

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
CENTRO ESTADUAL DE PESQUISA
EM SENSORIAMENTO REMOTO
E METEOROLOGIA - CEPSRM
CURSO DE PÓS-GRADUAÇÃO EM SENSORIAMENTO REMOTO

CLASSIFICAÇÃO DE IMAGENS DE SENSORIAMENTO REMOTO
BASEADA EM TEXTURA POR REDES NEURAS

por

Adriano Beluco

Porto Alegre, RS

Julho de 2002

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
CENTRO ESTADUAL DE PESQUISA EM SENSORIAMENTO REMOTO
E METEOROLOGIA - CEPSRM
CURSO DE PÓS-GRADUAÇÃO EM SENSORIAMENTO REMOTO

**CLASSIFICAÇÃO DE IMAGENS DE SENSORIAMENTO REMOTO
BASEADA EM TEXTURA POR REDES NEURAIAS**

ADRIANO BELUCO

LICENCIADO EM MATEMÁTICA

Orientador: Prof. Dr. Paulo Martins Engel

Dissertação apresentada como requisito parcial
para a obtenção do grau de Mestre em Senso-
riamento Remoto na área de concentração
classificação digital de imagens.

Porto Alegre, RS

30 de Julho de 2002

CLASSIFICAÇÃO DE IMAGENS DE SENSORIAMENTO REMOTO BASEADA EM TEXTURA POR REDES NEURAIAS

Autor: Adriano Beluco

Orientador: Prof. Dr. Paulo Martins Engel

RESUMO

O objetivo principal deste trabalho é propor uma metodologia de classificação de imagens de sensoriamento remoto que integre a importância de atributos de textura na seleção de feições, através da utilização de frequências espaciais de cada classe textural e sua direção, com a eficiência das redes neurais artificiais para classificá-las. O processo é composto por uma etapa de filtragem baseada nos filtros de Gabor, seguida de uma fase de classificação através de uma rede neural *Multi-Layer Perceptron* com algoritmo *BackPropagation*. A partir da transformada de Fourier são estimados os parâmetros a serem utilizados na constituição dos filtros de Gabor, adequados às frequências espaciais associadas a cada classe presente na imagem a ser classificada. Desta forma, cada filtro gera uma imagem filtrada. O conjunto de filtros determina um conjunto de imagens filtradas (canais texturais). A classificação pixel a pixel é realizada pela rede neural onde cada pixel é definido por um vetor de dimensionalidade igual ao número de filtros do conjunto.

O processo de classificação através da rede neural *Multi-Layer Perceptron* foi realizado pelo método de classificação supervisionada.

A metodologia de classificação de imagens de sensoriamento remoto proposta neste trabalho foi testada em imagens sintética e real de dimensões 256 x 256 pixels. A análise dos resultados obtidos é apresentada sob a forma de uma Matriz de Erros, juntamente com a discussão dos mesmos.

TEXTURE BASED REMOTE SENSING IMAGE CLASSIFICATION USING ARTIFICIAL NEURAL NETWORKS

Author: Adriano Beluco

Advisor: Prof. Dr. Paulo Martins Engel

ABSTRACT

This work aims at establishing a method for classification of remote sensing images integrating the importance of texture - via use of spacial frequencies of each textural class and its direction - with the efficiency of artificial neural networks in classifying them. The classification process consists of applying Gabor filters followed by neural network classification. This classification is based on a Multi Layer Perceptron neural network with Back propagation algorithm. The parameters of the Gabor filters are estimated by Fourier Transform, with spacial frequencies present in the classified image. Thus, each filter generates a corresponding image, and a set of filters consequently generates a set of filtered images or textural bands. The pixel by pixel classification is realized by the neural network, each pixel defined by a dimensional vector whose components are the number of filters in the set.

The classification process by Multi Layer Perceptron is realized with supervised classification.

The image classification method proposed is verified in synthetic and real images of 256 by 256 pixels. The analysis of results obtained is presented by an error matrix, followed by discussions.

*"Nunca subestime o seu poder de mudar a si mesmo,
e nunca superestime o seu poder de mudar os outros."*

S. Brown

*"Se um dia tudo lhe parecer perdido, lembre-se de que você nasceu sem nada,
e que tudo que conseguiu foi através de esforços
e os esforços nunca se perdem, somente dignificam as pessoas."*

Charles Chaplin

AGRADECIMENTOS

Gostaria de expressar todo o meu agradecimento, após este longo trabalho, a todos que contribuíram de alguma forma para a sua realização:

Ao meu irmão Alexandre, pelo apoio técnico e moral durante algumas fases do trabalho, além das "injeções" constantes de incentivo.

À minha mãe, Sirlei, pelo apoio e cobrança incondicionais durante todo o trabalho.

Ao meu orientador, professor Paulo M. Engel, pelo aconselhamento, paciência e disponibilidade indispensáveis para a constituição deste trabalho.

Ao professor Jorge Ducatti, como retribuição à confiança demonstrada quando do ingresso no curso de mestrado.

Ao professor Franz Semmelman, pelas oportunidades proporcionadas para o término da dissertação.

Aos colegas Neide Angelo e Viviane Todt, pela ajuda irrestrita às indicações bibliográficas de suma importância.

Aos colegas do curso de pós-graduação, pelas experiências vividas durante as disciplinas.

Aos funcionários do CEPSRM, Luciane, Magdalena, Mônica e Adriano, pela convivência e companheirismo durante todo o período.

À CAPES, pelo apoio financeiro indispensável aos estudantes de pós-graduação.

A todos os meus amigos, que sempre acreditaram na conclusão deste trabalho.

SUMÁRIO

Lista de Figuras	viii
Lista de Tabelas	xiv
Lista de Apêndices	xvi
Lista de Abreviaturas.....	xvii
Lista de Símbolos	xviii
Lista de Equações.....	xx
1. Introdução	1
1. 1. Objetivos	7
1. 2. Motivação.....	8
2. Revisão Bibliográfica.....	10
2. 1. Introdução	10
2. 2. Técnicas de classificação de imagens	11
2. 2. 1. Classificação supervisionada	12
2. 2. 1. 1. Método do paralelepípedo.....	13
2. 2. 1. 2. Método da distância mínima	14
2. 2. 1. 3. Método da Máxima Verossimilhança Gaussiana.....	15
2. 2. 2. Classificação não supervisionada	16
2. 3. Filtragem no domínio espacial	17
2. 4. Filtragem no domínio das freqüências	18
2. 5. Técnicas de filtragem em multicanais texturais	18
2. 6. Técnicas de redes neurais	21
2. 6. 1. Conceitos básicos	23
2. 6. 2. Histórico e paradigmas.....	31
2. 6. 2. 1. Perceptrons.....	32
2. 6. 2. 2. Adaline	33
2. 6. 2. 3. Hopfield.....	34
2. 6. 2. 4. Backpropagation	35
2. 6. 3. Diferenças entre redes neurais artificiais e biológicas.....	36
3. Material e Métodos	40
3. 1. Considerações gerais	40
3. 2. Textura	42
3. 3. Extração de feições texturais.....	44
3. 4. Filtragem através de Filtros de Gabor.....	48
3. 4. 1. Seleção de parâmetros dos Filtros de Gabor	52
3. 4. 1. 1. Determinação da freqüência espacial	52
3. 4. 1. 2. Dimensão e extensão espacial	53
3.5. Classificação de imagens obtidas por sensoriamento remoto com redes neurais	56
3. 5. 1. Rede neural multinível com <i>Backpropagation</i>	58
3. 5. 2. Algoritmo de treinamento <i>Backpropagation</i>	60
4. Experimentos	63
4. 1. Considerações gerais	63
4. 2. Experimentos envolvendo imagens sintéticas	64
4. 2. 1. Imagens sintéticas a partir de funções matemáticas	65
4. 2. 1. 1. Experimento A.....	65
4. 2. 1. 2. Experimento B.....	67
4. 2. 2. Imagens sintética: mosaico a partir de texturas de Brodatz	68
4. 2. 2. 1. Experimento A.....	71
4. 2. 2. 1. 1. Experimento A1.....	73
4. 2. 2. 1. 2. Experimento A2.....	75
4. 2. 2. 2. Experimento B.....	76
4. 2. 2. 2. 1. Experimento B1.....	78
4. 2. 2. 2. 2. Experimento B2.....	79
4. 2. 2. 3. Experimento C	81
4. 2. 2. 3. 1. Experimento C1	83
4. 2. 2. 3. 2. Experimento C2	84

4. 2. 2. 4. Experimento D	86
4. 2. 2. 4. 1. Experimento D1	88
4. 2. 2. 4. 2. Experimento D2	89
4. 3. Experimentos envolvendo imagens reais	91
4. 3. 1. Experimento A.....	98
4. 3. 1. 1. Experimento A1.....	100
4. 3. 1. 2. Experimento A2.....	101
4. 3. 2. Experimento B.....	103
4. 3. 2. 1. Experimento B1.....	105
4. 3. 2. 2. Experimento B2.....	107
4. 4. Resultados	108
4. 5. Considerações finais	109
5. Conclusões e Sugestões.....	111
APÊNDICES.....	114
BIBLIOGRAFIA	125

LISTA DE FIGURAS

Figura 1. 1	Satélite sondando a superfície da Terra.	1
Figura 1. 2	Convenção utilizada para o par de eixos (x,y).	2
Figura 1. 3	As variações de CD ao longo de uma linha de uma imagem (a) podem ser expressas como várias componentes senoidais (b), com diferentes amplitudes e freqüências (Crósta, 1992).	4
Figura 1. 4	Esquema de classificação com uma rede neural artificial. A rede mostrada corresponde a uma arquitetura contendo 11 unidades na camada de entrada, 20 unidades escondidas em 2 camadas e seis unidades na camada de saída. (Adaptado de Foody e Arora, 1997).	7
Figura 2. 1	Diagrama de espaço de atributos das classes A, B e C, mostrando a localização dos pixels pertencentes a um conjunto de treinamento e a pixels desconhecidos. Os limites de decisão de um classificador de paralelepípedo são mostrados ao redor da classe (Adaptado de Crósta, 1992).	14
Figura 2. 2	Espaço de atributos das classes A, B e C mostrando a posição das médias de cada classe e de dois pixels desconhecidos não classificados pelo método do paralelepípedo (Adaptado de Crósta, 1992).	15
Figura 2. 3	Espaço de atributos para as classes A, B e C mostrando na forma de contornos a densidade de pixels nos diversos conjuntos de treinamento (Adaptado de Crósta, 1992).	16
Figura 2. 4	Espaço de atributos mostrando a distribuição de densidade de todos os pixels em uma imagem, na forma de curvas de contorno. As maiores concentrações foram denominadas de A a H (Adaptado de Crósta, 1992).	17
Figura 2. 5	Vizinhança 3x3 ao redor de um ponto de coordenadas (x,y) em uma imagem (Marques e Vieira, 1999).	18
Figura 2. 6	Esquema simplificado de um neurônio.	22
Figura 2. 7	Esquema simplificado de uma conexão sináptica.	23
Figura 2. 8	Aprendizado em redes neurais supervisionadas.	26
Figura 2. 9	Aprendizado em redes neurais não supervisionadas.	27

Figura 2. 10	Modelo simplificado de um elemento processador.....	28
Figura 2. 11	Funções de transferência.....	28
Figura 2. 12	Rede neural de cinco camadas.....	29
Figura 2. 13	Rede neural auto-associativa.....	30
Figura 2. 14	Rede neural hetero-associativa.	30
Figura 2. 15	Modelo matemático simplificado de um neurônio de McCulloch.....	32
Figura 2. 16	Perceptrons de Roseblatt.....	33
Figura 2. 17	Madaline de Widrow & Hoff.....	34
Figura 2. 18	Perceptron multi-camadas.	35
Figura 3. 1	Diagrama de blocos do processo de filtragem e classificação supervisionada de imagens proposto no presente trabalho.....	40
Figura 3. 2	Respostas espectrais típicas da água (1), vegetação (2) e solo (3). Também estão indicadas as posições das bandas espectrais para alguns sensores remotos (Richards, 1986).....	43
Figura 3. 3	Componente real do filtro de Gabor (a) e componente imaginária do filtro de Gabor (b), ambas para valores de $\sigma = 1s$ e $k = 0.5$ ciclos/pixel. Adaptado de Angelo (2000).....	46
Figura 3. 4	Filtro de Gabor bidimensional apresentado em valores absolutos.	50
Figura 3. 5	Componente real do filtro de Gabor bidimensional no domínio espacial (a); resposta de frequência do filtro definido em a (b). Adaptado de Angelo (2000).....	51
Figura 3. 6	Topologia de uma RNA Multinível com Backpropagation (Lieberman, 1997).	59
Figura 4. 1	Imagem sintética de 256 x 256 pixels composta por funções senoidais bidimensionais.....	65
Figura 4. 2.	Imagens filtradas pelos filtros de Gabor para o experimento A , referente à imagem sintética a partir de funções matemáticas. A ordem das filtragens é expressa de (a) até (d).	66

Figura 4. 3	Imagem classificada referente ao experimento A.....	67
Figura 4. 4.	Imagens filtradas pelos filtros de Gabor para o experimento B , referente à imagem sintética a partir de funções matemáticas. A ordem das filtrações é expressa de (a) até (d).	67
Figura 4. 5	Imagem classificada referente ao experimento B.....	68
Figura 4. 6.	Imagem mosaico sintética de 256 x 256 pixels composta por microtexturas de Brodatz.	69
Figura 4. 7.	Imagem mosaico com localização das amostras representativas de cada classe para a seleção de feições.....	70
Figura 4. 8.	(a) Amostra da classe 1 na imagem mosaico; (b) Espectro de Fourier da amostra; (c) Freqüências da amostra que apresentam o maior nível de energia.	70
Figura 4. 9.	(a) Amostra da classe 2 na imagem mosaico; (b) Espectro de Fourier da amostra; (c) Freqüências da amostra que apresentam o maior nível de energia.	71
Figura 4. 10.	(a) Amostra da classe 3 na imagem mosaico; (b) Espectro de Fourier da amostra; (c) Freqüências da amostra que apresentam o maior nível de energia.	71
Figura 4. 11.	(a) Amostra da classe 4 na imagem mosaico; (b) Espectro de Fourier da amostra; (c) Freqüências da amostra que apresentam o maior nível de energia.	71
Figura 4. 12.	Imagens filtradas pelo filtro de Gabor para o experimento A, a partir da imagem mosaico.	73
Figura 4. 13.	Imagem classificada referente ao experimento A1, utilizando uma imagem sintética.	74
Figura 4. 14.	Imagem classificada referente ao experimento A2, utilizando uma imagem sintética.	75
Figura 4. 15.	Imagens filtradas pelo filtro de Gabor para o experimento B, a partir da imagem mosaico.	77
Figura 4. 16.	Imagem classificada referente ao experimento B1, utilizando uma imagem sintética.	78
Figura 4. 17.	Imagem classificada referente ao experimento B2, utilizando uma imagem sintética.	80

Figura 4. 18. Imagens filtradas pelo filtro de Gabor para o experimento C, a partir da imagem mosaico.	82
Figura 4. 19. Imagem classificada referente ao experimento C1, utilizando uma imagem sintética.	83
Figura 4. 20. Imagem classificada referente ao experimento C2, utilizando uma imagem sintética.	85
Figura 4. 21. Imagens filtradas pelo filtro de Gabor para o experimento D, a partir da imagem mosaico.	87
Figura 4. 22. Imagem classificada referente ao experimento D1, utilizando uma imagem sintética.	88
Figura 4. 23. Imagem classificada referente ao experimento D2, utilizando uma imagem sintética.	90
Figura 4. 24. Imagem digital de parte da cidade de Porto Alegre (RS) e seu entorno.	91
Figura 4. 25. Imagem real com localização das amostras representativas da classe água para a seleção de feições.	92
Figura 4. 26. (a) Amostra 1 da classe água na imagem real; (b) Espectro de Fourier da amostra; (c) Freqüências da amostra que apresentam o maior nível de energia.....	93
Figura 4. 27. (a) Amostra 2 da classe água na imagem real; (b) Espectro de Fourier da amostra; (c) Freqüências da amostra que apresentam o maior nível de energia.....	93
Figura 4. 28. (a) Amostra 3 da classe água na imagem real; (b) Espectro de Fourier da amostra; (c) Freqüências da amostra que apresentam o maior nível de energia.....	93
Figura 4. 29. (a) Amostra 4 da classe água na imagem real; (b) Espectro de Fourier da amostra; (c) Freqüências da amostra que apresentam o maior nível de energia.....	93
Figura 4. 30. (a) Amostra 5 da classe água na imagem real; (b) Espectro de Fourier da amostra; (c) Freqüências da amostra que apresentam o maior nível de energia.....	94
Figura 4. 31. Imagem real com localização das amostras representativas da classe urbana para a seleção de feições.	94
Figura 4. 32. (a) Amostra 1 da classe urbana na imagem real; (b) Espectro de Fourier da amostra; (c) Freqüências da amostra que apresentam o maior nível de energia.	95

Figura 4. 33. (a) Amostra 2 da classe urbana na imagem real; (b) Espectro de Fourier da amostra; (c) Freqüências da amostra que apresentam o maior nível de energia.	95
Figura 4. 34. (a) Amostra 3 da classe urbana na imagem real; (b) Espectro de Fourier da amostra; (c) Freqüências da amostra que apresentam o maior nível de energia.	95
Figura 4. 35. (a) Amostra 4 da classe urbana na imagem real; (b) Espectro de Fourier da amostra; (c) Freqüências da amostra que apresentam o maior nível de energia.	95
Figura 4. 36. (a) Amostra 5 da classe urbana na imagem real; (b) Espectro de Fourier da amostra; (c) Freqüências da amostra que apresentam o maior nível de energia.	96
Figura 4. 37. Imagem real com localização das amostras representativas da classe vegetação para a seleção de feições.	96
Figura 4. 38. (a) Amostra 1 da classe vegetação na imagem real; (b) Espectro de Fourier da amostra; (c) Freqüências da amostra que apresentam o maior nível de energia.	96
Figura 4. 39. (a) Amostra 2 da classe vegetação na imagem real; (b) Espectro de Fourier da amostra; (c) Freqüências da amostra que apresentam o maior nível de energia.	97
Figura 4. 40. (a) Amostra 3 da classe vegetação na imagem real; (b) Espectro de Fourier da amostra; (c) Freqüências da amostra que apresentam o maior nível de energia.	97
Figura 4. 41. (a) Amostra 4 da classe vegetação na imagem real; (b) Espectro de Fourier da amostra; (c) Freqüências da amostra que apresentam o maior nível de energia.	97
Figura 4. 42. (a) Amostra 5 da classe vegetação na imagem real; (b) Espectro de Fourier da amostra; (c) Freqüências da amostra que apresentam o maior nível de energia.	97
Figura 4. 43. Imagens filtradas pelo filtro de Gabor para o experimento A, a partir da imagem real.	99
Figura 4. 44. Imagem classificada referente ao experimento A1, utilizando uma imagem real.	100
Figura 4. 45. Imagem classificada referente ao experimento A2, utilizando uma imagem real.	102
Figura 4. 46. Imagens filtradas pelo filtro de Gabor para o experimento B, a partir da imagem real.	105

Figura 4. 47.	Imagem classificada referente ao experimento B1, utilizando uma imagem real.	106
Figura 4. 48.	Imagem classificada referente ao experimento B2, utilizando uma imagem real.	107
Figura A1	Função unidimensional contínua simples (a) e seu respectivo espectro de Fourier (b) (Marques e Vieira, 1999).	116
Figura A2	Função bidimensional contínua (a), seu espectro de Fourier (b) e o espectro representado como uma função de intensidade (c) (Marques e Vieira, 1999).	117
Figura A3	Imagem original (a), transformada de Fourier de a (b), imagem rotacionada (c) e transformada de Fourier de c (d) (Marques e Vieira, 1999).	122
Figura A4	Diversas microtexturas de Brodatz: Fabric 1 (a), Fabric 2 (b), Food (c), Gravel (d), Woolencloth (e), Plastic Bubs (f), Raffia (g) e Roughwall (h) (Brodatz, 1966).	124

LISTA DE TABELAS

Tabela 4. 1. Parâmetros para a constituição dos filtros de Gabor necessários para a realização do experimento A.	72
Tabela 4. 2. Dados da matriz de erro para a avaliação da classificação da imagem sintética composta por quatro texturas naturais de Brodatz, referente ao experimento A1.	74
Tabela 4. 3. Dados da matriz de erro para a avaliação da classificação da imagem sintética composta por quatro texturas naturais de Brodatz, referente ao experimento A2.	76
Tabela 4. 4. Parâmetros para a constituição dos filtros de Gabor necessários para a realização do experimento B.	76
Tabela 4. 5. Dados da matriz de erro para a avaliação da classificação da imagem sintética composta por quatro texturas naturais de Brodatz, referente ao experimento B1.	79
Tabela 4. 6. Dados da matriz de erro para a avaliação da classificação da imagem sintética composta por quatro texturas naturais de Brodatz, referente ao experimento B2.	80
Tabela 4. 7. Parâmetros para a constituição dos filtros de Gabor necessários para a realização do experimento C.	81
Tabela 4. 8. Dados da matriz de erro para a avaliação da classificação da imagem sintética composta por quatro texturas naturais de Brodatz, referente ao experimento C1.	84
Tabela 4. 9. Dados da matriz de erro para a avaliação da classificação da imagem sintética composta por quatro texturas naturais de Brodatz, referente ao experimento C2.	85
Tabela 4. 10. Parâmetros para a constituição dos filtros de Gabor necessários para a realização do experimento D.	86
Tabela 4. 11. Dados da matriz de erro para a avaliação da classificação da imagem sintética composta por quatro texturas naturais de Brodatz, referente ao experimento D1.	89
Tabela 4. 12. Dados da matriz de erro para a avaliação da classificação da imagem sintética composta por quatro texturas naturais de Brodatz, referente ao experimento D2.	90
Tabela 4. 13. Parâmetros para a constituição dos filtros de Gabor necessários para a realização do experimento A, utilizando a imagem real.	98

Tabela 4. 14. Dados da matriz de erro para a avaliação da classificação da imagem real de parte da cidade de Porto Alegre (RS) e seu entorno, referente ao experimento A1.	101
Tabela 4. 15. Dados da matriz de erro para a avaliação da classificação da imagem real de parte da cidade de Porto Alegre (RS) e seu entorno, referente ao experimento A2.	102
Tabela 4. 16. Parâmetros para a constituição dos filtros de Gabor necessários para a realização do experimento B, utilizando a imagem real.	104
Tabela 4. 17. Dados da matriz de erro para a avaliação da classificação da imagem real de parte da cidade de Porto Alegre (RS) e seu entorno, referente ao experimento B1.	106
Tabela 4. 18. Dados da matriz de erro para a avaliação da classificação da imagem real de parte da cidade de Porto Alegre (RS) e seu entorno, referente ao experimento B2.	108
Tabela 4. 19. Comparativo dos resultados dos experimentos realizados com imagens sintéticas e reais.	108

LISTA DE APÊNDICES

Apêndice 1	Transformada de Fourier.....	115
Apêndice 2	Micro texturas de Brodatz	124

LISTA DE ABREVIATURAS

ART	Adaptive Resonance Theory
BPN	BackPropagation Network
BSB	Brain-State-in-a-Box
CD	Contador Digital
EMQ	Erro Médio Quadrado
FFT	Fast Fourier Transform
IEEE	Institute of Electrical and Electronics Engineers
LMS	Least Mean Square Error
LTM	Memórias de Longo Termo
MATLAB	Matrix Laboratory
MLP	Multi-Layer Perceptron
PDP	Processamento Distribuído e Paralelo
PIXEL	Picture Element
RNA	Rede Neural Artificial
SOM	Self Organizing Map / Mapas Auto-organizáveis
TEXEL	Texture Element

LISTA DE SÍMBOLOS

$ \cdot $	Módulo ou valor absoluto
$I(i, j)$	Posição na sub-imagem I
Γ	Estatística Gamma de Hubert
$f(x, y)$	Função bidimensional em x, y
$F(u, v)$	Transformada de Fourier de uma função $f(x, y)$
σ	Extensão espacial do filtro de Gabor
σ_x	Extensão espacial do filtro de Gabor ao longo do eixo x
σ_y	Extensão espacial do filtro de Gabor ao longo do eixo y
Σ	Operador de somatório
Δ	Variação
k_x	Freqüência espacial do filtro de Gabor ao longo do eixo x
k_y	Freqüência espacial do filtro de Gabor ao longo do eixo y
π	Número pi
i	Unidade de número imaginário
$*$	Operador de convolução
\dim	Dimensão do filtro de Gabor
Dim	Dimensão única do filtro de Gabor
\exp	Função exponencial
\int_a^b	Integral definida no intervalo entre a e b
$\text{Re}(\cdot)$	Componente real de um número imaginário
$\text{Im}(\cdot)$	Componente imaginária de um número imaginário
$\phi(u, v)$	Ângulo de fase da transformada de Fourier
$\mathfrak{F}(\cdot)$	Operador transformada de Fourier
$\mathfrak{F}^{-1}(\cdot)$	Operador transformada inversa de Fourier
$\tanh(x)$	Função tangente hiperbólica em x
$\tanh'(x)$	Derivada da função tangente hiperbólica em x
Y_j	Saída do neurônio j
Y_{pk}	Saída desejada para o padrão p
O_{pk}	Saída obtida para o padrão p
X_p	Vetor de entrada do padrão p

X_{pi}	Vetor de entrada do padrão p para o neurônio i
W_{ji}	Peso entre os neurônios j e i
W_{ji}^h	Peso entre os neurônios j e i da camada oculta
t	Iteração atual
net_{pj}^h	Valor interno do neurônio j da camada oculta para o padrão p
net_{pk}^o	Valor interno do neurônio k da camada de saída para o padrão p
θ_j^h	Bias para o neurônio j da camada oculta
θ_k^o	Bias para o neurônio k da camada de saída
i_{pj}	Ativação do neurônio j da camada oculta para o padrão p
O_{pk}	Ativação do neurônio k da camada de saída para o padrão p
φ_{pk}^o	Termo de erro para a camada de saída para o padrão p
φ_{pj}^h	Termo de erro para a camada oculta para o padrão p
η	Coefficiente de aprendizado para o algoritmo de backpropagation

LISTA DE EQUAÇÕES

Equação 2. 1	Função de processamento de imagens no domínio espacial	17
Equação 2. 2	Processo de convolução no domínio espacial.....	18
Equação 2. 3	Processo de convolução no domínio de freqüências	18
Equação 2. 4	Função de transferência considerando a função de limiar	28
Equação 3. 1	Limite teórico inferior de incerteza conjunta para o domínio temporal e das freqüências.....	45
Equação 3. 2	Filtro de Gabor para o caso unidimensional	45
Equação 3. 3	Filtro de Gabor para o caso bidimensional	46
Equação 3. 4	Limite teórico inferior de incerteza conjunta para o domínio espacial e das freqüências.....	46
Equação 3. 5	Fator correspondente ao eixo das abscissas para o cálculo da área no domínio espacial	47
Equação 3. 6	Fator correspondente ao eixo das abscissas para o cálculo da área no domínio das freqüências.....	47
Equação 3. 7	Fator correspondente ao eixo das ordenadas para o cálculo da área no domínio espacial.....	47
Equação 3. 8	Fator correspondente ao eixo das ordenadas para o cálculo da área no domínio das freqüências	47
Equação 3. 9	Área ocupada pelo filtro no domínio espacial.....	47
Equação 3. 10	Área ocupada pelo filtro no domínio das freqüências	47
Equação 3. 11	Incerteza conjunta para os domínios espacial e das freqüências no caso do filtro de Gabor.....	48
Equação 3. 12	Função de Gabor bidimensional	49
Equação 3. 13	Componente real do filtro bidimensional de Gabor	49

Equação 3. 14	Componente imaginária do filtro bidimensional de Gabor	49
Equação 3. 15	Transformada de Fourier para o filtro bidimensional de Gabor	50
Equação 3. 16	Transformada de Fourier para a componente real do filtro de Gabor bidimensional	50
Equação 3. 17	Área no domínio espacial.....	51
Equação 3. 18	Área no domínio das freqüências	51
Equação 3. 19	Energia por pixel da imagem filtrada	52
Equação 3. 20	Dimensão do filtro de Gabor ao longo dos eixos coordenados.....	54
Equação 3. 21	Valor único da dimensão do filtro de Gabor segundo o máximo do recíproco das freqüências espaciais.....	55
Equação 3. 22	Valor da extensão espacial do filtro de Gabor em função de sua dimensão.....	55
Equação 3. 23	Valores internos dos neurônios da camada oculta.....	60
Equação 3. 24	Ativação dos neurônios da camada oculta	61
Equação 3. 25	Ativação dos neurônios da camada oculta considerando a função tangente hiperbólica.....	61
Equação 3. 26	Valores internos dos neurônios da camada de saída	61
Equação 3. 27	Ativação dos neurônios da camada de saída.....	61
Equação 3. 28	Termo de erro para as unidades da camada de saída.....	61
Equação 3. 29	Derivada da função tangente hiperbólica	62
Equação 3. 30	Derivada da função tangente hiperbólica considerando a ativação dos neurônios da camada de saída	62
Equação 3. 31	Termo de erro para as unidades da camada oculta	62
Equação 3. 32	Atualização dos pesos da camada de saída	62
Equação 3. 33	Atualização dos pesos da camada oculta	62
Equação 3. 34	Erro médio quadrado	62

Equação 3. 35	Erro instantâneo obtido com o padrão p.....	62
Equação A1	Transformada de Fourier	115
Equação A2	Transformada inversa de Fourier	115
Equação A3	Transformada de Fourier de uma função real	115
Equação A4	Transformada de Fourier de uma função real na forma exponencial	116
Equação A5	Transformada de Fourier para sinais bidimensionais contínuos	116
Equação A6	Transformada inversa de Fourier para sinais bidimensionais contínuos	117
Equação A7	Função contínua discretizada	118
Equação A8	Função discreta	118
Equação A9	Transformada discreta de Fourier para sinais unidimensionais	118
Equação A10	Transformada inversa de Fourier para sinais unidimensionais	118
Equação A11	Intervalos de espaçamento entre as amostras do sinal contínuo.....	118
Equação A12	Transformada discreta de Fourier para sinais bidimensionais	119
Equação A13	Transformada inversa de Fourier para sinais bidimensionais.....	119
Equação A14	Intervalo de amostragem no domínio espacial.....	119
Equação A15	Intervalo de amostragem no domínio de frequências	119
Equação A16	Transformada discreta de Fourier para sinais bidimensionais onde $M=N$	120
Equação A17	Transformada inversa discreta de Fourier para sinais bidimensionais onde $M=N$	120
Equação A18	Decomposição da transformada discreta de Fourier para sinais bidimensionais onde $M=N$	120
Equação A19	Decomposição da transformada inversa discreta de	

	Fourier para sinais bidimensionais onde $M=N$	120
Equação A20	Relação de equivalência sobre a translação de uma função bidimensional nos domínios espacial e das frequências	121
Equação A21	Relação de equivalência sobre a translação de uma função bidimensional nos domínios das frequências e espacial	121
Equação A22	Periodicidade da transformada discreta de Fourier	121
Equação A23	Valor médio de uma função bidimensional	122
Equação A24	Decomposição da transformada discreta de Fourier para sinais bidimensionais onde $M=N$ considerando $u = v = 0$	122
Equação A25	Relação entre o valor médio de uma função bidimensional e a sua transformada de Fourier.....	123
Equação A26	Processo de convolução contínua entre duas funções	123
Equação A27	Processo de convolução discreta entre duas funções	123
Equação A28	Equivalência entre o processo de convolução no domínio espacial e das frequências.....	123
Equação A29	Equivalência entre o processo de convolução no domínio das frequências e espacial.....	123

CAPÍTULO 1

INTRODUÇÃO

Grande parte dos dados físicos para cartógrafos, geólogos e oceanógrafos é coletada através de sensoriamento remoto, por meio de satélites especializados, que tiram fotos da Terra em intervalos fixos de tempo. Estas imagens podem ser feitas através da escolha do espectro de luz que se quer enxergar e alguns podem enviar sinais para captá-los em seu reflexo com a Terra, gerando milhares de possibilidades de informação sobre minerais, concentrações de vegetação, tipos de vegetação, entre outros. Outra forma de sensoriamento remoto é a aerofotometria, que se utiliza de vôos altos para tirar fotos de dentro de aviões adaptados, artifício muito usado em agricultura e instalações de fábricas e complexos industriais, porque produz uma resolução melhor do terreno em questão.



FIGURA 1. 1. Satélite sondando a superfície da Terra.

A princípio, é possível definir sensoriamento remoto como sendo a tecnologia que permite a aquisição de informações sobre objetos sem contato físico

com eles. Um de seus principais objetivos é o de distinguir e identificar as composições de diferentes materiais superficiais, sejam eles espécies de vegetação, padrões de uso do solo, rochas, etc. Essa distinção e identificação é possível devido ao fato dos materiais superficiais terem comportamentos específicos ao longo do espectro eletromagnético.

Dois tipos de sensores podem ser utilizados na obtenção de uma imagem: os *ativos*, que transmitem pulsos de energia e medem a energia refletida desses pulsos pelos objetos (satélites com sensores SAR, por exemplo); e os *passivos*, que medem radiações refletidas ou emitidas pelos objetos (satélites com sensores multi-espectrais).

As imagens de sensoriamento remoto, por sua natureza digital ou discreta, são constituídas por um arranjo de elementos sob a forma de uma malha ou grid. Cada cela desse grid tem sua localização definida em um sistema de coordenadas do tipo *linha e coluna*. Representados por x e y , respectivamente. Por convenção, a origem do grid é sempre no seu canto superior esquerdo (Figura 1.2). O nome dado a cada uma dessas celas é *pixel* (*picture element*). Cada pixel representa sempre uma área com as mesmas dimensões, tratando-se do mesmo sensor remoto.

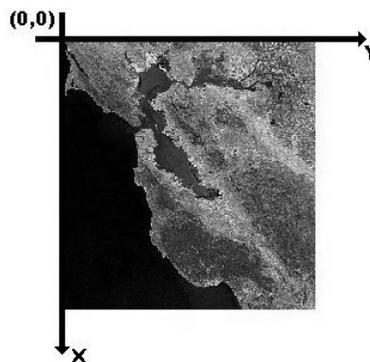


FIGURA 1. 2. Convenção utilizada para o par de eixos (x,y) .

Cada cela possui um atributo numérico z , que indica o nível de cinza dessa cela. Esse nível de cinza é conhecido como o contador digital (CD), que representa

a intensidade da energia eletromagnética (refletida ou emitida) medida pelo sensor, para a área da superfície da Terra correspondente ao tamanho do pixel. Cabe salientar que o CD de um pixel corresponde sempre à média da intensidade da energia refletida ou emitida pelos diferentes materiais presentes nesse pixel.

Uma imagem digital pode então ser vista como uma matriz, de dimensões x linhas por y colunas, com cada elemento possuindo um atributo z .

