

MOSAR FARIA BOTELHO

**APLICAÇÃO DE REDES NEURAIIS NA CLASSIFICAÇÃO DE IMAGENS DE
ALTA RESOLUÇÃO ESPACIAL E DADOS DO LASER SCANNER, USANDO
UMA ABORDAGEM ORIENTADA A REGIÕES**

Dissertação apresentada ao Curso de Pós-Graduação em Ciências Geodésicas, Setor de Ciências da Terra da Universidade Federal do Paraná, como requisito parcial à obtenção do Título de Mestre.

Orientador: Dr. Jorge A. Silva Centeno

CURITIBA
Fevereiro, 2004

SUMÁRIO

LISTA DE FIGURAS.....	IV
LISTA DE TABELAS	VII
RESUMO	IX
ABSTRACT	X
CAPÍTULO 1 - INTRODUÇÃO.....	1
1.1 - CONSIDERAÇÕES GERAIS	1
1.2 - OBJETIVOS	3
1.3 - JUSTIFICATIVA	4
CAPÍTULO 2 - REVISÃO DA LITERATURA	6
2.1 - SENSORIAMENTO REMOTO	6
2.1.1. - Propriedades espaciais das imagens.....	8
2.2 - LASER SCANNER	9
2.3 - ANÁLISE DIGITAL DE IMAGEM	12
2.4 - SEGMENTAÇÃO DE IMAGENS	13
2.4.1. - Detecção de Descontinuidade	15
2.4.2. - Similaridade	15
2.4.3. - Segmentação espacial-espectral	18
2.4.4. - Extração de características espaciais.....	19
2.4.4.1 - Descritores de Forma	20
2.4.4.2 - Descritores de Textura	20
2.5 - CLASSIFICAÇÃO	21
2.6 - INTEGRAÇÃO DE DADOS.....	23
2.7 - REDES NEURAIS ARTIFICIAIS	24
2.7.1. - O cérebro e o neurônio biológico.....	25
2.7.2. - Neurônio artificial e as redes neurais.....	27
2.7.2.1 - A função de transferência.....	28
2.7.3. - Arquitetura das redes neurais artificiais.....	29
2.7.4. - Treinamento da rede neural artificial	31
2.7.5. - Algoritmos para o treinamento da rede neural artificial	33

CAPÍTULO 3 - MATERIAL E MÉTODO.....	37
3.1 - MATERIAL	37
3.1.1. - Área de estudo.....	37
3.1.2. - Imagem QUICKBIRD II.....	40
3.1.3. - Grade regular altimétrica.....	41
3.1.4. - Programas utilizados	42
3.2 - METODOLOGIA	43
3.2.1. - Segmentação da imagem e sua rotulação.....	44
3.2.1.1 - Crescimento de Regiões implementado	44
3.2.1.2 - Segmentador espacial-espectral do “eCognition”	45
3.2.1.2.1 - Fundamentos	45
3.2.1.2.2 - Resultados da segmentação no eCognition.....	49
3.2.2. - Descrição dos segmentos	49
3.2.3. - Altimetria	49
3.2.4. - Descritores espaciais	50
3.2.4.1 - Área e Perímetro	50
3.2.4.2 - Coeficiente de compacidade	50
3.2.4.3 - Razão entre o Eixo Menor e Maior da Elipse	50
3.2.5. - Coleta de Amostras	51
3.2.6. - Implementação da rede neural artificial.....	53
3.2.6.1 - Algoritmo de aprendizado retro-propagação.	54
3.2.6.2 - Treinamento da rede neural artificial para a classificação.	58
3.2.6.3 - Metodologia para a definição da arquitetura.....	59
3.2.7. - Implementação do Método distância mínima	60
3.2.8. - Implementação do Método Máxima Verossimilhança	60
3.2.9. - Área de verificação	61
3.2.10. - Avaliação da imagem classificada	63
3.2.10.1 - Algoritmo de avaliação implementado	65
CAPÍTULO 4 - RESULTADOS E DISCUSSÕES.....	67
4.1 - SEGMENTAÇÃO DA IMAGEM.....	67
4.2 - VERIFICAÇÃO DA RELAÇÃO ENTRE CLASSES SEGUNDO OS PARÂMETROS ESPECTRAIS E ESPACIAIS.	70
4.2.1. - Separabilidade entre informações espectrais.	70

4.2.2. - Separabilidade entre informações espectrais e dados do laser scanner.....	72
4.2.3. - Separabilidade entre dados do laser scanner e descritores de forma.	74
4.2.4. - Conclusão da análise dos gráficos.....	76
4.2.5. - Separabilidade estatística entre informações espectrais e espaciais.	77
4.3 - CLASSIFICAÇÃO ORIENTADA A REGIÕES UTILIZANDO REDES NEURAIAS ARTIFICIAIS COM O PROGRAMA IMPLEMENTADO EM C++.....	85
4.3.1. - Testes usando diferentes combinações de variáveis.	86
4.4 - CLASSIFICAÇÃO ORIENTADA A REGIÕES UTILIZANDO DISTÂNCIA MÍNIMA.	87
4.4.1. - Testes usando diferentes combinações de variáveis.	88
4.5 - CLASSIFICAÇÃO ORIENTADA A REGIÕES UTILIZANDO MÁXIMA VEROSSIMILHANÇA.	89
4.6 - VERIFICAÇÃO DA INFLUÊNCIA DO LASER SCANNER PERANTE OS DESCRITORES DE FORMA UTILIZANDO OS CLASSIFICADORES IMPLEMENTADOS.....	91
4.7 - COMPARAÇÃO ENTRE CLASSIFICADORES	91
4.8 - DESEMPENHO DAS REDES NEURAIAS NA IDENTIFICAÇÃO DE CASAS.....	95
CAPÍTULO 5 - CONCLUSÕES E RECOMENDAÇÕES.....	97
CAPÍTULO 6 - REFERÊNCIA BIBLIOGRAFICA	99

LISTA DE FIGURAS

FIGURA 2.1 – VIZINHANÇA E CONECTIVIDADE	8
FIGURA 2.2 – HISTOGRAMA DE INTENSIDADE PARTICIONADO POR UM LIMIAR	16
FIGURA 2.3 – EXEMPLO DO ALGORITMO DE DIVISÃO E FUSÃO	17
FIGURA 2.4 – EXEMPLO DO ALGORITMO DE CRESCIMENTO DE REGIÕES.....	17
FIGURA 2.5 – NEURÔNIO BIOLÓGICO	26
FIGURA 2.6 - EXEMPLO DE NEURÔNIO ARTIFICIAL	28
FIGURA 2.7 – ARQUITETURA DA REDE ALIMENTADA ADIANTE COM CAMADA ÚNICA	30
FIGURA 2.8 – ARQUITETURA DE UMA REDE NEURAL FEED-FORWARD	30
FIGURA 2.9 – ARQUITETURA DA REDE RECORRENTE.....	31
FIGURA 2.10 – SUPERFÍCIE DO ERRO MÉDIO QUADRÁTICO COM MÍNIMO LOCAL E MÍNIMO GLOBAL.	32
FIGURA 3.1 - PRIMEIRO RECORTE DA IMAGEM QUICKBIRD II, DA REGIÃO EM ESTUDO	38
FIGURA 3.2 - SEGUNDO RECORTE DA IMAGEM QUICKBIRD II, DA REGIÃO EM ESTUDO	39
FIGURA 3.3 - TERCEIRO RECORTE DA IMAGEM QUICKBIRD II, DA REGIÃO EM ESTUDO	39
FIGURA 3.4 - IMAGEM ALTIMÉTRICA DA ALTURA DOS OBJETOS EM RELAÇÃO AO SOLO, ORIUNDA DE DADOS DO LASER SCANNER.	42
FIGURA 3.5 - DIAGRAMA DE BLOCOS DA METODOLOGIA A SER EMPREGADA NA DISSERTAÇÃO	43
FIGURA 3.6 - DIAGRAMA DE BLOCOS DO PROGRAMA IMPLEMENTADO EM LINGUAGEM C++.	44
FIGURA 3.7 - EXEMPLO DE OBJETOS COMPACTOS E SUAVES.	46
FIGURA 3.8 - INTERFACE DO ALGORITMO GERADOR DO ARQUIVO DE TREINAMENTO.	51
FIGURA 3.9 - INTERFACE DO ALGORITMO COLETOR DE AMOSTRAS COM AS AMOSTRAS COLETADAS NA IMAGEM QUICKBIRD II UTILIZADAS NO TREINAMENTO DOS CLASSIFICADORES.	52

