característica de textura extraída das imagens.

As tabelas 5.1, 5.2, 5.3 e 5.4 apresentam os resultados do cálculo de características de textura das imagens citadas anteriormente.

Característica	0°	45°	90°	135°	Média
Segundo Momento Angular	4.127e-04	3.363e-04	4.905e-04	3.261e-04	3.914e-04
Contraste	4.478e+02	6.129e+02	3.526e+02	6.336e+02	5.117e+02
Correlação	6.552e-01	5.281e-01	7.284e-01	5.122e-01	6.060e-01
Variância	6.504e+02	6.504e+02	6.502e+02	6.504e+02	6.504e+02
Homogeneidade	9.337e-02	7.987e-02	1.151e-01	7.782e-02	9.153e-02
Soma da Média	1.207e+02	1.207e+02	1.207e+02	1.207e+02	1.207e+02
Soma da Variância	1.618e+04	1.602e+04	1.628e+04	1.600e+04	1.612e+04
Soma da Entropia	2.228e+00	2.220e+00	2.230e+00	2.220e+00	2.224e+00
Entropia	3.660e+00	3.718e+00	3.619e+00	3.724e+00	3.680e+00
Diferença da Variância	2.941e-06	2.504e-06	3.582e-06	2.449e-06	2.869e-06
Diferença da Entropia	1.640e+00	1.706e+00	1.580e+00	1.714e+00	1.660e+00
Medidas de Correlação-1	-7.755e-02	-4.718e-02	-9.913e-02	-4.423e-02	-6.702e-02
Medidas de Correlação-2	5.057e-01	4.055e-01	5.607e-01	3.937e-01	4.664e-01

Tabela 5.1: Cálculo obtido para a imagem D9.PGM nas direções 0°, 45°, 90° e 135°.

Característica	0°	45°	90°	135°	Média
Segundo Momento Angular	8.078e-04	6.470e-04	7.806e-04	5.793e-04	7.037e-04
Contraste	1.913e+02	2.947e+02	1.975e+02	3.416e+02	2.563e+02
Correlação	7.816e-01	6.635e-01	7.745e-01	6.100e-01	7.074e-01
Variância	4.389e+02	4.390e+02	4.389e+02	4.390e+02	4.389e+02
Homogeneidade	1.411e-01	1.138e-01	1.353e-01	1.004e-01	1.227e-01
Soma da Média	7.576e+01	7.576e+01	7.578e+01	7.576e+01	7.576e+01
Soma da Variância	6.975e+03	6.873e+03	6.972e+03	6.827e+03	6.912e+03
Soma da Entropia	2.172e+00	2.163e+00	2.171e+00	2.157e+00	2.166e+00
Entropia	3.436e+00	3.519e+00	3.445e+00	3.550e+00	3.488e+00
Diferença da Variância	4.684e-06	3.710e-06	4.518e-06	3.329e-06	4.060e-06
Diferença da Entropia	1.444e+00	1.540e+00	1.454e+00	1.577e+00	1.504e+00
Medidas de Correlação-1	-1.250e-01	-7.995e-02	-1.204e-01	-6.296e-02	-9.707e-02
Medidas de Correlação-2	6.062e-01	5.040e-01	5.973e-01	4.539e-01	5.404e-01

Tabela 5.2: Cálculo obtido para a imagem D29.PGM nas direções 0°, 45°, 90° e 135°.

Característica	0°	45°	90°	135°	Média
Segundo Momento Angular	1.116e-03	5.319e-04	5.754e-04	5.751e-04	6.997e-04
Contraste	3.616e+01	1.517e+02	1.295e+02	1.289e+02	1.115e+02
Correlação	9.566e-01	8.180e-01	8.446e-01	8.453e-01	8.661e-01
Variância	4.177e+02	4.177e+02	4.177e+02	4.177e+02	4.177e+02
Homogeneidade	2.221e-01	1.222e-01	1.322e-01	1.297e-01	1.516e-01
Soma da Média	8.274e+01	8.273e+01	8.275e+01	8.273e+01	8.274e+01
Soma da Variância	8.120e+03	8.003e+03	8.029e+03	8.025e+03	8.044e+03
Soma da Entropia	2.185e+00	2.179e+00	2.181e+00	2.181e+00	2.181e+00
Entropia	3.176e+00	3.458e+00	3.427e+00	3.428e+00	3.372e+00
Diferença da Variância	9.482e-06	4.805e-06	5.187e-06	5.188e-06	6.166e-06
Diferença da Entropia	1.116e+00	1.414e+00	1.381e+00	1.380e+00	1.323e+00
Medidas de Correlação-1	-2.865e-01	-1.340e-01	-1.510e-01	-1.500e-01	-1.804e-01
Medidas de Correlação-2	8.088e-01	6.256e-01	6.547e-01	6.531e-01	6.856e-01

Tabela 5.3: Cálculo obtido para a imagem D37.PGM nas direções 0°, 45°, 90° e 135°.

Característica	0°	45°	90°	135°	Média
Segundo Momento Angular	1.014e-03	4.358e-04	4.535e-04	4.280e-04	5.828e-04
Contraste	3.552e+01	2.395e+02	2.311e+02	2.459e+02	1.880e+02
Correlação	9.728e-01	8.162e-01	8.227e-01	8.113e-01	8.557e-01
Variância	6.533e+02	6.525e+02	6.526e+02	6.525e+02	6.527e+02
Homogeneidade	2.414e-01	1.181e-01	1.224e-01	1.159e-01	1.494e-01
Soma da Média	1.095e+02	1.095e+02	1.095e+02	1.095e+02	1.095e+02
Soma da Variância	1.408e+04	1.387e+04	1.388e+04	1.386e+04	1.392e+04
Soma da Entropia	2.246e+00	2.247e+00	2.247e+00	2.246e+00	2.246e+00
Entropia	3.205e+00	3.554e+00	3.541e+00	3.561e+00	3.465e+00
Diferença da Variância	1.013e-05	4.078e-06	4.191e-06	4.039e-06	5.610e-06
Diferença da Entropia	1.104e+00	1.502e+00	1.493e+00	1.507e+00	1.402e+00
Medidas de Correlação-1	-3.264e-01	-1.441e-01	-1.506e-01	-1.404e-01	-1.904e-01
Medidas de Correlação-2	8.447e-01	6.512e-01	6.621e-01	6.450e-01	7.007e-01
Coef. Máx. de Correlação	1.655e-03	3.170e-03	1.284e-03	1.991e-03	2.025e-03

Tabela 5.4: Cálculo obtido para a imagem D38.PGM nas direções 0°, 45°, 90° e 135°.

A figura 5.4 mostra uma imagem de teste, de tamanho 640×640 pixels, capturada de uma região do Rio Nilo localizada no sul do Egito. A partir dessa imagem, foram recortadas um conjunto de quatro amostras de tamanho 16×16 pixels de cada uma das seguintes texturas: solo e vegetação, como mostram as figuras 5.5 e 5.6, respectivamente.

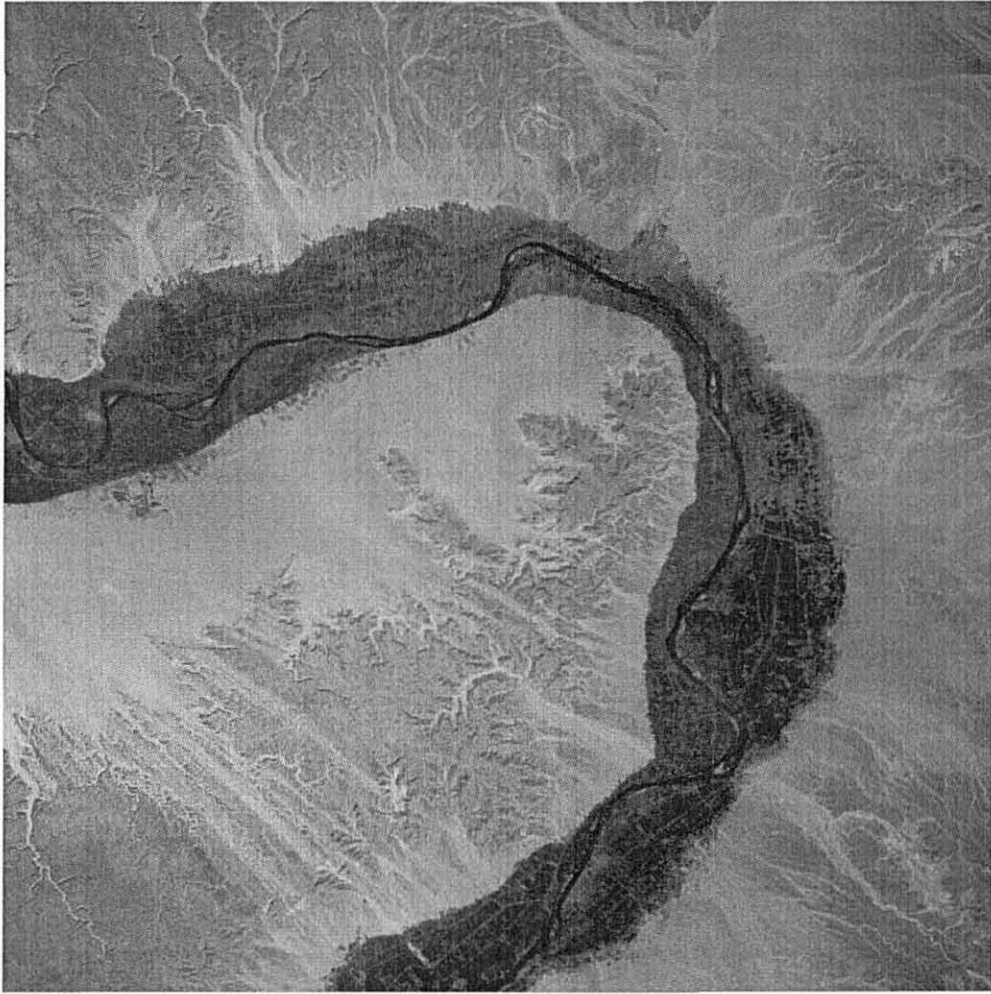


Figura 5.4: Imagem do *Valley of the Kings*, região do Rio Nilo.