Geralmente, na interpretação de imagens de sensoriamento remoto, são utilizadas três propriedades básicas: a tonalidade (atributo espectral), a textura (atributo espacial) e o contexto (atributo de contexto). A tonalidade refere-se à cor ou ao brilho dos objetos que compõem a cena; os tons estão relacionados às propriedades de reflectância dos materiais superficiais e dependem da porção do espectro eletromagnético coberto pela imagem considerada. A textura pode ser definida como uma combinação da magnitude e frequência da variação tonal em uma imagem, sendo produzida pelo efeito conjunto de todas as pequenas feições que compõem uma área em particular na imagem. Embora possam ser consideradas como propriedades independentes, tonalidade e textura possuem relações próximas, visto que sem variações de tonalidade, nenhuma mudança em textura poderia ser percebida. Contexto, por sua vez, refere-se à localização de detalhes de tons, textura e padrões (que são arranjos de tons e texturas) em relação a atributos conhecidos do terreno.

Uma das finalidades deste trabalho é o desenvolvimento de uma metodologia para classificação de imagens de sensoriamento remoto com base em atributos de textura.

Em uma imagem de sensoriamento remoto, considerando-se uma de suas bandas espectrais, observa-se que a intensidade de cinza, representada pelos con-

tadores digitais dessa imagem, varia significativamente em função da distância. Ao se plotar essas variações em um gráfico, obtém-se geralmente uma curva complexa (Figura 1.3a). Essa curva, no entanto, pode ser decomposta em um conjunto de componentes senoidais (Figura 1.3b), com diferentes freqüências e amplitudes. Essas senóides fornecem uma idéia da distribuição espacial de freqüências da imagem considerada. Na figura 1.3, temos um caso em apenas uma dimensão, porém imagens de sensoriamento remoto se apresentam de forma bidimensional.

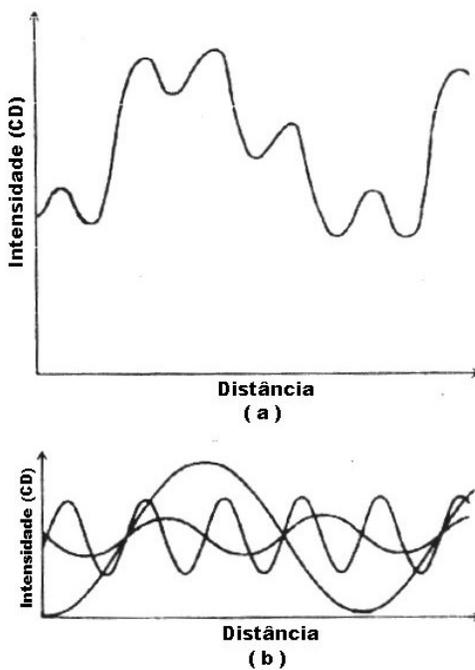


FIGURA 1. 3. As variações de CD ao longo de uma linha de uma imagem **(a)** podem ser expressas como várias componentes senóides **(b)**, com diferentes amplitudes e freqüências (Crósta, 1992).

Todas as imagens possuem limites entre áreas com diferentes respostas em relação à energia eletromagnética. Esses limites podem ser, por exemplo, entre diferentes coberturas do terreno, ou ainda, podem representar o contato entre áreas com diferentes condições de iluminação. No caso de uma imagem monocromática, esses limites representam mudanças de um intervalo de CDs para outro. Ao se plotar esses limites em um gráfico como o da figura (1.3), eles serão repre-

sentados por um gradiente bastante inclinado, podendo chegar a vertical. Limites desta natureza são conhecidos como bordas. Eles ocupam, geralmente, áreas pequenas na imagem, sendo mais estreitos que largos. Por variarem bastante em pequenas áreas, são chamados de feições de alta frequência, sendo representados por senóides de pequena amplitude. Já os limites que variam mais uniformemente com a distância são chamados de feições de baixa frequência.

A enorme mistura de frequências em uma imagem dificulta a interpretação de feições com frequências específicas. Para contornar esse problema e melhorar a aparência da distribuição espacial das informações, são utilizadas técnicas de filtragem espacial de frequências, que consistem em realçar seletivamente as feições que compõem as imagens de sensoriamento remoto. De uma forma geral, as filtragens espaciais são úteis no processamento de imagens para o realce de bordas, feições lineares de determinadas direções e padrões de textura.

Duas formas de se implementar as técnicas de filtragem espacial de frequência são o processo conhecido por convolução, que opera no domínio espacial de uma imagem, e a análise de Fourier, que opera no domínio das frequências que compõem a imagem. A filtragem no domínio de frequências é mais precisa, porém mais complexa. A convolução, por sua vez, é mais simples e fácil de se implementar.

As últimas gerações de sensores de observação da Terra vêm produzindo dados com grande potencial para o uso em investigações científicas e tecnológicas em grande escala. Ao mesmo tempo em que tais dados fornecem recursos para pesquisadores, também introduzem desafios para o processamento e interpretação desses dados, na forma de imagens. É no contexto destas necessidades que as

redes neurais artificiais são aplicadas em uma ampla variedade de aplicações de sensoriamento remoto.

O uso de redes neurais artificiais na interpretação de dados de sensoriamento remoto vem sendo motivado pela eficiência do cérebro humano para o processamento de grandes quantidades de dados de diversas fontes. Os neurônios no cérebro humano recebem sinais de outros neurônios e produzem respostas que são enviadas para outros neurônios. Algumas vezes, é conhecido que uma abordagem matemática, baseada em ações de neurônios biológicos, pode ser implementada para processar e interpretar diferentes tipos de dados digitais. Enquanto não é possível reproduzir a complexidade do cérebro humano, as redes neurais artificiais, baseadas numa arquitetura de elementos processadores simples como neurônios alcançam êxito em várias aplicações, incluindo o processamento e a classificação de dados obtidos por sensoriamento remoto.

As redes neurais podem ser interpretadas como transformadoras de dados (Pao, 1989), onde o objetivo é associar os elementos de um grupo de dados com os elementos de um segundo grupo. A utilização das abordagens neurais no sensoriamento remoto é adequada, principalmente pela grande habilidade demonstrada para:

- (i) desenvolver maior exatidão que algumas técnicas;
- (ii) trabalhar mais rapidamente;
- (iii) incorporar conhecimento *a priori* e limites físicos reais na análise.

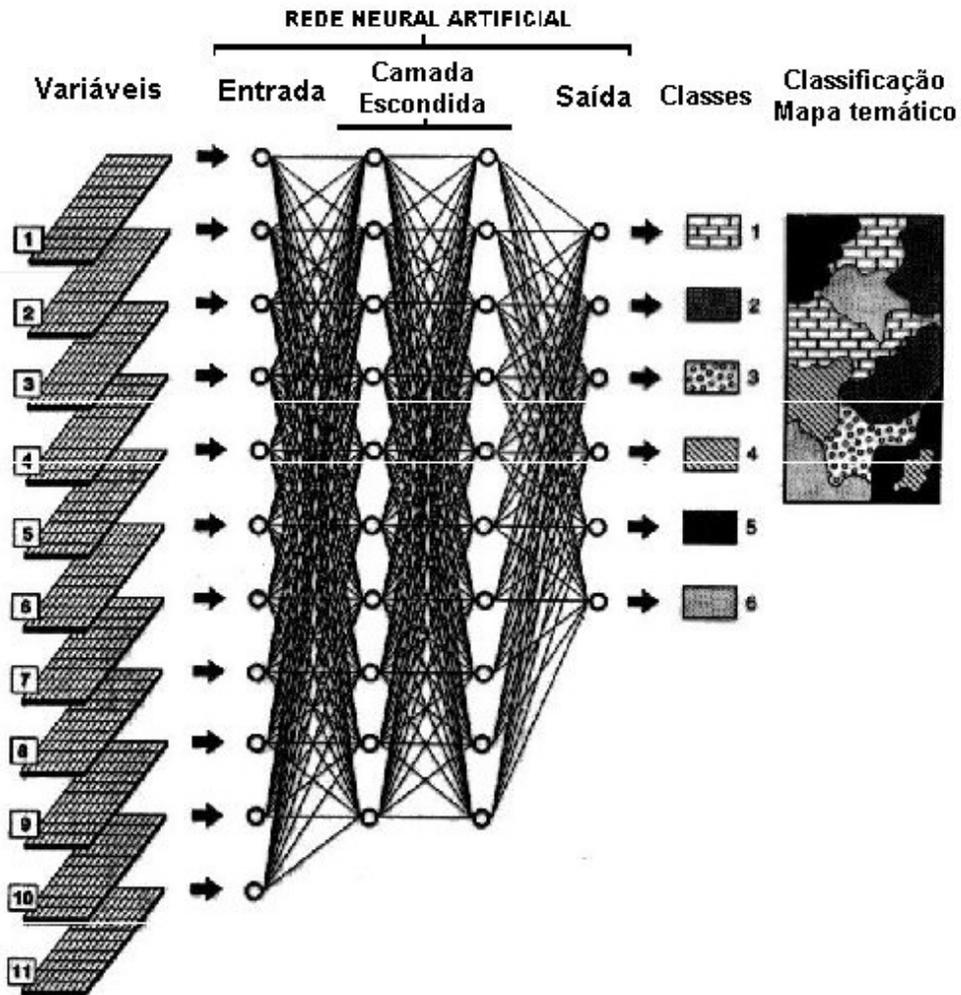


Figura 1. 4. Esquema de classificação com uma rede neural artificial. A rede mostrada corresponde a uma arquitetura contendo 11 unidades na camada de entrada, 20 unidades escondidas em 2 camadas e seis unidades na camada de saída. (Adaptado de Foody e Arora, 1997)

Dados os benefícios, uma das principais oportunidades oferecidas pelas redes neurais é a de permitir o eficiente manuseio de grandes quantidades de dados provenientes de sensoriamento remoto que estão sendo produzidos, sendo aplicadas, principalmente, para a classificação de imagens (Figura 1.4).

1. 1. Objetivos

O objetivo deste trabalho é propor uma metodologia de classificação de imagens de sensoriamento remoto que integre a importância de atributos de textura

na seleção de feições com a eficiência das redes neurais artificiais no processo de classificação. A seleção das características texturais é realizada através da utilização de frequências espaciais de cada classe textural e sua direção por um processo de filtragem da imagem com um banco de filtros de Gabor.

A classificação da imagem é realizada pela rede neural artificial multinível com algoritmo Backpropagation, num processo pixel a pixel. Para tanto, o treinamento da RNA foi realizada com dados retirados de amostras de cada textura da imagem original. A rotulação feita pela rede neural, e baseada nas texturas, gera a imagem classificada.

1. 2. Motivação

O trabalho foi desenvolvido a partir pesquisas realizadas por autores como Jain e Farroukhnia (1991), Raghu et al (1995) e Angelo (2000). Jain e Farroukhnia trabalharam numa proposta de filtragem multicanal utilizando bancos de filtro de Gabor na segmentação de textura. Raghu utiliza uma técnica de extração de características texturais baseada no processo de Jain e Farroukhnia fazendo uso de uma rede neural de dois estágios para a classificação da imagem. Angelo alia uma técnica de filtragem multicanal na segmentação da textura ao método de Máxima Verossimilhança Gaussiana para a classificação. Dessa forma, este trabalho baseia-se na técnica de filtragem multicanal de Angelo (2000) empregando para a classificação da imagem uma rede neural artificial multinível com algoritmo Backpropagation.

O texto está dividido em cinco capítulos. Este capítulo faz uma introdução ao assunto, construindo uma visão geral do trabalho desenvolvido. O capítulo 2 apresenta uma revisão bibliográfica de conteúdos e trabalhos desenvolvidos por

pesquisadores da área, dando destaque para tópicos de maior relevância para a compreensão do trabalho. No capítulo 3 é possível verificar a metodologia proposta e desenvolvida, bem como materiais utilizados. A descrição dos experimentos realizados é apresentada no capítulo 4. Os resultados obtidos, a análise sobre eles, conclusões e sugestões para futuros trabalhos são desenvolvidos no capítulo 5.

CAPÍTULO 2

REVISÃO BIBLIOGRÁFICA

2. 1. Introdução

Imagens de sensoriamento remoto vêm servindo de fontes de dados para estudos e levantamentos geológicos, ambientais, agrícolas, cartográficos, florestais, urbanos, oceanográficos, etc. Estas imagens passaram a representar uma das únicas formas viáveis de monitoramento ambiental em escalas locais e globais, devido à rapidez, eficiência, periodicidade e visão sinóptica que as caracterizam. Por isso, um grande número de países vem se envolvendo em programas espaciais voltados ao lançamento e operação de satélites de levantamento e monitoramento de recursos naturais, bem como o desenvolvimento de sofisticados sistemas sensores.

Uma das maiores conseqüências deste processo é o grande aumento na quantidade de dados, principalmente na forma de imagens digitais, que vem sendo gerados continuamente. Portanto, torna-se necessário o avanço das técnicas de processamento das imagens digitais, as quais contêm uma quantidade muito maior de informações do que a que o olho humano consegue perceber. O processamento digital de imagens de sensoriamento remoto trata especificamente das técnicas utilizadas para identificar, extrair, condensar e realçar a informação de interesse para determinados fins, a partir de um grande volume de dados que compõem essas imagens.

Por outro lado, a classificação de imagens multiespectrais de sensoriamento remoto diz respeito a associar cada pixel da imagem a um "rótulo" descrevendo um objeto real(vegetação, solo, etc.). Dessa maneira, os valores numéricos(CD) associados a cada pixel, definidos pela reflectância dos materiais que compõem

esse pixel, são identificados em termos de cobertura da superfície terrestre imageada. Quando esse tipo de operação é efetuado para todos os pixels de uma determinada área, o resultado é um mapa temático, mostrando a distribuição geográfica de um tema. Portanto, uma imagem de sensoriamento remoto classificada é uma forma de mapa digital temático (Crósta, 1992). Quando essa imagem é reorganizada para se ajustar a uma dada projeção cartográfica, torna-se um elemento importante a ser incorporado a um sistema geográfico de informações.

Metodologias que implementam atributos de textura no processamento e classificação de imagens digitais têm sido amplamente discutidas pela comunidade científica. Da mesma forma, algoritmos envolvendo a utilização de redes neurais artificiais no trabalho com imagens provenientes de sensoriamento remoto vêm sendo relatados em periódicos, tais como: *International Journal of Remote Sensing*, *Pattern Recognition*, *International Journal of Approximate Reasoning*, *IEEE Transactions on Geoscience and Remote Sensing*, *IEEE Transactions on Acoustics, Speech and Signal Processing*, *Optical Engineering*, entre outros. Devido à importância dada a estes tópicos, este capítulo visa apresentar algumas abordagens desenvolvidas, inserindo o tema deste trabalho num contexto mais amplo do processamento e classificação de imagens digitais.

2. 2. Técnicas de classificação de imagens

O processo de classificação multispectral consiste em atribuir cada pixel de uma imagem a uma das classes presentes. Essa atribuição pode ser feita com base em observações no pixel em si e na sua vizinhança.

A maioria das técnicas de classificação se baseia apenas no agrupamento de valores de intensidade espectral, representados pelos CDs presentes na imagem. O conteúdo de informação de uma imagem de sensoriamento remoto baseia-se na intensidade (nível de cinza) de cada pixel individual, bem como no arranjo espacial dos pixels (textura e forma). O que acontece é que as técnicas padrão de classificação levam em conta apenas as características espectrais, deixando de considerar as texturais. Portanto, a classificação apresenta resultados satisfatórios na indicação de tipos de cobertura caracterizadas apenas por variações de intensidade.

Quanto à classificação de imagens multiespectrais de sensoriamento remoto, são consideradas duas abordagens: a classificação supervisionada, onde o usuário identifica alguns dos pixels pertencentes às classes desejadas e o software executa a tarefa de localizar os demais pixels pertencentes àquelas classes, baseado em alguma regra estatística pré-estabelecida; a classificação não-supervisionada, onde o software decide, com bases em regras estatísticas, quais as classes a serem separadas e quais os pixels pertencentes a cada uma.

2. 2. 1. Classificação supervisionada

Uma área da imagem que o usuário identifica como representando uma das classes é chamada de amostra de treinamento. Várias amostras podem ser definidas para uma mesma classe, para assegurar que os pixels a ela pertencentes são representativos dessa classe. Todos os pixels dentro de uma amostra para uma dada classe constituem um conjunto de treinamento para aquela classe (Crósta, 1992). Os CDs dos pixels dos conjuntos de treinamento em cada uma das bandas espectrais são comparados com os CDs de cada pixel da imagem, para decidir a

qual classe estes pertencem. Este processo é chamado de Classificação Supervisionada.

Existem diversos métodos nos quais os pixels desconhecidos podem ser comparados com os pixels do conjunto de treinamento. O produto final é uma imagem classificada, representando um mapa temático (Schowengerdt, 1983). Alguns métodos mais conhecidos de classificação espectral supervisionada são citados a seguir.

2. 2. 1. 1.Método do Paralelepípedo

Este método considera uma área no espaço de atributos ao redor do conjunto de treinamento. Essa área tem a forma de um quadrado ou paralelepípedo, definido pelo contador digital máximo e mínimo do conjunto de treinamento. Todos os pixels situados dentro do paralelepípedo serão classificados como pertencentes à mesma classe.

Cada classe possui limites de decisão, representados pelas laterais do paralelepípedo (Crósta, 1992). Um dos problemas que afetam o classificador de paralelepípedo é a correlação normalmente existente entre bandas, responsável pela distribuição alongada em torno do eixo de 45° do espaço de atributos, onde os limites de decisão definidos por um retângulo ou quadrado acabam por abranger alguns pixels não pertencentes à classe. A figura 2.1 mostra um conjunto de treinamento com 23 pixels pertencentes à classe A, representado no espaço de atributos de 2 bandas.

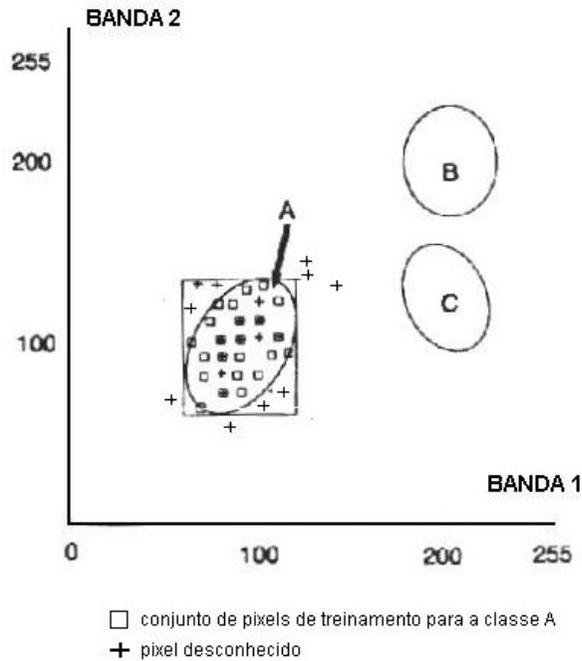


FIGURA 2.1. Diagrama de espaço de atributos das classes A, B e C, mostrando a localização dos pixels pertencentes a um conjunto de treinamento e a pixels desconhecidos. Os limites de decisão de um classificador de paralelepípedo são mostrados ao redor da classe A. (Adaptado de Crósta, 1992)

2. 2. 1. 2. Método da Distância Mínima

O método da distância mínima atribui cada pixel desconhecido à classe cuja média é mais próxima a ele. Este método considera a questão da proximidade entre classes com base em dados estatísticos. A figura 2.2 ilustra um exemplo: o pixel desconhecido 1 situa-se mais próximo da classe A, sendo a ela atribuído; da mesma forma em que o pixel 2 é atribuído à classe B. Um dos problemas deste método acontece quando um pixel que está à mesma distância das médias das classes A e B, necessitando parâmetros estatísticos mais sofisticados.

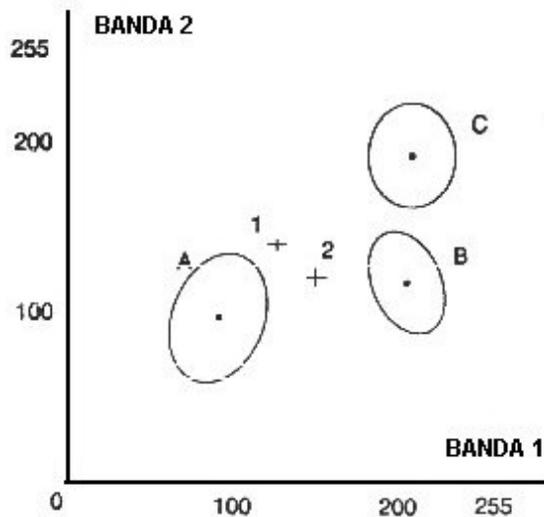


FIGURA 2.2. Espaço de atributos das classes A, B e C mostrando a posição das médias de cada classe e de dois pixels desconhecidos não classificados pelo método do paralelepípedo (Adaptado de Crósta, 1992).

2. 2. 1. 3. Método da Máxima Verossimilhança Gaussiana

Este método considera a ponderação das distâncias entre as médias através de parâmetros estatísticos. O método da máxima verossimilhança exige um número bastante elevado de pixels para cada conjunto de treinamento, permitindo uma base segura para o tratamento estatístico (Crósta, 1992). Uma forma de exemplificação gráfica do método é através de isolinhas ou curvas de contorno no espaço de atributos, representando a densidade de pixels do conjunto de treinamento para cada região (Figura 2.3), através de distribuições normais ou Gaussianas.

Os contornos em torno de cada classe representam a probabilidade de existir um pixel naquele conjunto de treinamento com uma determinada combinação de CDs, possibilitando a escolha de classificar apenas pixels desconhecidos que são muito semelhantes à média de determinada classe.

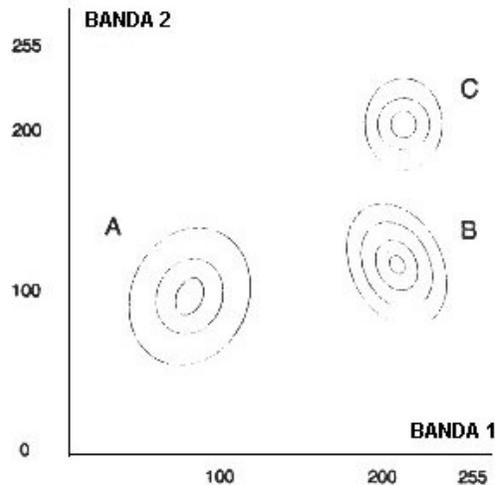


FIGURA 2.3. Espaço de atributos para as classes A, B e C mostrando na forma de contornos a densidade de pixels nos diversos conjuntos de (Adaptado de Crósta, 1992).

2. 2. 2. Classificação Não-supervisionada

Baseia-se no processo automático de identificação de classes dentro de um conjunto de dados. A figura 2.4 mostra um exemplo, onde a densidade de ocorrência de pixels é representada por curvas de contorno. É possível perceber que o espaço de atributos contém vários agrupamentos (*clusters*), os quais são caracterizados por regiões de alta densidade de curvas de contorno. No processo de classificação não-supervisionada, esses clusters são automaticamente identificados e usados como áreas de treinamento para a classificação.

O primeiro passo na classificação não-supervisionada é a realização de uma análise de agrupamento. Através dessa análise, os clusters são identificados no espaço de atributos. Então, uma decisão deve ser tomada em relação a quais concentrações devem ser tratadas como grupos separados. Uma forma utilizada é a de medir a altura dos picos em relação às depressões adjacentes, considerando apenas os picos cuja altura ultrapasse um determinado limiar.

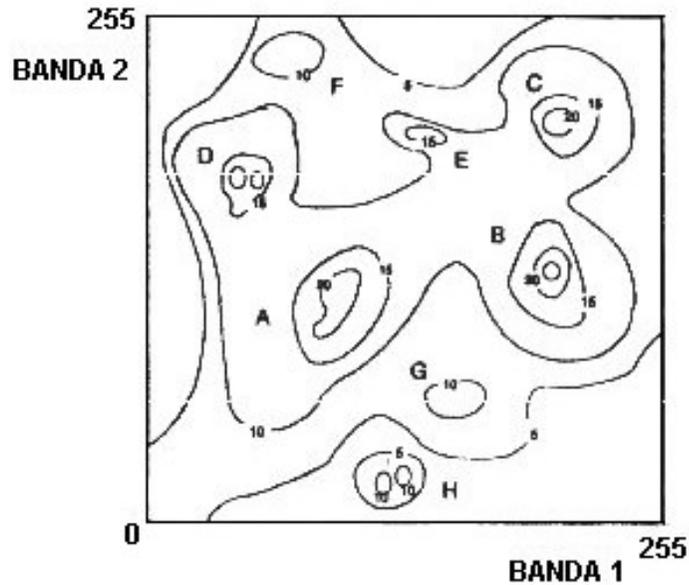


Figura 2.4. Espaço de atributos mostrando a distribuição de densidade de todos os pixels em uma imagem, na forma de curvas de contorno. As maiores concentrações foram denominadas de A a H (Adaptado de Crósta, 1992).

2. 3. Filtragem no domínio espacial

As técnicas de filtragem no domínio espacial atuam diretamente sobre a matriz de pixels que é a imagem digitalizada. Logo, as funções de processamento de imagens no domínio espacial podem ser expressas como:

$$g(x,y) = T [f(x,y)] \quad (2.1)$$

onde $g(x,y)$ é a imagem processada, $f(x,y)$ é a imagem original e T é um operador em f , definido em uma certa vizinhança de (x,y) . A vizinhança definida ao redor de (x,y) equivale a uma região 3×3 , na qual o pixel central é o de referência (Figura 2.5). O centro dessa região é movido pixel a pixel, iniciando no canto superior esquerdo da imagem e aplicando à localidade o operador T para calcular um novo valor g naquele ponto.

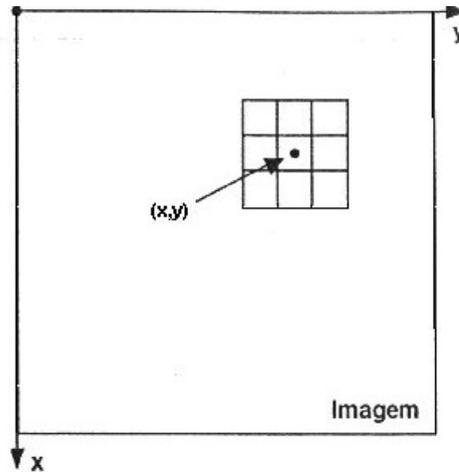


Figura 2.5. Vizinhança 3x3 ao redor de um ponto de coordenadas (x,y) em uma imagem (Marques e Vieira, 1999).

2. 4. Filtragem no domínio de freqüências

O teorema da convolução é usado como base matemática das técnicas de filtragem no domínio de freqüências. Seja $g(x,y)$ a imagem formada pela convolução da imagem $f(x,y)$ com um operador linear $h(x,y)$, então:

$$g(x,y) = f(x,y) * h(x,y) \quad (2.2)$$

Desta forma, pelo teorema da convolução (ver Apêndice 1: Transformada de Fourier), a seguinte relação no domínio de freqüências também é válida:

$$G(u,v) = F(u,v).H(u,v) \quad (2.3)$$

onde G , F e H são as transformadas de Fourier de g , f e h , respectivamente. A transformada $H(u,v)$ é a função de transferência do filtro.

2.5. Técnicas de filtragem em multicanais texturais

A extração de feições na imagem, nesta técnica, utiliza a filtragem com filtros de Gabor. Desta forma, os parâmetros determinantes do filtro de Gabor, isto é,

extensão, frequência e orientações espaciais, são determinadas a priori. A partir da extração destas feições procede-se à classificação da imagem, utilizando-se algum classificador.

Uma técnica de filtragem multi-canal para segmentação de textura que usa um banco de filtros simétricos de Gabor para caracterizar os canais é apresentada por Jain e Farrokhnia(1991). Eles propõem um esquema sistemático de seleção de filtros, baseado na reconstrução da imagem original a partir de suas imagens filtradas. Cada imagem filtrada(selecionada) é sujeita a uma transformação não-linear e, posteriormente, através de uma medida de “energia” em uma pequena janela ao redor de cada pixel na imagem para obter as feições texturais. O processo gera uma imagem de feição que corresponde a cada imagem filtrada. O tamanho da janela para cada imagem filtrada é determinado usando-se uma fórmula simples que envolve a frequência radial para a qual o filtro correspondente é tomado. Para identificar as classes texturais, usa-se um algoritmo de clusterização, baseado no algoritmo de “k-means”. Também é proposto um procedimento simples para inclusão de informação espacial no processo de clusterização. Uma modificação da estatística Γ de Hubert, que utiliza a correlação entre a distância euclidiana de dois padrões i, j e a distância euclidiana entre o centro dos dois aglomerados(*clusters*) que contêm estes padrões. Esta estatística é usada como um índice relativo para calcular o número de classes texturais. O desempenho do algoritmo é avaliado em imagens com um número grande de texturas naturais. Os resultados alcançados não dependeram de um conhecimento prévio das frequências espaciais presentes nas classes texturais. Além disto, pode-se diminuir a quantidade de imagens filtradas para a classificação da imagem original. Entretanto, o processo de escolha do número de classes na fase de clusterização aumenta consideravelmente o custo computacio-

nal do processo como um todo, principalmente para um número elevado de classes efetivamente presentes na imagem. Outro problema consistiu em admitir uma não sobreposição entre os diferentes canais de freqüência espacial quando parecem existir evidências fisiológicas e psico-físicas afirmando o contrário.

Uma aproximação computacional para analisar texturas visíveis é proposta por Bovik et al(1990). As texturas são modeladas como padrões que contêm um intervalo limitado de freqüências espaciais, onde as texturas mutuamente distintas diferem significativamente em suas freqüências características dominantes. A imagem original de textura é codificada em múltiplos canais de freqüência espacial e orientação, onde envelopes de variação suave de canais (amplitude e fase) são usados para separar regiões de textura com diferentes freqüências espaciais, orientação ou características de fase.

Um filtro de gabor bidimensional é usado como filtro de canal de freqüência para extrair feições texturais, devido às suas características específicas. Através da comparação das respostas de amplitudes dos canais, podem ser descobertas fronteiras entre as texturas e, portanto, pode-se separá-las.

Através dos resultados encontrados por Bovik et al, sugere-se que os filtros de Gabor podem ser utilizados para outras tarefas de seleção de feições simulando a percepção visual humana e incluindo, inclusive, localização temporal. Porém, um dos problemas é o alto custo computacional para compor um banco de filtros completo que represente a imagem original.

Raghu et al(1995), propõe uma estrutura de rede neural que combina as características de duas redes neurais: SOM(*Self Organizing Map*) e MLP (*MultiLayer Perceptron*) para a classificação de textura. As feições texturais são obtidas através do uso de um banco de filtros de Gabor com tamanhos, orientações e fre-

quências espaciais diferentes, constituindo um vetor de feições a N dimensões. A rede neural SOM age como um mecanismo de clusterização que projeta estes vetores de feições N-dimensionais sobre um espaço M-dimensional. O valor de M, neste estudo, foi tomado como dois, reduzindo, desta forma, a dimensionalidade do espaço das imagens filtradas.

As feições geradas após o uso da rede neural SOM são utilizadas para alimentar uma rede neural MLP com a finalidade de treinamento e subsequente classificação. Este mecanismo aumenta, portanto, a distância inter-classe (distância comum entre os vetores de classes diferentes) e ao mesmo tempo diminui a distância intra-classe (distância comum entre os vetores da mesma classe) no espaço de feições, reduzindo a complexidade de classificação. Os resultados demonstram que a utilização da rede neural SOM melhora a classificação, tanto na exatidão obtida, como no tempo de treinamento da rede neural MLP, comparando-se à classificação direta sem esta etapa.

2.6. Técnicas de redes neurais

Muitos pesquisadores consideram Redes Neurais como sistemas de plausibilidade biológica, porém, preferimos dizer que são sistemas que possuem inspiração biológica. As Redes Neurais representam um novo conceito de sistema de processamento onde os modelos são feitos com base no que se sabe dos princípios do processamento neurofisiológico.

O cérebro é composto por células altamente diferenciadas denominadas neurônios (Figura 2.6). Cada neurônio possui um corpo celular, ou soma, dentro do qual encontram-se a maioria dos seus orgânulos. Do soma de cada neurônio par-

tem prolongamentos que podem ser funcionalmente divididos em dendritos e axônios.

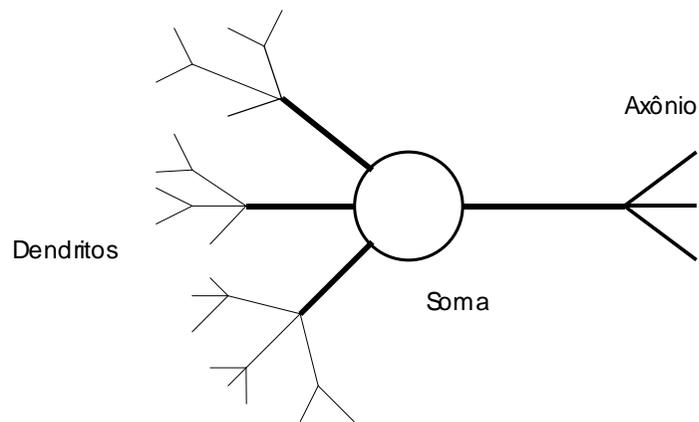


FIGURA 2.6. Esquema simplificado de um neurônio

O neurônio é uma célula altamente especializada, capaz de receber impulsos elétricos a partir de seus dendritos, processá-los no soma e, finalmente, transmiti-los através de seu axônio (geralmente único) para os dendritos (ou mesmo para o soma) de outros neurônios.

A conexão entre um axônio de um neurônio e um dendrito (ou soma) de outro é denominada sinapse (Figura 2.7). A sinapse é a unidade funcional básica para a construção de circuitos neurais biológicos e envolve a posição das membranas plasmáticas de dois neurônios de modo a formar uma junção pontual (o tamanho de uma junção sináptica é menor do que 1 mm) e orientada do neurônio pré-sináptico para o pós-sináptico.

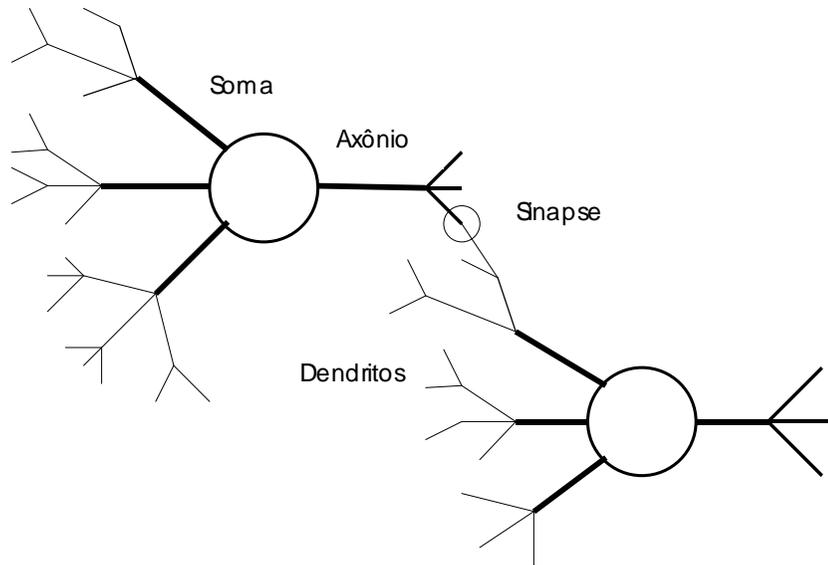


Figura 2. 7. Esquema simplificado de uma conexão sináptica

O cérebro humano possui cerca de 100 bilhões de neurônios, cada qual com aproximadamente 10 mil interconexões. Deste modo, estima-se em dezenas de trilhões o número de conexões existentes no cérebro.