FIGURA 3.10 - ARQUIVOS GERADOS PARA TREINAMENTO DOS CLASSIFICADORES CONSIDERANDO A INTERCALAÇÃO ENTRE CLASSES.....	53
FIGURA 3.11 - ARQUITETURA DO MODELO DE MULTICAMADA COM RETRO-PROPAGAÇÃO.	54
FIGURA 3.12 - SUPERFÍCIE DE ERRO, CORTE TRANSVERSAL.....	56
FIGURA 3.13 - INTERFACE DO ALGORITMO RETRO-PROPAGAÇÃO NA ARQUITETURA MULTICAMADA.	58
FIGURA 3.14 - VERDADE DE CAMPO VETORIZADA.....	62
FIGURA 3.15 - VERDADE DE CAMPO RASTERIZADA.	63
FIGURA 3.16 - INTERFACE DO MÓDULO DE AVALIAÇÃO DA CLASSIFICAÇÃO.	66
FIGURA 3.17 - INTERFACE COM A EXATIDÃO DO PONTO DE VISTA DO USUÁRIO E DO PRODUTO NA CLASSIFICAÇÃO DO OBJETO.	66
FIGURA 4.1 - RECORTE DA IMAGEM QUICKBIRD II ORIGINAL E SEGMENTADA PELO “ECOGNITION”.....	68
FIGURA 4.2 - IMAGEM DO PRIMEIRO RECORTE SEGMENTADO NO ECOGNITION E CODIFICADO PELO PROGRAMA	69
FIGURA 4.3 - IMAGENS DO SEGUNDO E TERCEIRO RECORTE SEGMENTADO NO ECOGNITION E CODIFICADO PELO PROGRAMA	69
FIGURA 4.4 - RELAÇÃO ENTRE BANDA 4 (INFRA-VERMELHO) E BANDA 3 (VISÍVEL).....	70
FIGURA 4.5 - RELAÇÃO ENTRE BANDA 1 (VISÍVEL) E BANDA 2 (VISÍVEL).	71
FIGURA 4.6 - RELAÇÃO ENTRE BANDA 3 (VISÍVEL) E BANDA 2 (VISÍVEL).	71
FIGURA 4.7 - RELAÇÃO ENTRE BANDA 4 E BANDA 2.....	72
FIGURA 4.8 - RELAÇÃO ENTRE BANDA 4 (INFRA-VERMELHO) E BANDA 1 (VISÍVEL).....	72
FIGURA 4.9 - RELAÇÃO ENTRE ALTITUDE E BANDA 4.	73
FIGURA 4.10 - RELAÇÃO ENTRE ALTITUDE E BANDA 3.	73
FIGURA 4.11 - RELAÇÃO ENTRE ALTITUDE E BANDA 2.	74
FIGURA 4.12 - RELAÇÃO ENTRE ALTITUDE E BANDA 1.	74
FIGURA 4.13 - RELAÇÃO ENTRE ALTITUDE E ÁREA	75
FIGURA 4.14 - RELAÇÃO ENTRE ALTITUDE E PERÍMETRO.	75
FIGURA 4.15 - RELAÇÃO ENTRE ALTITUDE E RAZÃO ENTRE EIXOS DA ELIPSE.	76

FIGURA 4.16 - RELAÇÃO ENTRE ALTITUDE E COEFICIENTE DE COMPACIDADE.	76
FIGURA 4.17 - MÉDIAS DAS CLASSES (TC – TELHADO CLARO; TE – TELHADO ESCURO; VB – VEGETAÇÃO BAIXA; ARVORE, VIAS) DE ACORDO COM AS INFORMAÇÕES ESPACIAIS E ESPECTRAIS.	78
FIGURA 4.18 - IMAGEM DO PRIMEIRO RECORTE CLASSIFICADA PELA REDE NEURAL ARTIFICIAL UTILIZANDO A COMPOSIÇÃO I+Z (INFORMAÇÃO ESPECTRAL E ALTIMETRIA DO LASER)	92
FIGURA 4.19 - IMAGEM DO SEGUNDO RECORTE CLASSIFICADA PELA REDE NEURAL ARTIFICIAL UTILIZANDO A COMPOSIÇÃO I+Z (INFORMAÇÃO ESPECTRAL E ALTIMETRIA DO LASER)	93
FIGURA 4.20 - IMAGEM DO TERCEIRO RECORTE CLASSIFICADA PELA REDE NEURAL ARTIFICIAL UTILIZANDO A COMPOSIÇÃO I+Z (INFORMAÇÃO ESPECTRAL E ALTIMETRIA DO LASER)	94
FIGURA 4.21 - IMAGEM COM A PERSPECTIVA DA EFICIENCIA DO PRODUTO.....	96

LISTA DE TABELAS

TABELA 3.1 - PRINCIPAIS CARACTERÍSTICAS DO SENSOR QUICKBIRD II.	40
TABELA 3.2 - COEFICIENTE “KAPPA” E O DESEMPENHO DA CLASSIFICAÇÃO	64
TABELA 4.1 - TABELA QUE FORNECE OS PARÂMETROS UTILIZADOS EM CADA NÍVEL DE SEGMENTAÇÃO NO ECOGNITION.	67
TABELA 4.2 - MÉDIA NORMALIZADA DAS CLASSES	77
TABELA 4.3 - DESVIO PADRÃO DAS CLASSES NORMALIZADAS DE ACORDO COM AS INFORMAÇÕES ESPECTRAIS E ESPACIAIS	79
TABELA 4.4 - MATRIZ DAS DISTÂNCIAS ENTRE CLASSES CONSIDERANDO TODAS AS VARIÁVEIS.....	80
TABELA 4.5 - DISTÂNCIA GLOBAL PARA CADA VARIÁVEL ESPECTRAL E ESPACIAL CONSIDERANDO AS CLASSES EM ESTUDO, (N=1)	81
TABELA 4.6 - DISTÂNCIA GLOBAL PARA DIFERENTES COMBINAÇÕES DE 2 VARIÁVEIS CONSIDERANDO AS CLASSES EM ESTUDO, (N=2)	81
TABELA 4.7 - DISTÂNCIA GLOBAL PARA DIFERENTES COMBINAÇÕES DE 3 VARIÁVEIS CONSIDERANDO AS CLASSES EM ESTUDO, (N=3).....	82
TABELA 4.8 - DISTÂNCIA GLOBAL PARA DIFERENTES COMBINAÇÕES DE 4 E 5 VARIÁVEIS CONSIDERANDO AS CLASSES EM ESTUDO, (N=4 E N=5).....	83
TABELA 4.9 - DISTÂNCIA GLOBAL PARA DIFERENTES COMBINAÇÕES DE 5 VARIÁVEIS CONSIDERANDO AS CLASSES EM ESTUDO, (N=5).....	83
TABELA 4.10 - DISTÂNCIA GLOBAL PARA DIFERENTES COMBINAÇÕES DE 6, 7 E 8 VARIÁVEIS CONSIDERANDO AS CLASSES EM ESTUDO, (N=6, N=7 E N=8)..	84
TABELA 4.11 - RESUMO COM AS MELHORES DISTÂNCIAS GLOBAIS PARA CADA NÍVEL DE COMBINAÇÕES ENTRE VARIÁVEIS	84
TABELA 4.12 - RESULTADOS DA REDE NEURAL PARA O SEGUNDO RECORTE REALIZADO NA IMAGEM.	86
TABELA 4.13 - RESULTADOS DA REDE NEURAL PARA O TERCEIRO RECORTE REALIZADO NA IMAGEM.	87
TABELA 4.14 - RESULTADOS DO MÉTODO DE MÍNIMA DISTÂNCIA PARA O SEGUNDO RECORTE REALIZADO NA IMAGEM.	88
TABELA 4.15 - RESULTADOS DO MÉTODO DE MÍNIMA DISTÂNCIA PARA O TERCEIRO RECORTE REALIZADO NA IMAGEM.	89

TABELA 4.16 - RESULTADOS DO MÉTODO DE MÁXIMA VEROSSIMILHANÇA PARA O SEGUNDO RECORTE REALIZADO NA IMAGEM.....	90
TABELA 4.17 - RESULTADOS DO MÉTODO DE MÁXIMA VEROSSIMILHANÇA PARA O TERCEIRO RECORTE REALIZADO NA IMAGEM.....	90
TABELA 4.18 - MELHORES RESULTADOS DO INDICE KAPPA PARA OS MÉTODOS DE CLASSIFICAÇÃO EMPREGADOS NO ESTUDO.	91
TABELA 4.19 - IDENTIFICAÇÃO DAS EDIFICAÇÕES NA CLASSIFICAÇÃO DAS IMAGENS.....	96

RESUMO

A classificação digital de imagens de sensoriamento remoto tem ganho, nas últimas décadas, reconhecimento como técnica para suprir a escassez dos mapeamentos temáticos. Atualmente, novos sensores, com maior resolução espacial, encontram-se disponíveis no mercado, aumentando o potencial uso de imagens de satélite. No entanto, estas novas imagens demandam novas técnicas de análise, pois oferecem um maior grau de detalhe, o que possibilita a identificação de objetos em função de suas propriedades espectrais e espaciais. Ao mesmo tempo, novos dados espaciais vem sendo coletados por sensores aerotransportados, como os dados do laser scanner. Estes dados altimétricos representam uma característica particular dos objetos presentes na superfície da Terra, sua elevação, o que os torna uma informação complementar valiosa na análise de imagens de sensoriamento remoto.