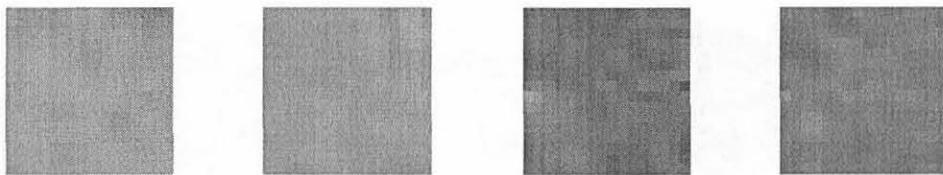


Figura 5.5: Amostras de solo da região do Rio Nilo.

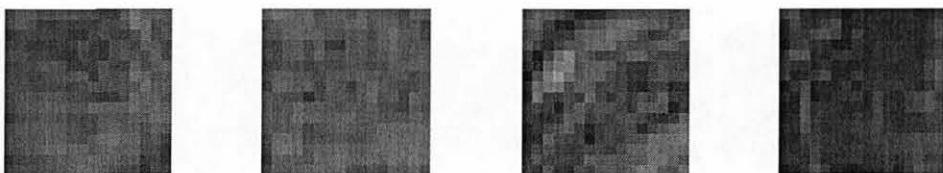


Figura 5.6: Amostras de vegetação da região do Rio Nilo.

Em seguida, o algoritmo foi executado separadamente para cada uma das amostras, fazendo-se um treinamento nas classes, ou seja, cada classe armazenou os valores médios das características calculadas pelo programa através das matrizes de co-ocorrência, como mostra a tabela 5.5.

Característica	Solo	Vegetação
Segundo Momento Angular	4.286e-02	1.147e-02
Contraste	4.583e+00	2.428e+01
Correlação	3.425e-01	3.909e-01
Variância	4.549e+00	2.427e+01
Homogeneidade	4.501e-01	2.701e-01
Soma da Média	1.236e+01	1.772e+01
Soma da Variância	1.523e+02	3.893e+02
Soma da Entropia	1.042e+00	1.385e+00
Entropia	1.598e+00	2.116e+00
Diferença da Variância	2.335e-03	3.404e-04
Diferença da Entropia	6.632e-01	9.823e-01
Medidas de Correlação-1	-1.005e-01	-1.906e-01
Medidas de Correlação-2	3.855e-01	5.863e-01

Tabela 5.5: Média das características calculadas a partir das matrizes de co-ocorrência das amostras de solo e vegetação recortadas da região do Rio Nilo.

Os valores mostrados na tabela 5.5 representam dois tipos distintos de textura, solo e vegetação. Porém, as médias encontradas para determinadas características tais como correlação, soma da média e soma da entropia apresentaram valores muito próximos, podendo ser descartados, visto que não auxiliariam o processo de classificação.

As seções seguintes descrevem cinco experimentos realizados para classificar imagens através de características de textura. Quatro técnicas de classificação foram utilizadas, o método da distância mínima, o método da máxima verossimilhança, o método baseado em intervalos e redes neurais.

5.1 CLASSIFICAÇÃO 1

O primeiro teste de classificação executado utiliza-se da imagem original, de tamanho 250×250 pixels, mostrada na figura 5.7(a). A partir dessa imagem, foram recortadas 30 amostras de textura de tamanho 3×3 pixels pertencentes a uma única classe, as quais foram utilizadas para o treinamento do sistema de classificação implementado. Ao final da apresentação da última amostra, foi considerado finalizado o processo de aprendizado. Através de experimentação, escolheu-se o conjunto de características que melhor representava a classe estatisticamente.

Para fins de comparação dos classificadores e para cada imagem sob consideração, todo o processo de extração, seleção e classificação, bem como a definição do tamanho da janela, foi realizado com os mesmos dados e parâmetros, alterando-se somente a técnica utilizada em cada caso.

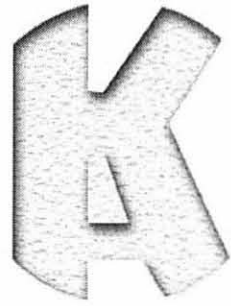
Segundo a abordagem proposta, a classificação utiliza sempre o mesmo tamanho de janela definido na etapa de treinamento. Verificou-se que essa igualdade proporciona um processo de classificação correto.

A tabela 5.6 mostra os dados de treinamento para o classificador da Distância Mínima. Foram selecionadas cinco características de um total de treze que tiveram capacidade de representar adequadamente os dados daquela textura. Os valores foram armazenados obtendo-se seqüencialmente a média entre um valor já existente e novos valores de amostras sendo apresentados ao processo de treinamento.

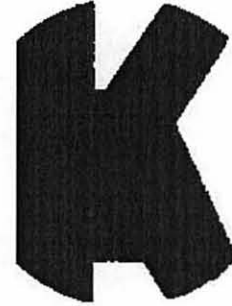
Característica	Média
Segundo Momento Angular	1.324e-01
Contraste	4.699e+00
Correlação	2.335e-01
Variância	3.933e+00
Homogeneidade	3.798e-01

Tabela 5.6: Distância Mínima - Letra.

Na tabela 5.7 são apresentados os limites inferiores e superiores da classe em questão utilizados para o classificador baseado em Intervalos. Este classificador deve ter um certo grau de tolerância, pois na maioria dos casos o espaço de atributos não é representado estatisticamente com 100% de acerto.



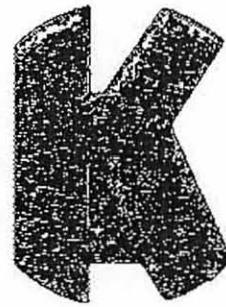
(a) Imagem original



(b) Distância Mínima



(c) Máxima Verossimilhança



(d) Intervalos



(e) Redes Neurais

Figura 5.7: Classificação - Letra.

Característica	Lim. Inf.	Lim. Sup.
Segundo Momento Angular	1.042e-01	3.333e-01
Contraste	1.979e+00	1.458e+01
Correlação	-2.986e-01	2.993e-01
Variância	1.781e+00	7.413e+00
Homogeneidade	1.852e-01	6.604e-01

Tabela 5.7: Intervalos - Letra.

No classificador da Máxima Verossimilhança, foram definidos limiares para o classificador, permitindo que valores de determinada característica que ultrapassassem em no máximo 1% os limites inferiores e superiores da classe fossem classificados dentro da mesma classe.

A seguir é mostrada uma parte do arquivo de treinamento gerado pelo sistema implementado e utilizado pelo SNNS para realizar o aprendizado da rede neural. No cabeçalho são indicados o número de padrões extraídos, o número de neurônios de entrada e o número de neurônios de saída. Na seqüência, cada par de linhas indica a extração de características de uma amostra; a primeira linha traz o cálculo de cinco características de textura e a segunda linha indica através do número 1 a classe a qual pertence tal amostra.

```
SNNS pattern definition file V3.2
generated at Mon Feb 18:08:50 2003
```

```
No. of patterns: 30
```

```
No. of input units: 5
```

```
No. of output units: 2
```

```
0.161458 6.291667 -0.050329 4.043403 0.429732
```

```
1 0
```

```
0.114583 11.104167 -0.265579 5.476996 0.248128
```

```
1 0
```

```
0.107639 14.583333 -0.278854 6.703993 0.216984
```

```
1 0
```

```
0.107639 8.270833 0.005193 5.336372 0.297120
```

```
1 0
  0.104167  12.979167 -0.102094  6.706163  0.197302
1 0
  0.119792  11.895833 -0.232025  5.740885  0.185185
1 0
  0.107639  8.875000  0.047135  5.970486  0.226301
1 0
  0.125000  6.312500  0.180062  5.171441  0.303139
1 0
  0.125000  5.500000  0.275889  4.834201  0.307353
1 0
  0.143229  2.500000  0.299294  2.667535  0.475000
1 0
.
.
. (continua)
```