2. 6. 1. Conceitos Básicos

Existem muitas referências sobre a aplicação de redes neurais na resolução de diversos problemas, porém, a maioria refere-se à aplicação de um único tipo de rede (BackPropagation) cujo desempenho já foi testado e aprovado em determinadas aplicações. Ao contrário do que se imagina, cada modelo de rede possui diferentes propriedades que, devido a características inerentes, as tornam adequadas à resolução de determinada classe de problemas. Estes conceitos são de grande valia para a compreensão dos paradigmas que serão apresentados, além de uniformizar os jargões utilizados na área.

O paradigma conexionista procura entender e emular as propriedades decorrentes do alto grau de paralelismo e conectividade. As redes neurais são mode-

los que utilizam o paradigma conexionista na resolução de determinado tipo de problemas.

Uma rede neural é composta por um elevado número de elementos processadores, também denominados unidades de processamento, amplamente conectados entre si. Cada uma das conexões interliga somente dois elementos processadores, geralmente em um único sentido, e possui um valor que determina o grau de conectividade entre estes, denominado peso da conexão.

Deste modo, todo o processamento é realizado distribuidamente entre os elementos processadores da rede, onde cada qual o realiza isolada e paralelamente, enviando seu resultado para outras unidades através das conexões entre eles. Por isso, as Redes Neurais também são conhecidas como sistemas de processamento distribuído e paralelo.

A forma pela qual os elementos processadores estão interligados é denominada topologia ou padrão de interconexão. Uma rede é dita rígida ou fixa quando sua topologia não pode variar; caso contrário, é denominada flexível ou plástica, uma vez que poderá crescer e diminuir de acordo com a necessidade do sistema. As redes rígidas são mais simples e, conseqüentemente, as mais utilizadas. Um dos maiores problemas em sua utilização é determinar o número de elementos processadores a serem utilizados e a topologia das conexões entre eles. Apesar dos inúmeros artigos sobre este assunto, o método mais utilizado continua sendo o de "tentativa e erro". Existem poucas redes plásticas conhecidas. Estas são particularmente úteis quando o problema não é conhecido ou quando ele varia no tempo.

Normalmente, uma arquitetura de rede é dita estática quando o peso das conexões entre os elementos processadores não pode variar, caso contrário é denominada plástica. As redes estáticas são utilizadas na resolução de problemas

conhecidos e equacionados. As arquiteturas de redes dinâmicas possuem mecanismos de "aprendizado", que possibilitam a modificação dos pesos de suas conexões de modo a melhorar seu desempenho. Geralmente, quando se quer implementar uma rede neural em hardware utilizam-se redes estáticas. Porém, até chegar ao modelo estático ideal, utiliza-se uma rede dinâmica em software.

A dinâmica ou plasticidade da rede neural pode ser feita de dois modos distintos: criação (ou remoção) de novos elementos e/ou conexões; modificação nos elementos e/ou conexões existentes (Rumelhart, 1986). Normalmente, modificam-se os pesos das conexões sem criar ou remover novos elementos. Porém, este pode ser considerado como um caso especial do primeiro, pois alterar o valor de uma conexão de ZERO para qualquer outro valor equivale a criar uma nova conexão, e zerá-la equivale a eliminá-la. Embora isto não represente um grande problema do ponto de vista matemático, pode representar um sério problema na prática, pois os recursos computacionais são limitados e o tempo de processamento é proporcional à quantidade de recursos necessários.

A capacidade das Redes Neurais em resolver um determinado problema encontra-se embutida na topologia (padrão de interconexão) da rede. Ou seja, o modo pelo qual os elementos processadores estão interconectados e os pesos destas conexões determinam o problema que uma Rede Neural é capaz de resolver.

O processo de síntese de uma Rede Neural consiste em inicializá-la de modo a ser capaz de resolver o problema ao qual se destina. A idéia dos métodos de síntese é gerar, de uma única vez, o padrão de interconexão capaz de resolvê-lo. Caso contrário, ela deve possuir algum processo de treinamento capaz de modificar gradualmente um padrão de interconexão inicial, de modo a adaptá-lo à resolução deste problema. A síntese de Redes Neurais é muito aplicada em problemas

de otimização, cuja equação de resolução seja bem conhecida, enquanto o treinamento se aplica aos demais problemas. É importante ressaltar que as redes treináveis não estão livres do problema de síntese. A inicialização dos pesos das conexões (síntese) pode ser um fator determinante do sucesso ou fracasso do treinamento destas na resolução de um problema.

Os mecanismos de aprendizado possibilitam a modificação do padrão de interconexão de uma Rede Neural, capacitando-a a resolver um determinado problema. Para treinar uma Rede Neural podem ser utilizados três mecanismos distintos de aprendizado: o aprendizado supervisionado, quando são fornecidos integralmente os resultados desejados; o aprendizado por reforço, quando apenas um parâmetro externo de comparação, ou medida relativa da adequação, para saber se estão agindo corretamente ou erroneamente; e o aprendizado não-supervisionado, quando a própria rede é capaz de ajustar o seu funcionamento.

A maioria absoluta das aplicações existentes compõe-se de redes neurais com aprendizado supervisionado, que pode ser considerado como a capacidade que a rede possui de modificar o seu desempenho a partir da comparação entre a resposta obtida e a resposta desejada (Figura 2.8).

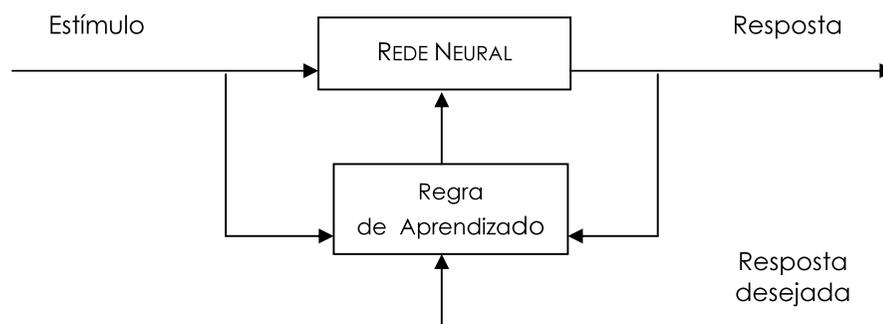


Figura 2. 8. Aprendizado em redes neurais supervisionadas

As redes não-supervisionadas, por sua vez, têm a capacidade de determi-

nar uma correlação entre os possíveis padrões de entrada e são particularmente úteis nos problemas em que as entradas variam com o tempo de forma conhecida. Podemos considerar este mecanismo de aprendizado como sendo a capacidade que a rede possui de abstrair correlações entre os estímulos de modo a obter as respostas desejadas (Figura 2.9).

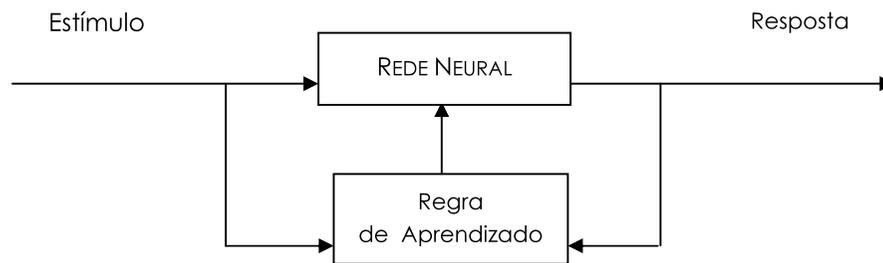


Figura 2. 9. Aprendizado em redes neurais não-supervisionadas

Os demais algoritmos de aprendizado que podem ser encontrados em Redes Neurais utilizam uma mistura desses dois mecanismos de aprendizado.

Costuma-se dizer que são necessárias 3 (três) fases para aplicar redes neurais à resolução de um problema qualquer. Na primeira fase, ou treinamento, ensina-se a rede a resolver um conjunto de padrões de saída associados à padrões de entrada. Na segunda fase, ou teste, são apresentados padrões de entrada à rede, e as saídas obtidas são comparadas às saídas desejadas. Na terceira fase, ou aplicação, também é conhecida como lembrança ("recall"), a rede aprovada na fase anterior é utilizada na resolução do problema.

De um modo geral, o problema de inicialização da rede neural é ignorado, ou seja, a determinação da topologia e a inicialização dos pesos das conexões são problemas desprezados ou minimizados pelos paradigmas existentes. Porém, a prática mostra que estes aspectos são fundamentais em seu desempenho. Assim, é possível chamar a resolução deste problema de fase zero, ou síntese.

O modelo de elemento processador normalmente possui N entradas e 1 única saída (Figura 2.10) e seu processamento consiste em transferir para sua saída um valor calculado a partir de outros valores presentes em suas entradas, através de uma função denominada função de transferência.

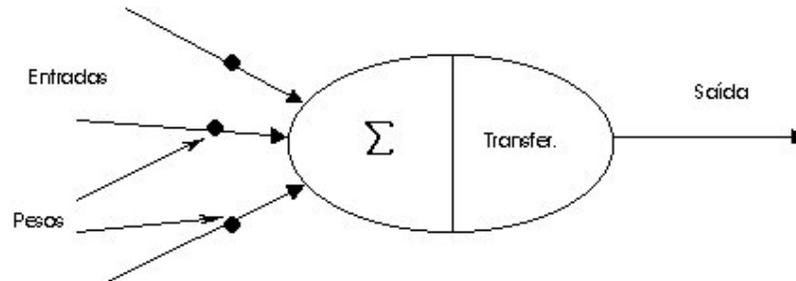


Figura 2. 10. Modelo simplificado de um elemento processador.

Usualmente, as entradas são combinadas por uma simples soma ponderada e transferidas para a saída através de uma função de limiar:

$$\text{saída} = \text{limiar}(\sum \text{entradas} \times \text{peso}) \quad (2.4)$$

Existem outras funções de transferência além da função de limiar (Figura 2.11a), como a função linear (Figura 2.11b) e a função sigmóide (Figura 2.11c).

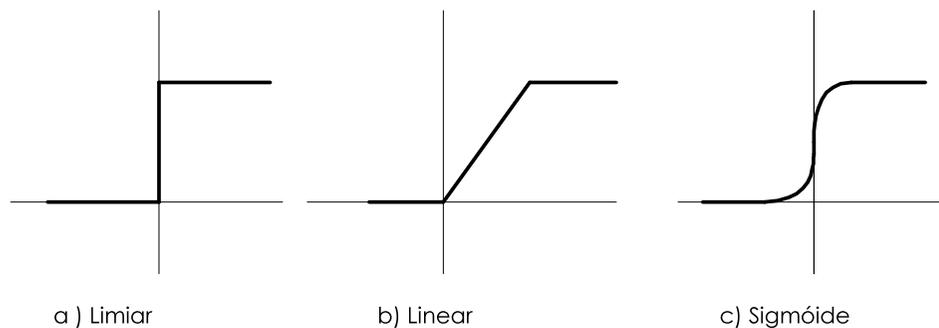


Figura 2. 11. Funções de transferência.

O valor da saída de um elemento processador em um instante de tempo é denominado potencial de ativação do elemento processador. O conjunto dos estados de ativação de cada um dos elementos processadores de uma rede neural é denominado função de ativação da rede neural ou às vezes também denominadas

estado de ativação da rede neural.

Os elementos processadores são, geralmente, agrupados em pelo menos duas camadas. Uma camada para apresentação dos dados à rede (camada de entrada) e outra para obtenção dos resultados (camada de saída). A maior parte dos modelos possui, além destas, uma ou mais camadas intermediárias, também denominadas camadas ocultas.

As camadas são processadas no sentido da camada de entrada para a de saída, de modo que a cada padrão de entrada corresponderá um padrão de saída. Normalmente, não existe conexão entre os elementos de uma mesma camada, cujos elementos se conectam com os elementos das camadas imediatamente anterior e posterior, formando as redes alimentadas *feed forward* (Figura 2.12). Porém, em alguns modelos, os elementos se conectam internamente na camada, ou, com camadas não adjacentes.

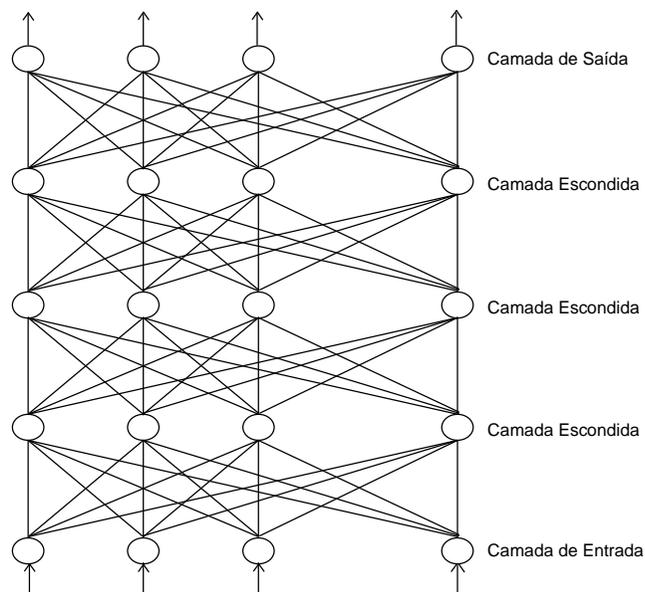


Figura 2. 12. Rede neural de 5 camadas.

O número de elementos em uma camada pode variar de um único (como em algumas camadas do modelo ART) a todos os elementos processadores de

uma rede neural (como na rede de Hopfield).

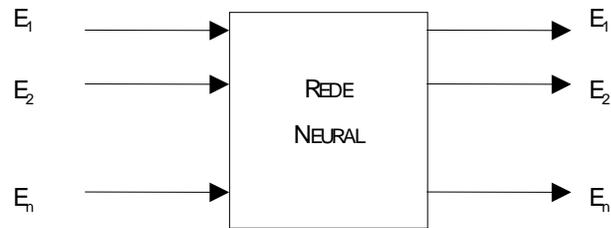


Figura 2. 13. Rede neural auto-associativa.

Uma rede neural pode ser considerada, basicamente, como um associador de padrões. Se os padrões de entrada forem iguais aos padrões de saída desejados, a rede é denominada auto-associativa (Figura 2.13), caso contrário, é denominada hetero-associativa (Figura 2.14).

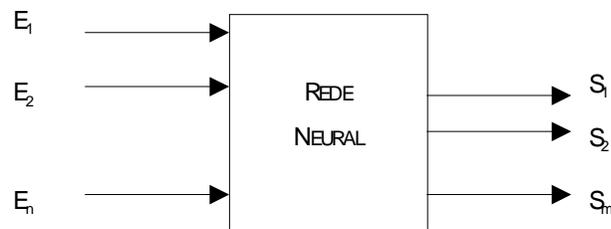


Figura 2. 14. Rede neural hetero-associativa.

É possível encontrar na bibliografia (Rumelhart, 1986) quem classifique as Redes Neurais em quatro ou mais paradigmas de funcionamento: "Auto-Associador", "Associador de Padrões", "Classificador" e "Detector de Regularidades". Na realidade, existem mais do que quatro modos de funcionamento, sendo possível qualquer forma de classificação. Porém, se desejamos nos restringir aos padrões apresentados e obtidos, as duas classificações acima representam os dois extremos possíveis.

Os padrões são processados de acordo com a topologia das redes neurais, que dividem-se em dois tipos distintos: redes sem realimentação (feed-foward, não-recursivas, direcionadas) e redes com realimentação (realimentadas, recursivas). Nas redes sem realimentação, a informação flui em alguma direção, de modo que,

ao ser processada por um elemento processador, não retorne a ele. Já nas redes com realimentação, a informação processada pelo elemento processador poderá ser novamente utilizada por este. Esta nomenclatura é análoga à utilizada em circuitos elétricos e eletrônicos.

As redes sem realimentação só possuem uma memória de curtíssimo prazo, sendo as únicas informações armazenadas relativas ao processamento imediatamente anterior de cada elemento. Nelas, os padrões de entrada são processados até originarem os padrões de saída, e pode-se garantir que a rede chegará ao estado final após o processamento de seus elementos.

Porém, em uma rede realimentada é necessário algum mecanismo capaz de garantir que ela oscilará até atingir um estado estável nas suas saídas em um determinado intervalo de tempo.

Normalmente, as redes hetero-associativas utilizam uma topologia sem realimentação (Back-Propagation), enquanto que as auto-associativas utilizam topologia com realimentação (Hopfield). Porém, existem contra-exemplos clássicos, como a rede ART, que apesar de hetero-associativa possui realimentação.

2.6.2. Histórico e Paradigmas

Em 1943, McCulloch desenvolveu um modelo matemático simplificado de um neurônio (Figura 2.15). Nele, o neurônio possuía apenas uma saída (s), que era uma função de limiar ("threshold") da soma de suas diversas entradas (e).

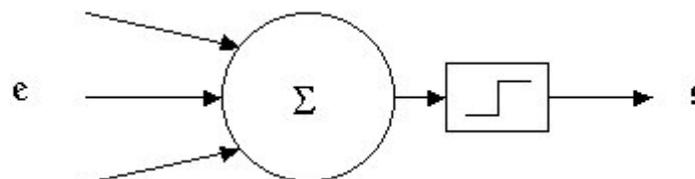


Figura 2.15. Modelo matemático simplificado de um neurônio de McCulloch.

Em 1947, McCulloch e Pitts criaram redes interligando vários elementos processadores (neurônios de McCulloch) e demonstraram que estes eram capazes de executar ações complexas quando conectados a outros elementos semelhantes.

Em 1949, Hebb sugere que a alteração da eficiência sináptica é a base do aprendizado, através do seguinte postulado: "Quando uma célula A está suficientemente próxima para excitar uma célula B e repetida ou persistentemente toma parte no disparo desta, algum processo de crescimento ou mudança metabólica ocorre em uma ou ambas as células de modo que a eficiência de A em excitar B é aumentada.". Então, segundo Hebb, uma determinada conexão somente é reforçada se tanto as células pré-sinápticas quanto as pós-sinápticas estiverem excitadas.

2. 6. 2. 1. Perceptrons

Em 1957, Rosemblatt mostra o modelo dos Perceptrons. Nele, os elementos processadores (neurônios) estão organizados em duas camadas (uma para entrada e outra para saída) totalmente conectadas entre si (Figura 2.16). Estas conexões eram direcionadas da camada de entrada para a de saída. O grupo de Rosemblatt enfocava o problema de como achar os pesos apropriados das conexões entre cada elemento processador (eficiência sináptica estabelecida entre o neurônio pré-sináptico e o pós-sináptico). Ele conseguiu criar um algoritmo que alterava estes pesos de modo a realizar a computação desejada, que foi denominado algoritmo de aprendizado do Perceptron.

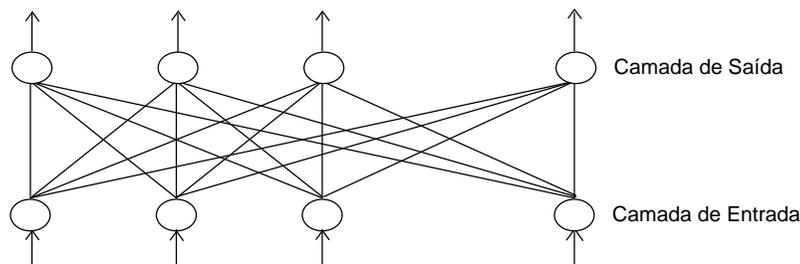


Figura 2. 16. Perceptron de Roseblatt.

Cada um dos elementos processadores possui uma função de limiar de ativação, e calcula sua saída a partir de uma soma ponderada de suas entradas.

Os Perceptrons funcionam como classificadores de padrões binários. A atualização dos pesos da única camada ajustável de conexões pode ser feita por diferentes regras de aprendizado, porém todas utilizando o mesmo princípio. Quando a saída desejada for maior do que a saída real, a taxa de aprendizado é adicionada ao peso, caso contrário, é subtraída. Se a saída real for igual à desejada, não há mudanças no peso.

2. 6. 2. 2. Adaline

Em 1959/60, Widrow & Hoff criam o Adaline, composto de elementos processadores "adaline" (acrônimo de Adaptive linear neuron), similar aos Perceptrons. Logo em seguida criam também o Madaline (Multilayer adaline), que possui uma arquitetura onde dois ou mais elementos processadores tipo "adaline" são utilizados como entrada e a tomada de decisão pode ser realizada por uma função AND, OR ou a de maior incidência entre os "adalines", criando a primeira arquitetura de 3 (três) camadas (Figura 2.17).

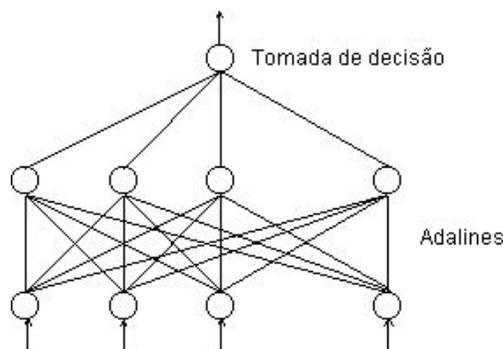


Figura 2. 17. Madaline de Widrow & Hoff.

O modelo Adaline é similar ao Perceptron. Os elementos processadores

também são binários, porém, variam em $\{-1,1\}$.

Em 1966, Minsky publica seu livro "Perceptrons", divulgando as limitações dos Perceptrons. Cada unidade de saída só poderia classificar entradas linearmente separáveis e não conseguiria, nem mesmo, aprender a resolver problemas simples como a função XOR (ou-exclusivo). Poucos pesquisadores - como Werbos, Anderson e Grossberg - continuaram as pesquisas na área, porém, sem o interesse outrora legado pelos demais pesquisadores da comunidade científica internacional.

2. 6. 2. 3. Hopfield

Em 1982, a partir da publicação dos trabalhos de Hopfield , que aplicava Redes Neurais em otimização, o interesse na área tornou a crescer. Ele apresentou um novo modelo de rede no qual todos os elementos processadores estavam totalmente interconectados entre si.

O modelo desenvolvido por Hopfield era essencialmente distinto de seus predecessores. Este tipo de topologia tornava a rede recursiva pois a saída de cada elemento servia como entrada para todas as demais unidades.

As Redes de Hopfield funcionam como uma memória associativa, onde cada padrão a ser memorizado deve ser escolhido a priori e associado a um mínimo da função de energia da rede. Porém, na prática, o efeito dos mínimos locais faz com que o padrão retornado não necessariamente seja o mesmo do padrão desejado e sim outro qualquer (correspondente a um mínimo local). Um outro efeito colateral ainda pior é que este mínimo local pode não corresponder a nenhum padrão desejado de memória.

2. 6. 2. 4. Back-Propagation

Em 1986, as pesquisas retomadas culminam quando Rumelhart, McClelland e o "PDP Research Group" redescobrem as Redes Neurais e publicam dois volumes do mais utilizado livro no assunto: Parallel Distributed Processing (Rumelhart). Nele, foi divulgado por Rumelhart um modelo desenvolvido por Werbos em 1974, e reinventado independentemente por Parker em 1982, denominado Back-Propagation. Apesar de parecer, em muito, com seu antecessor Perceptrons, difere deste por permitir a utilização de mais de duas camadas de elementos processadores (Figura 2.18) sendo, por isso, conhecido também como Perceptron multi-camadas. Ou seja, além das camadas de entrada e saída, podem ser acrescentadas uma ou mais camadas intermediárias (ou escondidas).

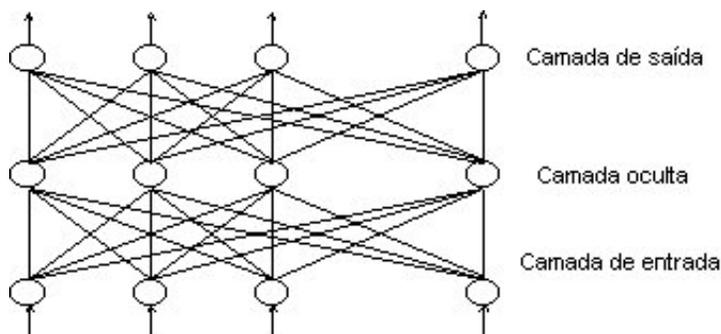


Figura 2. 18. Perceptron multi-camadas.

Neste modelo, os elementos processadores estão organizados em camadas, estando cada elemento de uma camada, totalmente conectado à camada subsequente. O padrão apresentado é propagado até a camada de saída (forward-propagation), onde é calculado o erro entre a saída real e a saída desejada de cada elemento processador da camada de saída.

A utilização deste algoritmo, denominado Backward-Error-Propagation, permitiu treinar redes neurais que possuem camadas intermediárias, resolvendo a

principal limitação dos Perceptrons. Este é, sem dúvida, o paradigma de Redes Neurais mais utilizado em todo o mundo e tem sido aplicado com sucesso nas mais diversas áreas.

Desde então, a maior parte das pesquisas na área tem se dedicado a utilizar variações dos modelos existentes em diversas aplicações. Esta tarefa foi facilitada com o advento de ferramentas (principalmente simuladores) para a experimentação destes modelos. Dentre os principais modelos que surgiram neste período podemos destacar o Brain-State-in-a-Box (BSB) de Anderson, o Self-Organization-Map (SOM) de Kohonen e a Adaptive Resonance Theory (ART) de Grossberg.

2. 6. 3. Diferenças entre Redes Neurais Artificiais e Biológicas

Talvez, a fascinação exercida pelas Redes Neurais deva-se, em parte, a seu relacionamento com o cérebro. É comum ouvir dizer que as Redes Neurais Artificiais explicam os mecanismos básicos e a dinâmica do cérebro humano. Porém, apesar do termo redes neurais ser proveniente das redes de células nervosas do cérebro, os presentes modelos diferem drasticamente de seus similares biológicos.

Com o propósito de uma melhor compreensão das funções cognitivas do cérebro, os pesquisadores das Redes Neurais Biológicas (experimentalistas) têm acumulado um grande número de informações sobre as propriedades moleculares e celulares dos neurônios e de circuitos neurais do cérebro.

Por outro lado, os pesquisadores de modelos de Redes Neurais Artificiais (modelistas) têm se esmerado na análise das propriedades decorrentes do elevado grau de paralelismo e da distribuição do processamento em elementos processadores simples (Rumelhart, 1986). McCulloch e Pitts mostraram que tudo que

pode ser computado em uma máquina de Turing pode ser computado com o auxílio de seu modelo formal de neurônio.

Infelizmente, os modelistas tradicionalmente ignoram a maior parte das informações obtidas pelos neurocientistas, substituindo a complexa estrutura do neurônio por um simples elemento processador de somas e as singulares conexões dos dendritos e axônios por uma matriz de conexões, enquanto estes (por este mesmo motivo) subestimam os resultados obtidos pelos modelistas.

Existe, portanto, uma necessidade crítica de se reduzir a distância entre os experimentalistas e os modelistas, de modo a possibilitar a criação de redes neurais mais realistas e capazes de simular um maior número de funções cognitivas.

O cérebro é um computador extremamente sofisticado e um dos seus principais desafios é entender como ele realiza o processamento de informações. De modo a transpor a distância existente entre as teorias computacionais e os dados biológicos é necessário primeiro entender como computações elementares podem ser realizadas no "hardware" do cérebro.

Os modelos de Redes Neurais consideram a existência de apenas uma conexão entre cada dois elementos processadores, o que equivale a um modelo biológico representando somente o que ocorre entre o axônio de um neurônio e o soma de outro. Os dendritos estão normalmente compondo uma estrutura denominada árvore dendrítica devido a suas ramificações semelhantes a arbustos

O processo de divergência do sinal elétrico que é enviado do axônio de um neurônio para os dendritos de um outro neurônio é da ordem de apenas 10% a 20%, e não de 100% (total), como nos modelos artificiais.

O processo de convergência de diversos sinais recebidos de neurônios distintos por um único neurônio não é somente temporal, mas também é espacial,

dependendo da topologia das sinapses. Muitos neurônios realizam uma soma não-linear de suas entradas, podendo haver, inclusive, processamento lógico (AND,OR,NOT) dentro da árvore dendrítica.

As árvores dendríticas não são lisas, ao contrário, apresentam pequenas estruturas da ordem de 1 a 2 micra, denominadas spines. Ao que tudo indica os spines são responsáveis por grande parte das sinapses e, conseqüentemente, pelo processamento de informações. Atualmente, acredita-se que os spines são responsáveis pelas memórias de longo-prazo (LTM) e pela plasticidade das eficiências sinápticas.

Os estímulos externos regulam a quantidade de neurotransmissores nas sinapses, possibilitando transformar eventos externos em informações neurais. Diferentes transmissores agrupados em diferentes combinações são regulados em diferentes instantes de tempo em diferentes sinapses.

Todo modelo de Rede Neural baseia-se na premissa de que o neurônio é a unidade básica de processamento da informação.

Um dos princípios gerais da biologia é o de que o comportamento de um organismo depende de níveis hierárquicos de organização. Podemos começar pela própria sinapse onde mecanismos moleculares realizam a comunicação. As sinapses formam pequenos circuitos envolvendo algumas formas de interação entre as sinapses, denominados micro-circuitos. Estes normalmente encontram-se nos dendritos e seu comportamento é determinado pela distribuição na árvore dendrítica. O próprio neurônio representa um outro nível de organização, e as interações entre neurônios de propriedades similares ou distintas formam circuitos locais. Estes interligam-se em circuitos interregionais, envolvendo múltiplas regiões em diferentes partes do cérebro.

O futuro da neuro-computação poderá ser altamente beneficiado pelas pesquisas biológicas. Estruturas encontradas nos sistemas biológicos podem, e devem, inspirar novas arquiteturas para modelos de Redes Neurais. Do mesmo modo que o desenvolvimento de Redes Neurais pode ser beneficiado pelos estudos biológicos, alguns modelos ou aplicações podem ilustrar e auxiliar a explicar diversas características dos sistemas biológicos. Atualmente, a anatomia e fisiologia das redes existentes em regiões específicas do cérebro é ainda imprecisa.

Apesar de todas as divergências, experimentalistas e modelistas têm procurado reduzir suas mútuas restrições, aumentando o nível de colaboração através de foros internacionais, participações em congressos e trabalhos conjuntos.

CAPÍTULO 3

METODOLOGIA

3. 1. Considerações gerais

O processo de classificação supervisionada de imagens proposto neste trabalho está esquematizado através da figura 3.1. Numa etapa anterior às descritas no diagrama de blocos, é necessária a seleção de amostras representativas das classes texturais dominantes na imagem. Tais amostras são constituídas de pequenas janelas retiradas das texturas de maior ocorrência da imagem. Aplica-se a transformada de Fourier a estas amostras e selecionam-se as freqüências representativas das classes, através da análise do espectro de energia no domínio de freqüências. Com base nestas freqüências, determinam-se os parâmetros necessários para a constituição dos filtros de Gabor para o processo de filtragem da imagem original.

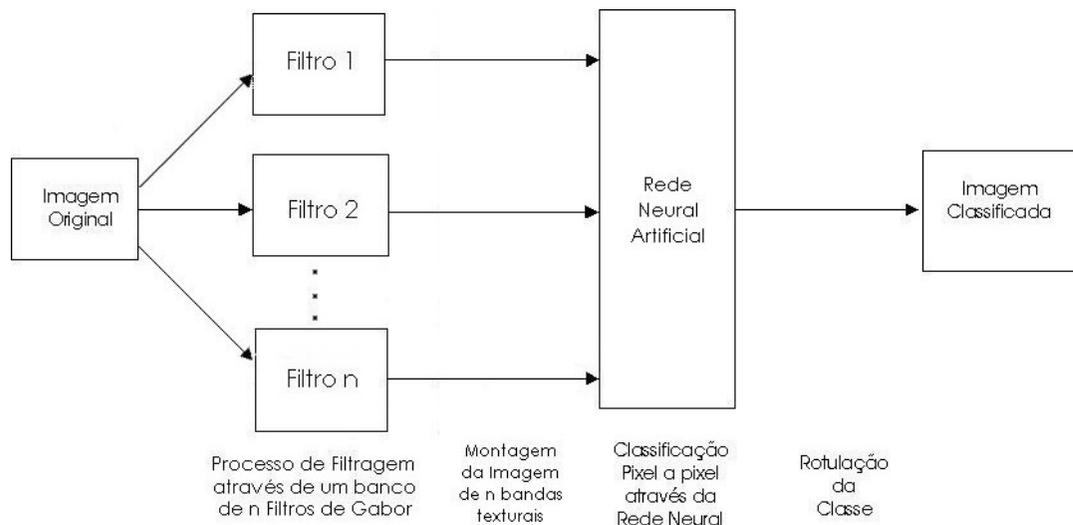


Figura 3. 1. Diagrama de blocos do processo de filtragem e classificação supervisionada de imagens proposto no presente trabalho.

O processo de filtragem é realizado pela convolução da imagem original com cada filtro, gerando um número de imagens filtradas igual ao número de filtros existentes. Tal processo é realizado por uma janela que faz a varredura da imagem. As dimensões da janela são específicas de cada filtro. Cada imagem filtrada identifica uma determinada frequência espacial, podendo ser designada como uma banda textural. Dessa forma, é montada uma imagem formada pelas "bandas texturais" obtidas na filtragem.

A classificação da imagem é realizada pela rede neural artificial num processo pixel a pixel. Para tanto, fez-se necessário o treinamento da RNA com dados retirados de amostras de cada textura da imagem original. A rotulação feita pela rede neural, e baseada nas texturas, gera a imagem classificada.

A metodologia desenvolvida nesta dissertação está baseada fundamentalmente nos trabalhos de Jain e Farroukhnia (1991), Raghu et al (1995) e Angelo (2000).

Jain e Farroukhnia (1991) propõem um algoritmo inspirado na técnica de filtragem multicanal para a segmentação de textura que usa um banco de filtros de Gabor para caracterizar os canais texturais. Para tal, utilizam um esquema sistemático de seleção de filtro baseado na reconstrução da imagem de entrada através das imagens filtradas. Raghu et al (1995) apresentam uma estrutura de rede neural de dois estágios, combinando as características de um mapa auto-organizável (SOM) a um perceptron multi-camadas (MLP). Numa etapa anterior, utilizam uma técnica de filtragem multicanal para a extração de características texturais. Os canais são constituídos por um banco de filtros de Gabor com diferentes dimensões, orientações e frequências. Angelo (2000) utilizou a técnica de filtragem multicanal, através de filtros de Gabor, juntamente com o método de Máxima Verossimilhança

Gaussiana, para propor um método de classificação supervisionado em imagens digitais com base em atributos de textura.

A aplicação de técnicas de filtragem multicanal para extração de características texturais vem sendo amplamente utilizada na classificação de imagens digitais provenientes de Sensoriamento Remoto. Da mesma forma, as Redes Neurais têm se mostrado de grande utilidade no processo de classificação destas imagens.

Este trabalho propõe a implementação da técnica de filtragem multicanal, através de um banco de filtros de Gabor para a extração de características texturais, combinada a um perceptron multi-camadas com algoritmo Backpropagation, com a função de classificação das imagens digitais de Sensoriamento Remoto.