O presente estudo aborda o tema da integração de dados derivados de um levantamento laser scanner com imagens de satélite de alta resolução espacial, Quickbird, para o mapeamento de áreas urbanas. Para isto, uma metodologia orientada à análise de regiões na imagem é proposta. Inicialmente, o ganho obtido com a inclusão dos dados do laser scanner é comparado com o ganho resultante do uso de parâmetros espaciais derivados das imagens, através da análise da separabilidade das classes e a classificação de áreas de treinamento. Os resultados mostram que a contribuição da informação altimétrica é muito maior do que aquela atribuída aos descritores de forma. Em uma segunda fase, a utilização de redes neurais artificiais como ferramenta para a integração dos dados espectrais e espaciais foi avaliada, sendo que esta metodologia comprovou ser mais eficiente do que outras abordagens tradicionais. O uso de redes neurais e os dados do laser scanner aumentam a qualidade do mapa temático em regiões onde a informação espectral não é suficiente para discriminar objetos diferentes.

ABSTRACT

Remote sensing digital image classification has been recognized, in the last decades, as a technique for thematic mapping. Today, new sensors, with enhanced spatial resolution, are available in the market, increasing the potential use of satellite images. However, the new images demand new analysis techniques, because they are richer in details, which should facilitate the identification of objects, based on its spectral and spatial properties. At the same time, other spatial data are being collected from airborne sensors, as laser scanner data. The altimetric information represents a special characteristic of the objects present on the earth's surface, its elevation, which turns them a valuable information for the analysis of remote sensing images.

The present study focuses on the integration of laser scanner data and remote sensing images with high spatial resolution, Quickbird imagery, for urban mapping. For this purpose, a region oriented approach is proposed. Initially, the gain obtained with the use of laser scanner is compared to the gain that results using of spatial parameters in the classification process. For this purpose, the separability of the classes was analysed and selected samples were classified. The experiments proved that the contribution of laser scanner data is larger than that attributed to spatial parameters, form parameters. In a second step, the use of artificial neural nets as a tool for the integration of spectral and spatial data was evaluated. This methodology proved to be more efficient than other traditional approaches. The use of neural nets and laser scanner data increased the quality of the thematic map in areas where spectral information is not enough to discriminate different objects.

CAPÍTULO 1 - INTRODUÇÃO

1.1 - CONSIDERAÇÕES GERAIS

Desde o início da década de 80 os avanços registrados no estudo de algoritmos ajudaram a intensificar as pesquisas e a aprimorar as técnicas e metodologias relacionadas ao Processamento Digital de Imagens como ferramenta de apoio à análise de imagens. Esse progresso serviu como alavanca para popularizar os sistemas de Processamento Digital de Imagens visando as exigências de se obter mais qualidade com eficácia, num tempo menor. Esses fatores vêm fazendo com que tarefas antes realizadas de forma não automatizada sejam substituídas, parcialmente, pela execução precisa das máquinas. No entanto, ainda não há máquinas que possa substituir o cérebro humano, principalmente na intuição e na heurística.

O interesse em Processamento Digital de Imagens está focado em aplicações existentes no campo da Classificação, Realce de Imagens e Compactação, que são áreas onde, embora já exista um bom número de técnicas, pode-se perceber também uma constante busca de novas soluções, mais abrangentes e eficazes.

Com o avanço da tecnologia dos sensores e o surgimento dos sistemas de Sensoriamento Remoto de alta resolução espacial, como o sistema IKONOS II, lançado em meados 1999, com resolução espacial de 1m e o QUICKBIRD II com 61 cm, há a possibilidade de analisar objetivos menores, como a identificação de ruas e edificações, fato que vêm provocando uma revolução nos processos de geração e manutenção de bases cartográficas, como também nos processos clássicos de classificação de imagens.

A classificação de imagens digitais de sensores de alta resolução tem por objetivo extrair feições espaciais de forma automática visando à compreensão humana. O processo de classificação de imagens pode ser realizado por dois processos distintos:

“pixel” a “pixel” ou por regiões. O processo “pixel” a “pixel” consiste na atribuição de cada “pixel” a uma determinada classe de um grupo de classes, e grande parte desta análise é realizada por meio da utilização das propriedades espectrais dos objetos presentes na cena. O processo por regiões consiste na atribuição de cada região a uma determinada classe de um grupo de classes. Esse processo tem como vantagem a utilização de descritores podendo ser forma, tamanho e textura da região. Sendo que a definição de regiões é dada por um grupo de “pixels” que são similares e adjacentes. Observa-se também que o aumento de variáveis de entrada no processo de classificação gera problemas como o aumento do tempo de processamento e dificuldade da proposta do modelo mais adequado.

Uma outra tecnologia recente é o laser scanner. Esta surge como uma tecnologia interessante para a extração de informações dos objetos da superfície do solo. Resumidamente, o método envolve a emissão de um pulso eletromagnético em direção a superfície tendo como produto as coordenadas plano-altimétricas e a intensidade do objeto atingido. Esta informação, a altura dos pontos na superfície da terra, pode também ser incluída no processo de classificação.

Para classificar dados de diferentes origens as Redes Neurais Artificiais (RNA's) são atualmente uma opção emergente. A utilização RNA's emerge como uma alternativa vantajosa aos classificadores de imagens baseados em conceitos estatísticos (Ex.: Máxima Verossimilhança), uma vez que nenhuma hipótese prévia sobre a distribuição dos dados a serem classificados é exigida e os dados de entrada podem ter escalas diferentes. Sua capacidade de distribuição dos dados de entrada é mais sutil do que os métodos estatísticos tradicionais e a capacidade de formar fronteiras de decisão não-lineares no espaço de características têm aumentado o seu uso.

O uso de RNA's, aliado ao aumento de variáveis derivadas da imagem, que oferece a abordagem de classificação baseado em regiões, é uma alternativa viável e interessante no sensoriamento remoto, especialmente na classificação de imagens de alta resolução espacial.

1.2 - OBJETIVOS

O objetivo principal deste trabalho é propor um método para extração de feições presentes em imagens ópticas de alta resolução espacial e dados do laser scanner baseado nos princípios da classificação orientada a regiões e o uso das RNA's como elemento integrador de variáveis espaciais e espectrais.

Como objetivos específicos tem-se:

- a) Implementar um programa que utilize as RNA's para a classificação da imagem, com variação das camadas intermediárias, taxas de aprendizado (variação do algoritmo - momento) e a possibilidade de alternância das classes na entrada da rede;
- b) Avaliar o potencial das RNA's para a integração de informação espacial e espectral na classificação;
- c) Avaliar o método proposto na identificação de objetos da cena como construções, estradas e vegetação, usando dados de imagem de alta resolução;
- d) Comparar esta metodologia com a abordagem de classificação orientada a regiões baseada no método convencional Máxima Verossimilhança;
- e) Comparar esta metodologia com a abordagem de classificação orientada a regiões baseada no método convencional Distância Mínima;
- f) Verificar a influência da altimetria oriunda do laser scanner na imagem final produzida;
- g) Comparar a influência da altimetria em relação aos descritores de forma: área, perímetro, coeficiente de compacidade e relação entre os eixos maior e menor da elipse.

1.3 - JUSTIFICATIVA

Após a liberação da tecnologia desenvolvida para sensores orbitais em 1994 pelo governo norte americano, outras empresas entraram no mercado tornando-o mais competitivo. Uma das vantagens da competitividade é o preço mais acessível. Os sensores de alta resolução espacial, como o KOMETA (1 m), IKONOS II (1 m), EROS A1 e A2 (1,8 m), QUICKBIRD II (0,61 m) e a ser lançado em março de 2004 o EROS B1 (0,82 m), têm revolucionado os paradigmas da interpretação de imagens, tornando viável a identificação de construções e outros objetos que anteriormente não eram visíveis, neste espaço, em áreas urbanas.

Esta identificação possibilita a atualização da base de dados em GIS, base para a manutenção atualizada dos mapas topográficos básicos e gerais, monitoramento contínuo de regiões, seleção de áreas para a implantação de infra-estrutura, reforma agrária etc.