Os dados seguintes demonstram as características da imagem de teste no formato SNNS. Estas são extraídas da matriz de co-ocorrência de cada pixel e seus respectivos vizinhos em uma janela $N \times N$. Pode-se verificar que nesse caso não se especifica a classe à qual pertence cada padrão extraído. Esse arquivo é submetido à rede neural após terminado o processo de treinamento para realizar a classificação.

```
SNNS pattern definition file V3.2
generated at Mon Feb 18:08:50 2003
```

```
No. of patterns: 62500
```

```
No. of input units: 5
```

```
No. of output units: 0
```



```
#1
 1.000000  0.000000  0.000000  1.000000  1.000000
#2
 1.000000  0.000000  0.000000  1.000000  1.000000
#3
 1.000000  0.000000  0.000000  1.000000  1.000000
#4
 1.000000  0.000000  0.000000  1.000000  1.000000
#5
 1.000000  0.000000  0.000000  1.000000  1.000000
#6
 0.347222  1.791667 -0.009350  1.922743  0.679167
#7
 0.125000  13.375000 -0.365906  5.818576  0.226873
#8
 0.107639  6.812500  0.298903  5.898871  0.242279
#9
 0.118056  4.770833  0.271169  4.350260  0.305760
#10
 0.118056  3.937500  0.276426  3.759983  0.331250
.
.
. (continua)
```

Enquanto os dados de teste são submetidos à rede neural, o resultado encontrado para cada pixel é armazenado em um arquivo de saída que determina à qual classe o mesmo pertence. Com isso, possibilita-se a geração de um arquivo PGM classificado, com as mesmas dimensões do original.

generated at Sat Feb 17:01:38 2003

No. of patterns : 62500

No. of input units : 5

No. of output units : 2

startpattern : 1

endpattern : 62500

#1.1

0.09966 0.90026

#2.1

0.09966 0.90026

#3.1

0.09966 0.90026

#4.1

0.09966 0.90026

#5.1

0.09966 0.90026

#6.1

0.09966 0.90026

#7.1

0.09966 0.90026

#8.1

0.09966 0.90026

#9.1

0.09966 0.90026

#10.1

0.09966 0.90026

.

. (continua)

A imagem mostrada na figura 5.8 apresenta a rede neural treinada para uma classificação envolvendo a imagem de teste. A estrutura da rede é formada por cinco neurônios de entrada, cinco neurônios na camada oculta e dois neurônios de saída. O método de treinamento é realizado abastecendo os neurônios de entrada com dados de cada amostra e indicando para esta mesma amostra qual é o neurônio de saída a ser ativado. Seguindo esse procedimento, a rede utiliza-se do método de retropropagação para aprender com cada nova entrada, ajustando os pesos de forma adequada.

Na etapa de classificação, apenas os neurônios de entrada recebem valores mantendo os pesos de suas interconexões inalterados. Esses ativam outros neurônios interconectados a eles até que alcancem um neurônio de saída, o qual determina se o pixel em questão pertence ou não à classe de textura.

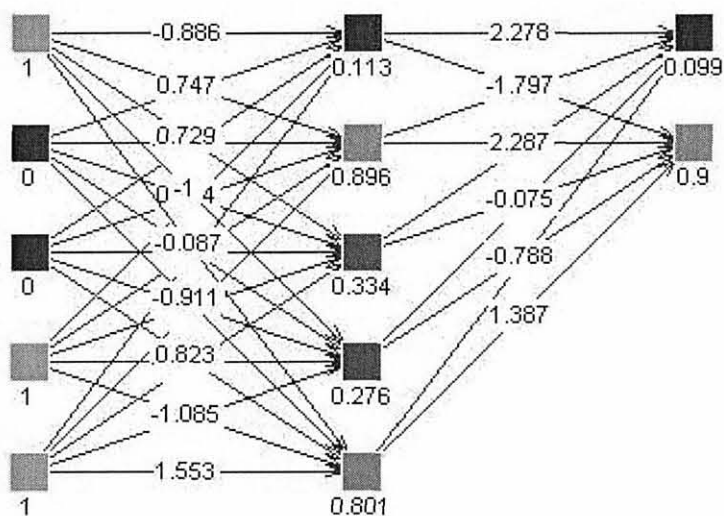


Figura 5.8: Rede Neural - Letra.

5.2 CLASSIFICAÇÃO 2

A imagem de teste mostrada na figura 5.9(a), de tamanho 450×434 pixels, extraída do satélite LANDSAT, é de uma região chamada Kodiak Island no Alaska, Estados Unidos. Foram extraídas 20 amostras de tamanho 3×3 pixels de cada uma das três classes, totalizando 60 amostras. O treinamento realizado segue os mesmos passos da classificação 1, descrita na seção 5.1. As tabelas 5.8 e 5.9 mostram os arquivos de treinamento utilizados para a classificação.

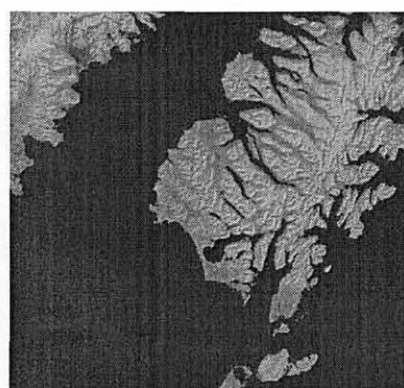
A Rede Neural treinada para a imagem de satélite é mostrada na figura 5.10.

Característica	Média	Média	Média
Segundo Momento Angular	3.126e-01	1.909e-01	1.235e-01
Contraste	9.045e-01	2.792e+00	3.210e+00
Correlação	-1.198e-01	-1.645e-01	1.793e-01

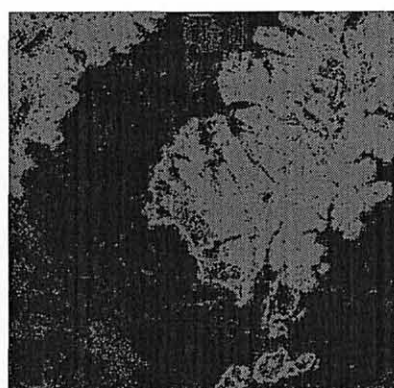
Tabela 5.8: Distância Mínima - Satélite.

Característica	Lim. Inf.	Lim. Sup.	Lim. Inf.	Lim. Sup.	Lim. Inf.	Lim. Sup.
Segundo Momento Angular	1.111e-01	1.476e-01	1.224e-01	2.457e-01	2.014e-01	1.000e+00
Contraste	3.417e+00	1.004e+01	1.229e+00	4.646e+00	2.917e-01	1.125e+00
Correlação	-2.864e-01	1.045e-02	-6.847e-02	3.212e-02	0.000e+00	0.000e+00
Variância	2.668e+00	4.893e+00	1.659e+00	3.237e+00	1.000e+00	1.467e+00
Homogeneidade	2.497e-01	4.578e-01	4.240e-01	6.354e-01	6.375e-01	1.000e+00
Soma da Média	6.583e+00	8.792e+00	5.312e+00	7.979e+00	2.000e+00	4.458e+00
Soma da Variância	3.946e+01	7.246e+01	2.440e+01	5.756e+01	4.000e+00	1.686e+01
Soma da Entropia	5.772e-01	6.023e-01	5.176e-01	6.901e-01	-4.343e-10	4.454e-01
Entropia	8.782e-01	9.660e-01	6.995e-01	9.347e-01	-4.343e-10	7.183e-01
Diferença da Variância	5.678e-03	1.423e-02	1.003e-02	2.083e-02	3.549e-02	8.507e-02
Diferença da Entropia	4.392e-01	5.647e-01	4.203e-01	5.050e-01	-4.343e-10	3.796e-01

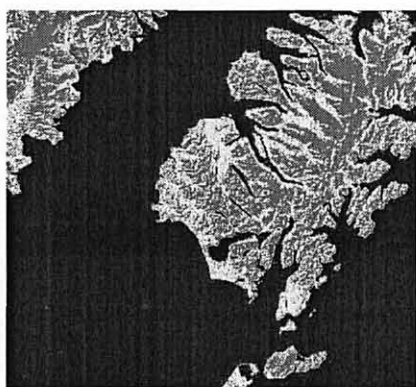
Tabela 5.9: Intervalos - Satélite.



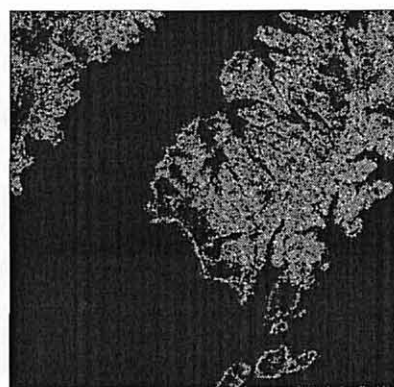
(a) Imagem original



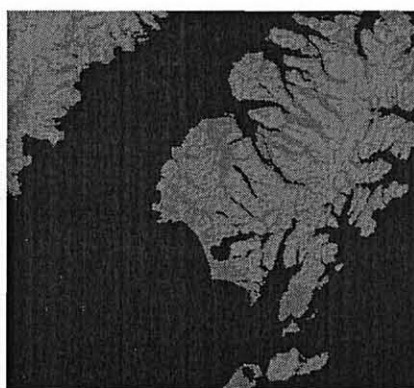
(b) Distância Mínima



(c) Máxima Verossimilhança



(d) Intervalos



(e) Redes Neurais

Figura 5.9: Classificação - Satélite.