3. 2. Textura

Atributo é o nome dado a cada uma das várias categorias de características associadas a determinado objeto, ou à mesma classe de cobertura do solo, capaz de ser distinguida das demais (Claro, 1995). Uma imagem proveniente de sensoriamento remoto pode ser descrita com base nos seus atributos, que podem ser agrupados em três categorias: espectrais, espaciais e de contexto. Os atributos espectrais são quantificados pelo comportamento espectral dos vários objetos ou alvos que compõem a cena, conforme registrado pelo sistema sensor (Swain e Davis, 1978). O comportamento espectral de alvos terrestres é descrito pela habilidade dos mesmos em refletir a energia solar incidente em regiões selecionadas do espectro eletromagnético. Tal atributo é o mais utilizado na classificação de imagens devido à fácil quantificação. A figura 3.1 mostra as respostas espectrais típicas da água, vegetação e solo (Richards, 1986).

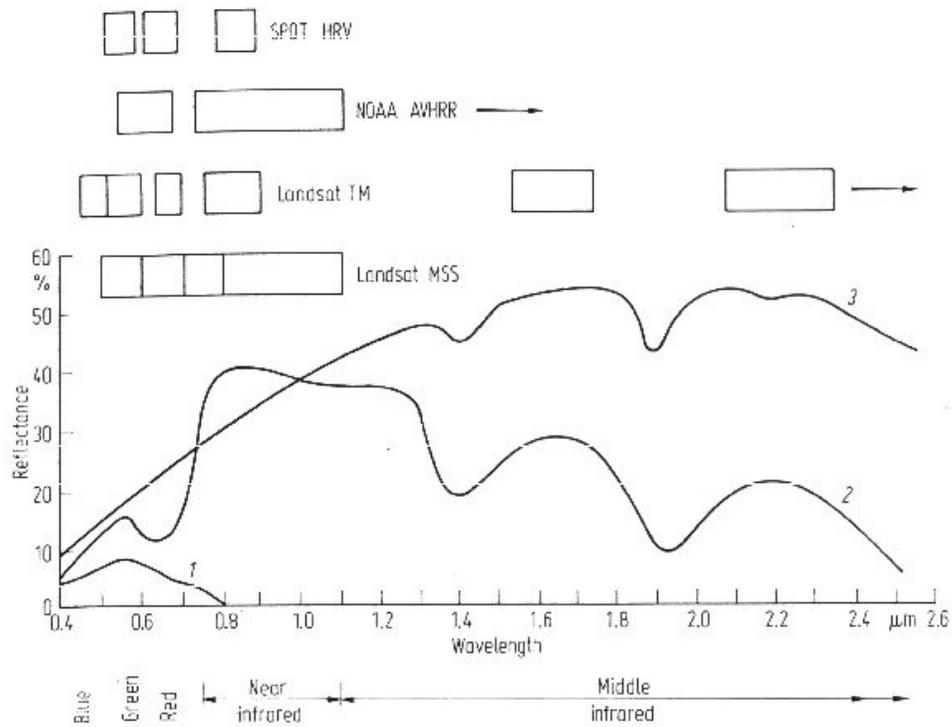


Figura 3. 2. Respostas espectrais típicas da água (1), vegetação (2) e solo (3). Também estão indicadas as posições das bandas espectrais para alguns sensores remotos. (Richards, 1986)

Por outro lado, os atributos espaciais estão relacionados com parâmetros como forma, dimensão e textura (Swain e Davis, 1978). A textura é o atributo espacial que tem sido mais estudado em sensoriamento remoto, e vem sendo utilizada como elemento auxiliar no processo de classificação, especialmente na distinção entre duas classes com comportamento espectral similar, podendo ser descrita, de forma qualitativa, como “uniforme”, “grosseira”, “regular”, “fina” ou “média” (Claro, 1995).

Dougherty (1992) define textura como uma variação tonal repetitiva e organizada em uma pequena região da imagem. No entanto, a caracterização que Dougherty constrói sobre a textura não é suficiente para que ela possa ser quantificada. Para tal, a imagem deve sofrer algum processo que permita medir ou estimar algum

parâmetro sobre a mesma. Nesta pesquisa, este processo é realizado pela utilização de um banco de Filtros de Gabor.

3. 3. Extração de feições texturais

No processo de classificação de uma imagem, utilizando atributos de textura, é necessário levar em consideração algumas variáveis de aspecto relevante, tais como a orientação, a freqüência e a dimensão do elemento de textura.

Estudos em sistemas de visão biológicos mostram que a habilidade para reconhecer objetos no mundo real depende parcialmente da percepção de consistência e diferença textural. Representações por meio de transformadas de Fourier, como se sabe, são completas, ou seja, durante a sua transformação nenhuma informação da imagem original é perdida. No entanto, a transformada de Fourier não deixa explícito as informações necessárias para uma rápida discriminação de diferentes regiões de textura, pois apenas apresenta as freqüências espaciais presentes na imagem original, não identificando sua posição na mesma. Desta forma, a representação de uma imagem pela transformada de Fourier não mostra a variação na distribuição do brilho ao longo de uma região na imagem. A análise por técnicas de Fourier não revela distribuições locais de brilho ou combinações no brilho da imagem, não sendo, então, apropriadas para a análise de feições locais que caracterizam texturas. Desta forma, torna-se necessária a utilização de outro tipo de filtro que obtenha um ótimo desempenho simultaneamente no domínio espacial e no domínio de freqüências. Através da utilização da desigualdade de Schwarz, Gabor(1946) demonstrou, de forma análoga à relação de incerteza de Heisenberg,

que a caracterização de um sinal, simultaneamente no domínio temporal e no domínio de frequências espaciais, apresenta um limite inferior definido pelo produto de sua extensão pela sua duração.

$$\Delta t \cdot \Delta w \geq \frac{1}{4\pi} \quad (3.1)$$

Através dos trabalhos de Heisenberg e Weyl, Gabor derivou uma relação de incerteza, encontrando uma família de sinais que alcançavam o limite teórico inferior de incerteza conjunta nos domínios temporal e das frequências. Essa função tornou-se conhecida como Função de Gabor:

$$\mathbf{g}(\mathbf{x}, \sigma, \mathbf{k}) = \exp\left\{-\frac{1}{2}\left(\frac{\mathbf{x}}{\sigma}\right)^2 + 2\pi i \mathbf{k} \mathbf{x}\right\} \quad (3.2)$$

onde σ é a extensão do envelope Gaussiano que limita espacialmente o filtro de Gabor, e k é a frequência da onda senoidal associada ao filtro, segundo Mehrotra et al (1992).

A figura (3.3) mostra as partes real e imaginária deste filtro unidimensional.

O trabalho de Gabor foi complementado por Daugman, em 1988, para problemas envolvendo duas dimensões. A característica mais importante da função de Gabor é a de simular com grande exatidão o comportamento dos perfis de campos receptivos encontrados nas células do córtex cerebral de mamíferos, as quais são as responsáveis pela captura de propriedades de localização espacial, orientação, frequência espacial, etc.

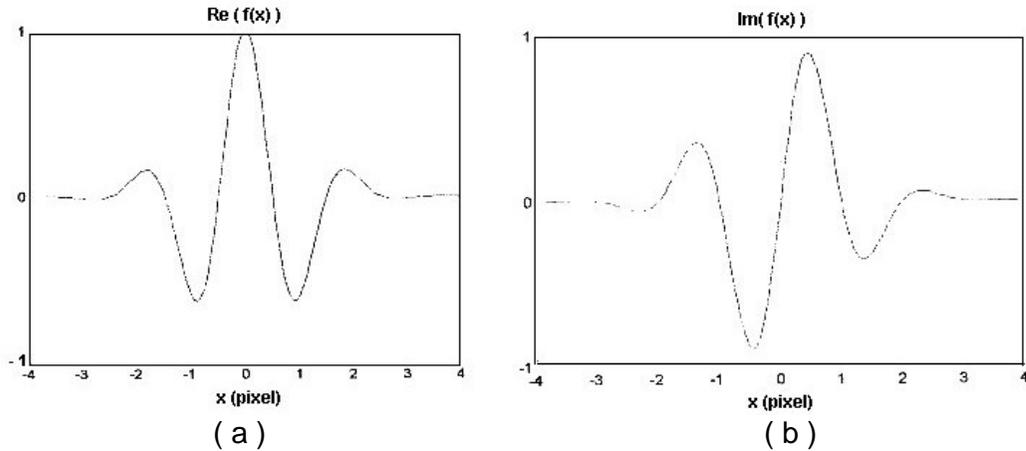


Figura 3. 3. Componente real do filtro de Gabor (a) e componente imaginária do filtro de Gabor (b), ambas para valores de $\sigma = 1s$ e $k = 0.5$ ciclos/pixel. Adaptado de Angelo (2000).

Os filtros de Gabor bidimensionais são definidos por:

$$g(x, y, \sigma_x, \sigma_y, k_x, k_y) = \exp \left\{ -\frac{1}{2} \left[\left(\frac{x}{\sigma_x} \right)^2 + \left(\frac{y}{\sigma_y} \right)^2 \right] + 2\pi i (k_x x + k_y y) \right\} \quad (3.3)$$

Uma propriedade essencial da abordagem bidimensional dos filtros de Gabor é o fato desses filtros alcançarem o limite teórico inferior de incerteza conjunta nos domínios espacial e das freqüências, que é de $1/16\pi^2$ (Daugman, 1988). Para as variáveis (x,y) , centradas em x_0 e y_0 no domínio espacial, e (u,v) centradas em u_0 e v_0 no domínio das freqüências, obtém-se a partir do princípio de incerteza conjunta que

$$(\Delta x) \cdot (\Delta y) \cdot (\Delta u) \cdot (\Delta v) \geq \frac{1}{16\pi^2} \quad (3.4)$$

Tal princípio afirma que, ao se tomar uma área de filtragem no domínio espacial e a aumentarmos, a respectiva área de filtragem no domínio de freqüências será diminuída, de forma que o produto de suas áreas nos dois domínios é limitado por um valor, conhecido por limite inferior de incerteza conjunta. É possível mostrar

que para os filtros de Gabor este produto é exatamente igual ao limite mínimo

$\frac{1}{16\pi^2}$. Os valores de (Δx) , (Δy) , (Δu) e (Δv) são dados por:

$$(\Delta x) = \left[\frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - x_o)^2 f(x, y) \cdot f^*(x, y) dx dy}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) f^*(x, y) dx dy} \right]^{\frac{1}{2}} \quad (3.5)$$

$$(\Delta u) = \left[\frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (u - u_o)^2 F(u, v) \cdot F^*(u, v) du dv}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(u, v) F^*(u, v) du dv} \right]^{\frac{1}{2}} \quad (3.6)$$

$$(\Delta y) = \left[\frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (y - y_o)^2 f(x, y) \cdot f^*(x, y) dx dy}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) f^*(x, y) dx dy} \right]^{\frac{1}{2}} \quad (3.7)$$

$$(\Delta v) = \left[\frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (v - v_o)^2 F(u, v) \cdot F^*(u, v) du dv}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(u, v) F^*(u, v) du dv} \right]^{\frac{1}{2}} \quad (3.8)$$

sendo $f^*(x, y)$ e $F^*(u, v)$, o complexo conjugado de $f(x, y)$ e $F(u, v)$, respectivamente, de tal forma que o produto $f(x, y) \cdot f^*(x, y)$ e $F(u, v) \cdot F^*(u, v)$ corresponde à distribuição de energia nos respectivos domínios. Além disso, o produto de $(\Delta x) \cdot (\Delta y)$ determina a área ocupada pelo filtro no domínio espacial e $(\Delta u) \cdot (\Delta v)$ determina a área ocupada pelo filtro no domínio das frequências.

Substituindo as equações (3.5), (3.6), (3.7) e (3.8) na expressão 3.3 que define o filtro bidimensional de Gabor, obtém-se

$$(\Delta x) \cdot (\Delta y) = \frac{\sigma_x \sigma_y}{2} \quad (3.9)$$

e

$$(\Delta u) \cdot (\Delta v) = \frac{1}{8\pi^2 \sigma_x \sigma_y} \quad (3.10)$$

logo,

$$(\Delta x).(\Delta y).(\Delta u).(\Delta v) = \frac{1}{16\pi^2} \quad (3.11)$$

igual ao valor mínimo de incerteza conjunta.

A extração de textura, escala e eixos de modulação podem ser efetuadas sem prejudicar a extração de informações sobre localização espacial e relações métricas, tomando por base a minimalidade do limite de incerteza conjunta para os filtros de Gabor (Angelo, 2000).

3. 4. Filtragem através de Filtros de Gabor

Os filtros de Gabor são um conjunto de funções senoidais complexas, bidimensionais, moduladas por uma função Gaussiana também bidimensional e que apresentam certas propriedades extremamente úteis para fins de classificação de imagens. Neste processo, são criados bancos de filtros de gabor com a finalidade de extrair as feições texturais presentes numa imagem. Daugman(1988) mostra que os filtros de Gabor são adequados à análise textural, no sentido em que minimizam a incerteza conjunta, nos domínios espacial e no domínio de freqüências espaciais, permitindo, desta forma, selecionar feições cujas características ficam determinadas por freqüências espaciais diretamente no domínio espacial, sem a necessidade de desenvolver operações no domínio de Fourier ou das freqüências.

A estrutura local da textura, nos vários canais utilizados, é descrita pela freqüência e pela orientação, que são definidas pela função senoidal no filtro de Gabor. O envelope Gaussiano define as resoluções espaciais com que as estruturas de textura são caracterizadas. A forma geral da Função de Gabor bidimensional é dada através da equação:

$$g(x, y, \sigma_x, \sigma_y, k_x, k_y) = A \cdot \exp \left\{ -\frac{1}{2} \left[\left(\frac{x}{\sigma_x} \right)^2 + \left(\frac{y}{\sigma_y} \right)^2 \right] + 2\pi i (k_x x + k_y y) \right\} \quad (3.12)$$

onde A é um fator de escala e (σ_x, σ_y) definem a extensão espacial da função de Gabor. A orientação da grade senoidal é fornecida por $\tan^{-1} \left(\frac{k_y}{k_x} \right)$. A frequência da onda senoidal é dada por $k = (k_x, k_y)$, onde k_x, k_y são suas componentes ao longo dos eixos x e y, respectivamente. O módulo de k é dado por $|k| = \sqrt{k_x^2 + k_y^2}$. Tomando-se as partes reais e imaginárias da equação 3.12, tem-se as componentes par e ímpar da função de Gabor:

$$\text{Re}[g(x, y, \sigma_x, \sigma_y, k_x, k_y)] = \exp \left\{ -\frac{1}{2} \left[\left(\frac{x}{\sigma_x} \right)^2 + \left(\frac{y}{\sigma_y} \right)^2 \right] \right\} \cos(2\pi(k_x x + k_y y)) \quad (3.13)$$

$$\text{Im}[g(x, y, \sigma_x, \sigma_y, k_x, k_y)] = \exp \left\{ -\frac{1}{2} \left[\left(\frac{x}{\sigma_x} \right)^2 + \left(\frac{y}{\sigma_y} \right)^2 \right] \right\} \text{sen}(2\pi(k_x x + k_y y)) \quad (3.14)$$

Um típico filtro de Gabor é ilustrado na figura 3. 4:

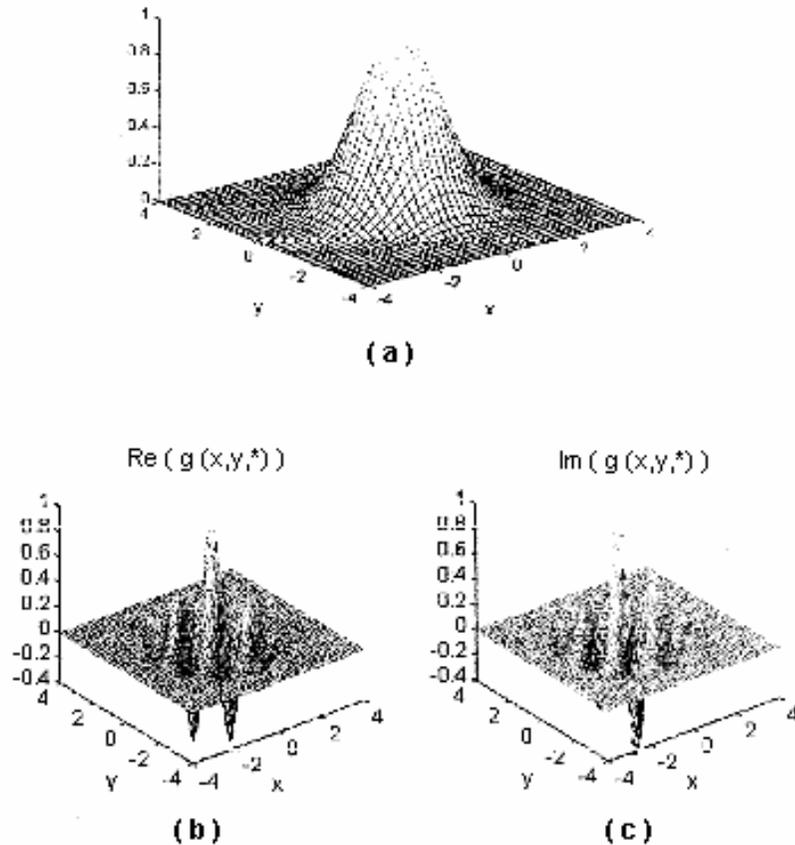


Figura 3. 4. Filtro de Gabor bidimensional apresentado em valores absolutos (a). Componente real do filtro de Gabor (b). Componente imaginária do filtro de Gabor, ambos para valores $\sigma = \sigma_x = \sigma_y = 4$, $k_x = -0.5$ e $k_y = 0.5$ ciclos/pixel. O valor * em parênteses corresponde a $(\sigma_x, \sigma_y, k_x, k_y)$. (Angelo, 2000)

A transformada de Fourier da função de Gabor (equação 3.12) é dada por:

$$G(u, v, \sigma_x, \sigma_y, k_x, k_y) = \sigma_x \sigma_y \exp\left\{-\frac{1}{2}\left[(u - k_x)^2 \sigma_x^2 + (v - k_y)^2 \sigma_y^2\right]\right\} \quad (3.15)$$

No domínio das freqüências, o filtro de Gabor pode ser obtido aplicando-se a transformada de Fourier à equação 3.13.

$$G(u, v) = \sigma_x \sigma_y \left[\exp\left\{-\frac{\sigma_x^2}{2}\left[(u - k_x)^2 + (v - k_y)^2\right]\right\} + \exp\left\{-\frac{\sigma_y^2}{2}\left[(u + k_x)^2 + (v + k_y)^2\right]\right\} \right] \quad (3.16)$$

É possível observar que, no domínio das freqüências espaciais, a função 3.16, através das propriedades da transformada de Fourier, torna-se duas funções

Gaussianas centradas em $-(k_x, k_y)$ e $+(k_x, k_y)$, respectivamente (figura 3.4). Pelo princípio de incerteza citado no item anterior, tem-se que as áreas efetivamente ocupadas pelo filtro definido em (3.3), podem ser obtidas a partir das equações (3.5), (3.6), (3.7) e (3.8).

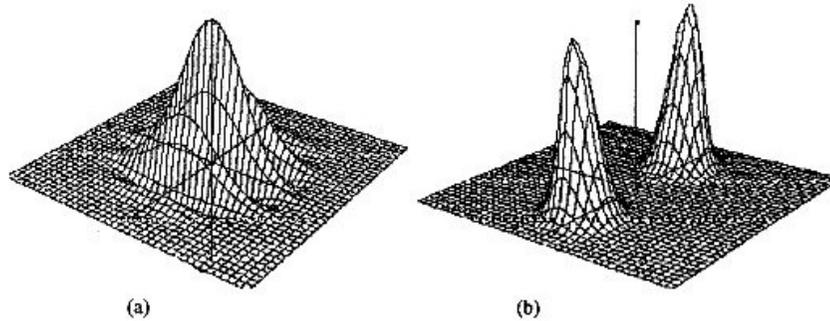


Figura 3. 5. Componente real do filtro de Gabor bidimensional no domínio espacial (a); resposta de freqüência do filtro definido em a (b). Adaptado de Angelo (2000).

$$\text{Área no domínio espacial} = (\Delta x) \cdot (\Delta y) = \frac{\sigma_x \sigma_y}{2} \quad (3.17)$$

e

$$\text{Área no domínio das freqüências} = (\Delta u) \cdot (\Delta v) = \frac{1}{8\pi^2 \sigma_x \sigma_y} \quad (3.18)$$

Desta forma, obtém-se $(\Delta x) \cdot (\Delta y) (\Delta u) \cdot (\Delta v) = \frac{1}{16\pi^2}$, limite mínimo de incerteza conjunta.

A imagem processada por um filtro de Gabor é obtida através da convolução da imagem com a função de Gabor, que pode ser implementada tanto no domínio das freqüências como no domínio espacial, usando-se para isso (3.15) ou (3.12), respectivamente (Raghu et al, 1995). Após o processo de convolução obtém-se uma imagem filtrada formada por valores imaginários. A imagem assim formada, por conter uma componente imaginária, não pode ser diretamente representada. Segundo Angelo, uma alternativa consiste na representação através de valo-

res absolutos ou o quadrado deste, conhecido como energia. A energia de cada pixel é dada por:

$$f(x,y) = |I(x,y) * g(x,y)|^2 \quad (3.19)$$

onde * significa convolução, $I(x,y)$ representa a imagem original e $g(x,y)$ o filtro de Gabor, tomado com os parâmetros k_x , k_y , σ_x , σ_y , previamente escolhidos.

3. 4. 1. Seleção de parâmetros dos Filtros de Gabor

A determinação dos parâmetros essenciais para a aplicação dos filtros de Gabor resume-se à escolha da freqüência espacial do filtro e à seleção da dimensão e extensão espacial. Os processos de determinação de tais parâmetros são propostas por Jain e Farroukhnia (1991) e Raghu et al(1995), e foram experimentados por Angelo (2000). Alguns desses experimentos foram reproduzidos e inseridos no próximo capítulo.

3. 4. 1. 1. Determinação da freqüência espacial

Somente um número muito grande de freqüências espaciais pode caracterizar, de forma rigorosa, uma textura. Porém, é inviável o emprego de todas as freqüências espaciais presentes devido ao enorme custo computacional que representa.

Angelo (2000) optou por representar cada textura por um número reduzido de freqüências espaciais que melhor preservassem as características da textura original, mantendo o nível de informação necessário à caracterização das diferentes regiões da imagem, através de dois critérios principais:

1º) mais alto nível de energia da frequência espacial e

2º) orientações destas frequências.

Porém, a existência do segundo critério faz-se necessária apenas nos casos em que as frequências obtidas através do primeiro critério para duas texturas distintas sejam muito próximas entre si, onde o poder discriminatório ficaria bastante reduzido. Desta forma, o uso de certas frequências espaciais presentes em mais de uma classe de textura apenas aumenta o custo computacional do processo de classificação, contribuindo muito pouco para o poder discriminatório do método. Assim, estas frequências devem ser eliminadas, mantendo-se apenas aquelas que são características de somente uma classe.

3. 4 .1. 2. Dimensão e extensão espacial

Numa imagem, o valor do contador digital e a localização espacial dos pixels são considerados como feições de 1ª ordem. Segundo Raghu et al (1995), estas feições não estão relacionadas com o contexto. As feições de ordem mais alta estimam as características ou propriedades ao longo de uma região. Portanto, em análise textural, as feições de 1ª ordem não tem um papel preponderante, pois a textura é uma propriedade de contexto, só podendo ser interpretada através da análise conjunta de um grupo de pixels numa determinada região da imagem, cujo grupo de pixels que carrega uma determinada propriedade ou estatística local e que se repete ao longo da região textural caracteriza o TEXEL (Angelo, 2000).

A identificação e a extração de algumas feições de textura capazes de caracterizar eficientemente as diferentes classes presentes na imagem torna-se indispensável para que se tenha uma boa análise de textura. É razoável supor que cada

classe presente na imagem apresenta dimensões de TEXEL próprias, surgindo, dessa forma, a necessidade de se encontrar uma excelente resolução de TEXEL, cujas dimensões sejam suficientemente grandes a fim de capturar as características de cada uma das texturas envolvidas.

Em seu trabalho, Angelo (2000) analisa a escolha, para cada amostra de classe, das freqüências espaciais mais representativas de suas características texturais, utilizando a Transformada de Fourier, bem como a determinação dos valores para dim_x e dim_y que definem as dimensões do filtro de Gabor, de forma a abranger, no mínimo, um período completo da freqüência espacial implementada no filtro de Gabor. Salientando o fato de o período ser igual ao inverso da freqüência, os valores de dim_x e dim_y , correspondentes às dimensões do filtro de Gabor, podem ser tomados como o inverso da freqüência que descreve o padrão que está sendo analisado:

$$dim_x = \frac{1}{k_x} \text{ e } dim_y = \frac{1}{k_y} \quad (3.20)$$

Porém, podem acontecer freqüências muito baixas ou muito altas. Tais freqüências podem gerar filtros degenerados quando uma das dimensões x ou y deste filtro tender ao infinito, o que ocorre quando uma das freqüências tender a 0 (zero). Outro filtro degenerado ocorre quando apenas um único pixel é tomado para uma de suas dimensões. Isto ocorre quando uma das freqüências é igual a $\frac{N-1}{N}$ (onde N representa a dimensão da amostra tomada, tanto ao longo do eixo x quanto do eixo y). O maior problema reside nos filtros de dimensões infinitas.

Para contornar este problema, Angelo7 (2000) optou pela adoção de um único valor para a dimensão do filtro, isto é

$$Dim = dim_x = dim_y$$

tal que

$$Dim = \max\left(\frac{1}{k_x}, \frac{1}{k_y}\right) \quad (3.21)$$

No caso de filtros em que uma das freqüências envolvidas k_x ou k_y tende ao valor 0 (zero), e que gera no filtro uma dimensão tendendo ao infinito, atribui-se como dimensão correspondente a esta freqüência no filtro, a maior dimensão da amostra (obtida da imagem) que gerou esta freqüência. Desta forma, fica assegurado que a dimensão do filtro é suficiente para capturar a freqüência em questão.

A equação 3.21 gera um filtro com dimensões constantes o que não afeta a seleção de feições texturais, uma vez que as dimensões do filtro serão sempre maiores que a dimensão mínima necessária para caracterizar as freqüências utilizadas. Visto que o filtro de Gabor é ponderado por um envelope Gaussiano e que a forma da função que define este envelope é definida pelo parâmetro σ , denominado neste contexto de “extensão espacial” do filtro, é aconselhável para cobrir 99,73% da dimensão do filtro, utilizar-se de uma extensão espacial tal que

$$\sigma = \frac{Dim}{6} \quad (3.22).$$

No entanto, pode haver um problema neste processo. Considere dois filtros distintos projetados para freqüências distintas e, portanto, com extensões espaciais diferentes. Pode acontecer que no processo de filtragem da imagem, o filtro com extensão espacial menor (determinado para capturar uma freqüência mais alta) capture freqüências mais baixas em outra região textural, de forma parcial e incompleta. Desta forma, podem ser geradas informações errôneas que irão degradar o resultado do classificador.

Angelo (2000) adotou o seguinte procedimento para evitar este problema: toma-se como extensão espacial para todos os filtros de Gabor, um valor único e igual à máxima extensão espacial:

$$\sigma = \text{máximo}(\sigma_i)$$

onde σ_i = extensão espacial de cada filtro e $i = (1, 2, \dots, \text{número de filtros})$. O aumento da extensão espacial σ no domínio espacial aumenta a área de filtragem e, conseqüentemente, através do princípio de incerteza, reduzir-se-á a área de filtragem no domínio das freqüências, o que ocasiona uma diminuição das freqüências capturadas pelo filtro e dessa forma é minimizado o efeito de sobreposição de freqüência o que melhora a classificação, mas exige um número maior de filtros.

Embora o uso de extensão espacial relacionada a cada par de freqüência espacial (k_x e k_y) de cada filtro também ofereça um bom resultado, na maioria dos casos, o uso de extensão espacial única minimiza as possíveis distorções mencionadas, apresentando dessa forma, um resultado mais homogêneo. Entretanto, para imagens texturais formadas por um número elevado de freqüências espaciais este bom resultado somente ocorre se for utilizado um número muito grande de filtros de Gabor, como foi evidenciado em alguns experimentos de Angelo (2000).

3.5. Classificação de imagens obtidas por sensoriamento remoto com redes neurais

As redes neurais podem ser interpretadas como transformadoras de dados (Pao, 1989), onde o objetivo é associar os elementos de um grupo de dados com os elementos de um segundo conjunto. Quando aplicado à classificação, por exemplo, existe o interesse na transformação de dados do espaço de características

para o espaço de classes. Por pertencerem à mesma classe de técnicas que o reconhecimento de padrões e a regressão linear, e visto a sua importância, as Redes Neurais Artificiais (RNA) têm sido freqüentemente utilizadas no sensoriamento remoto, principalmente por permitir o manuseio de grandes quantidades de dados.

O perceptron multicamadas (MLP), que é uma das RNA mais utilizadas no sensoriamento remoto, pode separar dados que são não linearmente separáveis e, geralmente, consiste de 3 ou mais camadas. Para o início do processo de aprendizado da Rede Neural Artificial faz-se necessária a seleção de um conjunto de amostras das classes de padrões (conjunto de treinamento) a serem aprendidos e as saídas desejadas correspondentes. Para cada classe, devem ser selecionadas amostras representativas e em número suficiente.

A primeira camada é a de entrada, que tem por finalidade repassar os padrões de entrada para os neurônios intermediários, não havendo, em geral, a realização de nenhum cálculo. Nesta camada, os componentes de entrada são apresentados na forma de um vetor de características, constituído de bandas de um conjunto de dados, a textura de uma imagem ou outros parâmetros.

A segunda camada é chamada de camada intermediária ou "oculta". O número de elementos dessa camada é um fator significativo tanto para a acurácia dos resultados da rede como para a convergência do processo de treinamento.

Os pesos, representativos das conexões entre os neurônios, devem ser inicializados com valores aleatórios pequenos, por exemplo, $\pm 0,5$ (Todt, 1998). Também é importante a determinação da função de ativação a ser utilizada na rede. Segundo Todt (1998), a função que apresenta os melhores resultados para o aprendizado é a função tangente hiperbólica, pela rápida convergência demonstrada. No entanto, ao se usar a função tangente hiperbólica, é preciso escalonar as entra-

das da rede para o intervalo $[-1 ; 1]$, antes da propagação, para que as entradas e saídas sejam compatíveis, ocupando o mesmo espaço vetorial. Da mesma forma, se for utilizada a função sigmóide, é necessária o escalonamento das entradas para o intervalo $[0 ; 1]$.

A camada de saída é o terceiro tipo de camada e é responsável pela apresentação dos dados de saída. Na classificação de imagens, por exemplo, o número de elementos na camada de saída pode ser igual ao número de classes existentes.

Cada elemento da rede é interconectado aos elementos em ambas camadas anteriores e posteriores através de conexões associadas aos seus pesos.

3. 5. 1. Rede Neural Multicamadas com BackPropagation

Modelo capaz de resolver problemas de reconhecimento de padrões complexos, apresentado por Rumelhart et al (1986), através de funções que mapeiam valores de entrada da rede com valores de saída desejados. Basicamente, esta rede apresenta três ou mais camadas de neurônios: uma de entrada (input layer), uma ou mais intermediárias (hidden layer) e uma de saída (output layer). A sua topologia é completamente interconectada para frente (feedforward), sem retroalimentação, onde nenhuma conexão passa diretamente por sobre uma outra mais a frente (figura 3.5).

A definição do número de neurônios das camadas de entrada e saída é realizada de acordo com o problema o qual a rede solucionará. O número de neurônios intermediários, ou mesmo o número de camadas intermediárias, é definido de forma iterativa, não havendo, portanto, uma regra que defina a priori o seu número. Se o número de neurônios definido for muito alto, isto pode fazer com que

apenas alguns neurônios especializem-se em características úteis, porém, se o contrário ocorrer e o número de neurônios for insuficiente, pode acontecer da rede não conseguir aprender os padrões desejados.

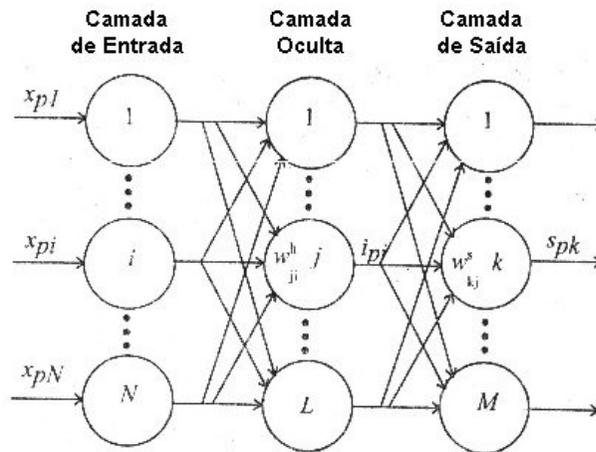


Figura 3. 6. Topologia de uma RNA Multinível com Back-propagation (Lieberman, 1997).

A característica de grande importância deste modelo é o aprendizado supervisionado baseado em duas etapas: propagação e adaptação.

O treinamento (*supervised-learning*) consiste em fornecer à rede um conjunto de estímulos (padrões de entrada) e a saída desejada correspondente, onde o primeiro padrão de entrada é propagado até a saída (fase de propagação). Já na fase de adaptação, um sinal de erro na saída é computado e transmitido de volta para cada neurônio da camada intermediária que contribuiu para a saída obtida. Sendo assim, cada neurônio da camada intermediária recebe somente uma parte do erro total, conforme a contribuição relativa que o neurônio teve na saída gerada. Este processo repete-se, camada por camada, até que cada neurônio da rede receba o seu correspondente. Tal processo recebe o nome *Backpropagation*, pois o aprendizado baseia-se na propagação retrógrada do erro para níveis superiores da rede.

De acordo com o erro recebido pelo neurônio associado, os pesos existen-

tes nas conexões entre os neurônios são atualizados. Esta regra de aprendizado é uma generalização da regra LMS (Least Mean Square Error), também conhecida como Regra Delta (Todt, 1998).

Com as devidas alterações de pesos, o processo de aprendizado permanece até o momento em que a saída obtida pela rede neural, para cada um dos padrões de entrada, seja próxima o bastante da saída desejada, de forma que a diferença entre ambas seja aceitável. Esta diferença é obtida através do cálculo do Erro Médio Quadrado (equação 3.35). Considera-se diferença aceitável aquela que for menor ou igual a um erro previamente estipulado (1% ou 0,5%, em média).