A extração de objetos em imagens pode ser visual ou assistida por computadores. O cérebro humano extrai um grande número de informações dos objetos de uma imagem digital e os integra, classificando-os para as respectivas classes de forma muito eficaz. Na tentativa de reproduzir o comportamento humano, cientistas tentam substituir ou reproduzir as decisões humanas na interpretação por meio de regras estatísticas ou modelos matemáticos implementados em computador. A extração de informações por meio de técnicas de processamento digital de imagens vem sendo utilizada analisando-se o “pixel”. Novas abordagens, com o intuito de melhorar a acurácia da classificação digital das imagens, estão em evidência, tais como a análise utilizando como elemento mínimo da cena regiões e não “pixel”.

Quanto melhor a resolução espacial das imagens, mais detalhes são observados, podendo assim extrair mais informações dos objetos da cena. Portanto surge a necessidade de novos algoritmos para a classificação dessas imagens. Algoritmos que levem em consideração mais informações do que apenas as

informações espectrais.

Este trabalho justifica-se pela utilização das RNA's na classificação dos padrões presentes em imagens de alta resolução espacial, utilizando a integração de variáveis espectrais e espaciais, ou seja, forma, textura e altura dos objetos, além dos dados do laser scanner, como também a comparação dos vários parâmetros de forma e textura, a fim de se obter uma qualidade e acurácia maior na geração de mapas temáticos oriundos da classificação.

A contribuição deste trabalho está na tentativa de automatizar decisões, facilitar a identificação automática de objetos, por exemplo, casas, ruas, alterações no uso do solo em áreas urbanas e melhorar o método de classificação para imagens de alta resolução espacial.

CAPÍTULO 2 - REVISÃO DA LITERATURA

No trabalho, são abordadas técnicas das áreas da Inteligência Artificial, do Sensoriamento Remoto e classificadores convencionais frente à análise orientada a objetos. Assim sendo, alguns conceitos introdutórios serão descritos nos itens 2.1, 2.2, 2.3 e 2.4.

2.1 - SENSORIAMENTO REMOTO

Na área do sensoriamento remoto, um aspecto que tem experimentado maior aprimoramento é a resolução espacial. Esta evolução tem revolucionado os paradigmas da interpretação de imagens, tornando viável a identificação de construções com muito mais detalhes, bem como outros objetos que anteriormente não eram com sensores de resolução média, como Landsat, em áreas urbanas. Após a liberação da tecnologia americana em 1994, que impedia a captura de imagens com uma melhor resolução espacial, houve grande incentivo no setor privado a investir fortemente nesta tecnologia. Empresas privadas investiram na área e atualmente são comercializadas imagens com resolução espacial de 3m com o EARLYBIRD lançado em 1997, 1,8m com o EROS A1 e A2 em 2000, 1m com o IKONOS II em 1999 ou 0,61m com o QUICKBIRD II lançado em outubro de 2001. Em março de 2004 inicia o lançamento das versões do EROS de B1 a B6 que terão resolução espacial de 0,82m.

TOUTIN (2003, p.43), avaliando a alta resolução espacial do IKONOS II. Este autor verifica a potencialidade deste sensor como ferramenta, tanto em sua planimetria como em sua altimetria. Ele conclui que as imagens deste sensor possibilitam a visualização em 3D, uma vez que tais imagens podem ser utilizadas como estéreo par, encontrando uma acurácia de 2 a 3m.

A competitividade no mercado favorece a evolução das tecnologias dos sensores e favorece também os usuários dessas imagens, pois os preços destas tornam-se mais acessíveis à medida que novos sensores e mais imagens de alta resolução estão

a disposição mercado.

Os usuários das imagens digitais necessitam de uma análise qualitativa e quantitativa. A execução da avaliação quantitativa é realizada eficientemente por meio de sistemas computacionais, pois pode ser processada de maneira digital. A metodologia para a análise quantitativa das imagens digitais é intitulada de classificação digital de imagens (RICHARDS, 1993, p. 79).

A identificação de feições em imagens capturadas por sensoriamento remoto é feita através da habilidade do analista em diferenciar padrões, utilizando seu conhecimento intuitivo e o de fotointerpretação sobre a área de estudo. O intérprete fica limitado de executar esta mesma função com as imagens digitais, pois estas são multiespectrais. Além disto, o intérprete não tem acuidade visual para discriminar o limite da resolução radiométrica dos sensores atuais. Portanto a interpretação visual de imagens digitais gera produtos com algumas limitações (BRONSVELD et al, 1994 p. 349).

O anseio de diversos pesquisadores está na expectativa de utilizar imagens do sensoriamento remoto para a atualização cadastral, como base para a manutenção atualizada dos mapas topográficos básicos e gerais, monitoramento contínuo de regiões, seleção de áreas para a implantação de infra-estrutura, reforma agrária e em um futuro próximo realizar o próprio cadastro técnico municipal. Nesta linha de pesquisa pode-se citar, como exemplo, KIEMA and BÄHR (2000), CHENG (2002), SMITH, PRIESTNALL and ASAL (2000), TOUTIN (2003).

Informações espectrais são convencionalmente usadas na classificação de imagem multiespectral. Segundo KIEMA e BÄHR (2000 p. 29 – 31) os resultados obtidos usando métodos usuais são insatisfatórios, particularmente para aplicações envolvendo estruturas e mapas de natureza urbana, devido existir semelhança na característica da refletância espectral entre feições como construções com materiais similares.

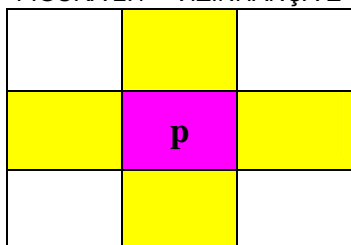
2.1.1. - Propriedades espaciais das imagens

Uma imagem digital pode ser vista como uma matriz cujas linhas e colunas identificam um ponto na imagem. A cada instante, uma pequena área do terreno é coberta e o total de energia desta unidade de área é armazenado sob forma de um número digital através do processo de varredura. Tratando-se de uma imagem colorida ou multiespectral, várias matrizes são necessárias para armazenar os valores correspondentes à informação multiespectral (FACON, 1993 p.24).

As propriedades espaciais de uma imagem podem ser derivadas a partir dos conceitos de regiões, “vizinhança” e “conectividade”, definidas a seguir.

Vizinhança – Segundo a Figura 2.1, um “pixel” p , com coordenadas (x, y) , possui 4 vizinhos horizontais e verticais, onde suas coordenadas são $(x+1, y)$, $(x-1, y)$, $(x, y+1)$, $(x, y-1)$. Estes “pixels” formam então a denominada vizinhança-de-4 do “pixel” p , que é representado por $N_4(p)$ (GONZALES e WOODS, 2000 p.26).

FIGURA 2.1 – VIZINHANÇA E CONECTIVIDADE



Conectividade - Dois “pixels” são conexos quando compartilham uma fronteira. Uma região é conexa, se é possível traçar um caminho de um “pixel” qualquer dentro da região até qualquer outro “pixel” da região, sem sair dela (GONZALES e WOODS, 2000 p.27):

Região - é um conjunto de pontos “conectados” (adjacentes), ou seja, de qualquer ponto da mesma pode-se chegar a qualquer outro ponto por um caminho inteiramente contido na região. As regiões que se deseja detectar, geralmente são regiões “homogêneas”, ou seja, apresentam alguma propriedade local aproximadamente constante em toda sua extensão. Algumas propriedades locais

comumente utilizadas são o valor do contador digital, textura ou cor (ANDRADE, 2003 p.25).

2.2 - LASER SCANNER

De acordo com SMITH, PRIESTNALL e ASAL (2000 p.45), o laser scanner é um sistema de sensoriamento remoto ativo usado para medir a distância entre o sensor e a superfície dos objetos. Seu funcionamento baseia-se na utilização de um feixe de laser (LASER: Light Amplification by Stimulated Emission of Radiance) que é disparado em direção aos objetos, refletido por eles e novamente captados pelo sistema. A partir do sinal de retorno e o tempo decorrido entre a emissão e o registro do retorno, informações como a distância entre o sensor e o objeto podem ser calculadas. Este tipo de sistema também é chamado de LIDAR (Light Detection And Ranging) ou LADAR (Laser Detection And Ranging).

O sistema laser scanner pode ser dividido em três componentes principais: a unidade de medição laser propriamente dita, encarregada de emitir e receber o sinal laser, um sistema de varredura óptico-mecânico e uma unidade de registro de medições de apoio (SMITH, PRIESTNALL e ASAL, 2000 p.46-47).