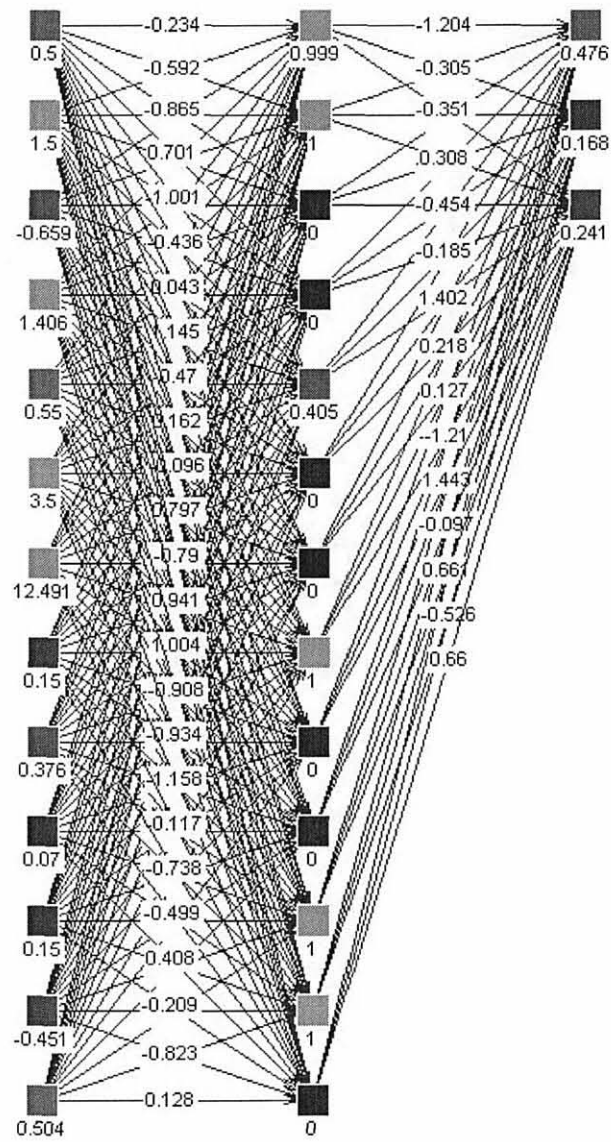


Figura 5.10: Rede Neural - Satélite.

5.3 CLASSIFICAÇÃO 3

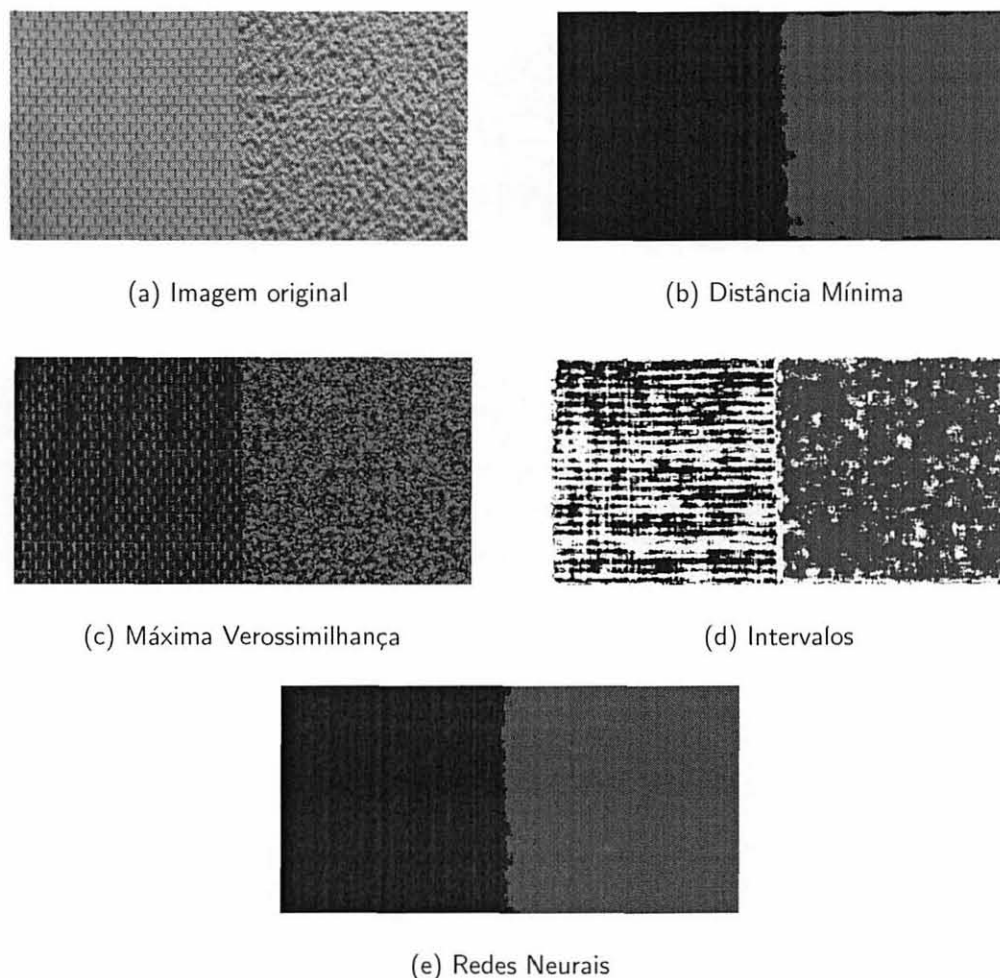


Figura 5.11: Classificação - Mosaico I.

A imagem de teste mostrada na figura 5.11(a), de tamanho 512×256 pixels, contém a união de duas texturas complexas do álbum Brodatz [6], nomeadas D6(fio de alumínio entrelaçado) e D57(papel amassado), respectivamente. Para a realização do teste, foram extraídas 30 amostras de textura de tamanho 15×15 pixels de cada uma das classes, totalizando 60 amostras.

Os resultados do treinamento pela Distância Mínima são mostrados na tabela 5.10 e pelos Intervalos na tabela 5.11.

A Rede Neural treinada para a imagem do mosaico I é mostrada na figura 5.12.

Característica	Média	Média
Segundo Momento Angular	3.549e-03	2.676e-03
Contraste	1.480e+02	4.211e+02
Correlação	7.110e-01	7.541e-01
Variância	2.565e+02	8.578e+02
Homogeneidade	1.369e-01	6.602e-02

Tabela 5.10: Distância Mínima - Mosaico I.

Característica	Lim. Inf.	Lim. Sup.	Lim. Inf.	Lim. Sup.
Segundo Momento Angular	3.479e-03	4.766e-03	2.642e-03	2.784e-03
Contraste	1.067e+02	1.730e+02	2.801e+02	4.328e+02
Correlação	6.928e-01	7.288e-01	7.160e-01	7.959e-01
Variância	1.931e+02	2.892e+02	6.021e+02	9.372e+02
Homogeneidade	1.315e-01	1.677e-01	6.241e-02	9.153e-02
Soma da Média	7.457e+01	9.768e+01	9.200e+01	1.129e+02
Soma da Variância	5.956e+03	1.018e+04	1.039e+04	1.529e+04
Soma da Entropia	1.804e+00	1.864e+00	2.019e+00	2.087e+00
Entropia	2.429e+00	2.502e+00	2.571e+00	2.587e+00
Diferença da Variância	1.020e-05	1.890e-05	2.955e-06	4.313e-06
Diferença da Entropia	1.282e+00	1.362e+00	1.486e+00	1.566e+00
Medidas de Correlação-1	8.961e-01	5.282e+04	3.247e-01	9.653e-01
Medidas de Correlação-2	8.961e-01	5.282e+04	3.247e-01	9.653e-01

Tabela 5.11: Intervalos - Mosaico I.

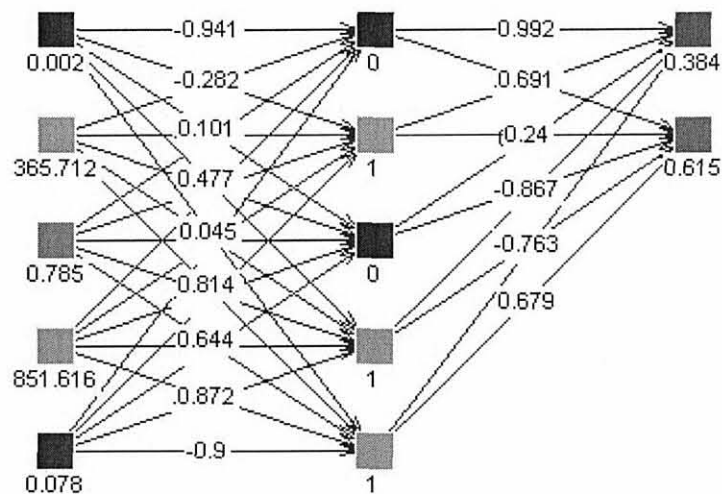


Figura 5.12: Rede Neural - Mosaico I.