3. 5. 2. Algoritmo de treinamento *Backpropagation*

Seja X_p um vetor de entrada que representa o padrão p ; W_{ji} o peso de conexão entre o neurônio j da camada escondida e o neurônio i da camada de entrada; Y_{pk} a saída desejada e O_{pk} a saída obtida pela rede para o padrão p do neurônio k de saída; é possível descrever o algoritmo de *Backpropagation* através do seguinte processo:

- a) Inicializar os pesos e os parâmetros da rede (coeficiente de aprendizado e parâmetro do momento).
- b) Aplicar o vetor de entrada X_p à rede neural.
- c) Calcular os potenciais de ativação dos neurônios (j) da camada oculta

(h):

$$\text{net}_{pj}^h = \sum_{i=1}^N W_{ji}^h \cdot X_{pi} + \theta_j^h \quad (3.23)$$

onde N é o número de neurônios da camada de entrada e θ_j é um valor de limiar do

neurônio j . θ_j , conhecido por *bias*, tem a função de contribuir na convergência da rede. Na verdade, ele funciona como um peso a mais a ser adicionado ao somatório. Cada neurônio da camada intermediária e da camada de saída possui o seu próprio *bias*.

d) Calcular as saídas dos neurônios da camada escondida:

$$i_{pj} = f(\text{net}_{pj}^h) \quad (3.24)$$

onde f é a aplicação de uma função de ativação. Em geral, utiliza-se a função sigmóide ou a função tangente hiperbólica, como visto anteriormente.

Para a função tangente hiperbólica, os neurônios geram valores de saída dentro do intervalo $[-1;+1]$.

Dessa forma:

$$i_{pj} = \tanh(\text{net}_{pj}^h) \quad (3.25)$$

e) Calcular os potenciais de ativação para os neurônios(k) da camada de saída(o):

$$\text{net}_{pk}^o = \sum_{j=1}^L W_{kj}^o \cdot i_{pj} + \theta_k^o \quad (3.26)$$

onde L é o número de neurônios da camada intermediária e θ_k^o é o valor de limiar do neurônio k .

f) Calcular os valores de ativação dos neurônios de saída de forma semelhante à camada escondida:

$$O_{pk} = f(\text{net}_{pk}^o) \quad (3.27)$$

g) Calcular os termos de erro (gradiente local do erro) para as unidades de saída:

$$\phi_{pk}^o = (Y_{pk} - O_{pk}) \cdot f' \quad (3.28)$$

onde f' é a derivada da função de ativação escolhida. No caso da função tangente hiperbólica, tem-se:

$$f' = \tanh'(x) = 1 - (\tanh(x))^2 \quad (3.29)$$

E, como para a camada de saída, $\tanh(x) = O_{pk}$, obtém-se:

$$f' = \tanh'(x) = 1 - (O_{pk})^2 \quad (3.30)$$

h) Calcular os termos de erro para as unidades escondidas:

$$\phi_{pj}^h = [1 - (i_{pj})^2] \sum_k \phi_{pk}^\theta \cdot W_{kj}^\theta \quad (3.31)$$

i) Atualizar os pesos na camada de saída

$$W_{kj}^o(t+1) = W_{kj}^o(t) + \eta \cdot \phi_{pk}^o \cdot i_{pj} \quad (3.32)$$

onde η é o parâmetro da taxa de aprendizado, que determina a velocidade do aprendizado.

j) Atualizar os pesos na camada escondida:

$$W_{ji}^h(t+1) = W_{ji}^h(t) + \eta \cdot \phi_{pj}^h \cdot X_{pi} \quad (3.33)$$

k) Repetir os passos **b** a **j** para todos os P padrões.

O processo se encerra quando o erro médio quadrado (EMQ), dos P padrões, for menor ou igual a um erro mínimo previamente estipulado, ou ao atingir-se um número determinado de iterações.

$$\text{Erro Médio Quadrado} = \frac{1}{P} \sum_{p=1}^P E_p \quad (3.34)$$

onde E_p é o erro instantâneo obtido ao apresentar o padrão X_p na entrada da rede.

Este erro é dado por

$$E_p = \frac{1}{2} \cdot \sum_{k=1}^M (Y_{pk} - O_{pk})^2 \quad (3.35)$$

CAPÍTULO 4

EXPERIMENTOS

4. 1. Considerações gerais

O processo de filtragem e classificação de imagens por redes neurais baseado em características texturais, proposto neste trabalho, foi testado em imagens sintéticas e reais. Os conjuntos de experimentos envolvendo as imagens sintéticas têm o objetivo de determinar a acurácia do processo. Já a utilização de imagens reais permite avaliar o comportamento da metodologia proposta quando aplicada a imagens obtidas por sensoriamento remoto. Ambos os experimentos são descritos e analisados neste capítulo, com o intuito de permitir a devida compreensão e avaliação da metodologia proposta no capítulo anterior.

As imagens sintéticas usadas na primeira etapa de experimentos foram construídas de duas formas: primeiramente, imagens sintéticas geradas por funções senoidais bidimensionais; posteriormente, imagens mosaicos a partir de texturas reais extraídas do álbum fotográfico de Brodatz (Brodatz, 1966). Numa segunda etapa, utiliza-se uma imagem digitalizada em formato raster da região metropolitana do município de Porto Alegre (RS), caracterizada por classes de textura distintas.

Tanto os experimentos que fazem uso de imagens sintéticas quanto de imagens reais são desenvolvidos através de fases bem definidas: extração de frequências características a cada classe textural; formação dos filtros de Gabor correspondentes a cada frequência espacial selecionada; convolução da imagem com cada filtro de Gabor; treinamento da rede neural artificial para a constituição do classificador baseado em redes neurais; e, finalmente, a classificação da imagem.

A linguagem de programação abordada para o processamento dos algoritmos é aquela utilizada pelo software Matlab for Windows, por tratar-se de um ambiente propício à demanda computacional exigida.

Os experimentos descritos neste capítulo foram realizados com o uso dos seguintes recursos:

- PC Pentium II, 400 MHz, 164 Mb de RAM, monitor SVGA colorido (800 x 600) e 20 Gb de disco rígido;
- Scanner Genius Color Page Vivid III;
- Sistema Operacional Windows 98;
- Software MATLAb for Windows, desenvolvido pela Math Works, Inc., versão 5.3;
- Software Idrisi 32 for Windows versão I32.02, desenvolvido pela Clark Labs;
- Software MultSpec Application, desenvolvido pela Purdue Research, versão 1.2.

Tais recursos se mostraram suficientes para o desenvolvimento do trabalho.

4. 2. Experimentos envolvendo imagens sintéticas

Os experimentos realizados com imagens sintéticas, e apresentados neste item, têm o propósito de analisar as possibilidades de seleção dos parâmetros envolvidos na concepção tanto dos filtros de Gabor quanto do classificador baseado em redes neurais.

A primeira etapa de experimentos é realizada com imagens sintéticas geradas a partir de funções matemáticas, com o intuito de avaliar a metodologia proposta em condições conhecidas. Numa segunda etapa, utilizam-se imagens sintéticas

compostas por microtexturas de Brodatz (vide Apêndice 2), a fim de estudar o comportamento da metodologia em situações mais próximas de imagens reais. A partir dos testes realizados nas etapas anteriores, é possível avaliar o desempenho do classificador neural quando aplicado a imagens reais, já numa terceira etapa de experimentos.

4. 2. 1. Imagens sintéticas a partir de funções matemáticas

A imagem sintética utilizada para os experimentos, num primeiro momento, foi construída a partir de funções senoidais bidimensionais com diferentes orientações. Esta imagem, ilustrada na figura 4.1, é composta por quatro classes, onde cada classe é formada por uma frequência única: 4/128 e 8/128 ciclos/pixel na orientação do eixo y; e 12/128 e 20/128 ciclos/pixel na orientação do eixo x. Tal imagem tem dimensão 256 x 256 pixels.

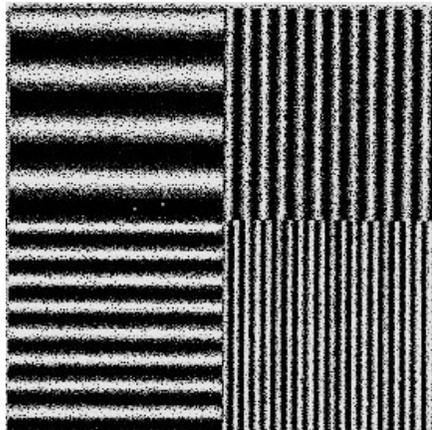


Figura 4.1.. Imagem sintética de 256 x 256 pixels composta por funções senoidais bidimensionais.

4. 2. 1. 1. Experimento A

Este primeiro experimento tem o objetivo de testar o efeito da extensão espacial dos filtros de Gabor no processo de filtragem. Para a constituição dos quatro filtros de Gabor necessários, foram utilizados como parâmetros as frequências

4/128 e 8/128 ciclos/pixel na orientação do eixo y e 12/128 e 20/128 na orientação do eixo x. Para a extensão espacial, considerou-se um valor único, igual ao maior valor das extensões espaciais correspondentes a cada freqüência: 128/24. Cada filtro gera uma imagem filtrada, como pode ser visto na figura 4. 2.

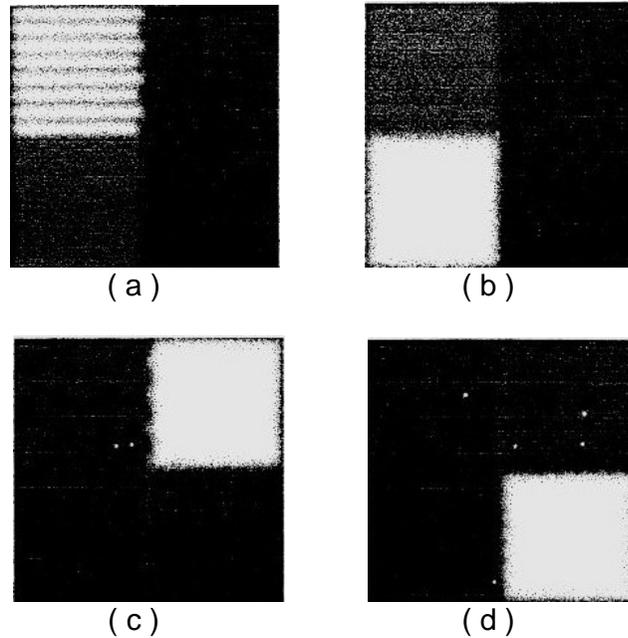


Figura 4.2. Imagens filtradas pelos filtros de Gabor para o experimento **A**, referente à imagem sintética a partir de funções matemáticas. A ordem das filtrações é expressa de (a) até (d).

O número de imagens filtradas será o parâmetro que define o número de neurônios da camada de entrada da RNA. O número de classes a serem classificadas na imagem sintética da figura 4.1 determina o número de neurônios da camada de saída em quatro. Já o número de neurônios da camada oculta varia para cada experimento. Nos experimentos A e B foram utilizados 6 e 8 neurônios na camada oculta, respectivamente.

Na figura 4.3 podemos ver a imagem classificada pela RNA a partir das imagens filtradas na figura 4.2.

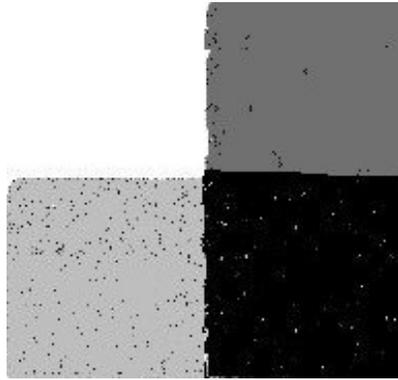


Figura 4.3. Imagem classificada referente ao experimento A .

4. 2. 1. 2. Experimento B

Da mesma forma que o experimento anterior, este tem o objetivo de analisar a influência da extensão espacial dos filtros de Gabor, bem como suas consequências após a filtragem da imagem sintética. Neste caso, o valor adotado para a extensão espacial no experimento A foi duplicado: 128/12. Os valores utilizados para as frequências espaciais foram mantidos. Dessa forma, obtém-se quatro novas imagens filtradas (figura 4. 4).

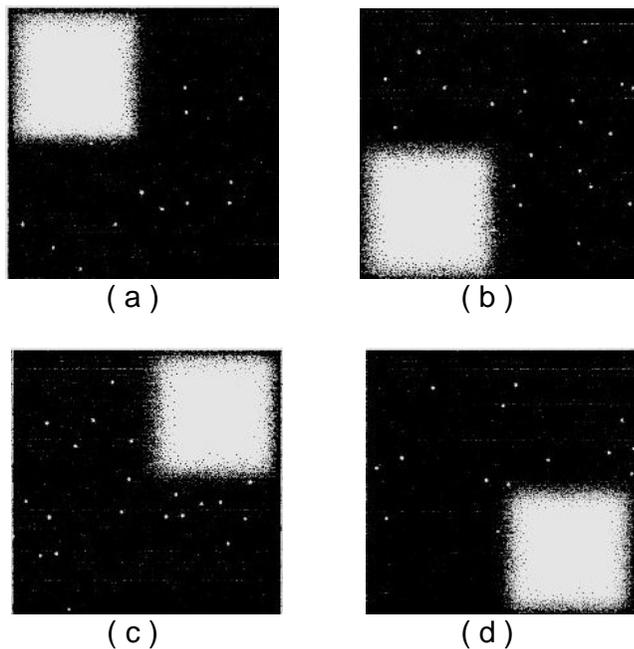


Figura 4. 4. Imagens filtradas pelos filtros de Gabor para o experimento B, referente à imagem sintética a par-

tir de funções matemáticas. A ordem das filtragens é expressa de (a) até (d).

O número de imagens filtradas será o parâmetro que define o número de neurônios da camada de entrada da RNA. O número de classes a serem classificadas na imagem sintética da figura 4.1 determina o número de neurônios da camada de saída em quatro. Já o número de neurônios da camada oculta varia para cada experimento.

Na figura 4.5 podemos ver a imagem classificada pela RNA a partir das imagens filtradas na figura 4.4.

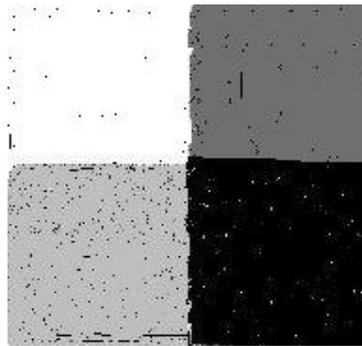


Figura 4.5. Imagem classificada referente ao experimento B .

4. 2. 2. Imagem sintética: mosaico a partir de texturas de Brodatz

Na segunda etapa de experimentos foi utilizada uma imagem mosaico, composta por micro texturas a partir de imagens do álbum de Brodatz (Brodatz, 1966). Por motivo de conveniência computacional, as dimensões da imagem mosaico são de 256 x 256, o que permite a utilização do algoritmo da Transformada Rápida de Fourier (FFT), cuja implementação minimiza o tempo de processamento e o espaço de memória exigido pela máquina (Angelo, 2000). Tal conveniência é observada, apenas, quando da ocorrência de imagens no formato $2^n \times 2^n$ pixels, onde n é um número inteiro.

A imagem mosaico está ilustrada na figura 4.6. Visto que o processo de filtragem se trata de uma abordagem supervisionada, é necessária a extração de amostras da cada classe. A figura 4.7 mostra a localização de cada uma das 4 amostras utilizadas, cada qual contendo as freqüências representativas de sua classe. As dimensões das janelas utilizadas para a coleta das amostras devem ser suficientes para caracterizarem as texturas de cada classe presente na imagem. As janelas visualizadas na figura 4.7 , referentes à coleta de amostras das classes da imagem mosaico da figura 4.6 , possuem as mesmas dimensões: 32x32.

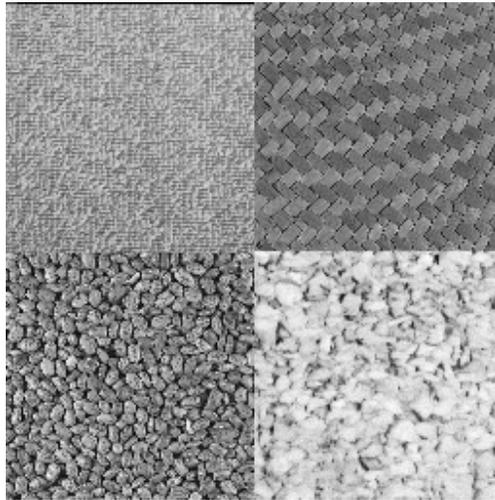


Figura 4. 6. Imagem mosaico sintética de 256 x 256 pixels composta por microtexturas de Brodatz.

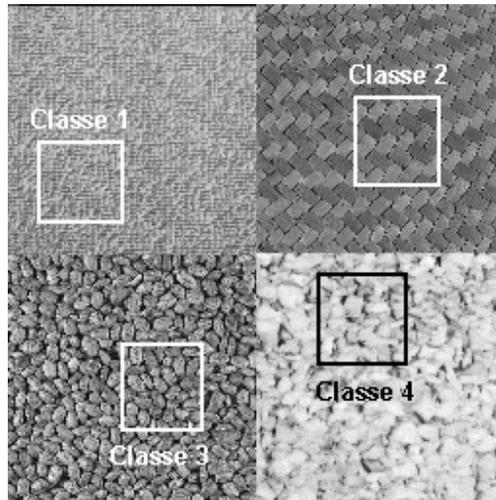


Figura 4. 7. Imagem mosaico com localização das amostras representativas de cada classe para a seleção de feições.

Nas figuras 4.8 a 4.11 são apresentadas as freqüências mais representativas das amostras, bem como os respectivos espectros de Fourier, onde os valores de u são dados em ciclos/dimensão x da amostra e os valores de v são dados em ciclos/dimensão y da amostra. A seleção de freqüências para a determinação dos filtros de Gabor em cada experimento é realizada através do conhecimento de um "especialista", a partir das freqüências de cada amostra que apresentem os maiores níveis de energia. Este processo de seleção de freqüências é feito com base no espectro de Fourier de cada amostra de cada classe da imagem.

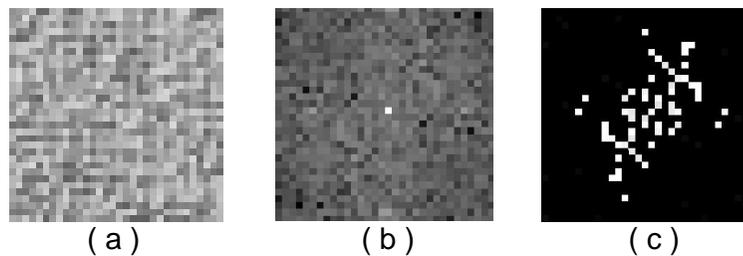


Figura 4. 8. (a) Amostra da classe 1 na imagem mosaico; (b) Espectro de Fourier da amostra; (c) Freqüências da amostra que apresentam o maior nível de energia.

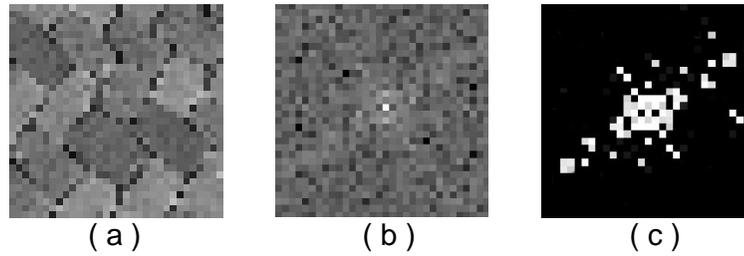


Figura 4. 9. (a) Amostra da classe 2 na imagem mosaico; (b) Espectro de Fourier da amostra; (c) Freqüências da amostra que apresentam o maior nível de energia.

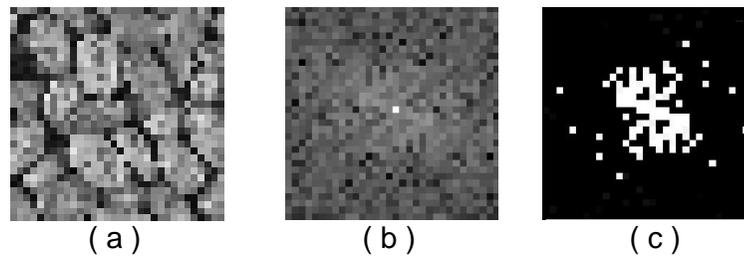


Figura 4. 10. (a) Amostra da classe 3 na imagem mosaico; (b) Espectro de Fourier da amostra; (c) Freqüências da amostra que apresentam o maior nível de energia.

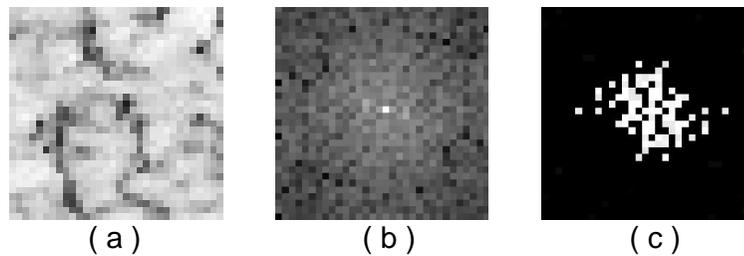


Figura 4. 11. (a) Amostra da classe 4 na imagem mosaico; (b) Espectro de Fourier da amostra; (c) Freqüências da amostra que apresentam o maior nível de energia.

4. 2. 2. 1. Experimento A

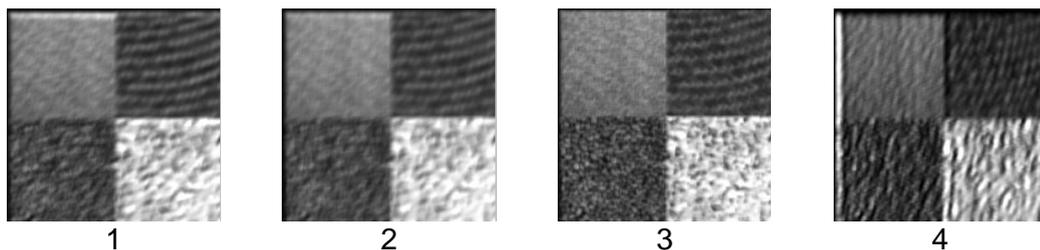
Os parâmetros necessários para a constituição dos filtros de Gabor para este experimento foram obtidos a partir das amostras determinadas na figura 4.5. Estes parâmetros podem ser visualizados na tabela 4. 1.

Filtro	Dimensão Do filtro (pixels)	Frequência ao longo do eixo X (k_x)	Frequência ao longo do eixo Y (k_y)	Extensão Espacial (σ)
1	36	0,0532	0,0279	6,0000
2	34	0,0414	0,0296	5,6667
3	19	0,0549	0	3,1667
4	48	0,0211	0,0718	8,0000
5	29	0,0346	-0,1715	4,8333
6	44	0,0228	-0,0059	7,3333
7	44	0,0228	-0,1563	7,3333
8	18	0,0566	0,0718	3,0000
9	40	0,0253	-0,0282	6,6667
10	79	0,0127	0	13,1667
11	90	0	0,0602	15,0000
12	90	0,0112	0	15,0000
13	115	0,0087	0,0201	19,1667
14	115	0	0,0300	19,1667
15	115	0,0087	0,0391	19,1667

Tabela 4. 1. Parâmetros para a constituição dos filtros de Gabor necessários para a realização do experimento A.

Os filtros de Gabor definidos na tabela 4.1 geram, por sua vez, 15 imagens filtradas, como podem ser vistas na figura 4.12.

A classificação da imagem pela RNA nos experimentos A1 e A2 ocorrerá através de um processo pixel a pixel onde os parâmetros de entrada da RNA serão os valores do contador digital (CD) de um pixel nas suas diversas "bandas texturais" (imagens filtradas).



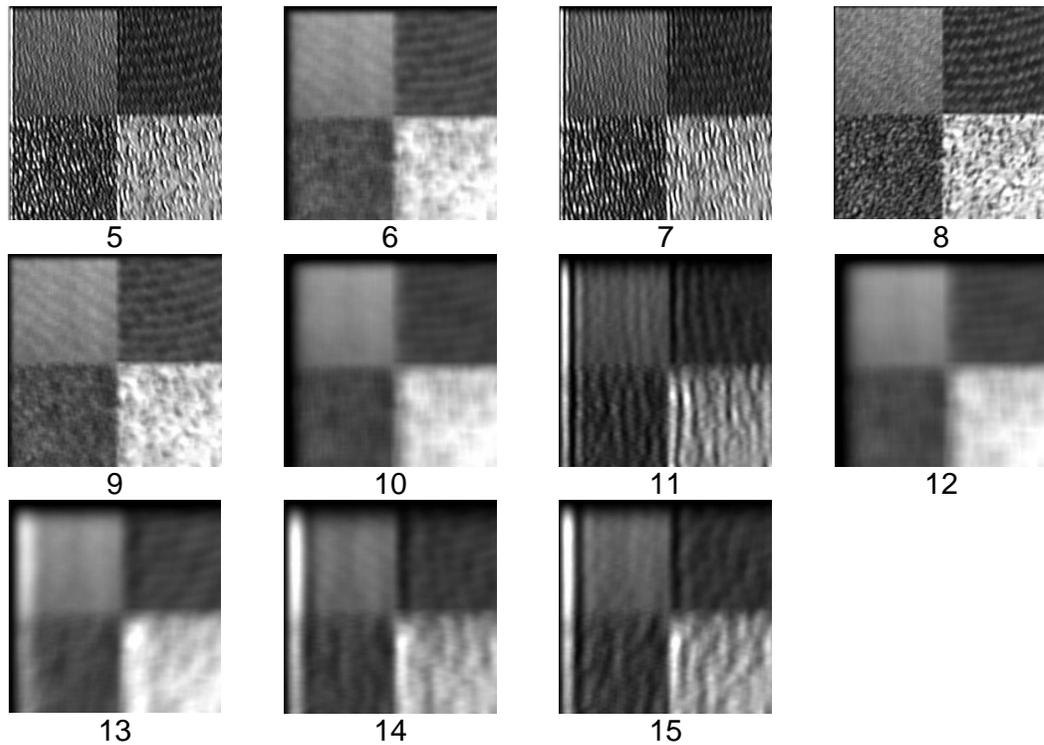


Figura 4. 12. Imagens filtradas pelo filtro de Gabor para o experimento A, a partir da imagem mosaico.

4. 2. 2. 1. 1. Experimento A1

Neste experimento, foram utilizadas como parâmetros de entrada as imagens filtradas de através dos filtros de Gabor relativos à tabela 4.1, que geraram as imagens referentes à figura 4.12.

O número de imagens filtradas será o parâmetro que define o número de neurônios da camada de entrada da RNA. O número de classes a serem classificadas nas imagem sintética da figura 4.6 determina o número de neurônios da camada de saída em quatro. Já o número de neurônios da camada intermediária varia para cada experimento.

A partir disso, a rede neural artificial empregada no experimento A1 faz uso de 15 neurônios na camada de entrada, 18 neurônios na camada intermediária e 4 neurônios na camada de saída.

A figura 4.13 ilustra a imagem obtida após a classificação através da RNA, pixel a pixel, da imagem sintética.

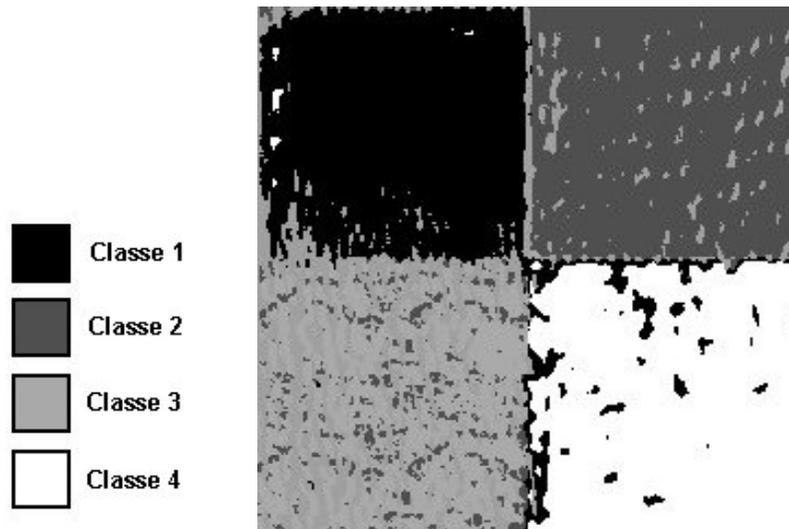


Figura 4. 13. Imagem classificada referente ao experimento A1, utilizando uma imagem sintética.

A tabela 4.2 avalia os resultados da imagem classificada da figura 4.13 através de uma matriz de erro (Angelo, 2000). Na mesma, é possível vislumbrar uma taxa de exatidão geral do processo de classificação.

	EXATIDÃO (%)	ERRO DE OMISSÃO (%)	NÚMERO DE PIXELS REFERÊNCIA	IMAGEM CLASSIFICADA			
				CLASSE 1	CLASSE 2	CLASSE 3	CLASSE 4
IMAGEM REFERÊNCIA				1	2	3	4
CLASSE 1	89.13	10.87	16384	14603	260	809	759
CLASSE 2	87.99	12.01	16384	212	14416	939	817
CLASSE 3	87.30	12.70	16384	225	16	14303	1840
CLASSE 4	90.18	9.82	16384	811	474	324	14775
Número de pixels da Imagem Classificada			65536	15851	15166	16375	18191
Erro de Comissão (%)				7.62	4.58	12.65	20.85
EXATIDÃO GERAL (%)				88.65			

Tabela 4. 2. Dados da matriz de erro para a avaliação da classificação da imagem sintética composta por quatro texturas naturais de Brodatz, referente ao experimento A1.

4. 2. 2. 1. 2. Experimento A2

No experimento A2, foram utilizadas como parâmetros de entrada as mesmas imagens filtradas do experimento A1, relativas à figura 4.12.

Dessa forma, a rede neural artificial empregada no experimento A2 faz uso de 15 neurônios na camada de entrada, 23 neurônios na camada intermediária e 4 neurônios na camada de saída.

A figura 4.14 ilustra a imagem obtida após a classificação através da RNA, pixel a pixel, da imagem sintética.

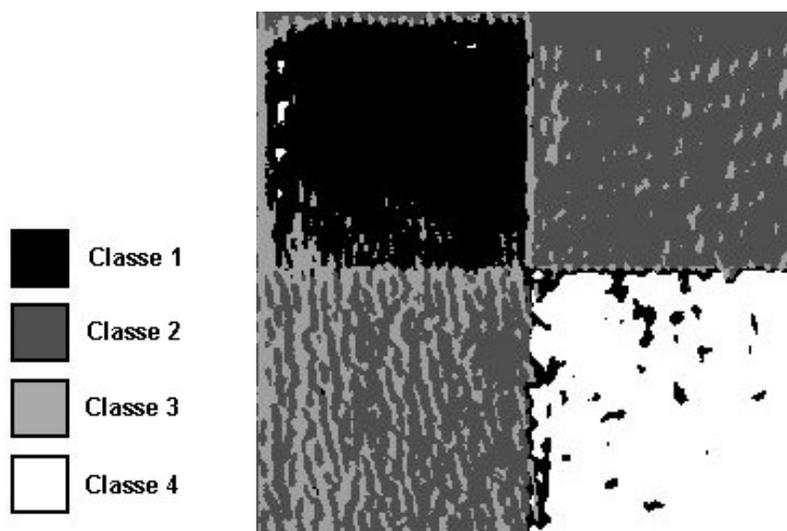


Figura 4. 14. Imagem classificada referente ao experimento A2, utilizando uma imagem sintética.

Na tabela 4.3, é possível avaliar os resultados da imagem classificada da figura 4.14 através de uma matriz de erro (Angelo, 2000).

	EXATIDÃO (%)	ERRO DE OMISSÃO (%)	NÚMERO DE PIXELS REFERÊNCIA	IMAGEM CLASSIFICADA			
				CLASSE 1	CLASSE 2	CLASSE 3	CLASSE 4
IMAGEM REFERÊNCIA				1	2	3	4
CLASSE 1	85.39	14.61	16384	13991	544	938	911
CLASSE 2	91.77	8.23	16384	305	15036	201	842

CLASSE 3	67.49	32.51	16384	230	2949	11057	2148
CLASSE 4	87.09	12.91	16384	1712	276	127	14269
Número de pixels da Imagem Classificada			65536	16238	18805	12323	18170
Erro de Comissão (%)				13.71	23.00	7.73	23.81
EXATIDÃO GERAL (%)			82.94				

Tabela 4. 3. Dados da matriz de erro para a avaliação da classificação da imagem sintética composta por quatro texturas naturais de Brodatz, referente ao experimento A2.

4. 2. 2. 2. Experimento B

Utilizando as mesmas frequências do experimento A, o processo de classificação foi implementado alterando-se a extensão espacial dos filtros de Gabor, ou seja, assumindo um valor único e igual ao maior valor das extensões espaciais existentes (tabela 4.4).

Filtro	Dimensão do filtro (pixels)	Frequência ao longo do eixo X (k_x)	Frequência ao longo do eixo Y (k_y)	Extensão Espacial (σ)
1	115	0,0532	0,0279	19,1667
2	115	0,0414	0,0296	19,1667
3	115	0,0549	0	19,1667
4	115	0,0211	0,0718	19,1667
5	115	0,0346	-0,1715	19,1667
6	115	0,0228	-0,0059	19,1667
7	115	0,0228	-0,1563	19,1667
8	115	0,0566	0,0718	19,1667
9	115	0,0253	-0,0282	19,1667
10	115	0,0127	0	19,1667
11	115	0	0,0602	19,1667
12	115	0,0112	0	19,1667
13	115	0,0087	0,0201	19,1667
14	115	0	0,0300	19,1667
15	115	0,0087	0,0391	19,1667

Tabela 4. 4. Parâmetros para a constituição dos filtros de Gabor necessários para a realização do experimento B.

Os filtros de Gabor definidos na tabela 4.4 geram, por sua vez, 15 imagens filtradas, diferentes das obtidas para o experimento A e que podem ser vistas na figura 4.15. A classificação da imagem pela RNA nos experimentos B1 e B2 ocorrerá através de um processo pixel a pixel onde os parâmetros de entrada da RNA serão os valores do contador digital (CD) de um pixel nas suas 15 "bandas texturais" (imagens filtradas).