Segundo SHERSTHA et al (1997), uma unidade típica possui um espelho montado em frente ao ponto de emissão do feixe laser. Este espelho roda com determinada frequência com a finalidade de varrer uma linha transversal ao deslocamento do avião. Os pulsos possuem uma duração de nanosegundos que são direcionados pelo espelho através de uma abertura ou janela em direção ao chão. Quando o laser scanner é montado no avião, com a linha do scanner perpendicular a direção do voo, um objeto é varrido e formado dentro de uma faixa coberta. A largura desta faixa e o espaço entre os pontos do objeto depende da altura e velocidade do voo e do domínio limite dos ângulos do scanner. Estes fatores variam devido ao tipo de scanner usado e o plano de voo. Tipicamente a velocidade do voo é de 200 a 250 km/h e a altura varia entre 350 a 1000 m. A taxa de pulso pode ser entre 2000 a 5000 pulsos

por segundo. Em adição, o limite do ângulo de visão do scanner é na ordem de aproximadamente 20°. O processo utilizando um laser scanner com estes parâmetros e com acurácia das medidas razoável podem render medidas pontuais menores que um metro. Isto é o suficiente para criar um DSM (Modelo Digital do Solo) satisfatório para aplicações ambientais.

Existem diferentes opções para efetuar o redimensionamento do feixe laser. Entre as opções disponíveis estão o uso de um espelho oscilante, que produz linhas paralelas ou em zig-zag, o Palmer Scanner, que utiliza um espelho que gira em torno do seu centro e produz um padrão elíptico e o espelho em forma de polígono rotatório, que produz linhas paralelas. (SHERSTHA et al, 1997).

SMITH, PRIESTNALL e ASAL (2000, p. 48), citam vários autores que verificaram a qualidade do LIDAR DSM em diferentes aplicações: VAUGHN ET AL (1996) afirma que o laser altimétrico, pode medir a altura elipsoidal com acurácia melhor que 10cm, se cuidadosamente calibrado, utilizando um GPS com plataforma de trajetória precisa e um vôo baixo. Também o CSATHO ET AL. (1996) descreve o uso do sistema laser da NASA, o qual investigou uma área de 130 a 200 m em largura para estudos da região polar. O objetivo deste estudo foi obter um modelo de elevação digital (DEM) e mapas topográficos a partir de dados de altimetria laser. A densa distribuição dos pontos medidos para uma acurácia de 10 a 20 cm foi usada para criar o DEM. Outro DEM foi criado da mesma área usando o método de fotogrametria tradicional. Uma comparação entre eles indicou uma forte concordância. FOOD E GUTELLIUS (1997) afirma que o laser scanner é mais econômico e mais produtivo em áreas onde os métodos tradicionais tal como fotogrametria aérea ou terrestre são impraticáveis ou impossíveis. O sistema laser scanner foi benéfico se usado em lugares como áreas costeiras, florestas, solo úmido (wetlands) e corredores de transporte. Também pode ser usado com sucesso em mapeamento topográfico e em áreas com topografia irregular devido a distúrbios como terremoto ou região vulcânica. WRIGHT ET AL. (1996) explica que o laser scanner, chamado mapeador topográfico (ATM-II)

desenvolvido pela NASA, tem a capacidade de produzir mapas de alta resolução. Foi usado no mapeamento do gelo ártico em Greeland, Svalbard e Iceland em outubro de 1996. Este instrumento foi também usado para pesquisar 600 km das praias da costa oeste entre Delaware e Carolina do Sul nos EUA. Os resultados indicaram que uma acurácia de aproximadamente 10 cm em elevação foi encontrada. NAESSET (1997) utilizou o laser scanner para a determinação das alturas das árvores na Noruega com propósito de planejamento florestal. As pesquisas mostram que usando a tecnologia para este propósito tem-se resultados igual ou melhor que aqueles obtidos usando o método de fotogrametria aérea.

De acordo com HUISING E PEREIRA (1998, p.245-261), o departamento de pesquisa de Rijkswaterstaat em Netherlands desenvolveu um processo de laser scanner para obter medidas topográficas para o estudo de fontes de erros no sistema. O estudo indica que os erros podem variar entre 5 cm a 200 cm, e mencionou que a principal fonte deles é devido a incerteza na determinação da altitude e posição da plataforma. O mesmo estudo predicou que alguns erros podem resultar de medidas relacionadas ao alcance do laser. Entretanto, o estudo recomendou que seguindo a estratégia adequada nas observações para remover erros a uma acurácia melhor pode ser atingida.

De acordo MASA HARU e HASEGAWA (2000 p.557), um método eficiente para extrair dados de prédios de alta elevação, obtidos por um laser scanner, é aplicando o método de segmentação de regiões (o qual será descrito nas seções que seguirão). A idéia básica é que cada prédio tem uma altura diferente do vizinho. Entretanto pode-se distinguir cada prédio segmentando o DEM, com a condição de que um “pixel”, com diferença de altura, comparado com seu vizinho, pertença a mesma região se estiverem dentro de uma área pré-determinada. O estudo constatou alguns problemas como: grandes construções não puderam ser distinguidas de superfícies de estradas, estando relacionado à seleção da área de segmentação. Outro problema é que os objetos extraídos não são sempre prédios ou casas mas por exemplo árvores. O estudo, apesar dos problemas mencionados aponta que a segmentação dos

dados do laser scanner é um bom método para gerar modelo de imagem em 3D, no sentido que o método pode gerá-la através de um processo automático.

Verifica-se, pelos resultados dos pesquisadores acima citados, que os dados do laser scanner oferecem um grande potencial para a coleta de informações a respeito da geometria dos objetos na superfície da terra, mas também é constatado que a análise isolada destes dados não garante os melhores resultados.

2.3 - ANÁLISE DIGITAL DE IMAGEM

De acordo com RIBEIRO e CENTENO (2002 p.55), a melhoria de informação visual para a interpretação humana e o processamento de dados de cenas para percepção automática por meio de máquinas, é o principal incentivador ao estudo dos métodos de processamento digital de imagens.

SONKA, HLAVAC e BOYLE (1998, p.42-43), afirmam que, na percepção visual, a principal meta da análise por computadores é encontrar uma relação entre a imagem e o modelo do mundo real. Vários níveis de representação de informação visual são definidos para esta finalidade; a visão do computador inclui um desígnio (esboço) da estrutura de dados e algoritmos usados para criação da representação e suas relações. A estrutura de dados pode ser dividida em quatro níveis: “pixel” da imagem, imagem segmentada, topologia e modelo relacional. No estudo, apenas as duas primeiras estruturas de dados serão consideradas, “pixel” da imagem e “imagem segmentada”. A primeira é considerada como o mais baixo nível de representação. Nela, a imagem consiste dos dados originais: matrizes com números inteiros representando o valor digital do “pixel” (brilho). Para a segunda, partes da imagem são reunidas em grupos que, se assume, pertencem a mesma região ou objeto na cena. O resultado é o agrupamento de “pixels” formando regiões homogêneas.

Dentre as ferramentas do processamento digital de imagens que se preocupam com análise digital da imagem, pode destacar as técnicas para extração de informação como a segmentação e a classificação digital.

2.4 - SEGMENTAÇÃO DE IMAGENS

Segundo FACON (1993, p.87), o ser humano, no nível da visão, efetua agrupamentos sobre o que ele percebe, baseados na proximidade, similaridade e continuidade. A decomposição de uma cena visual em “unidades significantes” constitui uma etapa decisiva na compreensão da imagem.

O processo de segmentação de imagens tem por objetivo fragmentar uma imagem em regiões homogêneas, considerando algumas de suas características intrínsecas como, por exemplo, o valor do contador digital dos “pixels”, a forma ou a textura de uma região.

Segundo KOK et al (2000, p.115), o objetivo da segmentação em sensoriamento remoto é definir regiões na imagem, correspondendo a objetos na superfície terrestre. As regiões que se deseja detectar, em geral, são regiões “homogêneas”, ou seja, apresentam alguma propriedade local aproximadamente constante em toda sua extensão.

Formalmente, pode-se descrever uma região da seguinte maneira. Seja R a região da imagem, pode-se imaginar a segmentação como o processo de particionar R em n regiões R_1, R_2, \dots, R_n , tal que:

$$(a) \bigcup_{i=1}^n R_i = R \quad (2.1)$$

(b) R_i é uma região conexa, $i = 1, 2, \dots, n$.

$$(c) R_i \cap R_j = \text{Vazio} \quad \forall i \neq j \quad (2.2)$$

(d) $P(R_i) = \text{VERDADEIRO}$ para $i = 1, 2, \dots, n$ e

$$(e) P(R_i \cup R_j) = \text{FALSO} \quad \forall i \neq j \quad (2.3)$$

em que $P(R_i)$ é um predicado lógico sobre os pontos do conjunto R_i e ϕ é o conjunto vazio. O exemplo mais imediato de predicado é o valor digital do “pixel”.

A condição (a) indica que a segmentação deve ser completa, ou seja, cada “pixel” deve pertencer a uma região. A condição (b) requer que os “pixels” em uma

região sejam conexos. A condição (c) indica que as regiões devem ser disjuntas. A condição (d) trata das propriedades que devem ser satisfeitas pelos “pixels” em uma região segmentada, por exemplo, $P(R_i)$ =verdadeiro se todos os “pixels” em R_i possuírem a mesma característica. Finalmente, a condição (e) indica que as regiões R_i e R_j são diferentes no sentido do predicado P (GONZALES e WOODS, 2000 p.326).