Característica	Média	Média	Média	Média	Média
Segundo Momento Angular	1.403e-02	1.403e-02	1.852e-02	2.372e-02	1.361e-02
Contraste	1.150e+02	4.582e+01	1.578e+01	1.918e+02	1.918e+02
Correlação	-3.522e-02	-3.522e-02	-8.884e-02	-1.285e-02	1.465e-01
Variância	5.675e+01	5.675e+01	2.198e+01	8.803e+00	1.133e+02
Homogeneidade	9.761e-02	9.761e-02	1.635e-01	2.630e-01	7.997e-02
Soma da Média	3.226e+01	3.226e+01	1.851e+01	1.229e+01	3.948e+01
Soma da Variância	1.065e+03	1.065e+03	3.382e+02	1.407e+02	1.708e+03
Soma da Entropia	1.359e+00	1.359e+00	1.201e+00	1.101e+00	1.420e+00
Entropia	1.865e+00	1.865e+00	1.771e+00	1.677e+00	1.874e+00
Diferença da Variância	9.489e-05	9.489e-05	2.972e-04	1.010e-03	4.249e-05
Diferença da Entropia	1.152e+00	1.152e+00	1.045e+00	8.518e-01	1.251e+00

Tabela 5.12: Distância Mínima - Mosaico II.

5.4 CLASSIFICAÇÃO 4

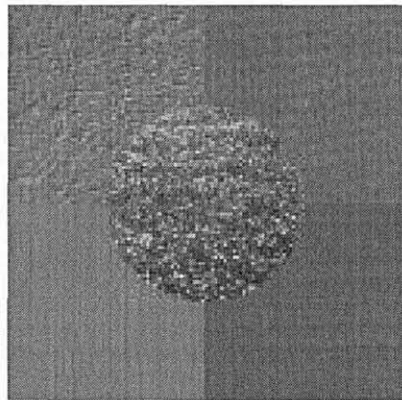
A imagem de teste mostrada na figura 5.13(a), de tamanho 128×128 pixels, contém um mosaico de 5 texturas do álbum Brodatz [6]. Mosaicos são muito utilizados para testes para verificar a precisão dos classificadores. Para a realização do teste, 20 amostras de textura de tamanho 7×7 pixels foram extraídas de cada uma das classes, totalizando 100 amostras.

As tabelas 5.12 e 5.13 mostram, respectivamente, os arquivos de treinamento utilizados para a classificação da Distância Mínima e de Intervalos.

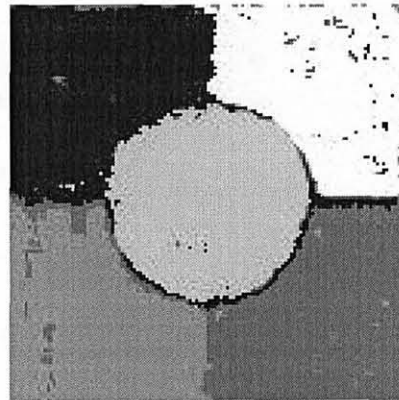
A Rede Neural treinada para a imagem do mosaico II é mostrada na figura 5.14.

Característica	Lim. Inf.	Lim. Sup.	Lim. Inf.	Lim. Sup.	Lim. Inf.	Lim. Sup.	Lim. Inf.	Lim. Sup.	Lim. Inf.	Lim. Sup.
Segundo Momento Angular	1.373e-02	1.548e-02	1.373e-02	1.548e-02	1.640e-02	2.445e-02	2.190e-02	3.156e-02	1.313e-02	1.414e-02
Contraste	7.090e+01	1.308e+02	7.090e+01	1.308e+02	2.768e+01	5.731e+01	8.778e+00	2.370e+01	1.683e+02	2.655e+02
Correlação	-8.142e-02	1.707e-01	-8.142e-02	1.707e-01	-1.294e-01	1.156e-01	-2.323e-01	5.717e-01	-9.948e-02	2.306e-01
Variância	4.112e+01	6.653e+01	4.112e+01	6.653e+01	1.332e+01	2.727e+01	6.535e+00	1.429e+01	8.823e+01	1.251e+02
Homogeneidade	8.563e-02	1.632e-01	8.563e-02	1.632e-01	1.297e-01	2.194e-01	2.018e-01	3.085e-01	5.943e-02	8.620e-02
Soma da Média	2.589e+01	3.355e+01	2.589e+01	3.355e+01	1.409e+01	2.056e+01	1.174e+01	1.605e+01	3.607e+01	4.279e+01
Soma da Variância	6.866e+02	1.157e+03	6.866e+02	1.157e+03	1.897e+02	4.199e+02	1.289e+02	2.586e+02	1.400e+03	1.936e+03
Soma da Entropia	1.318e+00	1.386e+00	1.318e+00	1.386e+00	1.126e+00	1.276e+00	1.025e+00	1.228e+00	1.365e+00	1.442e+00
Entropia	1.834e+00	1.872e+00	1.834e+00	1.872e+00	1.693e+00	1.818e+00	1.585e+00	1.707e+00	1.867e+00	1.885e+00
Diferença da Variância	7.344e-05	1.386e-04	7.344e-05	1.386e-04	2.352e-04	5.194e-04	7.793e-04	1.125e-03	3.620e-05	6.007e-05
Diferença da Entropia	1.092e+00	1.207e+00	1.092e+00	1.207e+00	9.606e-01	1.078e+00	7.645e-01	9.184e-01	1.221e+00	1.295e+00

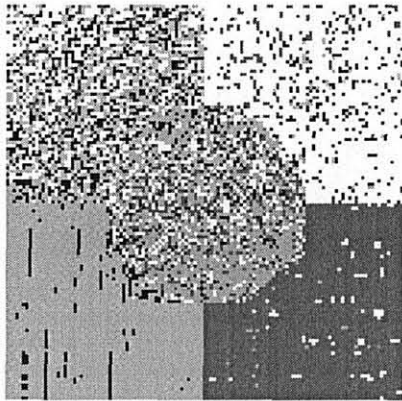
Tabela 5.13: Intervalos - Mosaico II



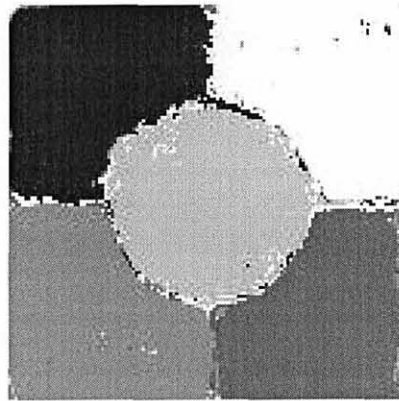
(a) Imagem original



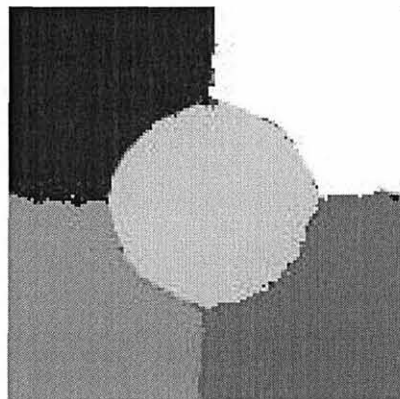
(b) Distância Mínima



(c) Máxima Verossimilhança



(d) Intervalos



(e) Redes Neurais

Figura 5.13: Classificação - Mosaico II.

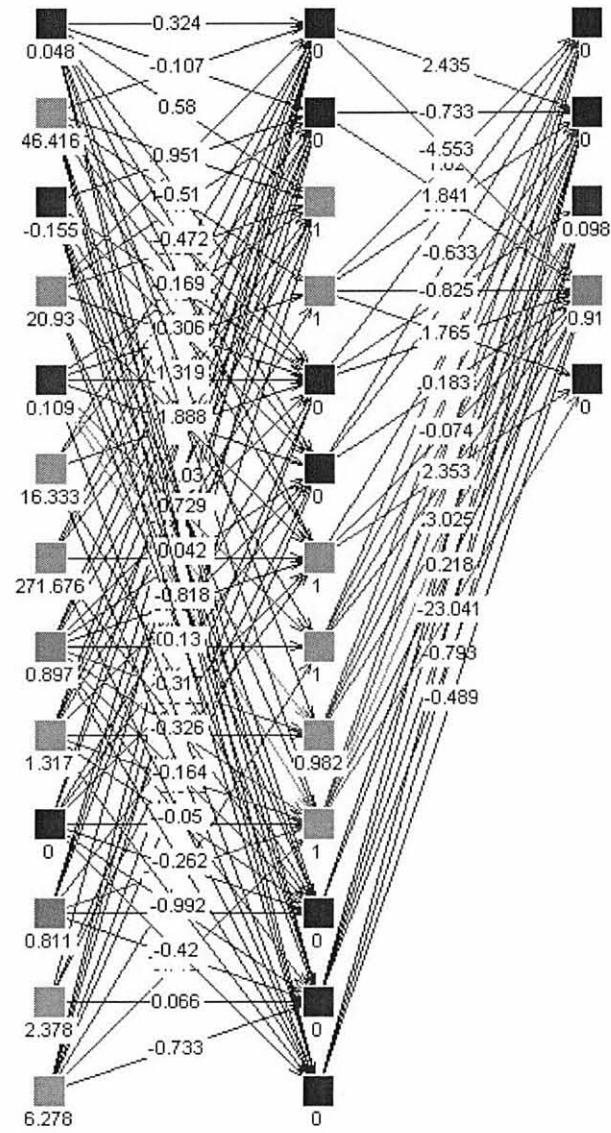


Figura 5.14: Rede Neural - Mosaico II.