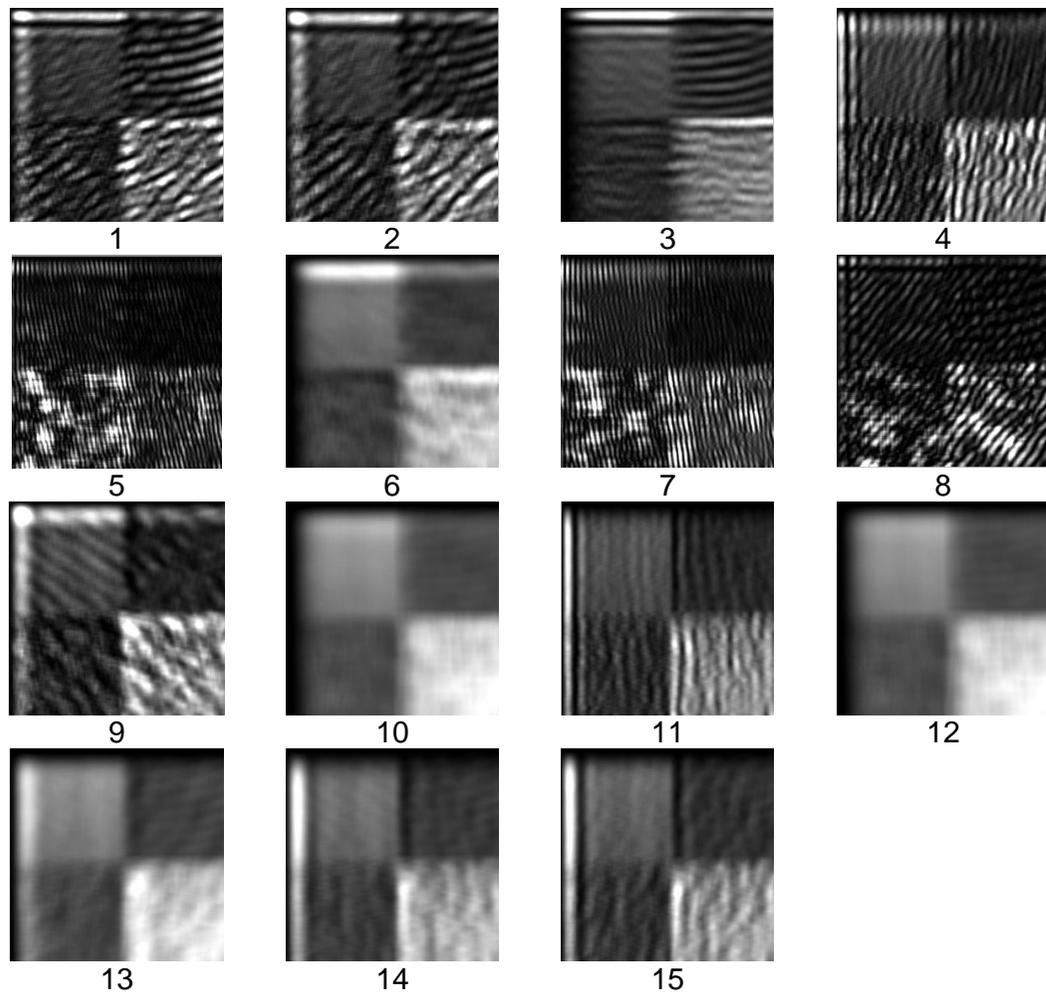


Figura 4. 15. Imagens filtradas pelo filtro de Gabor para o experimento B, a partir da imagem mosaico.

4. 2. 2. 2. 1. Experimento B1

Neste experimento, foram utilizadas como parâmetros de entrada as imagens filtradas de através dos filtros de Gabor relativos à tabela 4.4, que geraram as imagens referentes à figura 4.15.

Da mesma forma que na série de experimentos A, o número de imagens filtradas será o parâmetro que define o número de neurônios da camada de entrada da RNA. O número de classes a serem classificadas nas imagem sintética da figura 4.8 determina o número de neurônios da camada de saída em quatro. Já o número de neurônios da camada intermediária varia para cada experimento.

Portanto, a rede neural artificial empregada no experimento B1 faz uso de 15 neurônios na camada de entrada, 18 neurônios na camada intermediária e 4 neurônios na camada de saída.

A figura 4.16 ilustra a imagem obtida após a classificação através da RNA, pixel a pixel, da imagem sintética.

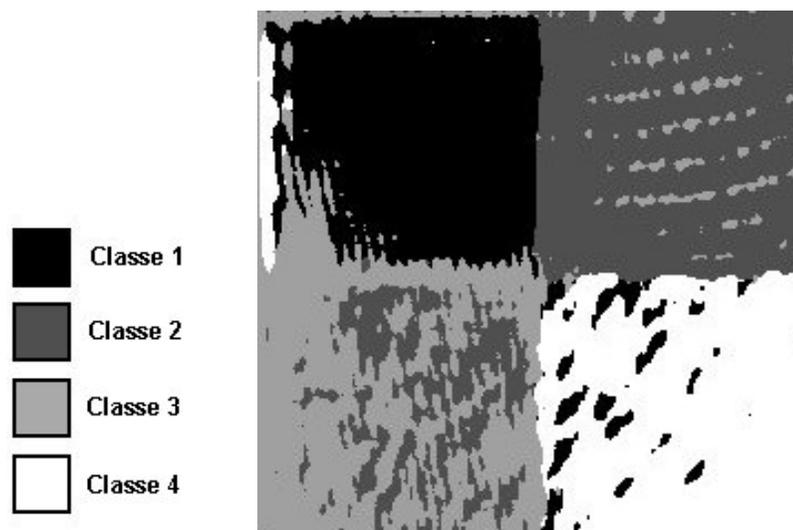


Figura 4. 16. Imagem classificada referente ao experimento B1, utilizando uma imagem sintética.

A tabela 4.5 avalia os resultados da imagem classificada da figura 4.16 através de uma matriz de erro (Angelo, 2000). Na mesma, é possível vislumbrar uma taxa de exatidão geral do processo de classificação.

IMAGEM REFERÊNCIA	EXATIDÃO (%)	ERRO DE OMISSÃO (%)	NÚMERO DE PIXELS REFERÊNCIA	IMAGEM CLASSIFICADA			
				CLASSE 1	CLASSE 2	CLASSE 3	CLASSE 4
CLASSE 1	81.56	18.44	16384	13362	1630	807	585
CLASSE 2	91.27	8.73	16384	212	14953	590	629
CLASSE 3	79.13	20.87	16384	33	3386	12965	0
CLASSE 4	85.02	14.98	16384	1583	707	165	13929
Número de pixels Da Imagem Classificada			65536	15190	20676	14527	15143
Erro de Comissão (%)				11.16	34.93	9.53	7.41
EXATIDÃO GERAL (%)			84.24				

Tabela 4. 5. Dados da matriz de erro para a avaliação da classificação da imagem sintética composta por quatro texturas naturais de Brodatz, referente ao experimento B1.

4. 2. 2. 2. Experimento B2

No experimento B2, foram utilizadas como parâmetros de entrada as mesmas imagens filtradas do experimento B1, relativas à figura 4.15.

Porém, a rede neural artificial empregada no experimento B2 utiliza 15 neurônios na camada de entrada, 23 neurônios na camada intermediária e 4 neurônios na camada de saída.

A figura 4.17 ilustra a imagem obtida após a classificação através da RNA, pixel a pixel, da imagem sintética.

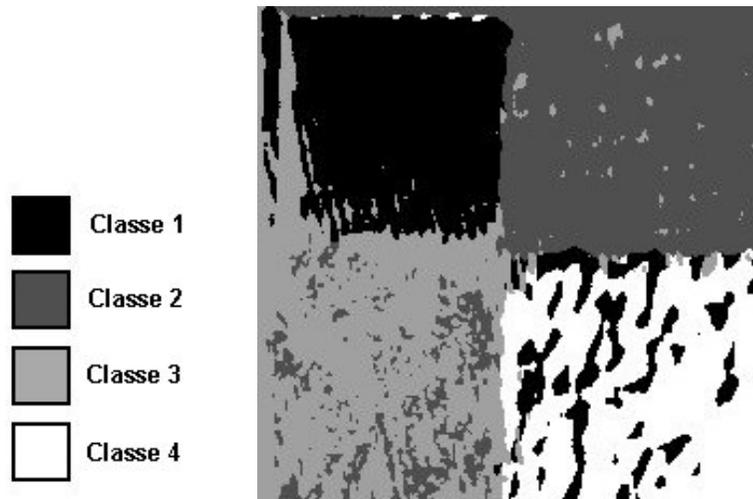


Figura 4. 17. Imagem classificada referente ao experimento B2, utilizando uma imagem sintética.

Na tabela 4.6, é possível avaliar os resultados da imagem classificada da figura 4.17 através de uma matriz de erro (Angelo, 2000).

	EXATIDÃO (%)	ERRO DE OMISSÃO (%)	NÚMERO DE PIXELS REFERÊNCIA	IMAGEM CLASSIFICADA			
				CLASSE 1	CLASSE 2	CLASSE 3	CLASSE 4
IMAGEM REFERÊNCIA				CLASSE 1	CLASSE 2	CLASSE 3	CLASSE 4
CLASSE 1	72.49	27.51	16384	11877	764	2224	1519
CLASSE 2	96.26	3.74	16384	184	15772	216	212
CLASSE 3	80.77	19.23	16384	35	3032	13234	83
CLASSE 4	76.66	23.34	16384	3517	67	240	12560
Número de pixels da Imagem Classificada			65536	15613	19635	15914	14374
Erro de Comissão (%)				22.80	23.58	16.36	11.07
EXATIDÃO GERAL (%)				81.55			

Tabela 4. 6. Dados da matriz de erro para a avaliação da classificação da imagem sintética composta por quatro texturas naturais de Brodatz, referente ao experimento B2.

4. 2. 2. 3. Experimento C

O terceiro experimento tem o objetivo de testar a influência da extensão espacial dos filtros de Gabor no processo de classificação, com o aumento do número de filtros. Os dados relativos a estes filtros são ilustrados na tabela 4.7.

Filtro	Dimensão do filtro (pixels)	Frequência ao longo do eixo X (k_x)	Frequência ao longo do eixo Y (k_y)	Extensão Espacial (σ)
1	36	0,0532	0,0279	6,0000
2	34	0,0414	0,0296	5,6667
3	19	0,0549	0	3,1667
4	48	0,0211	0,0718	8,0000
5	29	0,0346	-0,1715	4,8333
6	44	0,0228	-0,0059	7,3333
7	44	0,0228	-0,1563	7,3333
8	18	0,0566	0,0718	3,0000
9	40	0,0253	-0,0282	6,6667
10	79	0,0127	0	13,1667
11	90	0	0,0602	15,0000
12	90	0,0112	0	15,0000
13	115	0,0087	0,0201	19,1667
14	115	0	0,0300	19,1667
15	115	0,0087	0,0391	19,1667
16	79	0	0,0423	13,1667
17	43	0,0235	0,0267	7,1667
18	85	0,0118	0,0133	14,1667
19	145	0,0069	0,0109	24,1667
20	102	0,0443	0	17,0000
21	57	0,0177	0,0291	9,5000
22	29	0,0353	-0,0133	4,8333
23	30	0,0338	0,2500	5,0000
24	50	0,0203	-0,0878	8,3333
25	85	0	0,0133	14,1667

Tabela 4. 7. Parâmetros para a constituição dos filtros de Gabor necessários para a realização do experimento C.

Os filtros de Gabor definidos na tabela 4.7 geram, por sua vez, 25 imagens filtradas ilustradas na figura 4.18. A classificação da imagem pela RNA nos experimentos C1 e C2 ocorrerá através de um processo pixel a pixel onde os parâmetros de entrada da RNA serão os valores do contador digital (CD) de um pixel nas suas 25 imagens filtradas.

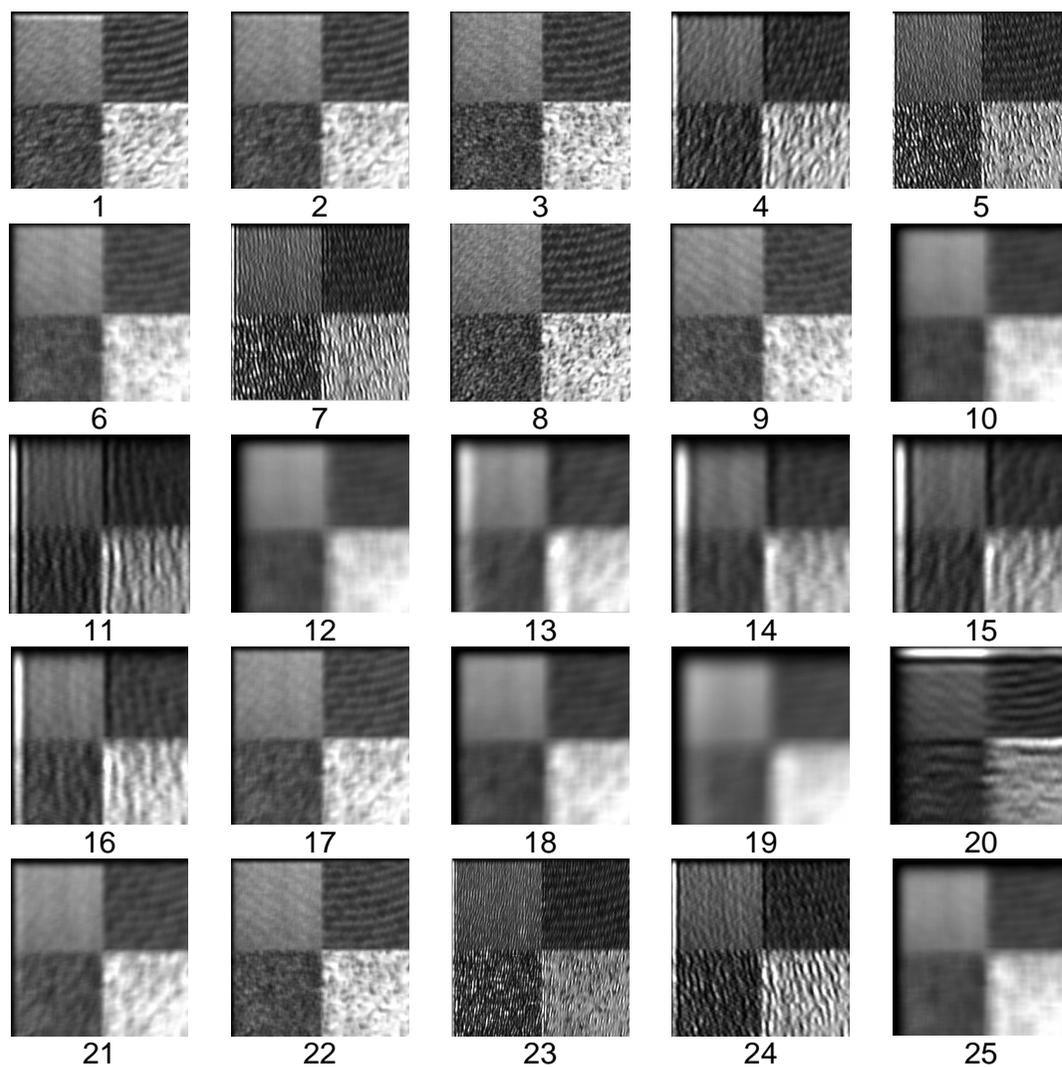


Figura 4. 18. Imagens filtradas pelo filtro de Gabor para o experimento C, a partir da imagem mosaico.

4. 2. 2. 3. 1. Experimento C1

No experimento C1, foram empregadas como parâmetros de entrada as 25 imagens filtradas de através dos filtros de Gabor relativos à tabela 4.7, que geraram as imagens referentes à figura 4.18.

A rede neural artificial utilizada para a classificação da imagem sintética no experimento C1 faz uso de 25 neurônios na camada de entrada, 28 neurônios na camada intermediária e 4 neurônios na camada de saída.

A figura 4.19 ilustra a imagem obtida após a classificação através da RNA, pixel a pixel, da imagem sintética.

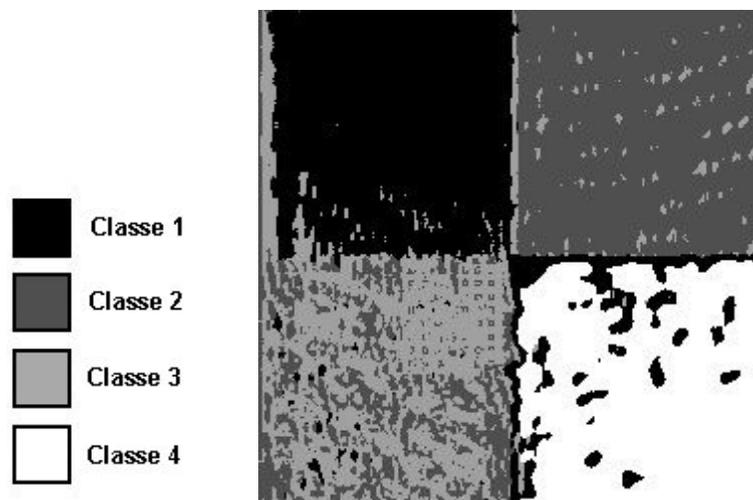


Figura 4. 19. Imagem classificada referente ao experimento C1, utilizando uma imagem sintética.

A tabela 4.8 avalia os resultados da imagem classificada da figura 4.19 através de uma matriz de erro (Angelo, 2000). Na mesma, é possível vislumbrar uma taxa de exatidão geral do processo de classificação.

IMAGEM REFERÊNCIA	EXATIDÃO	ERRO DE	NÚMERO	IMAGEM CLASSIFICADA			
	(%)	OMISSÃO	DE PIXELS	CLASSE	CLASSE	CLASSE	CLASSE
		(%)	REFERÊNCIA	1	2	3	4
CLASSE 1	89.46	10.54	16384	14657	421	753	553
CLASSE 2	90.80	9.20	16384	281	14877	792	434
CLASSE 3	69.01	30.99	16384	466	182	11307	4429
CLASSE 4	82.97	17.03	16384	2541	170	80	13593
Número de pixels da Imagem Classificada			65536	17945	15650	12932	19009
Erro de Comissão (%)				20.07	4.72	9.92	33.06
EXATIDÃO GERAL (%)			83.06				

Tabela 4. 8. Dados da matriz de erro para a avaliação da classificação da imagem sintética composta por quatro texturas naturais de Brodatz, referente ao experimento C1.

4. 2. 2. 3. 2. Experimento C2

No presente experimento, foram utilizadas como parâmetros de entrada as mesmas imagens filtradas do experimento C1, relativas à figura 4.18.

Da mesma forma que a rede neural artificial empregada no experimento C1, o experimento C2 também utiliza 25 neurônios na camada de entrada e 4 neurônios na camada de saída, porém faz uso de 32 neurônios na camada intermediária.

A figura 4.20 ilustra a imagem obtida após a classificação através da RNA, pixel a pixel, da imagem sintética.

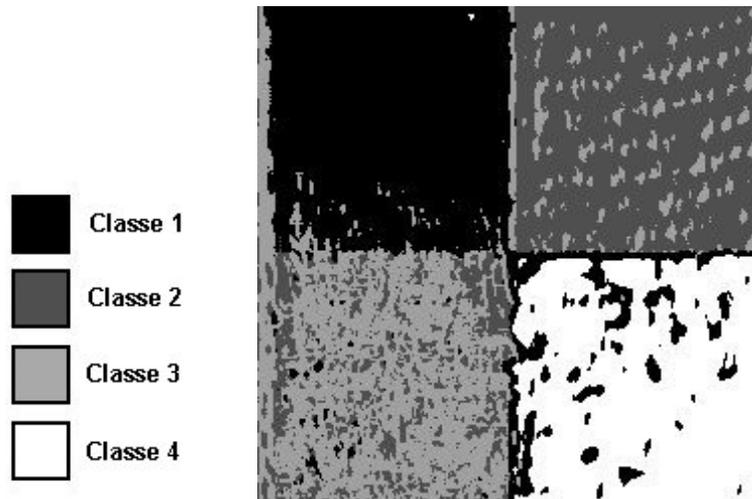


Figura 4. 20. Imagem classificada referente ao experimento C2, utilizando uma imagem sintética.

A tabela 4.9 ilustra uma avaliação dos resultados da imagem classificada da figura 4.20 através de uma matriz de erro (Angelo, 2000).

IMAGEM REFERÊNCIA	EXATIDÃO (%)	ERRO DE OMISSÃO (%)	NÚMERO DE PIXELS REFERÊNCIA	IMAGEM CLASSIFICADA			
				CLASSE 1	CLASSE 2	CLASSE 3	CLASSE 4
CLASSE 1	82.77	17.23	16384	13561	520	1235	1068
CLASSE 2	79.35	20.65	16384	721	13001	1094	1568
CLASSE 3	67.44	32.56	16384	474	125	11050	4735
CLASSE 4	77.95	22.05	16384	2765	406	441	12772
Número de pixels da Imagem Classificada			65536	17521	14052	13820	20143
Erro de Comissão (%)				24.17	6.41	16.91	44.99
EXATIDÃO GERAL (%)			76.88				

Tabela 4. 9. Dados da matriz de erro para a avaliação da classificação da imagem sintética composta por quatro texturas naturais de Brodatz, referente ao experimento C2.

4. 2. 2. 4. Experimento D

Concluindo a fase de experimentos com a imagem mosaico, o experimento D continua os testes sobre a influência da extensão espacial dos filtros de Gabor. Desta vez, o valor da extensão espacial é incrementado em 1.5 vezes o valor máximo adotado no experimento anterior (tabela 4.10). Os filtros utilizados neste experimento são os mesmos relativos ao experimento C.

Filtro	Dimensão do filtro (pixels)	Frequência ao longo do eixo X (k_x)	Frequência ao longo do eixo Y (k_y)	Extensão Espacial (σ)
1	220	0,0532	0,0279	36,6667
2	220	0,0414	0,0296	36,6667
3	220	0,0549	0	36,6667
4	220	0,0211	0,0718	36,6667
5	220	0,0346	-0,1715	36,6667
6	220	0,0228	-0,0059	36,6667
7	220	0,0228	-0,1563	36,6667
8	220	0,0566	0,0718	36,6667
9	220	0,0253	-0,0282	36,6667
10	220	0,0127	0	36,6667
11	220	0	0,0602	36,6667
12	220	0,0112	0	36,6667
13	220	0,0087	0,0201	36,6667
14	220	0	0,0300	36,6667
15	220	0,0087	0,0391	36,6667
16	220	0	0,0423	36,6667
17	220	0,0235	0,0267	36,6667
18	220	0,0118	0,0133	36,6667
19	220	0,0069	0,0109	36,6667
20	220	0,0443	0	36,6667
21	220	0,0177	0,0291	36,6667
22	220	0,0353	-0,0133	36,6667
23	220	0,0338	0,2500	36,6667
24	220	0,0203	-0,0878	36,6667
25	220	0	0,0133	36,6667

Tabela 4. 10. Parâmetros para a constituição dos filtros de Gabor necessários para a realização do experimento D.

Os filtros de Gabor definidos na tabela 4.10 geram 25 imagens filtradas ilustradas na figura 4.21. A classificação da imagem pela RNA nos experimentos D1 e D2 ocorrerá através de um processo pixel a pixel onde os parâmetros de entrada da RNA serão os valores do contador digital (CD) de um pixel nas 25 imagens filtradas.

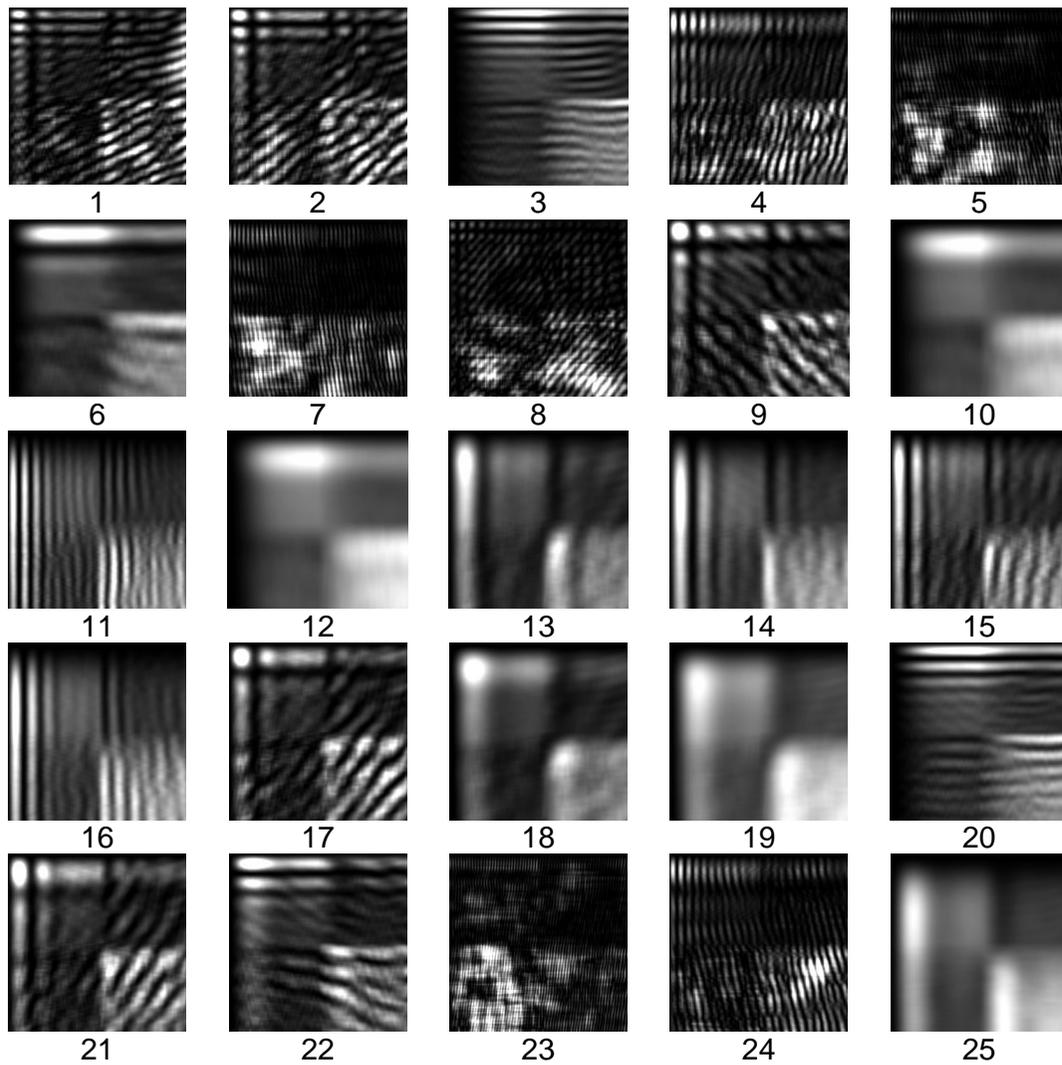


Figura 4. 21. Imagens filtradas pelo filtro de Gabor para o experimento D, a partir da imagem mosaico.

4. 2. 2. 4. 1. Experimento D1

No experimento D1, foram empregadas como parâmetros de entrada as 25 imagens filtradas de através dos filtros de Gabor relativos à tabela 4.10, que geraram as imagens referentes à figura 4.21.

A rede neural artificial utilizada para a classificação da imagem sintética neste experimento usa 25 neurônios na camada de entrada, 28 neurônios na camada intermediária e 4 neurônios na camada de saída.

A figura 4.22 ilustra a imagem obtida após a classificação através da RNA, pixel a pixel, da imagem sintética.

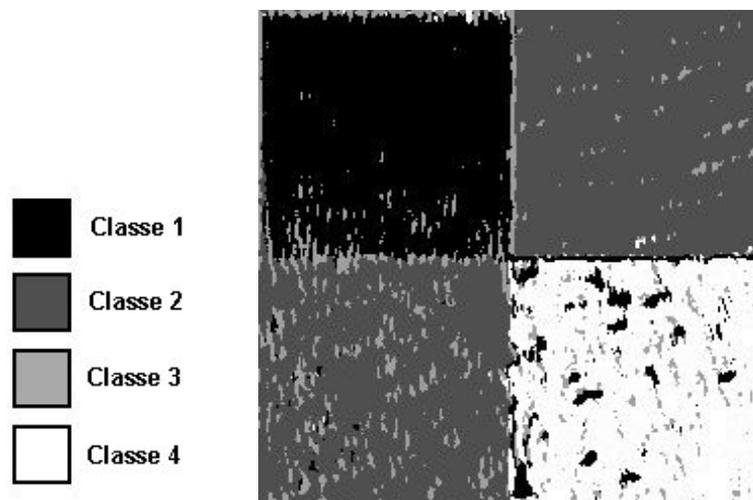


Figura 4. 22. Imagem classificada referente ao experimento D1, utilizando uma imagem sintética.

A tabela 4.11 avalia os resultados da imagem classificada da figura 4.22 através de uma matriz de erro (Angelo, 2000). Na mesma, é possível vislumbrar uma taxa de exatidão geral do processo de classificação.

IMAGEM REFERÊNCIA	EXATIDÃO	ERRO DE	NÚMERO	IMAGEM CLASSIFICADA			
	(%)	OMISSÃO	DE PIXELS	CLASSE	CLASSE	CLASSE	CLASSE
		(%)	REFERÊNCIA	1	2	3	4
CLASSE 1	91.02	8.98	16384	14913	181	724	566
CLASSE 2	93.70	6.30	16384	224	15352	419	389
CLASSE 3	86.43	13.57	16384	181	1105	14160	938
CLASSE 4	85.05	14.95	16384	1267	650	533	13934
Número de pixels da Imagem Classificada			65536	16585	17288	15836	15827
Erro de Comissão (%)				10.21	11.82	10.23	11.55
EXATIDÃO GERAL (%)			89.05				

Tabela 4. 11. Dados da matriz de erro para a avaliação da classificação da imagem sintética composta por quatro texturas naturais de Brodatz, referente ao experimento D1.

4. 2. 2. 4. 2. Experimento D2

Em tal experimento são utilizadas como parâmetros de entrada as mesmas imagens filtradas do experimento D1, relativas à figura 4.21.

Da mesma forma que a rede neural artificial empregada no experimento D1, o experimento D2 também utiliza 25 neurônios na camada de entrada e 4 neurônios na camada de saída, porém faz uso de 32 neurônios na camada intermediária.

A figura 4.23 ilustra a imagem obtida após a classificação através da RNA, pixel a pixel, da imagem sintética.

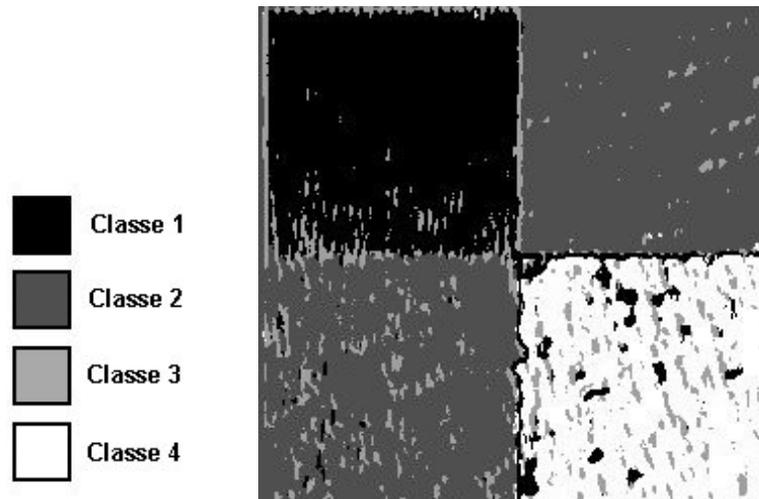


Figura 4. 23. Imagem classificada referente ao experimento D2, utilizando uma imagem sintética.

A tabela 4.12 ilustra uma avaliação dos resultados da imagem classificada da figura 4.23 através de uma matriz de erro (Angelo, 2000).

IMAGEM REFERÊNCIA	EXATIDÃO (%)	ERRO DE OMISSÃO (%)	NÚMERO DE PIXELS REFERÊNCIA	IMAGEM CLASSIFICADA			
				CLASSE 1	CLASSE 2	CLASSE 3	CLASSE 4
CLASSE 1	88.29	11.71	16384	14465	551	776	592
CLASSE 2	92.61	7.39	16384	266	15174	469	475
CLASSE 3	85.46	14.54	16384	530	911	14002	941
CLASSE 4	83.14	16.86	16384	1345	388	1030	13621
Número de pixels da Imagem Classificada			65536	16606	17024	16277	15629
Erro de Comissão (%)				13.07	11.29	13.89	12.26
EXATIDÃO GERAL (%)			87.37				

Tabela 4. 12. Dados da matriz de erro para a avaliação da classificação da imagem sintética composta por quatro texturas naturais de Brodatz, referente ao experimento D2.

4. 3. Experimentos envolvendo imagens reais

A imagem digital utilizada nesta etapa de experimentos está em formato raster e abrange parte da cidade de Porto Alegre (RS) e região metropolitana, como pode ser visto na figura 4.24. As dimensões desta imagem são de 256 x 256 pixels e o fator preponderante de seu uso para o experimento é a existência de três classes texturais distintas: água, urbana e vegetação. A região urbana é caracterizada pela região da cidade de Porto Alegre, a classe de água é representada pela região da laguna do Guaíba e a vegetação refere-se à área composta pelas ilhas do delta do Jacuí.



Figura 4. 24. Imagem digital de parte da cidade de Porto Alegre (RS) e entorno.

Os filtros necessários para a seleção de feições na imagem acima foram baseados nas amostras das classes água (figura 4. 25), urbana (figura 4. 31) e vegetação (figura 4. 37).

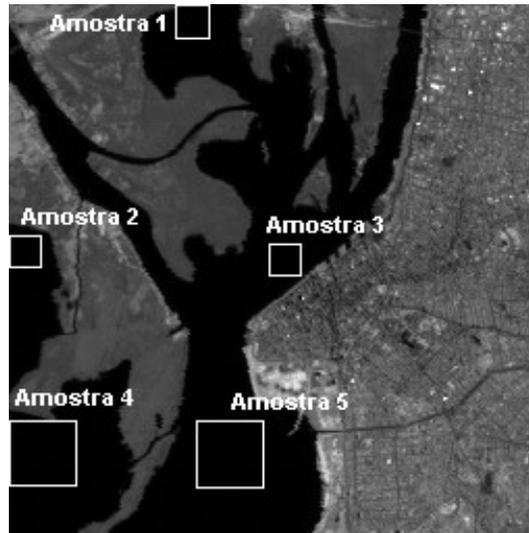


Figura 4. 25. Imagem real com localização das amostras representativas da classe água para a seleção de feições.

As dimensões das janelas utilizadas para a coleta das amostras devem ser suficientes para caracterizarem as texturas de cada classe presente na imagem. As janelas visualizadas nas figuras 4.25, 4.31 e 4.37, referentes à coleta de amostras das classes da imagem da figura 4.24 , possuem dimensões variadas: 16x16 ou 32x32.

Nas figuras 4.26 a 4.30 são apresentadas as freqüências mais representativas das amostras da classe água, bem como os respectivos espectros de Fourier, onde os valores de u são dados em ciclos/dimensão x da amostra e os valores de v são dados em ciclos/dimensão y da amostra. A seleção de freqüências para a determinação dos filtros de Gabor em cada experimento é realizada através do conhecimento de um "especialista", a partir das freqüências de cada amostra que apresentem os maiores níveis de energia. Este processo de seleção de freqüências é feito com base no espectro de Fourier de cada amostra de cada classe da imagem.

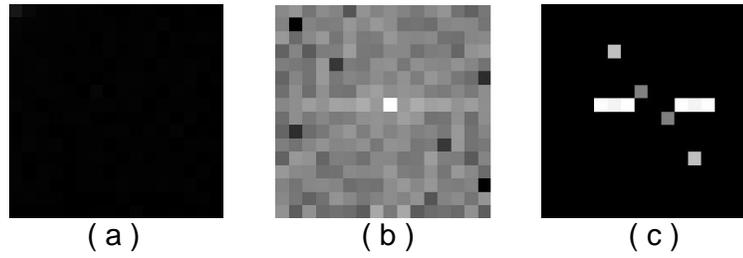


Figura 4. 26. (a) Amostra 1 da classe água na imagem real; (b) Espectro de Fourier da amostra; (c) Freqüências da amostra que apresentam o maior nível de energia.