Segundo GONZALES e WOODS (2000, p.295), o nível até o qual essa subdivisão da imagem deve ser realizado depende do problema que está sendo resolvido, ou seja, a segmentação deve parar quando os objetos de interesse na aplicação tiverem sido isolados.

Os algoritmos de segmentação são geralmente baseados na análise da descontinuidade ou da similaridade entre “pixels”. Na primeira categoria a abordagem é particionar a imagem baseando-se em mudanças bruscas nos valores digitais. Os principais algoritmos nessa categoria são a detecção de pontos isolados e detecção de linhas e bordas na imagem. A abordagem da segunda categoria é particionar a imagem baseando-se na homogeneidade dos níveis de cinza, ou seja, detecção de regiões uniformes. As principais técnicas nessa categoria são: a limiarização, crescimento de regiões e divisão e fusão de regiões (FACON, 1993 p. 96; GONZALES e WOODS, 2000 p. 295).

Para a geração de uma imagem temática, as regiões encontradas pela segmentação devem ser classificadas. Nesta classificação, além do valor digital médio da região (espectral), outros descritores podem ser analisados, como a forma ou tamanho da região.

Segundo SCHNEIDER E STEINWENDNER (1999, p.743) que utilizaram a segmentação e a classificação de imagens de satélite para a obtenção de mapas de uso da terra, o resultado da classificação depende do resultado da segmentação. No entanto, ainda está sendo pesquisado um segmentador genérico para qualquer imagem, que possua resultados positivos.

2.4.1. - Detecção de Descontinuidade

Segundo GONZALES e WOODS (2000, p. 297), a maneira mais comum de procura por descontinuidades utiliza varredura da imagem por uma máscara, de maneira similar à filtragem. Para cada “pixel”, avalia-se o contraste da região, por meio da vizinhança definida pela máscara. Regiões com alto contraste caracterizam os limites espaciais dos segmentos.

A detecção de bordas é realizada por um algoritmo que estima o gradiente do valor do contador digital da imagem original, para gerar uma imagem gradiente ou imagem de intensidade de borda. O próprio algoritmo calcula o limiar para encontrar as bordas e, encontrando um “pixel” com valor superior ao limiar estabelecido, inicia-se o processo de perseguição desta borda, encerrando-se apenas quando esta borda toca outra já detectada ou a fronteira da imagem. Neste processo gera-se uma imagem binária com valores 1 referentes às bordas e valores 0 referentes às regiões não-bordas. (FILHO e NETO, 1999 p.51).

A detecção de pontos pode ser obtida de forma direta na imagem. A idéia é que o valor contador digital de um ponto isolado seja completamente diferente do valor digital de seus vizinhos.

2.4.2. - Similaridade

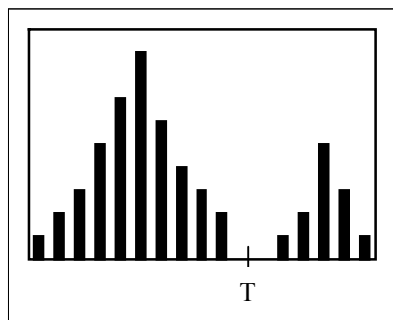
O objetivo desta categoria é particionar a imagem baseando-se na uniformidade entre “pixels” vizinhos. As principais técnicas são: a limiarização, crescimento de regiões e divisão e fusão de regiões.

a) Limiarização

Segundo GONZALEZ E WOODS (2000, p.316) a mais simples de todas as técnicas de limiarização é a do particionamento do histograma da imagem por um limiar T (Figura 2.2). A segmentação é efetuada, varrendo a imagem, “pixel” a “pixel”, e rotulando cada “pixel” dependendo se o valor digital daquele “pixel” for

maior ou menor que o limiar T . O sucesso desse método depende de quão bem o histograma pode ser particionado. As técnicas de limiarização não levam em conta o contexto, ou seja, a vizinhança dos “pixels” não é analisada (FACON, 1993 p.89).

FIGURA 2.2 – HISTOGRAMA DE INTENSIDADE PARTICIONADO POR UM LIMIAR



FONTE: Adaptado de GONZALES e WOODS. (2000, p.316)

b) Divisão e Fusão de Regiões

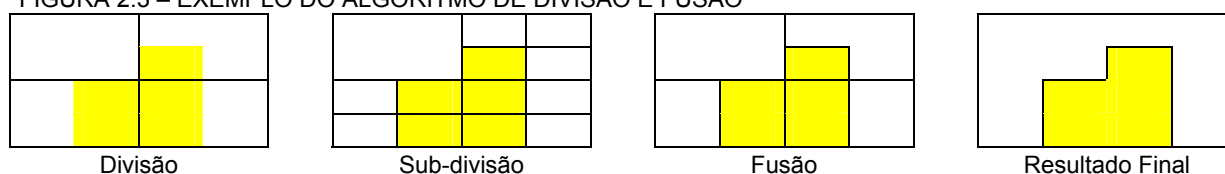
Segundo GONZALES e WOODS (2000 p. 329), o método de divisão e fusão de regiões consiste em subdividir a imagem em um conjunto de regiões arbitrárias e disjuntas e então realizar a fusão das regiões adjacentes e similares.

No caso de uma imagem quadrada (R), uma abordagem para segmentá-la é subdividir esta imagem em quadrantes cada vez menores, de maneira que todos os “pixels” em uma região compartilhem a mesma propriedade, enquanto regiões adjacentes distintas possuam propriedades diferentes. Caso uma região não esteja composta por “pixels” uniformes, então essa é novamente subdividida em quatro regiões. Este processo é repetido até atingir um ponto onde todas as regiões sejam uniformes (Quadtree).

Se apenas a divisão (subdivisão) fosse usada, a partição final provavelmente conteria regiões adjacentes com propriedades idênticas. Esse problema pode ser resolvido efetuando a fusão destas regiões numa segunda etapa. Duas regiões adjacentes R_j e R_k são fundidas apenas se $P(R_j \cup R_k) = \text{verdadeiro}$, onde $P(R)$ identifique a propriedade considerada nos critérios de uniformidade.

O processo de divisão e fusão de regiões pode ser visualizado pelo exemplo da Figura 2.3.

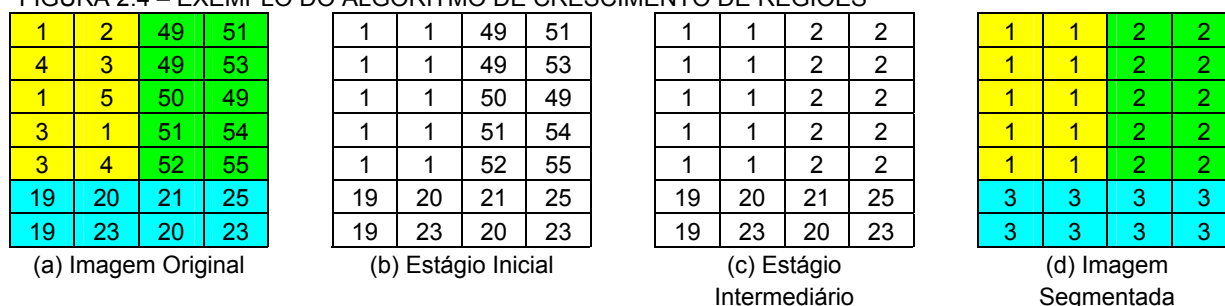
FIGURA 2.3 – EXEMPLO DO ALGORITMO DE DIVISÃO E FUSÃO



c) Crescimento de regiões por agregação de “Pixels”

Esta é uma técnica de agrupamento de dados, na qual somente os “pixels” adjacentes, podem ser agrupados. São agrupados os “pixels” com valores de similaridade inferior ao limiar definido pelo usuário, baseado em um teste de hipótese estatístico realizado com as médias entre as regiões. As regiões menores são posteriormente agregadas as regiões maiores, segundo um limiar de agregação (tamanho mínimo aceitável para uma sub-região), definido também pelo usuário (GONZALES e WOODS, 2000 p.327).

FIGURA 2.4 – EXEMPLO DO ALGORITMO DE CRESCIMENTO DE REGIÕES



A Figura 2.4 (a) mostra uma matriz da imagem original onde observa-se três grupos distintos. Adotando-se uma diferença absoluta de valores menores que três entre níveis de intensidade de cinza. Inicia-se o processo escolhendo um ponto semente, obtendo assim o estágio inicial (b) com a agregação dos “pixels” adjacentes e que possuem essa mesma propriedade. Quando o algoritmo não encontra essa propriedade nos “pixels” adjacentes escolhe-se uma nova semente, sendo esta a fase intermediária (c), e o processo é repetido até que não encontre mais regiões a segmentar que exijam novas sementes, gerando assim a imagem segmentada (d).