5.5 CLASSIFICAÇÃO 5

A imagem de teste mostrada na figura 5.15(a), de tamanho 179×200 pixels, contém um conjunto de grãos de feijão. Para a realização do teste, 5 amostras de textura de tamanho 3×3 pixels foram extraídas de cada uma das classes, totalizando 10 amostras.

As tabelas 5.14 e 5.15 mostram, respectivamente, os arquivos de treinamento utilizados para a classificação da Distância Mínima e de Intervalos. A figura 5.16 mostra a rede neural treinada. Para esse teste, foram selecionadas apenas três características de textura: segundo momento angular, contraste e correlação, cujos valores foram suficientes e significativos para a classificação.

Característica	Média	Média
Segundo Momento Angular	7.100e-01	2.524e-01
Contraste	1.875e-01	7.982e-01
Correlação	-8.385e-09	-8.726e-02

Tabela 5.14: Distância Mínima - Grãos de Feijão.

Característica	Lim. Inf.	Lim. Sup.	Lim. Inf.	Lim. Sup.
Segundo Momento Angular	4.201e-01	1.000e+00	1.693e-01	3.116e-01
Contraste	0.000e+00	3.750e-01	5.625e-01	1.271e+00
Correlação	-2.235e-08	0.000e+00	-1.452e-01	2.505e-01

Tabela 5.15: Intervalos - Grãos de Feijão.

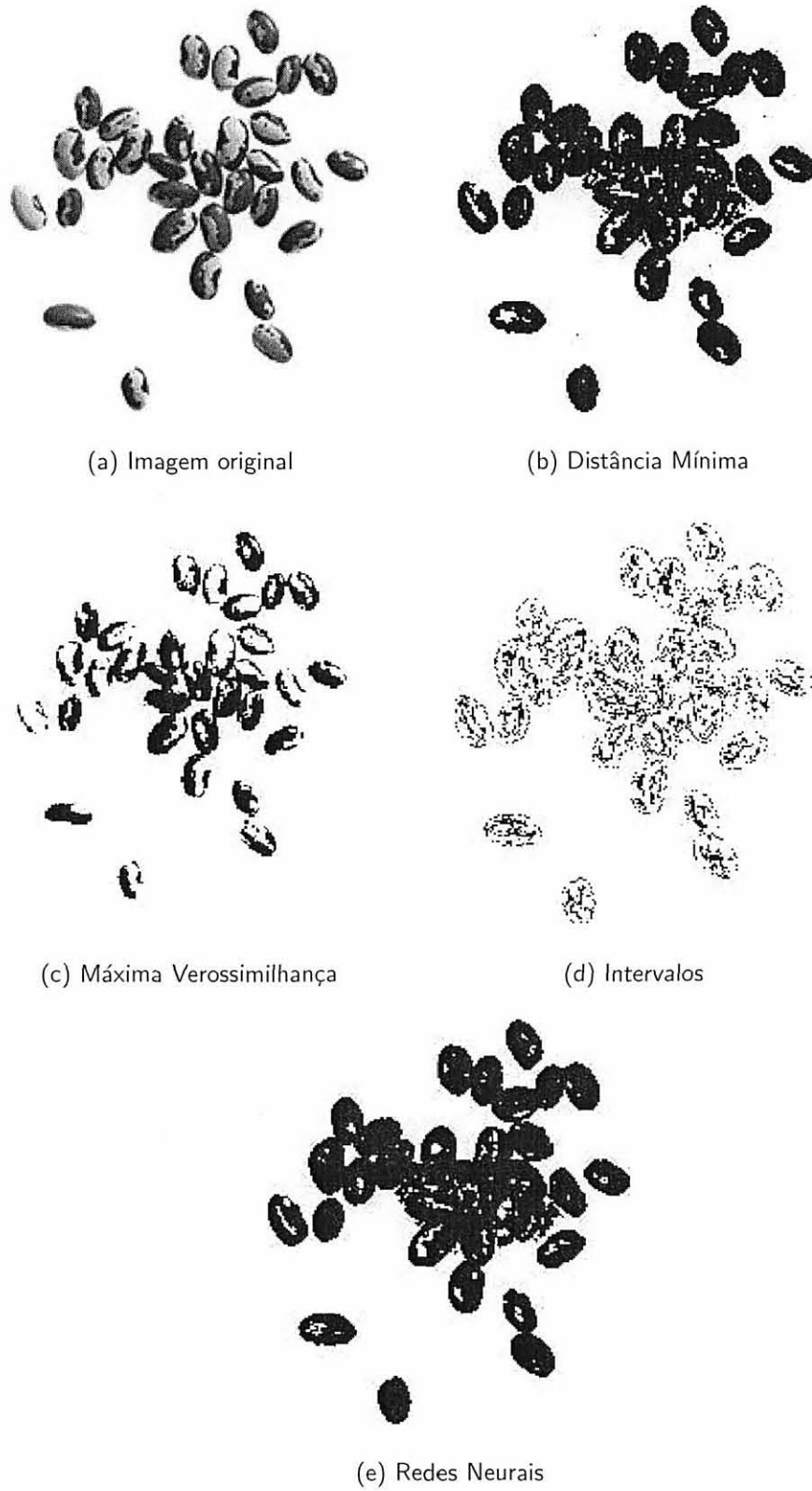


Figura 5.15: Classificação - Grãos de Feijão.

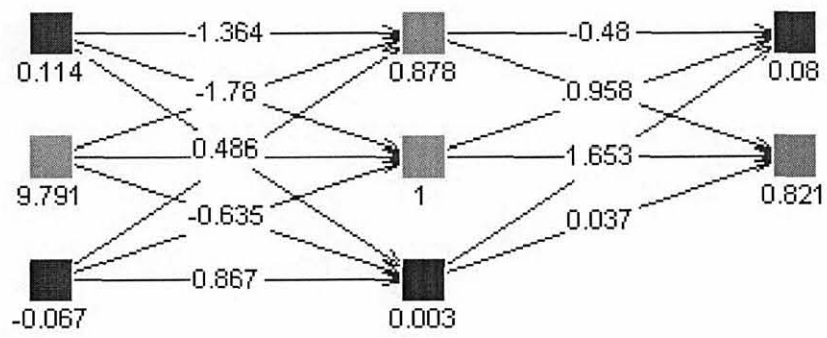


Figura 5.16: Rede Neural - Grãos de Feijão.

5.6 ANÁLISE DOS RESULTADOS

De acordo com [25], 90% da variabilidade na precisão de classificação usando textura dependem do tamanho da janela usado para calcular textura, enquanto os algoritmos particulares de textura determinam somente 10%. Essa é uma particularidade importante, pois a maioria dos estudos envolvendo textura abordam o desenvolvimento e a avaliação de algoritmos sem se preocupar com o efeito do tamanho da janela [12].

As figuras 5.17(a)-(h), de tamanho 450×435 pixels, demonstram a mudança na variabilidade de textura para uma imagem capturada pelo satélite LANDSAT do Mar do Mediterrâneo, Grécia. A partir da imagem original, foram geradas algumas imagens utilizando a medida de contraste como característica de textura. O procedimento foi definir um tamanho de janela de classificação, calcular a matriz de co-ocorrência e o contraste para as N janelas da imagem, uma janela para cada pixel da imagem, e ativar o pixel central da mesma com o valor encontrado.

Pode-se notar que a variabilidade da textura tende a se acentuar com o aumento do tamanho da janela, devido à grande quantidade de informação que se define dentro de uma janela, que é considerada uma possível classe textural.

Foi avaliado que o tamanho da janela, em pixels, depende da textura em questão para realizar uma classificação satisfatória. Janelas de tamanho pequeno não capturam o padrão particular da maioria das classes, enquanto janelas maiores podem incluir pixels de mais do que uma classe, resultando em classificações mais pobres. O tamanho ideal de janela seria aquele que produz a menor variabilidade na textura.

Como a imagem de textura não é uniforme em uma cena, ou mesmo em uma simples classe, amostras de textura simplificam a interpretação da cena inteira. Por esse motivo, tanto o treinamento como a classificação seguem esse mesmo padrão neste trabalho.

De acordo com os experimentos realizados, o classificador que pode ser considerado mais robusto é o de Redes Neurais. Sua grande capacidade de percepção em relação às classes texturais apresentadas a ele e às variações entre elas leva a essa conclusão.

O método baseado em Intervalos apresentou dificuldades em classificar pixels que não se

enquadram entre os limites inferior e superior de cada classe e também em classificar pixels onde os limites das classes se sobrepõem. Nesse último caso, a definição de critérios de decisão adequados que possam corretamente se ajustar e identificar cada uma das classes é uma tarefa difícil.