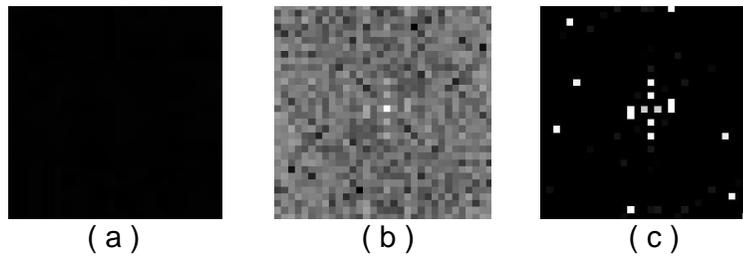


Figura 4. 27. (a) Amostra 2 da classe água na imagem real; (b) Espectro de Fourier da amostra; (c) Freqüências da amostra que apresentam o maior nível de energia.

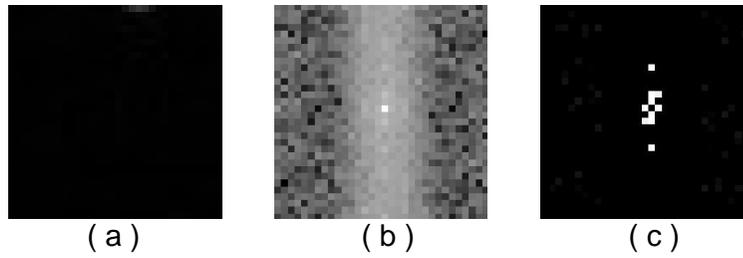


Figura 4. 28. (a) Amostra 3 da classe água na imagem real; (b) Espectro de Fourier da amostra; (c) Freqüências da amostra que apresentam o maior nível de energia.

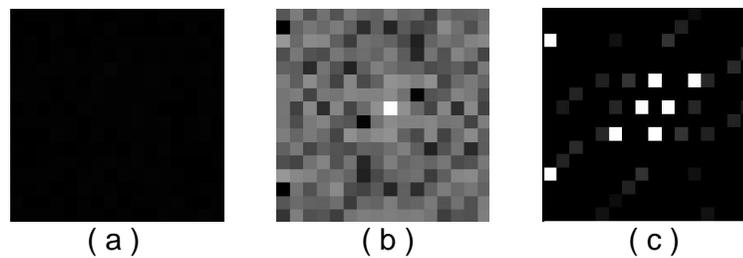


Figura 4. 29. (a) Amostra 4 da classe água na imagem real; (b) Espectro de Fourier da amostra; (c) Freqüências da amostra que apresentam o maior nível de energia.

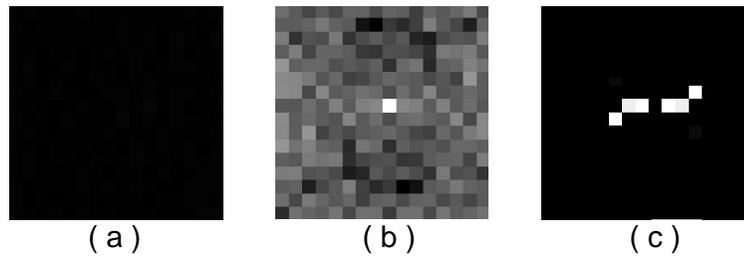


Figura 4. 30. (a) Amostra 5 da classe água na imagem real; (b) Espectro de Fourier da amostra; (c) Freqüências da amostra que apresentam o maior nível de energia.

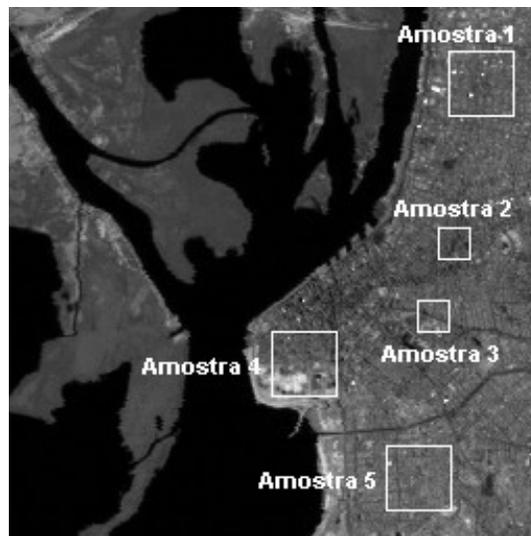


Figura 4. 31. Imagem real com localização das amostras representativas da classe urbana para a seleção de feições.

Nas figuras 4.32 a 4.36 são apresentadas as freqüências mais representativas das amostras da classe urbana, localizadas na figura 4.31, bem como os respectivos espectros de Fourier, onde os valores de u são dados em ciclos/dimensão x da amostra e os valores de v são dados em ciclos/dimensão y da amostra.

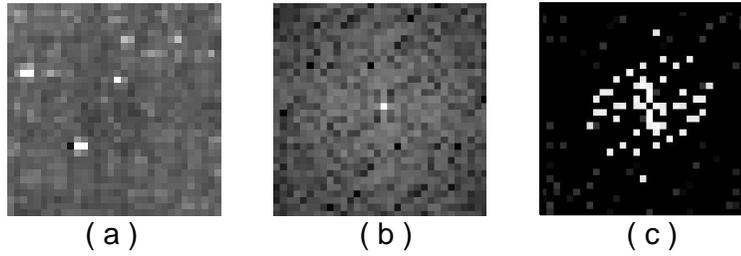


Figura 4. 32. (a) Amostra 1 da classe urbana na imagem real; (b) Espectro de Fourier da amostra; (c) Freqüências da amostra que apresentam o maior nível de energia.

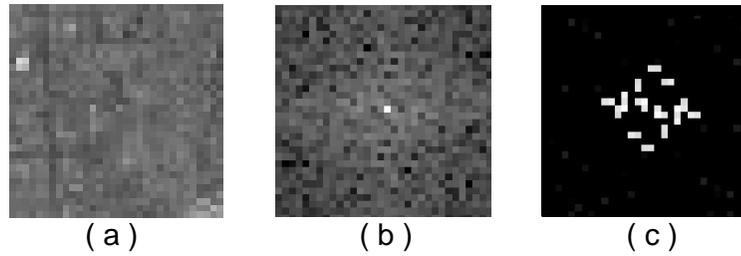


Figura 4. 33. (a) Amostra 2 da classe urbana na imagem real; (b) Espectro de Fourier da amostra; (c) Freqüências da amostra que apresentam o maior nível de energia.

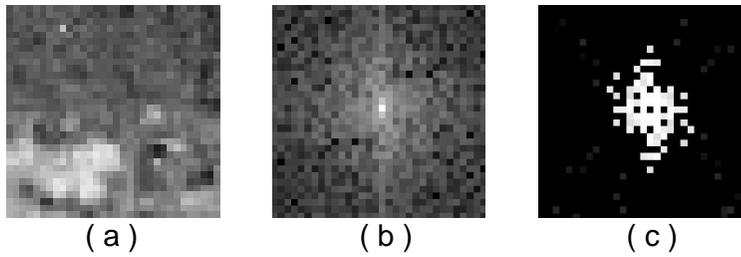


Figura 4. 34. (a) Amostra 3 da classe urbana na imagem real; (b) Espectro de Fourier da amostra; (c) Freqüências da amostra que apresentam o maior nível de energia.

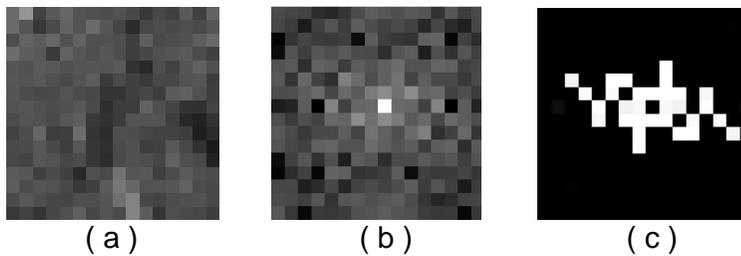


Figura 4. 35. (a) Amostra 4 da classe urbana na imagem real; (b) Espectro de Fourier da amostra; (c) Freqüências da amostra que apresentam o maior nível de energia.

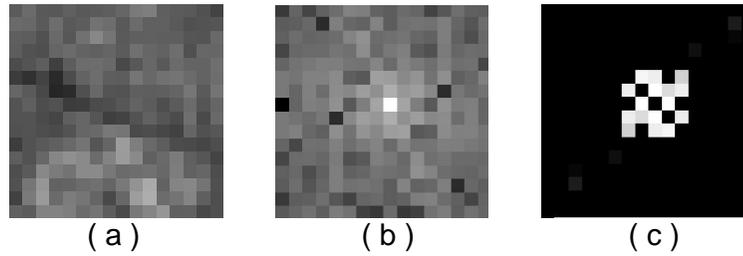


Figura 4. 36. (a) Amostra 5 da classe urbana na imagem real; (b) Espectro de Fourier da amostra; (c) Freqüências da amostra que apresentam o maior nível de energia.

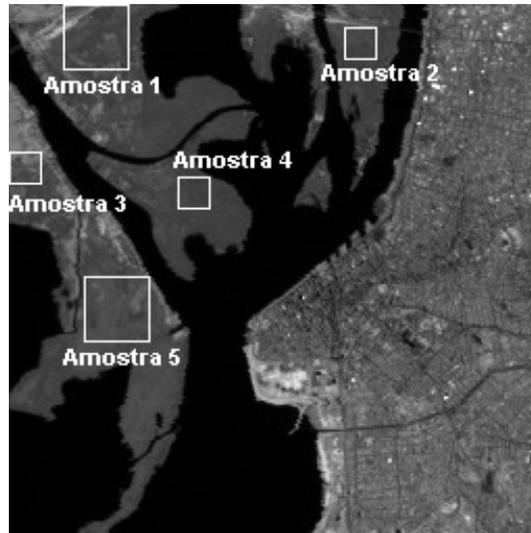


Figura 4. 37. Imagem real com localização das amostras representativas da classe vegetação para a seleção de feições.

Nas figuras 4.38 a 4.42 são apresentadas as freqüências mais representativas das amostras da classe água, localizadas na figura 4.37, bem como os respectivos espectros de Fourier, onde os valores de u são dados em ciclos/dimensão x da amostra e os valores de v são dados em ciclos/dimensão y da amostra.

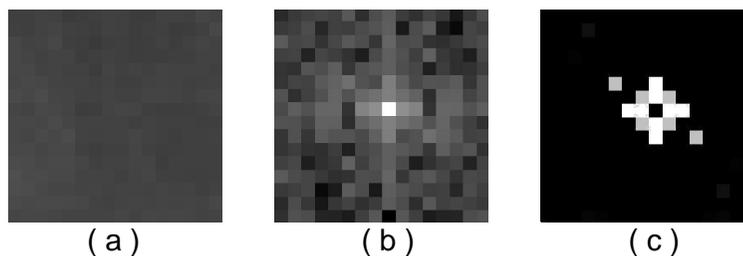


Figura 4. 38. (a) Amostra 1 da classe vegetação na imagem real; (b) Espectro de Fourier da amostra; (c) Freqüências da amostra que apresentam o maior nível de energia.

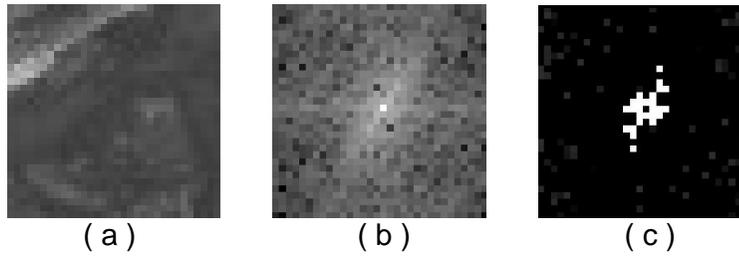


Figura 4. 39. (a) Amostra 2 da classe vegetação na imagem real; (b) Espectro de Fourier da amostra; (c) Freqüências da amostra que apresentam o maior nível de energia.

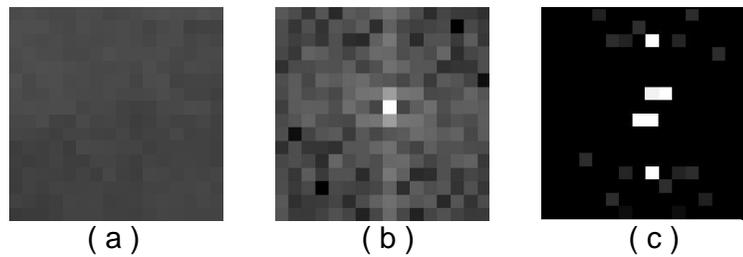


Figura 4. 40. (a) Amostra 3 da classe vegetação na imagem real; (b) Espectro de Fourier da amostra; (c) Freqüências da amostra que apresentam o maior nível de energia.

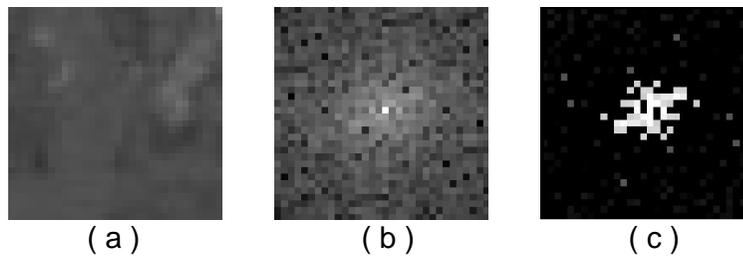


Figura 4. 41. (a) Amostra 4 da classe vegetação na imagem real; (b) Espectro de Fourier da amostra; (c) Freqüências da amostra que apresentam o maior nível de energia.

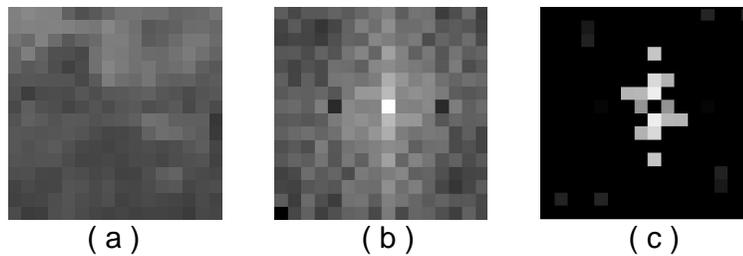


Figura 4. 42. (a) Amostra 5 da classe vegetação na imagem real; (b) Espectro de Fourier da amostra; (c) Freqüências da amostra que apresentam o maior nível de energia.

4. 3. 1. Experimento A

Os parâmetros necessários para a constituição dos filtros de Gabor para este experimento foram obtidos a partir das amostras determinadas na figura 4.25 (amostras da classe água), figura 4.31 (amostras da classe urbana) e figura 4.37 (amostras da classe vegetação). Estes parâmetros podem ser visualizados na tabela 4. 13.

Filtro	Dimensão Do filtro (pixels)	Frequência ao longo do eixo X (k_x)	Frequência ao longo do eixo Y (k_y)	Extensão Espacial (σ)
1	27	0,0380	0,0600	4,5000
2	36	0,0285	0,0280	6,0000
3	17	0,0853	0,0617	2,8333
4	52	0,0194	-0,0296	8,6667
5	47	0,0215	0,0731	7,8333
6	27	0,0503	0,0376	4,5000
7	25	0	0,0405	4,1667
8	64	0	0,0157	10,6667
9	21	0,0621	0,0477	3,5000
10	108	0,0830	0,0093	18,0000
11	96	0,0105	0	16,0000
12	58	0,0173	0	9,6667
13	53	0,0477	0,0190	8,8333
14	15	0,0678	-0,0333	2,5000
15	49	0,0207	0,0384	8,1667
16	25	0,0655	0,0401	4,1667
17	39	0	0,0258	6,5000
18	29	0	0,0346	4,8333

Tabela 4. 13. Parâmetros para a constituição dos filtros de Gabor necessários para a realização do experimento A, utilizando a imagem real.

Os filtros de Gabor definidos na tabela 4.13 geram 18 imagens filtradas ilustradas na figura 4.43. A classificação da imagem pela RNA nos experimentos A1 e A2 ocorrerá através de um processo pixel a pixel onde os parâmetros de entrada da RNA serão os valores do contador digital (CD) de um pixel nas 18 "bandas texturais" (imagens filtradas).

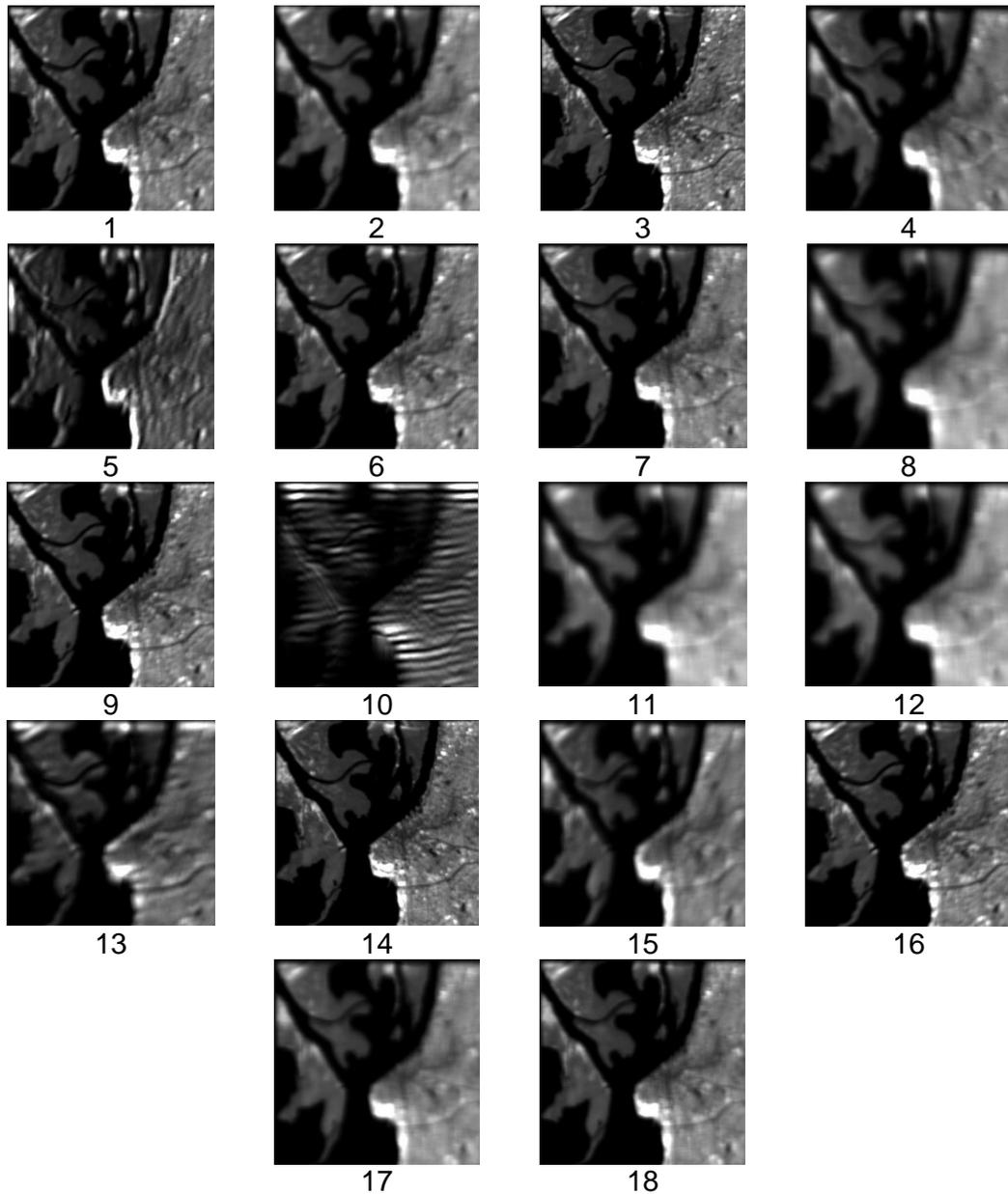


Figura 4. 43. Imagens filtradas pelo filtro de Gabor para o experimento A, a partir da imagem real.

4. 3. 1. 1. Experimento A1

No experimento A1, foram empregadas como parâmetros de entrada as 18 imagens filtradas de através dos filtros de Gabor relativos à tabela 4.13, que geraram as imagens referentes à figura 4.43.

A rede neural artificial utilizada para a classificação da imagem sintética neste experimento usa 18 neurônios na camada de entrada, 20 neurônios na camada intermediária e 3 neurônios na camada de saída.

A figura 4.44 ilustra a imagem obtida após a classificação através da RNA, pixel a pixel, da imagem sintética.



Figura 4. 44. Imagem classificada referente ao experimento A1, utilizando uma imagem real.

A tabela 4.14 avalia os resultados da imagem classificada da figura 4.44 através de uma matriz de erro (Angelo, 2000). Na mesma, é possível vislumbrar uma taxa de exatidão geral do processo de classificação.

IMAGEM REFERÊNCIA	EXATIDÃO	ERRO DE	NÚMERO	IMAGEM CLASSIFICADA		
	(%)	OMISSÃO	DE PIXELS	CLASSE	CLASSE	CLASSE
		(%)	REFERÊNCIA	ÁGUA	URBANA	VEGETAÇÃO
CLASSE ÁGUA	94.43	5.57	22799	21530	375	894
CLASSE URBANA	71.95	28.05	24576	85	17683	6808
CLASSE VEGET.	91.36	8.64	18161	58	1512	16591
Número de pixels Da Imagem Classificada			65536	21673	19570	24293
Erro de Comissão (%)				0.63	7.68	42.41
EXATIDÃO GERAL (%)			85.15			

Tabela 4. 14. Dados da matriz de erro para a avaliação da classificação da imagem real de parte da cidade de Porto Alegre(RS) e seu entorno, referente ao experimento A1.

4. 3. 1. 2. Experimento A2

Em tal experimento são utilizadas como parâmetros de entrada as mesmas imagens filtradas do experimento A1, relativas à figura 4.43.

Da mesma forma que a rede neural artificial empregada no experimento A1, o experimento A2 também utiliza 18 neurônios na camada de entrada e 3 neurônios na camada de saída, porém faz uso de 25 neurônios na camada intermediária.

A figura 4.45 ilustra a imagem obtida após a classificação através da RNA, pixel a pixel, da imagem sintética.



Figura 4. 45. Imagem classificada referente ao experimento A2, utilizando uma imagem real.

A tabela 4.15 avalia os resultados da imagem classificada da figura 4.45 através de uma matriz de erro (Angelo, 2000). Na mesma, é possível vislumbrar uma taxa de exatidão geral do processo de classificação.

IMAGEM REFERÊNCIA	EXATIDÃO (%)	ERRO DE OMISSÃO (%)	NÚMERO DE PIXELS REFERÊNCIA	IMAGEM CLASSIFICADA		
				CLASSE ÁGUA	CLASSE URBANA	CLASSE VEGETAÇÃO
CLASSE ÁGUA	93.04	6.96	22799	21212	532	1055
CLASSE URBANA	69.85	30.15	24576	102	17166	7308
CLASSE VEGET.	89.55	10.45	18161	87	1810	16264
Número de pixels da Imagem Classificada			65536	21401	19508	24627
Erro de Comissão (%)				0.83	9.53	46.05
EXATIDÃO GERAL (%)			83.38			

Tabela 4. 15. Dados da matriz de erro para a avaliação da classificação da imagem real de parte da cidade de Porto Alegre(RS) e seu entorno, referente ao experimento A2.

4. 3. 2. Experimento B

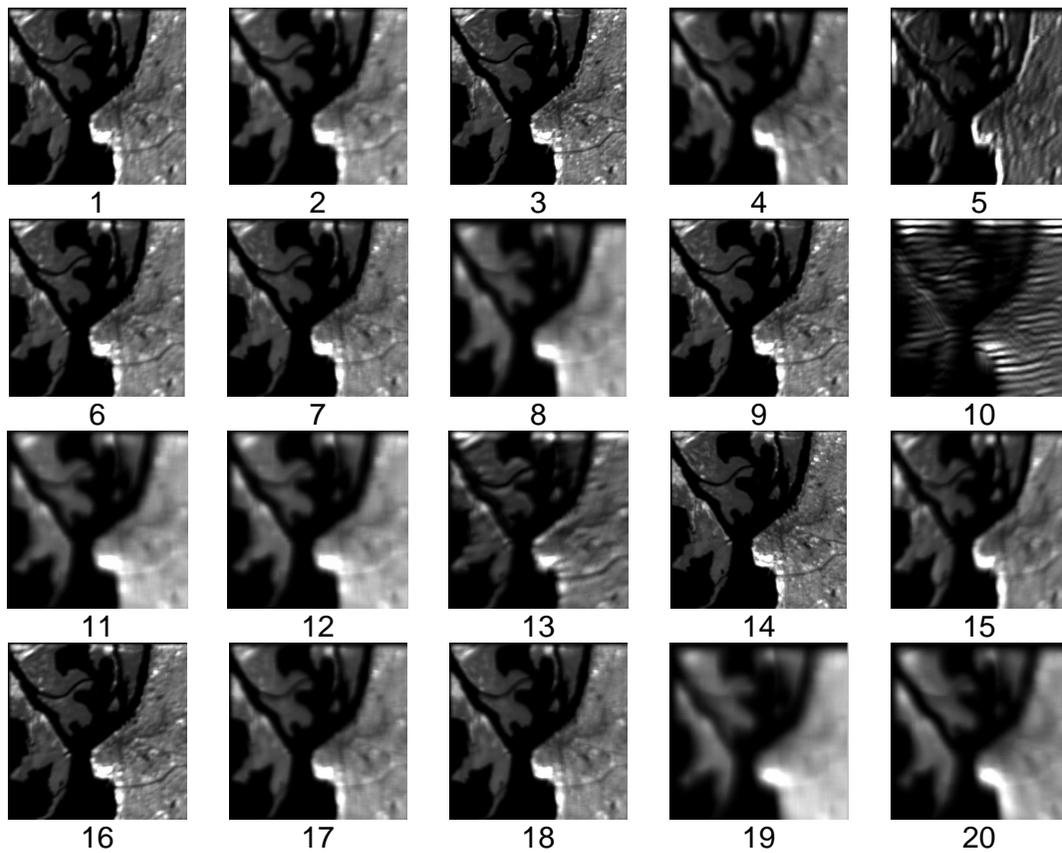
Este experimento tem como objetivo evidenciar o efeito introduzido pelo incremento de frequências espaciais (aumento no número de filtros de Gabor). Para tal, foram utilizadas as mesmas amostras definidas nas figuras 4.25, 4.31 e 4.37, para a extração de frequências. Os parâmetros dos filtros de Gabor utilizados neste experimento estão relacionados na tabela 4.16.

Filtro	Dimensão Do filtro (pixels)	Frequência ao longo do eixo X (k_x)	Frequência ao longo do eixo Y (k_y)	Extensão Espacial (σ)
1	27	0,0380	0,0600	4,5000
2	36	0,0285	0,0280	6,0000
3	17	0,0853	0,0617	2,8333
4	52	0,0194	-0,0296	8,6667
5	47	0,0215	0,0731	7,8333
6	27	0,0503	0,0376	4,5000
7	25	0	0,0405	4,1667
8	64	0	0,0157	10,6667
9	21	0,0621	0,0477	3,5000
10	108	0,0830	0,0093	18,0000
11	96	0,0105	0	16,0000
12	58	0,0173	0	9,6667
13	53	0,0477	0,0190	8,8333
14	15	0,0678	-0,0333	2,5000
15	49	0,0207	0,0384	8,1667
16	25	0,0655	0,0401	4,1667
17	39	0	0,0258	6,5000
18	29	0	0,0346	4,8333
19	94	0,0107	0	15,6667
20	94	0,0107	-0,0150	15,6667
21	22	0,0455	-0,0364	3,6667
22	17	0,0588	0	2,8333
23	15	0,0678	-0,0333	2,5000
24	44	0,0232	-0,0055	7,3333
25	34	0,0298	0	5,6667
26	26	0,0393	0,0410	4,3333
27	8	0,1250	-0,1053	1,3333

28	56	0,0317	0,0179	9,3333
29	39	0,0256	0	6,5000
30	40	0	0,0579	6,6667
31	40	0	0,0789	6,6667
32	50	0	0,0400	8,3333

Tabela 4. 16. Parâmetros para a constituição dos filtros de Gabor necessários para a realização do experimento B, utilizando a imagem real.

Os filtros de Gabor definidos na tabela 4.16 geram 32 imagens filtradas ilustradas na figura 4.46. A classificação da imagem pela RNA nos experimentos B1 e B2 ocorrerá através de um processo pixel a pixel onde os parâmetros de entrada da RNA serão os valores do contador digital (CD) de um pixel nas 32 "bandas texturais" (imagens filtradas).



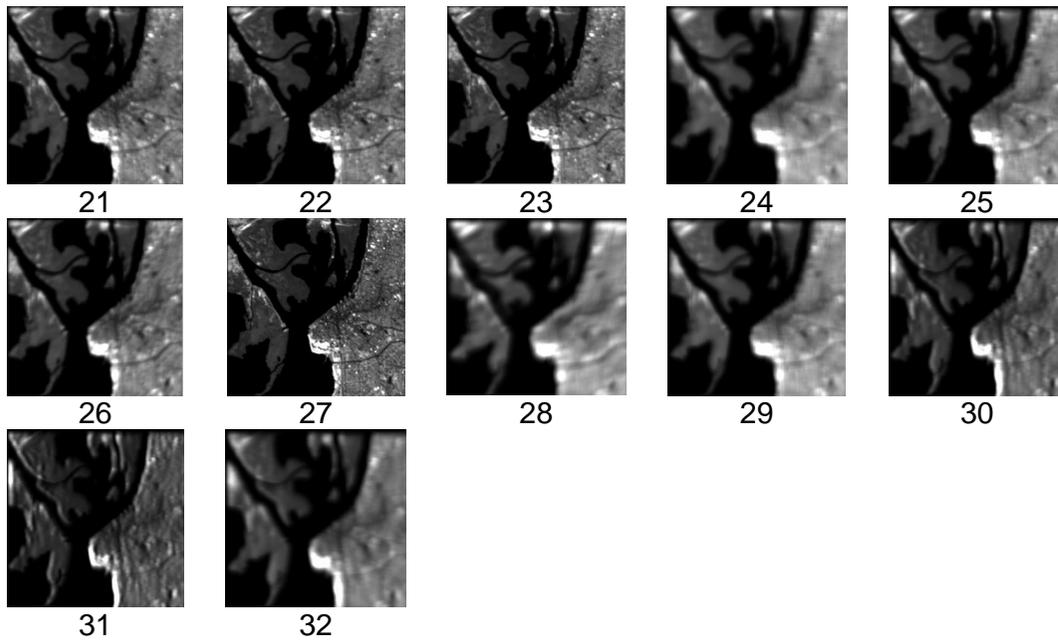


Figura 4. 46. Imagens filtradas pelo filtro de Gabor para o experimento B, a partir da imagem real.

4. 3. 2. 1. Experimento B1

No experimento B1, foram empregadas como parâmetros de entrada as 32 imagens filtradas de através dos filtros de Gabor relativos à tabela 4.16, que geram as imagens referentes à figura 4.46.

A rede neural artificial utilizada para a classificação da imagem sintética neste experimento usa 32 neurônios na camada de entrada, 35 neurônios na camada intermediária e 3 neurônios na camada de saída.

A figura 4.47 ilustra a imagem obtida após a classificação através da RNA, pixel a pixel, da imagem sintética.



Figura 4. 47. Imagem classificada referente ao experimento B1, utilizando uma imagem real.

A tabela 4.17 avalia os resultados da imagem classificada da figura 4.47 através de uma matriz de erro (Angelo, 2000). Na mesma, é possível vislumbrar uma taxa de exatidão geral do processo de classificação.

IMAGEM REFERÊNCIA	EXATIDÃO (%)	ERRO DE OMISSÃO (%)	NÚMERO DE PIXELS REFERÊNCIA	IMAGEM CLASSIFICADA		
				CLASSE ÁGUA	CLASSE URBANA	CLASSE VEGETAÇÃO
CLASSE ÁGUA	93.65	6.35	22799	21351	275	1173
CLASSE URBANA	72.28	27.72	24576	446	17764	6366
CLASSE VEGET.	82.59	17.41	18161	1211	1951	14999
Número de pixels Da Imagem Classificada			65536	23008	19990	22538
Erro de Comissão (%)				7.27	9.06	41.51
EXATIDÃO GERAL (%)			82.57			

Tabela 4. 17. Dados da matriz de erro para a avaliação da classificação da imagem real de parte da cidade de Porto Alegre(RS) e seu entorno, referente ao experimento B1.

4. 3. 1. 2. Experimento B2

Em tal experimento são utilizadas como parâmetros de entrada as mesmas imagens filtradas do experimento B1, relativas à figura 4.46.

Da mesma forma que a rede neural artificial empregada no experimento B1, o experimento B2 também utiliza 32 neurônios na camada de entrada e 3 neurônios na camada de saída, porém faz uso de 40 neurônios na camada intermediária.

A figura 4.48 ilustra a imagem obtida após a classificação através da RNA, pixel a pixel, da imagem sintética.

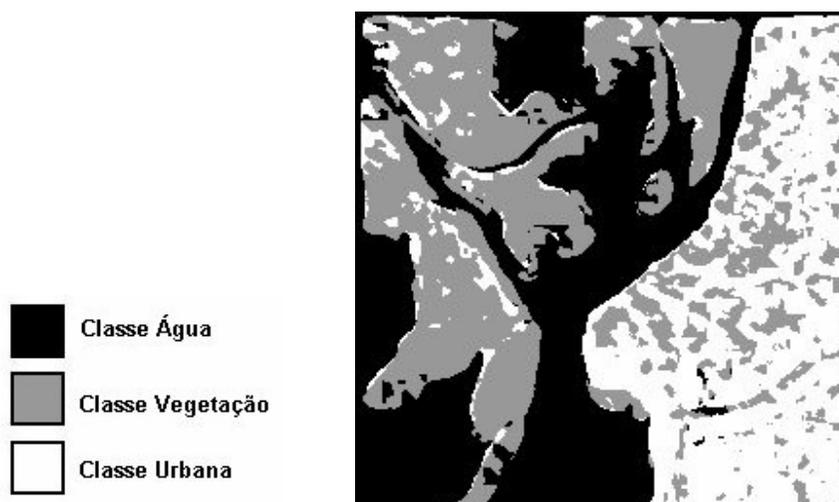


Figura 4. 48. Imagem classificada referente ao experimento B2, utilizando uma imagem real.

A tabela 4.18 avalia os resultados da imagem classificada da figura 4.48 através de uma matriz de erro (Angelo, 2000). Na mesma, é possível vislumbrar uma taxa de exatidão geral do processo de classificação.

IMAGEM REFERÊNCIA	EXATIDÃO (%)	ERRO DE OMISSÃO (%)	NÚMERO DE PIXELS REFERÊNCIA	IMAGEM CLASSIFICADA		
				CLASSE ÁGUA	CLASSE URBANA	CLASSE VEGETAÇÃO
CLASSE ÁGUA	91.18	8.82	22799	20789	775	1235
CLASSE URBANA	70.66	29.34	24576	2009	17366	5201

CLASSE VEGET.	87.57	12.43	18161	712	1546	15903
Número de pixels Da Imagem Classificada			65536	23510	19687	22339
Erro de Comissão (%)				11.93	9.44	35.44
EXATIDÃO GERAL (%)			82.49			

Tabela 4. 18. Dados da matriz de erro para a avaliação da classificação da imagem real de parte da cidade de Porto Alegre(RS) e seu entorno, referente ao experimento B2.