2.4.3. - Segmentação espacial-espectral

O aumento da resolução espacial de uma imagem, demanda uma análise que considere o conhecimento do intérprete humano, de forma a refletir melhor a subjetividade na decisão a respeito da classificação dos “pixels”.

Uma das grandes desvantagens das técnicas de agrupamento convencionais é não considerar a característica espacial dos segmentos. Em imagens de alta resolução este tipo de algoritmo possui eficiência reduzida (ANTUNES, 2003 p.30).

Com o objetivo de superar as limitações dos segmentadores convencionais em imagens multiespectrais de alta resolução espacial, estudos se voltam a analisar os objetos de estrutura topológica vetorial agregados as informações espectrais. Desta forma, o aspecto espacial possui um importante papel e pode ser descrito por relações entre os objetos vizinhos. (CHENG, 2002 p.41).

O conceito de evolução de rede fractal (*FNEA: fractal net evolution approach*), considera a imagem de sensoriamento remoto como de natureza fractal. Ao segmentar-se a imagem simultaneamente de forma fina e de forma grosseira, dá-se origem a diferentes escalas de segmentação. O procedimento de segmentação é similar ao crescimento de regiões que agrupa “pixels” ou sub-regiões em regiões maiores. A segmentação da imagem resulta em segmentos ou objetos, que baseados em parâmetros espectrais e espaciais, podem ainda ser reagrupados em objetos maiores denominados superobjetos ou regiões. Estas regiões possuem relações com os segmentos que as constituem num nível hierárquico mais baixo (ANTUNES, 2003 p.32).

No mercado existem alguns programas que vem utilizando o algoritmo de segmentação *FNEA*. Cita-se como exemplo, o programa “eCognition”, onde o usuário pode representar o conhecimento utilizando relações derivadas experimentalmente e modificar tais relações de maneira a satisfazer os fins da análise da imagem. Neste programa, os objetos resultantes da segmentação representam a informação da imagem de uma forma abstrata. Além das informações espectrais outras informações ou

atributos adicionais, como forma, textura relacionamento entre objetos e sua vizinhança, podem ser usados para a classificação (ECOGNITION, 2001 p.1-7).

Para representar o conhecimento do usuário, o “eCognition” dispõe de redes conectadas, chamadas de classes hierárquicas de operadores. Estas classes são capazes de modificar suas conexões ao sofrerem alterações nas informações fornecidas pelo usuário. Regiões representativas de cada classe são fornecidas como amostras e, a partir delas, uma descrição das classes é obtida. O usuário pode selecionar a informação útil, interagindo com o programa através da escolha dos atributos relevantes, derivados da imagem, ou informações advindas de mapas. Quanto mais informações relevantes o usuário introduzir no sistema, maior é o conhecimento armazenado nas classes hierárquicas.

Após a segmentação, os segmentos são caracterizados por descritores de forma, textura e etc. ANTUNES (2003, p. 30) salienta que o procedimento de segmentação deve ser moldado de acordo com a resolução da imagem e com a escala esperada para os objetos. As regiões derivadas da segmentação são ricas em informação espacial, pois elas podem ser descritas em termos de forma, tamanho, ou textura, entre outros parâmetros espaciais. Faz-se necessário na análise digital, parametrizar estes descritores de maneira quantitativa.

2.4.4. - Extração de características espaciais

Os segmentos devem ser representados e descritos em um formato apropriado para o processamento subsequente. A representação de uma região envolve duas escolhas:

1. Da região em termos de sua característica externa (fronteira), ou
2. Da região em termos de suas características internas (os “*pixels*” que compõem a região).

Uma representação externa é escolhida quando a atenção estiver voltada para as características de forma (espaciais). Na representação interna, a atenção é voltada às

propriedades como cor e textura (GONZALES e WOODS, 2000 p. 345).

2.4.4.1 - Descritores de Forma

Em termos visuais, a forma é um dos elementos mais importantes para a interpretação de imagens digitais, sendo uma variável que permite percepção associativa dos objetos da imagem. No entanto, de acordo com ANDRADE (2003, p.35), embora seja uma poderosa ferramenta para a discriminação de objetos, a análise da forma não é fácil de ser implementada na análise de imagens “raster” por computadores.

Os limites dos objetos das imagens são desconhecidos inicialmente, e para analisar a forma é necessário obtê-los. Por isto, os algoritmos de análise de forma dependem em alto grau da qualidade do resultado de uma classificação (supervisionada ou não supervisionada), ou da segmentação da imagem.

ANDRADE (2003 p.105) comenta que parâmetros de forma como o coeficiente de compacidade (CC), a relação entre o eixo menor e maior da elipse (Re), a Área (A) e o Perímetro (P) são os coeficientes que mais contribuem para a descrição de um objeto na classificação de imagem digital gerada por uma Rede Neural. A seguir, estes descritores serão apresentados.

2.4.4.2 - Descritores de Textura

Textura é a variação dos valores digitais (de uma imagem), causada pela ocorrência de objetos menores que a resolução espectral ao longo de uma região (JAIN et al, 1995 p.235). Assim, a textura é caracterizada pela distribuição espacial de níveis em uma vizinhança e não pode ser definida por um ponto.

A textura é um elemento importante na visão humana, provendo em uma cena a profundidade e orientação da superfície. Contém informações importantes sobre a disposição estrutural de superfícies e seus relacionamentos com o ambiente ao redor.

Por causa de sua importância e utilidade para o reconhecimento de padrões e visão computacional, foram realizadas pesquisas com resultados importantes nas últimas décadas (SCHRÖDER et al., 2000 p.423-432; FRANKLIN et al. ,2001 p.849-855; FERRO E WARNER ,2002 p.51-63 e CLAUSI E ZHAO ,2002 p.763-774).

SALI E WOLFSON (1992, p.3395-3408) utilizaram textura na segmentação de imagens para a classificação supervisionada, e chegaram a resultados que indicam que uma combinação de características texturais pode resultar numa melhoria mais efetiva na segmentação de imagens. Os autores utilizaram estatísticas de segunda ordem através de matriz de Co-ocorrência dos níveis de cinza, por permitirem a detecção de muitas propriedades de textura e por terem obtido resultados preliminares satisfatórios. Utilizaram também características baseadas na multi-resolução para o algoritmo de segmentação, pois a textura de terrenos possui um comportamento fractal e modelos fractais seriam apropriados para descrever a textura de suas imagens. As estatísticas de primeira ordem utilizadas foram a média e a variância, por serem muito úteis na discriminação de textura. As estatísticas de segunda ordem utilizadas foram baseadas na matriz de Co-ocorrência, como momento da diferença inversa, contraste, segundo momento angular, entropia e diferença da entropia.

Segundo PEREIRA (2001, p.09), a co-ocorrência explora a dependência do nível espacial de cinza de uma imagem digital. Esta é construída baseada na orientação e na distância entre os “pixels” da imagem e então extraídas estatísticas significativas da matriz como a representação de textura. Esta matriz possui uma abordagem estatística.

2.5 - CLASSIFICAÇÃO

Um modo de tornar a interpretação de imagens mais simples consiste em separar grupos de elementos da imagem (“pixels” ou regiões) com características similares em classes, geralmente associadas a tipos de uso ou cobertura do solo. Tomando como hipótese que os elementos da imagem que cobrem um mesmo alvo

têm características similares, algoritmos computacionais podem ser usados para dividi-los em grupos. Este processo é denominado classificação de imagem.

Os métodos de classificação podem ser “pixel” a “pixel” ou por regiões, dependendo do nível de abstração.

O método “pixel” a “pixel” admite como o menor elemento a ser classificado o “pixel” e o método por regiões admite como menor elemento o segmento, ou seja, a região que pode representar um objeto na cena ou parte deste.

Para a execução da classificação de uma imagem digital são necessárias as seguintes etapas: Definição das classes e amostragem; fase de destino – classificação e a avaliação da classificação.

Na definição das classes e amostragem são determinadas as classes de interesse e para estas classes são calculados parâmetros estatísticos que as descrevem a partir de amostras representativas cuidadosamente escolhidas. A amostragem é o primeiro passo fundamental para uma boa classificação supervisionada, pois nesta as características das classes são definidas. É importante considerar que a qualidade das amostras é fundamental para o processo.

Após a realização da seleção e estudo das classes que entraram na classificação, inicia-se a Fase de Destino – Classificação que trata da distribuição de cada elemento da imagem a uma das classes previamente selecionadas. Segundo JENSEN (1996, p.245), em Sensoriamento Remoto, vários métodos são utilizados, tais como: mínima distância, método do paralelepípedo e o método da máxima verossimilhança (ou geral Gaussiana).