O classificador da Máxima Verossimilhança requer a utilização de uma grande quantidade de amostras representativas, preferencialmente selecionadas por um especialista na aplicação. A escolha inadequada das amostras causa confusão nas distribuições de probabilidade das classes, aumentando as taxas de erro na classificação.

O classificador da Distância Mínima comportou-se de forma adequada aos testes realizados, apresentando baixo custo de implementação. Sua estrutura de funcionamento permitiu treinar e classificar imagens com uma precisão razoável. Porém, sua desvantagem em alguns casos é não possuir informações suficientes sobre a forma da distribuição das classes.

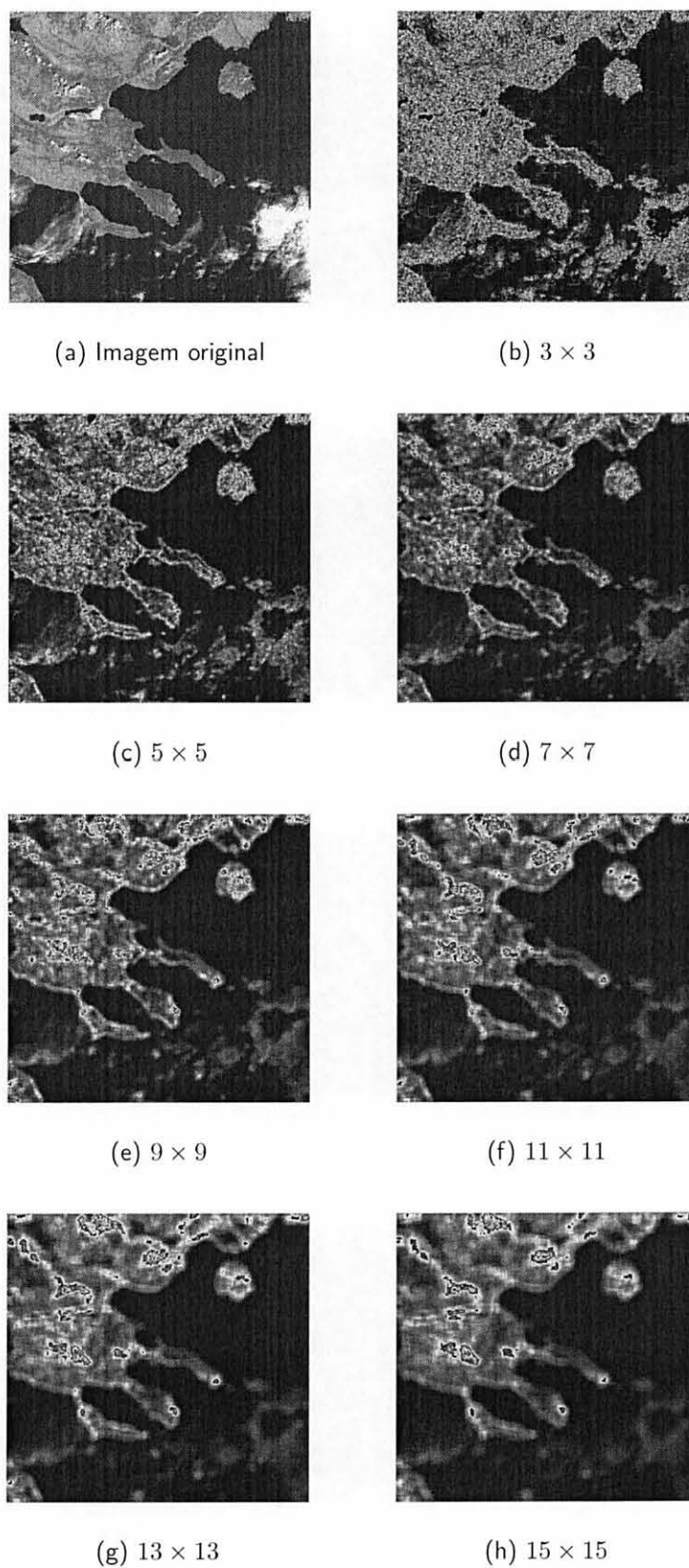


Figura 5.17: Efeito do tamanho (em pixels) da janela na classificação de texturas.

CAPÍTULO 6

CONCLUSÕES

Este trabalho apresentou um estudo sobre a análise e classificação de imagens utilizando textura. Foram abordadas algumas técnicas relevantes para a realização de tal procedimento.

Métodos estatísticos e de redes neurais para a classificação supervisionada de imagens digitais baseados na característica de textura foram discutidos. Em ambos os casos, a textura foi descrita em termos de medidas estatísticas indicando a distribuição espacial das variações de intensidade ou cor de objetos na imagem. Matrizes de co-ocorrência foram utilizadas para estimar as medidas estatísticas, cujos cálculos são computacionalmente simples e rápidos.

Observou-se que o tamanho da janela utilizado nos algoritmos para analisar a textura influenciou o resultado do processo de classificação. A redução de dimensionalidade, ou seja, a diminuição do número de características selecionadas da imagem contribuíram significativamente para um melhor desempenho do classificador em termos de memória utilizada e tempo computacional.

A extração de características de textura e posterior classificação através de janelas de pixels foi estatisticamente adequada. O tamanho da janela foi crucial no processo de classificação, sendo altamente sensível para expressar as propriedades de cada classe em consideração, devendo ser utilizado com cautela.

Há um tamanho ótimo de janela para cada uma das características. Qualquer alteração no tamanho da janela que não seja o ótimo torna o processo computacionalmente mais custoso sem prover qualquer melhoria na precisão, podendo também levar a interpretações incorretas sobre a textura. No tamanho ótimo de janela, a maioria das características fornece precisão de classificação estatisticamente similar.

A utilização da média de características de textura produz bons resultados, pois pode identificar as variações espaciais em diversas direções na textura.

A distância $d = 1$ entre pixels foi utilizada para o cálculo da matriz de co-ocorrência visto

que, para distâncias maiores, as características de textura são altamente correlacionadas.

Experimentos demonstraram que a abordagem é bastante genérica e robusta, podendo ser aplicada a uma grande variedade de imagens.

A partir da pesquisa realizada, novas propostas surgem para trabalhos futuros. As técnicas implementadas podem ser aplicadas a bases de imagens, permitindo a recuperação de imagens contendo padrões de textura específicos. Novas medidas estatísticas devem ser pesquisadas para caracterizar mais eficientemente a informação de textura, refinando e aumentando a precisão da classificação. Outra proposta é a utilização de conjuntos nebulosos (*fuzzy*) no processo de classificação, já que nem sempre é possível se definir fronteiras rígidas entre as diferentes classes.

REFERÊNCIAS BIBLIOGRÁFICAS

- [1] AKSOY, S. Textural features for image database retrieval. Master's thesis, University of Washington, Seattle, WA, 1998.
- [2] 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).
- [3] 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.
- [4] BARALDI, A. E PARMIGIANNI, F. An investigation of the textural characteristics associated with gray level co-occurrence matrix statistical parameters. *IEEE Transactions on Geoscience and Remote Sensing* 33, 2 (1995), 293–302.
- [5] BITTENCOURT, H. R. Reconhecimento estatístico de padrões: O caso da discriminação logística aplicada à classificação de imagens digitais obtidas por sensores remotos. In *I Congresso Brasileiro de Computação* (Porto Alegre, RS, 2001).
- [6] BRODATZ, P. Brodatz textures, 2002. www.ux.his.no/~tranden/brodatz.html.
- [7] CHAVES NETO, A. *Apostila de Introdução ao Reconhecimento de Padrões*. Departamento de Estatística - UFPR, Curitiba, Brasil, 2000.
- [8] CHEN, Y., NIXON, M. E THOMAS, D. Statistical geometric features for texture classification. *Pattern Recognition* 28 (1995), 537–552.
- [9] CHEN, Y. Q. *Novel techniques for image texture classification*. PhD thesis, University of Southampton, Department of Electronics and Computer Science, 1995.

- [10] DE CAMPOS, T. E. Técnicas de seleção de características com aplicações em reconhecimento de faces. Master's thesis, IME - Instituto de Matemática e Estatística da Universidade de São Paulo, Maio 2001.
- [11] DUBOIS, D., PRADE, H. E YAGER, R. R. *Fuzzy Information Engineering*. Wiley Computer Publishing, USA, 1997.
- [12] FERRO, C. J. S. Scale and texture in digital image classification. Master's thesis, West Virginia University, Morgantown, WV, 1998.
- [13] GONZALEZ, R. C. E WOODS, R. E. *Processamento de Imagens Digitais, tradução do original Digital Image Processing*. Edgard Blucher, São Paulo, 2000.
- [14] 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.
- [15] HEBB, D. O. *The Organization of Behavior*. Wiley Computer Publishing, New York, 1949.
- [16] IIVARINEN, J. E VISA, A. An adaptive texture and shape based defect classification. In *Proc. International Conf. on Pattern Recognition* (1998), pp. 117–123. <http://citeseer.nj.nec.com/iivarinen98adaptive.html>.
- [17] JAIN, A. E ZONGKER, D. Feature selection: Evaluation, application, and small sample performance. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 19, 2 (Fevereiro 1997), 153–158.
- [18] 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.
- [19] JULESZ, B. Experiments in the visual perception of texture. *Scientific American* 4, 232 (Abril 1975), 34–43.
- [20] KARTALOPOULOS, S. V. *Understanding Neural Networks and Fuzzy Logic: Basic Concepts and Applications*. Wiley-IEEE Press, 1996.