4. 4. Resultados

Os resultados obtidos através dos experimentos relatados neste capítulo foram organizados na forma da tabela 4.19. Tais resultados são discutidos no capítulo 5.

	EXPERIMENTO	NÚMERO DE FILTROS	EXTENSÃO ESPACIAL	REDE NEURAL			EXATIDÃO GERAL
				Entrada	Oculto	Saída	
IMAGEM SINTÉTICA	A1	15	específica	15	18	4	88,65%
	A2	15		15	23	4	82,94%
	B1	15	única	15	18	4	84,24%
	B2	15		15	23	4	81,55%
	C1	25	específica	25	28	4	83,06%
	C2	25		25	32	4	76,88%
	D1	25	única	25	28	4	89,05%
	D2	25		25	32	4	87,37%
REAL	A1	18	específica	18	20	3	85,15%
	A2	18		18	25	3	83,38%
	B1	32	específica	32	35	3	82,57%
	B2	32		32	40	3	82,49%

Tabela 4. 19. Comparativo dos resultados dos experimentos realizados com imagens sintéticas e reais.

4. 5. Considerações finais

A filtragem através dos filtros de Gabor, bem como a classificação das imagens filtradas decorrentes pelas redes neurais artificiais, foi processada no software MATLAB for Windows, desenvolvido pela Math Works, Inc., versão 5.3.

No âmbito geral da etapa de aplicação dos filtros de Gabor na imagem original, este processo gera um número de imagens filtradas igual ao número de filtros aplicados. Portanto, cada pixel passa a ser descrito por um vetor cuja dimensionalidade é igual ao número de filtros utilizados. Da mesma forma que em imagens multi-espectrais, as imagens resultantes da filtragem podem ser denominadas de "canais ou bandas texturais"(Angelo, 2000).

A constituição das RNAs utilizadas na classificação das imagens segue a topologia do modelo apresentado na figura 2.9. Para tal, em cada experimento são especificados os números de neurônios utilizados em cada camada (entrada, intermediária e saída). Para o aprendizado da RNA utilizada em cada situação foi usado um coeficiente de aprendizado igual a 0,01 e um número de iterações em torno de 30 000.

Para a ativação do processo de aprendizado de cada RNA torna-se necessário a seleção de um conjunto de amostras das classes de padrões a serem aprendidos, bem como suas saídas correspondentes. Esse conjunto de amostras é chamado de conjunto de treinamento da rede.

Os pesos de cada neurônio, tanto da camada intermediária quanto da camada de saída, foram inicializados com valores aleatórios pequenos(entre -1 e 1). A função de ativação selecionada para as RNAs utilizadas foi a função tangente hiperbólica, pois converge mais rapidamente. Dessa forma, foi necessário normali-

zar as entradas das redes para o intervalo $[-1;1]$. Esta normalização das entradas é necessária para que as entradas e saídas da rede sejam compatíveis, ocupando o mesmo espaço vetorial. O algoritmo de BackPropagation utilizado nas RNAs baseia-se no Gradiente descendente do Erro (minimização do erro médio quadrado).

É importante citar que a camada de entrada da rede neural tem como objetivo, apenas, o de repassar os padrões de entrada para os neurônios intermediários.

O processo de classificação realizado pelas RNAs é feito diretamente a partir dos níveis de cinza dos pixels de cada uma das imagens decorrentes do processamento pelos filtros de Gabor, que extraíram as frequências espaciais associadas às diferentes classes presentes nas imagens. Essa forma de classificação pixel a pixel, através dos níveis de cinza é desenvolvido por Liberman(1997). Cabe citar que seria possível implementar um processo de classificação que empregasse parâmetros de textura, tais como média, desvio padrão e uniformidade(Liberman, 1997).

Portanto, cada neurônio da camada de entrada recebe a informação do mesmo pixel referente a cada uma das imagens filtradas de Gabor

O número de neurônios da camada de saída da rede será igual ao número de classes existentes na imagem. Havendo 4 classes definidas, teremos uma camada de saída com quatro neurônios. A associação da saída a cada classe é determinada pelo neurônio que resultar no valor mais alto.

CAPÍTULO 5

CONCLUSÕES E SUGESTÕES

O objetivo deste trabalho é o de apresentar uma metodologia de classificação de imagens digitais obtidas por sensoriamento remoto através de textura utilizando redes neurais. Tal processo consiste, basicamente, de duas etapas:

- A aplicação de filtros de Gabor na imagem, com o propósito de selecionar as frequências características a cada classe presente na imagem;
- A classificação das imagens resultantes da filtragem através de uma rede BPN.

A metodologia proposta foi aplicada, inicialmente, em imagens sintéticas compostas por mosaicos naturais, retiradas de um álbum fotográfico (Brodatz, 1966). Esta etapa serviu para o estudo das configurações e resultados dos filtros de Gabor e dos algoritmos BPN. Posteriormente, tais dados foram utilizados para a composição dos experimentos com imagens reais obtidas por sensoriamento remoto.

Os experimentos apresentados no capítulo 4 são subdivididos em três outras etapas. Uma primeira etapa, dedicada ao estudo dos parâmetros de formação dos filtros de Gabor; numa segunda etapa, a realização de testes com filtros em imagens formadas por texturas naturais, bem como a classificação das imagens filtradas através de RNAs; finalmente, a aplicação da metodologia de filtragem e classificação em imagens reais.

Quanto aos resultados sobre as imagens sintéticas formadas por mosaicos de Brodatz é impossível deixar de citar que a escolha das amostras para a obtenção de frequências espaciais relativas a cada classe é determinante para o proces-

so em geral. A princípio, quanto mais representativa for a amostra de cada classe, melhor seria a seleção de feições. Porém, quando as imagens tratam de mosaicos naturais, além de acontecer um número muito extenso de freqüências espaciais em cada classe, existe a possibilidade da ocorrência de uma mesma freqüência em classes distintas, mas com amplitudes distintas, resultando em erro no momento da classificação.

No que tange ao processo de filtragem da imagem original, os resultados obtidos nos experimentos mostram que, ao utilizar um pequeno número de freqüências espaciais, algumas freqüências não consideradas influenciam nos resultados intra-classe. Na tentativa de amenizar esse contratempo, foi usado um valor único para a extensão espacial, o que ocasionou uma deterioração no processo geral de classificação. Portanto, sugere-se o uso de um número maior de freqüências espaciais aliado a um parâmetro de extensão espacial adequado, diminuindo a perda de informação (Angelo, 2000).

Os problemas acima explicitados podem ser verificados nas imagens classificadas a partir dos mosaicos de Brodatz, onde ocorrem dificuldades na separação entre as classes 2 e 3. O mesmo acontece entre as classes urbana e vegetação nos experimentos envolvendo a imagem real de Porto Alegre e seu entorno.

Quanto à aplicação da metodologia na imagem de Porto Alegre, o número de classes foi limitado às três mais representativas, possibilitando que a presença de outras texturas não tão evidentes incorressem em erros no momento da classificação.

Da mesma forma, apesar da rede neural MLP com algoritmo BackPropagation ser um excelente classificador, o mesmo não ocorre quando as entradas escolhidas para esta rede não estiverem bem distribuídas. No intuito de reduzir o erro

acarretado pelas entradas, as mesmas passam por um processo de escalonamento linear, proposto por Liberman(1997).

Outro fator analisado nos experimentos é a influência do número de neurônios na camada intermediária sobre a imagem classificada resultante do processo. Tanto nas imagens sintéticas quanto nas imagens reais foram utilizados diferentes números de neurônios. Em quase todas as situações, foi observado que um número maior de neurônios na camada intermediária implica num decréscimo na exatidão geral da classificação. Como não existe uma regra para a determinação da quantidade de neurônios a ser utilizada na camada intermediária, a diferença na exatidão da classificação é ocasionada pela possibilidade da existência de neurônios excedentes à rede.

De uma forma geral, a metodologia composta por uma etapa de extração de freqüências espaciais de cada classe presente na imagem (filtragem de Gabor), aliada a um processo de classificação baseado em redes neurais MLP com algoritmo BackPropagation obteve bons resultados, instigando a continuidade de pesquisas envolvendo o uso de RNAs na classificação de imagens obtidas por sensoria-mento remoto.

Algumas sugestões de estudos podem ser propostos, entre eles a utilização da filtragem através de filtros de Gabor para a seleção de freqüências espaciais de cada classe da imagem, porém utilizando parâmetros estatísticos (média e desvio padrão, por exemplo) como entradas da rede neural. Também é possível o uso de diferentes tipos de redes neurais com a função de classificador.

Também não se descarta a possibilidade de uso, concomitantemente, de atributos texturais, espectrais e de contexto, para uma melhor exatidão do processo de classificação.

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
CENTRO ESTADUAL DE PESQUISA EM SENSORIAMENTO REMOTO
E METEOROLOGIA - CEPSRM
CURSO DE PÓS-GRADUAÇÃO EM SENSORIAMENTO REMOTO

**CLASSIFICAÇÃO DE IMAGENS DE SENSORIAMENTO REMOTO
BASEADA EM TEXTURA POR REDES NEURAI**

- APÊNDICES -

por

Adriano Beluco

Porto Alegre, RS

Julho de 2002

APÊNDICE 1

TRANSFORMADA DE FOURIER

Neste apêndice são apresentados, de forma sucinta, os principais conceitos e propriedades da transformada de Fourier, ferramenta matemática para a especificação e projeto de filtros no domínio de frequências.

Transformada de Fourier para sinais unidimensionais contínuos

Seja $f(x)$ uma função contínua de variável real x . A transformada de Fourier de $f(x)$, indicada por $\mathfrak{T}\{f(x)\}$, é definida pela equação

$$\mathfrak{T}\{f(x)\} = F(u) = \int_{-\infty}^{+\infty} f(x) \exp[-j2\pi ux] dx \quad (A1)$$

onde $j = \sqrt{-1}$.

Dado $F(u)$, $f(x)$ pode ser obtida calculando-se a transformada inversa de Fourier.

$$\mathfrak{T}^{-1}\{F(u)\} = f(x) = \int_{-\infty}^{+\infty} F(u) \exp[j2\pi ux] du \quad (A2)$$

As equações A1 e A2, chamadas conjuntamente de “par de Fourier”, existem se $f(x)$ for contínua e integrável e $F(u)$ for integrável.

A transformada de Fourier de uma função real é, geralmente, um valor complexo, ou seja:

$$F(u) = R(u) + j I(u) \quad (A3)$$

onde $R(u)$ e $I(u)$ são, respectivamente, os componentes reais e imaginários de $F(u)$. $F(u)$ também pode ser representada exponencialmente através da equação:

$$F(u) = |F(u)| e^{j\phi(u)} \quad (A4)$$

onde a componente de magnitude $|F(u)|$ é denominada “espectro de Fourier” de $f(x)$ e $\phi(u)$ seu ângulo de fase. A figura A1 mostra um exemplo de função unidimensional simples (pulso retangular de amplitude A) e seu respectivo espectro de Fourier.

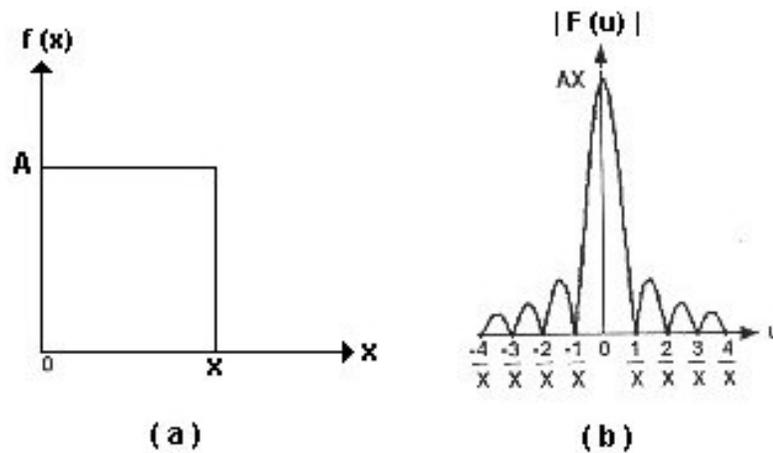


Figura A1. Função unidimensional contínua simples (a) e seu respectivo espectro de Fourier (b) (MARQUES e VIEIRA, 1999).

Transformada de Fourier para sinais bidimensionais contínuos

O conceito de transformada de Fourier pode ser estendido para uma função de duas variáveis $f(x,y)$. Se $f(x,y)$ é contínua e integrável e $F(u,v)$ é integrável, então o par de Fourier a seguir existe:

$$\mathfrak{F}\{f(x,y)\} = F(u,v) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x,y) \exp[-j2\pi(ux + vy)] dx dy \quad (A5)$$

e

$$\mathfrak{F}^{-1}\{F(u,v)\} = f(x,y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} F(u,v) \exp[j2\pi(ux + vy)] dx dy \quad (A6)$$

Da mesma forma que o caso unidimensional, $F(u,v)$ é complexa e sua amplitude, $|F(u,v)|$, é denominada espectro de Fourier.

A figura A2 mostra uma função bidimensional contínua e seu respectivo espectro de Fourier, representado de duas maneiras: em perspectiva tridimensional (figura A2 b) e como uma função de intensidade, na qual o brilho é proporcional à amplitude de $|F(u,v)|$ (figura A2 c).

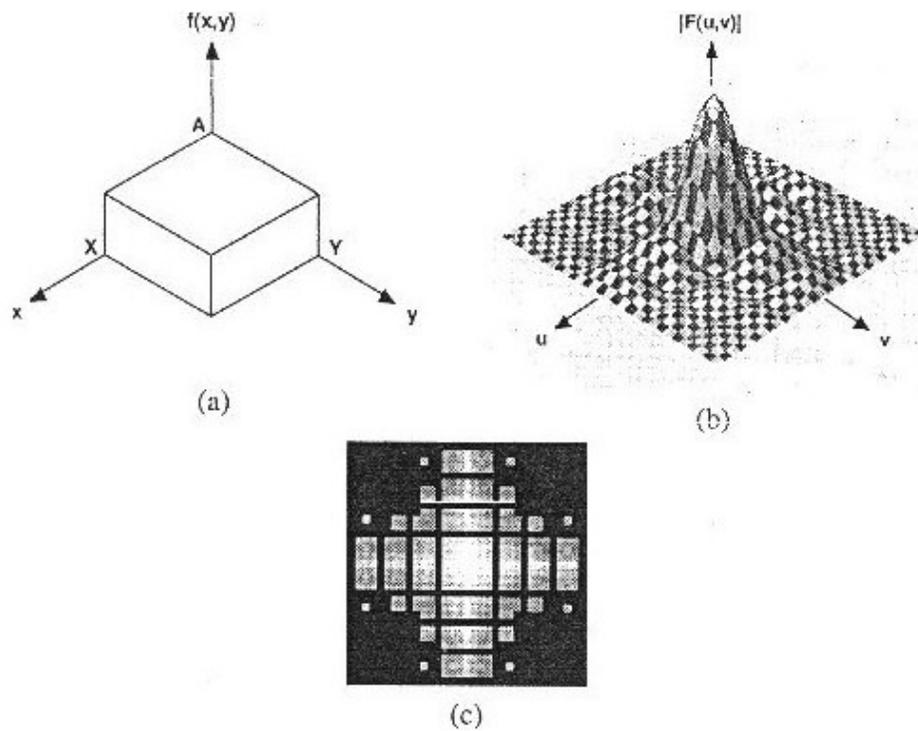


Figura A2. Função bidimensional contínua (a), seu espectro de Fourier (b) e o espectro representado como uma função de intensidade (c) (MARQUES e VIEIRA, 1999).

Transformada de Fourier para sinais unidimensionais discretos

Supondo que uma função contínua $f(x)$ seja discretizada, produzindo a seqüência

$$\{f(x_0), f(x_0 + \Delta x), f(x_0 + 2\Delta x), \dots, f(x_0 + [N-1]\Delta x)\} \quad (A7)$$

através da amostragem de N pontos, espaçados entre si de Δx .

Denominando a função discreta $f(n)$,

$$f(n) = f(x_0 + n\Delta x) \quad (A8)$$

onde n pode assumir os valores discretos $0, 1, 2, \dots, N - 1$, podemos interpretar a seqüência $\{f(0), f(1), f(2), \dots, f(N-1)\}$ como qualquer seqüência de N amostras consecutivas do sinal original contínuo $f(x)$, uniformemente espaçadas.

A transformada de Fourier de $f(n)$ será

$$F(u) = \frac{1}{N} \sum_{n=0}^{N-1} f(n) \exp\left[\frac{-j2\pi un}{N}\right] \quad (A9)$$

para $u = 0, 1, 2, \dots, N-1$ e a transformada inversa de Fourier de $F(u)$ será

$$f(n) = \frac{1}{N} \sum_{u=0}^{N-1} F(u) \exp\left[\frac{j2\pi un}{N}\right] \quad (A10)$$

para $n = 0, 1, 2, \dots, N-1$.

Os valores de $u = 0, 1, 2, \dots, N-1$ na transformada discreta de Fourier(A9) correspondem a amostras dos valores da transformada do sinal contínuo nos pontos $0, \Delta u, 2\Delta u, \dots, (N-1)\Delta u$. Ou seja, $F(u)$ representa $F(u\Delta u)$. Portanto, os intervalos de espaçamento entre as amostras do sinal e de sua transformada estão relacionados através da expressão

$$\Delta u = \frac{1}{N\Delta x} \quad (A11)$$

Transformada de Fourier para sinais bidimensionais discretos

Seja uma função contínua bidimensional $f(x,y)$, discretizada em M amostras ao longo de x e N amostras ao longo de y . Neste caso, sua transformada discreta de Fourier será

$$F(u, v) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) \exp \left[-j2\pi \left(\frac{ux}{M} + \frac{vy}{N} \right) \right] \quad (A12)$$

para $u = 0, 1, 2, \dots, M-1$ e $v = 0, 1, 2, \dots, N-1$.

A transformada inversa é dada por

$$f(x, y) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u, v) \exp \left[j2\pi \left(\frac{ux}{M} + \frac{vy}{N} \right) \right] \quad (A13)$$

O processo de amostragem da função bidimensional contínua $f(x,y)$ pode ser visto como uma grade 2D, com impulsos unitários espaçados de Δx e Δy , nos eixos x e y , respectivamente. A função $f(x,y)$ discretizada (equações A12 e A13) representa as amostras de $f(x,y)$ contínua original em pontos espaçados entre si de Δx e Δy , nos respectivos eixos. Os intervalos de amostragem nos domínios espacial e de frequências estão relacionados entre si por

$$\Delta u = \frac{1}{M\Delta x} \quad (A14)$$

e

$$\Delta v = \frac{1}{N\Delta y} \quad (A15)$$

Para o caso particular em que $M = N$, as equações A12 e A13 podem ser descritas como

$$F(u, v) = \frac{1}{N} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) \exp \left[-j2\pi \left(\frac{ux + vy}{N} \right) \right] \quad (A16)$$

e

$$f(x, y) = \frac{1}{N} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} F(u, v) \exp \left[j2\pi \left(\frac{ux + vy}{N} \right) \right] \quad (A17)$$

Propriedades da transformada de Fourier para sinais bidimensionais discretos

Existem diversas propriedades da transformada de Fourier para sinais bidimensionais de grande interesse para o processamento de imagens, como a separabilidade, translação, periodicidade, entre outras.

Separabilidade

O par de Fourier das equações A16 e A17 pode ser decomposto em:

$$F(u, v) = \frac{1}{N} \sum_{x=0}^{N-1} \exp \left[-\frac{j2\pi ux}{N} \right] \sum_{y=0}^{N-1} f(x, y) \exp \left[-\frac{j2\pi vy}{N} \right] \quad (A18)$$

para $u, v = 0, 1, 2, \dots, N-1$ e

$$f(x, y) = \frac{1}{N} \sum_{u=0}^{N-1} \exp \left[\frac{j2\pi ux}{N} \right] \sum_{v=0}^{N-1} F(u, v) \exp \left[\frac{j2\pi vy}{N} \right] \quad (A19)$$

para $x, y = 0, 1, 2, \dots, N-1$.

Tal propriedade nos mostra que se pode obter $f(x, y)$ e $F(u, v)$ através da aplicação sucessiva das transformadas discretas unidimensionais de Fourier diretas ou inversas.

Translação

As propriedades de translação do par de Fourier bidimensional são resumidas nas relações

$$f(x, y) \exp\left[\frac{j2\pi(u_o x + v_o y)}{N}\right] \Leftrightarrow F(u - u_o, v - v_o) \quad (A20)$$

e

$$f(x - x_o, y - y_o) \Leftrightarrow F(u, v) \exp\left[\frac{-j2\pi(ux_o + vy_o)}{N}\right] \quad (A21)$$

onde o sinal de equivalência indica a correspondência entre uma função e sua transformada de Fourier e vice-versa.

Periodicidade

A transformada discreta de Fourier e sua inversa são periódicas, com período N . Ou seja,

$$F(u, v) = F(u + N, v) = F(u, v + N) = F(u + N, v + N) \quad (A22)$$

Rotação

A propriedade da rotação estabelece que, se uma imagem $f(x, y)$ for rotacionada de um certo ângulo θ_o , sua transformada $F(u, v)$ será rotacionada do mesmo ângulo, como mostra a figura A3.

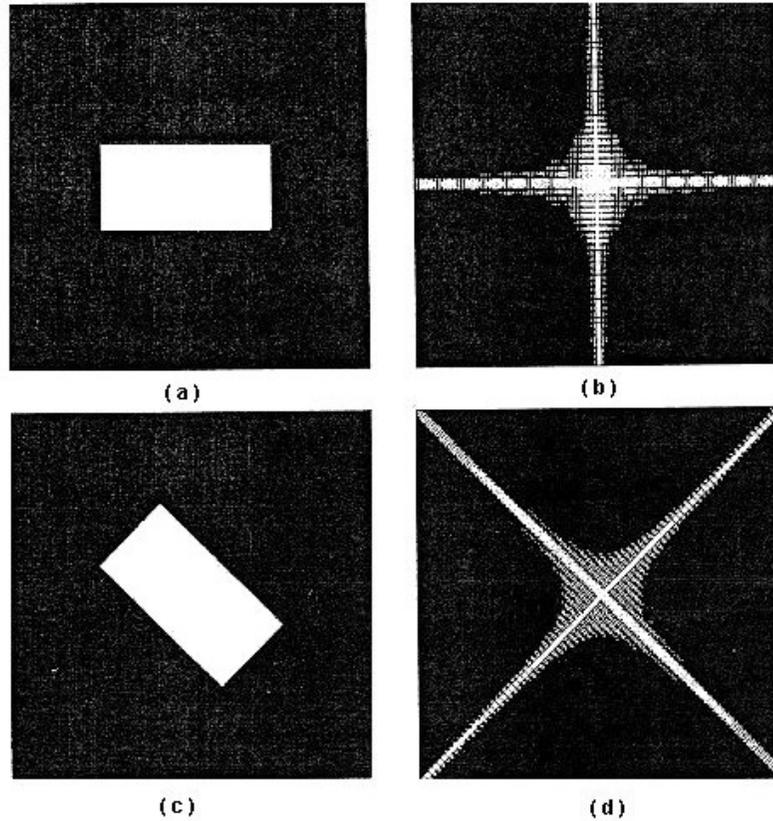


Figura A3. Imagem original (a), transformada de Fourier de a (b), imagem rotacionada (c) e transformada de Fourier de c (d) (MARQUES e VIEIRA, 1999).

Valor médio

O valor médio de uma função bidimensional $f(x,y)$ é dado por

$$\bar{f}(x,y) = \frac{1}{N^2} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x,y) \quad (A23)$$

Substituindo $u = v = 0$ na equação A18, obtemos

$$F(0,0) = \frac{1}{N} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x,y) \quad (A24)$$

Logo, o valor médio de uma função bidimensional está relacionado à sua transformada de Fourier através da relação

$$\bar{f}(x,y) = \frac{1}{N} F(0,0) \quad (A25)$$

Convolução

O processo de convolução entre duas funções $f(x,y)$ e $g(x,y)$ é representado por $f(x,y)*g(x,y)$ e definida é para os casos contínuo e discreto através de

$$f(x,y) * g(x,y) = \int \int_{-\infty}^{+\infty} f(\alpha, \beta) g(x - \alpha, y - \beta) d\alpha d\beta \quad (A26)$$

e

$$f(x,y) * g(x,y) = \sum_{M=0}^{M-1} \sum_{N=0}^{N-1} f(m,n) g(x-m, y-n) \quad (A27)$$

Onde as seguintes equivalências constituem importante resultado derivado do teorema da convolução:

$$f(x,y)*g(x,y) \Leftrightarrow F(u,v)G(u,v) \quad (A28)$$

e

$$f(x,y)g(x,y) \Leftrightarrow F(u,v)*G(u,v) \quad (A29)$$

Transformada Rápida de Fourier

Trata-se de um algoritmo cujo principal objetivo é reduzir o custo computacional do cálculo da transformada de Fourier de N pontos, substituindo o processo convencional de cálculo, no qual o número de multiplicações e adições é proporcional a N^2 por um arranjo que combina diversas transformadas parciais, cada qual com pequeno número de pontos, em que o número de adições e multiplicações é proporcional a $N \log_2 N$.

APÊNDICE 2

Microtexturas de Brodatz

São pequenas imagens com resolução 128 x 128 pixels, derivadas de texturas idealizadas por Brodatz, apresentadas na obra *Textures: A photogrametric album for artists and designers*.

Algumas microtexturas foram utilizadas como imagens de teste para as rotinas de classificação de imagens digitais apresentadas no capítulo 4. Abaixo, é possível visualizar algumas microtexturas de Brodatz.

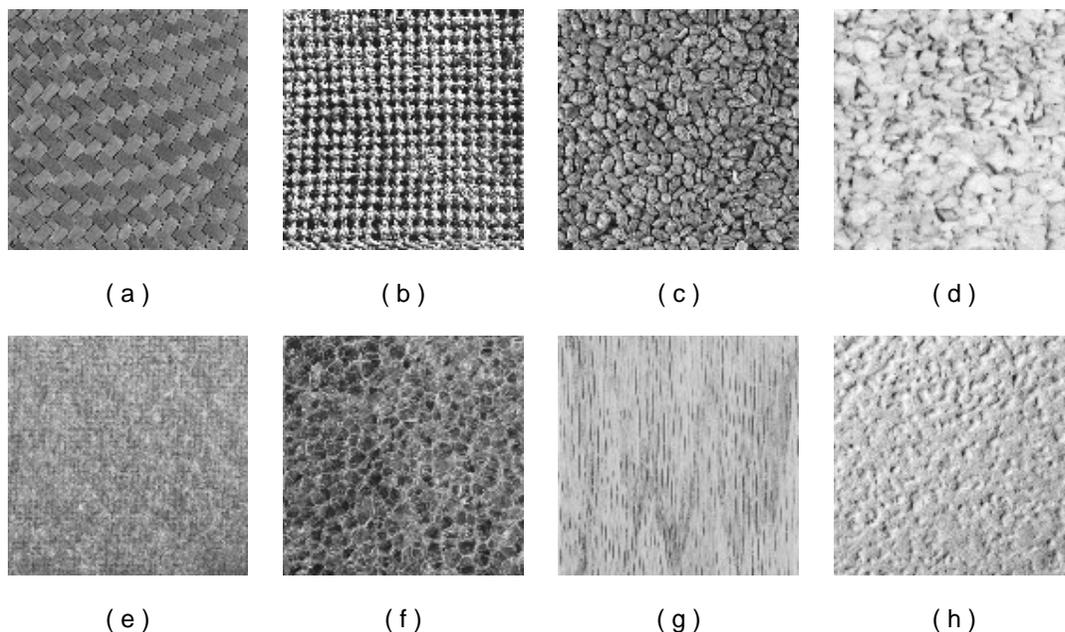


FIGURA A4. Diversas microtexturas de Brodatz: Fabric1 (a), Fabric2 (b), Food (c), Gravel (d), Woolencloth (e), Plastic Bubs (f), Wood (g) e Roughwall (h). Adaptado de BRODATZ(1966).

BIBLIOGRAFIA

ANGELO, N.P. **Aplicação de filtros de Gabor no processo de classificação de imagens digitais com base em atributos de textura.** Porto Alegre: CEP SRM/CPGSR-UFRGS, 2000.

ATKINSON, P. M.; TATNALL, A. R. L. Neural networks in remote sensing. **International Journal of Remote Sensing**, V. 18, n° 4, p. 699-710, 1997.

BARBOSA, V. C.; MACHADO, R. J.; LIPORACE, F. A neural system for deforestation monitoring on Landsat images of the amazon region. **International Journal of Approximate Reasoning**, V. 11, p. 321-359, 1994.

BISCHOF, H.; SCHNEIDER, W.; PINZ, J. Multispectral classification of Landsat images using neural networks. **IEEE Transactions on Geoscience and Remote Sensing**, V. 30, n° 3, 1992.

BOVIK, A.C.; CLARK, M.; GEISLER, W.S.. Multichannel texture analysis using localized spatial filters. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, V. 12, n° 1, p. 55-73 1990.

BRODATZ, P. T. **Textures: A photogrametric album for artists and designers.** New York: Dover, 1966.

CLARO, Luís O. T. **Texturas de imagens utilizando conceitos de morfologia matemática.** Porto Alegre: CEP SRM/CPGSR-UFRGS, 1995.

CRACKNELL, A. P.; HAYES, L. W. B. **Introduction to Remote Sensing**. London: Taylor & Francis, 1991.

CRÓSTA, Álvaro. P. **Processamento Digital de Imagens de Sensoriamento Remoto**. Campinas: IG/UNICAMP, 1992.

DAUGMAN, J. D. Complete discrete 2-d Gabor transforms by neural networks for image analysis and compression. **IEEE Trans. Acoustics, Speech and Signal Processing**, V. 36, p. 1169-1179, 1988.

DOUGHERTY, E. **An introduction to morphological image processing**. New York: SPIE, 1992.

DOUGHERTY, E., PELZ, J., NEWELL, J. Morphological texture-based maximum-likelihood pixel classification based on local granulometric moments. **Pattern Recognition**, Vol. 25, 1992.

FOODY, G. M.; ARORA, M. K. An evaluation of some factors affecting the accuracy of classification by an artificial neural network. **International Journal of Remote Sensing**, V. 18, n^o 4, p. 799-810, 1997.

FREEMAN, J. A.; SKAPURA, D. M. **Neural Networks, Algorithms, Applications and Programming Techniques**. New York: Addison-Wesley, 1991.

FUKUNAGA, K. **Introduction to statistical pattern recognition**. San Diego: Academic Press Inc., 1990.

GARCIA, Gilberto. **Sensoriamento Remoto: princípios e interpretação de imagens**. São Paulo: Nobel, 1982.

GONZALEZ, R. F.; WOODS, R. E. **Digital image processing**. New York: Addison-Wesley, 1992.

HARALICK, Robert M. Statistical and structural approach to texture. **Proceedings of the IEEE**, V. 67, n° 5, p. 786-803, 1979.

HAYKIN, Simon. **Neural Networks, a Comprehensive Foundation**. New York: Prentice-Hall, 1994.

JAIN, Anil K.; FARROKHNIA, Farshid. Unsupervised texture segmentation using Gabor filters. **Pattern Recognition**, V. 24, n° 12, p. 1167-1186, 1991.

JENSEN, J. R. **Introductory Digital Image Processing**. New Jersey: Prentice-Hall, 1986.

KAMINSKY, E. J.; BARAD, H.; BROWN, W. Textural neural network and version space classifiers for remote sensing. **International Journal of Remote Sensing**, V. 18, n° 4, p. 727-740, 1997.

KANELLOPOULOS, I.; WILKINSON, G. G. Strategies and best practice for neural network image classification. **International Journal of Remote Sensing**, V. 18, n° 4, p. 711-726, 1997.

LIBERMAN, Felipe. **Classificação de Imagens Digitais por Textura usando redes Neurais**. Porto Alegre: CPGCC-UFRGS, 1997.

MARQUES Filho, Ogê ; VIEIRA Neto, Hugo. **Processamento Digital de Imagens**. Rio de Janeiro: Brasport, 1999.

MEHROTRA, R.; NAMUDURI, K. R.; RANGANATHAN, N. Gabor filter-based edge detection. **Pattern Recognition**, V. 25, n° 12, p. 1479-1494, 1992.

MURAI, H.; OMATU, S. Remote sensing image analysis using a neural network and knowledge-based processing. **International Journal of Remote Sensing**, V. 18, n° 4, p. 811-828, 1997.

NOVO, E. M. **Sensoriamento Remoto: Princípios e Aplicações**. São Paulo: Edgard Blücher, 1989.

PAO, Y. H. **Adaptive Pattern Recognition and Neural Networks**. New York: Addison-Wesley, 1989.

RAGHU, P. P.; POONGODI, R.; YEGNANARAYANA, B. A combined neural network for texture classification. **Neural Networks**, V. 8, n° 6, p. 975-987, 1995.

RICHARDS, J. A. **Remote Sensing Digital Image Analysis – An Introduction**. New York: Springer-Verlag, 1986.

RUMELHART, D. E., McCLELLAND, J. L. **Parallel Distributed Processing**. Cambridge: MIT Press, 1986.

SABINS Jr., Floyd F. **Remote Sensing: principles and interpretation**. New York: Freeman & Co., 1987.

SCHOWENGERDT, Robert A. **Techniques for Image Processing and Classification in Remote Sensing**. New York: Academic Press, 1983.

SLATER, P. N. **Remote Sensing: optics and optical systems**. Massachusetts: Addison-Wesley, 1980.

STEIN, A. et al. Integrating spatial statistics and remote sensing. **International Journal of remote Sensing**, V. 19, nº 9, p. 1793-1814, 1998.

SWAIN, P. H.; DAVIS, S.M. **Remote Sensing: The Quantitative Approach**. New York: McGraw-Hill, 1978.

TAFNER, M. A.; XEREZ, Marcos de; RODRIGUEZ Filho, I. W. **Redes Neurais Artificiais - introdução e princípios de neurocomputação**. Blumenau: FURB/EKO, 1996.

TODT, Viviane. **Avaliação do Desempenho de Classificadores Neurais para Aplicações em Sensoriamento Remoto**. Porto Alegre: CEPSSRM/CPGSR-UFRGS, 1998.

VAN HULLE, M.M.; TOLLENAERE, T. A modular artificial neural network for texture processing. **Neural Networks**, V. 6, p. 7-32, 1993.

WELDON, T. P.; HIGGINS, W. E. Designing Multiple Gabor Filters for Multi-Texture Image Segmentation. **Optical Engineering**, V. 38, n° 9, p. 1478-1489, 1999.

WELDON, T. P.; HIGGINS, W. E. Design of Multiple Gabor Filters for Texture Segmentation. **IEEE Int. Conf.. Acoustics, Speech and Signal Processing**, V. 4, p. 2245-2248, 1996.

Na Internet:

COLLECTION OF MICROTEXTURES.

<http://www-dbv.cs.uni-bonn.de/image> (Consulta em março de 1999)

MULTIRESOLUTION DESIGN OF MULTIPLE GABOR FILTERS FOR TEXTURE SEGMENTATION

<http://ws351.uncc.edu/tpw/diss/diss.html> (Consulta em outubro de 1999)