- [21] KARU, K., JAIN, A. E. & BOLLE, R. Is there any texture in the image? *Pattern Recognition* 29 (1996), 1437–1446.
- [22] KLIR, G. J. & FOLGER, T. A. *Fuzzy Sets, Uncertainty and Information*. Prentice-Hall, Inc., New Jersey, 1988.
- [23] KUNDU, P. & CHAUDHURI, B. B. Fuzzy geometric feature-based texture classification. *Pattern Recognition Letters* 14 (1993), 825–832.
- [24] LIEW, S. C., LIM, H., KWONG, L. K. & TAY, G. K. Texture analysis of SAR images. In *International Geoscience and Remote Sensing Symposium* (1995), pp. 1412–1414.
- [25] MARCEAU, D. J., HOWARTH, P. J., DUBOIS, J.-M. M. & GRATTON, D. J. Evaluation of the grey-level co-occurrence matrix method for land-cover classification using SPOT imagery. *IEEE Transactions on Geoscience and Remote Sensing* (March 1990), 513–519.
- [26] MATERKA, A. & STRZELECKI, M. Texture analysis methods - a review, 1998.
- [27] MCCULLOCH, W. S. & PITTS, W. A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics* 5 (1943), 115–133.
- [28] MINSKY, M. L. *Theory of Neural Analog Reinforcement Systems and Application to the Brain Model Problem*. PhD thesis, Princeton University, New Jersey - USA, 1954.
- [29] MINSKY, M. L. & PAPER, S. A. *Perceptrons : An Introduction to Computational Geometry*. MIT Press, Cambridge, MA, 1969.
- [30] MUDIGONDA, N. R., RANGAYAN, R. M. & DESAUTELS, J. E. L. Gradient and texture analysis for the classification of mammographic masses. *IEEE Transactions on Medical Imaging* 19, 10 (October 2000), 1032–1043.
- [31] MUHAMAD, A. K. & DERAVI, F. Neural network texture classifiers using direct input co-occurrence matrices. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (1993), 117–120.

- [32] MURINO, V., OTTONELLO, C. E PAGNAN, S. Noisy texture classification: a higher order statistics approach. *Pattern Recognition* 34 (1998), 383–393.
- [33] NADLER, M. E SMITH, E. P. *Pattern Recognition Engineering*. John Wiley - Sons, 1993.
- [34] NASCIMENTO, J. P. R., MADEIRA, H. M. F. E PEDRINI, H. Classificação de imagens utilizando descritores estatísticos de textura. In *XI Simpósio Brasileiro de Sensoriamento Remoto* (Belo Horizonte-MG, Abril 2003).
- [35] PICAN, N., TRUCCO, E., ROSS, M., LANE, D., PETILLOT, Y. E RUIZ, I. T. Texture analysis for seabed classification: Co-occurrence matrices vs self-organizing maps. *IEEE Transactions on Computers* (Junho 1998), 424–428.
- [36] PIERCE, L., VHIE, H., DOBSON, M. C. E ULABY, F. Texture features for classification with ERS/JERS composites. *IEEE Transactions on Computers* (1998), 348–350.
- [37] RICHARDS, J. A. *Remote Sensing Digital Image Analysis: An Introduction*. Springer-Verlag, Berlin, 1986.
- [38] ROSENBLATT, F. The perceptron: A perceiving and recognizing automaton (project PARA). Technical report 85-460-1, Cornell Aeronautical Laboratory, 1957.
- [39] ROSENFELD, A. E AVINASH, C. *Digital Picture Processing*. Academic Press, San Diego, 1982.
- [40] RUMELHART, D., HINTON, G. E. E WILLIAMS, R. J. Learning internal representations by error propagation. In *Parallel Distributed Processing: Explorations in the Microstructure of Cognition* (1986).
- [41] SCHOWENGERDT, R. A. *Techniques for Image Processing an Classification in Remote Sensing*. Academic Press, Arizona, 2000.
- [42] SHABAN, M. A. E DIKSHIT, O. Textural classification of high resolution digital satellite imagery. *IEEE Transactions on Computers* (1998), 2590–2592.

- [43] SHARMA, M. Performance evaluation of image segmentation and texture extraction methods in scene analysis. Master's thesis, University of Exeter, 2001.
- [44] SHARMA, M., MARKOU, M. E SINGH, S. Evaluation of texture methods for image analysis. *Pattern Recognition Letters* (2000), 1–8.
- [45] SHAW, I. S. E SIMÕES, M. G. *Controle e Modelagem Fuzzy*. Edgard Blücher Ltda., São Paulo, 1999.
- [46] SNNS. *Stuttgart Neural Network Simulator*. University of Stuttgart, 2001. <http://www-ra.informatik.uni-tuebingen.de/SNNS/>.
- [47] SOLBERG, A. H. S. E JAIN, A. K. A study of the invariance properties of textural features in SAR images. *IEEE Transactions on Computers* (Julho 1995), 670–672.
- [48] SONKA, M., HLAVAC, V. E BOYLE, R. *Image Processing, Analysis and Machine Vision*. Brooks, USA, 1998.
- [49] SPRING. *Sistema de Processamento de Informações Georeferenciadas*. Instituto Nacional de Pesquisas Espaciais (INPE), 2001. http://www.dpi.inpe.br/spring/usuario/c_clapix.htm.
- [50] TAFNER, M. A., DE XEREZ, M. E FILHO, I. W. R. *Redes neurais artificiais: introdução e princípios de neurocomputação*. Editora Eko, Blumenau, 1995.
- [51] THEODORIDIS, S. *Pattern Recognition*, 1 ed. Academic Press, USA, 1999.
- [52] TOBIAS, O. J., SEARA, R., SOARES, F. A. P. E BERMUDEZ, J. C. M. Automatic visual inspection using the co-occurrence approach. In *IEEE Midwest Symposium on Circuits and Systems* (Florianópolis, SC, Agosto 1996), vol. 38, pp. 154–157.
- [53] 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.

- [54] WALKER, R. F., JACKWAY, P. T. E LONGSTAFF, I. Recent developments in the use of the co-occurrence matrix for texture recognition. *IEEE Transactions on Computers* (1977), 63–65.
- [55] WESZKA, J. E ROSENFELD, A. An application of texture analysis to materials inspection. *Pattern Recognition* (1976), 195–199.
- [56] YAMAGATA, Y. E YASUOKA, Y. Classification of wetland vegetation by texture analysis method using ERS-1 and JERS-1 images. In *International Geoscience and Remote Sensing Symposium* (Julho 1993), pp. 1614–1616.

APÊNDICE A

FORMATO DAS IMAGENS

O formato PGM (*Portable Gray Map*) foi utilizado nas imagens, sendo constituído pelos seguintes campos:

- Identificador do tipo de formato (designado por *magic number*), de acordo com seu tipo: P2 (ASCII) e P5 (Binário).
- Espaço em branco
- Largura da imagem dada, em pixels e em notação decimal, formatada em caracteres ASCII
- Espaço em branco
- Altura da imagem dada, em pixels e em notação decimal, formatada em caracteres ASCII
- Espaço em branco
- Nível de cinza máximo da imagem, em notação decimal dos níveis de cinza
- Espaço em branco
- Valores de cinza dos pixels da imagem, em número igual à altura da imagem vezes a sua largura.

A ordenação dos valores dos pixels corresponde à varredura das imagens linha a linha, de cima para baixo, e da esquerda para a direita em cada uma das linhas.

Quando o conteúdo da imagem for representado em ASCII, os valores correspondentes aos pixels serão apresentados em notação decimal e separados por espaços em branco, marcas de tabulação ou marcas de fim de linha. As variantes binárias do tipo PGM armazenam os valores correspondentes aos pixels das imagens em caracteres (*bytes*) contíguos, sem qualquer

```

P2 19 7 15
15 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 15
 0 2 2 2 2 0 0 6 6 6 6 0 0 9 9 9 9 9 0
 0 2 0 0 2 0 0 6 0 0 0 0 0 9 0 9 0 9 0
 0 2 2 2 2 0 0 6 0 6 6 0 0 9 0 9 0 9 0
 0 2 0 0 0 0 0 6 0 0 6 0 0 9 0 0 0 9 0
 0 2 0 0 0 0 0 6 6 6 6 0 0 9 0 0 0 9 0
15 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 15

```

Figura A.1: Exemplo de arquivo PGM.

separador entre valores consecutivos. O tipo PGM faz corresponder a cada pixel um caracter. Aplicam-se ainda as seguintes regras:

- É permitida a inserção de comentários em qualquer parte do arquivo. O início de um comentário é assinalado por um caracter "#", e todo o texto desde este caracter até o fim da respectiva linha é interpretado como texto do comentário.
- O comprimento máximo de cada linha está limitado a 70 caracteres.

A figura A.1 mostra um exemplo de arquivo PGM e a figura A.2 a imagem em níveis de cinza correspondente ao arquivo no formato PGM, com largura de 19 pixels e altura de 7 pixels.



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