O método da distância mínima atribui cada elemento desconhecido à classe cuja média é mais próxima, ou seja, a atribuição de uma classe ao elemento depende da distância em relação à classe a qual este foi atribuído.

O método do paralelepípedo considera uma área no espaço de atributos ao redor do conjunto de treinamento. Essa área tem a forma de um retângulo ou paralelepípedo, definido pelos valores máximo e mínimo do conjunto de treinamento.

Todos os elementos dentro do paralelepípedo serão classificados como pertencentes à classe representada nesta área.

O último método, mais comumente utilizado é o da máxima verossimilhança, sendo que o mesmo utiliza parâmetros estatísticos. Este método pressupõe que os elementos de cada classe seguem uma distribuição normal multivariada, e o método avalia as probabilidades que um determinado elemento tem de pertencer às classes identificadas e o designa àquela cuja probabilidade é maior (NISHIDA, 1998 p.58).

Na última fase, a fase de avaliação da classificação, a qualidade do produto final é avaliada através de um índice que indica a acurácia, a confiabilidade do usuário no produto. Nesta fase, a verdade de campo é comparada com o produto produzido.

2.6 - INTEGRAÇÃO DE DADOS

Quando a informação espectral contida nas bandas de imagem de satélite, não são suficientes para discriminar classes de cobertura do solo, é possível efetuar um estudo integrado, utilizando, além dos dados espectrais, outras informações derivadas de outras fontes, como por exemplo, mapas ou levantamentos laser scanner.

De acordo com RIBEIRO (2003, p.13), o intuito da integração de dados é de ampliar a quantidade e qualidade de informações a respeito da área em estudo. Os dados de diferentes fontes vêm complementar as informações que podem ser extraídas das imagens adquiridas por sensoriamento remoto, visando melhorar a acurácia dos mapas temáticos e auxiliar na extração de informações da cena imageada.

Cientistas da área das Ciências da Terra, tais como (WARNER e SHANK., 1997, p. 1285 - 1294); (HARRIS e MURRAY., 1990, p. 1631 - 1641); (CENTENO et al., 1999, p. 19 - 28); (RIBEIRO e CENTENO, 2002 p. 47-57) e (ANDRADE, 2003 112p.), podem ser citados como referência em estudos de integração de dados.

De acordo com KIEMA e BÄHR (2000 p.32), informação espectral é convencionalmente empregada na classificação de imagem multiespectral. A inadequação disto, na extração de objetos urbanos, tem sido reconhecida em muitos

estudos (FUNG E CHAN, 1994); (HAALA E BRENNER, 1999) etc. Em aplicações onde a alta acurácia é requerida, é imperativo realçar a base da feição do objeto. Isto precisa ser ampliado para incluir características da feição espacial (estrutura, tamanho, forma, topologia) e espectral (assinatura espectral, textura). Estas podem ser incluídas também, explicitamente através da integração de diferentes fontes de dados multi-sensor, ou implicitamente através do uso do método de segmentação apropriado. Devido os motivos citados acima, os autores levantam a hipótese de em trabalhos futuros serem incorporadas informações contextuais na extração de feições.

Um exemplo de integração é a classificação empregada em imagens multi-espectrais é a classificação hierárquica utilizando fusão de dados. A fusão de dados também pode ser usada no conceito de canal adicional. KIEMA (2000, p. 489) define classificação hierárquica como uma técnica estruturada, através da qual os diferentes dados a serem fundidos são aplicados de tal forma, que sucessivamente dividem a área de trabalho em classes de objetos mais detalhados. Este método depende de fatores como a estrutura do dado a ser fundido e as características específicas da imagem que precisa ser preservada. A integração de dados é definida como uma armação formal na qual são expressas meios e ferramentas para a aliança de dados originados de diferentes fontes. O objetivo é obter informação de maior qualidade. No método de canal adicional, os diferentes dados a serem fundidos são introduzidos como canais separados em uma integração dentro de um dado expandido. O estudo aponta que a principal desvantagem da classificação hierárquica é a propagação de erros nos passos subseqüentes da mesma. Por outro lado a principal vantagem do canal adicional é a flexibilidade dos dados processados.

Outro método de classificação, muito utilizado para integração de dados de diferentes fontes é a rede neural artificial a qual será dissertada na próxima seção.

2.7 - REDES NEURASIS ARTIFICIAIS

O objetivo do homem em criar máquinas que simulem a inteligência do

cérebro humano, foi o principal motivador para o desenvolvimento desta área da inteligência artificial.

Uma rede neural pode ser definida como um arranjo de unidades (neurônios interconectados) que trabalham em paralelo para classificar e generalizar dados de entrada em classes de saída. São atribuídos pesos a interconexões e com os dados de entrada é determinado o nível de ativação da unidade na rede. A atividade desta unidade influencia no nível de ativação de outras unidades na rede, devido os pesos entre as unidades serem modificados até que a rede esteja apta a operar resultados desejados. O aprendizado da rede é determinado por um processo repetitivo que tenta adequar os pesos a uma saída correta esperada, para cada amostra de treinamento. É relevante salientar que um resultado significativo requer um conjunto de treinamento com amostras puras e representativas das classes (RIBEIRO e CENTENO, 2002 p.58).

Nas seções que seguirão, serão discutidos conceitos fundamentais das redes neurais, apresentando primeiro o funcionamento do cérebro biológico e depois sua modelação matemática através das redes neurais artificiais.

2.7.1. - O cérebro e o neurônio biológico

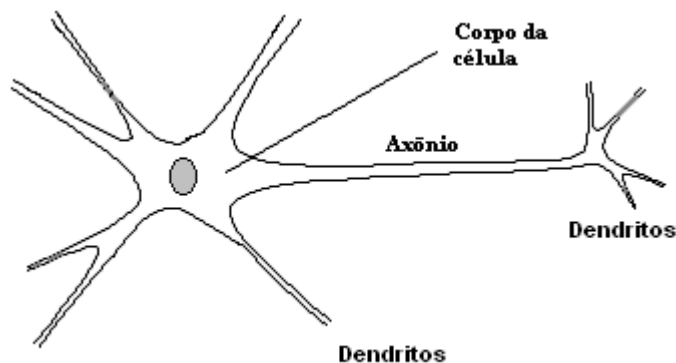
De acordo com HAYKIN (2001, p.26), o cérebro é um sistema de processamento de informação altamente complexo, não-linear e paralelo. Ele tem a capacidade de organizar seus constituintes estruturais, conhecidos por neurônios, de forma a realizar certos processamentos (p. ex., reconhecimento de padrões, percepção e controle motor) muito mais rápido que o mais ágil computador hoje existente. Sua capacidade de manipular dados, e a maneira que permite constituir imagens mentais de pessoas, lugares e eventos baseado em fragmentos de informação, como sons ou cheiros familiares, ou simplesmente a lembrança de uma data de calendário, através da organização estrutural de seus neurônios. Esta organização estrutural (espécie de rede) estimula pesquisadores a tentar imitar o comportamento do cérebro.

Como o cérebro é baseado na estruturação de seus neurônios, serão mostradas

as principais características de um neurônio biológico, para uma posterior compreensão, a respeito do funcionamento das redes neurais artificiais.

O neurônio biológico é uma célula que pode ser dividida em três seções: o corpo da célula, os dendritos e o axônio, cada uma com funções específicas, porém complementares (Figura 2.5).

FIGURA 2.5 – NEURÔNIO BIOLÓGICO



Como se percebe no modelo do neurônio biológico (Figura 4.10), fibras nervosas estão conectados ao corpo da célula, chamadas de dendritos. Partindo do corpo da célula existe uma fibra nervosa mais grossa chamada axônio, da qual surgem ramificações e sub-ramificações. No fim destas ramificações estão os pontos de transmissão para os outros neurônios, chamados de junções sinápticas ou sinapses. A sinapse é quem controla a transmissão do fluxo de informação entre a saída do neurônio e a entrada do neurônio (vizinho). Os pontos de recepção onde as sinapses vão conectar-se aos outros neurônios podem ser os dendritos ou mesmo o corpo da célula. O axônio de um neurônio possui algumas centenas de sinapses que irão conectar-se com outros neurônios (VIEIRA, 2000 p.253).

A célula torna-se ativa quando o neurônio transmissor libera substâncias químicas, causando uma queda no potencial elétrico no corpo da célula receptora. Se este potencial alcança o limite de ativação da célula, um pulso é enviado através do axônio (ABELÉM, 1994 p.18). Portanto, impulsos são passados ao neurônio através de seus dendritos. Estes impulsos são combinados no corpo da célula e, dependendo da intensidade destes impulsos de entrada, um impulso pode ser passado aos outros

neurônios através do axônio.

2.7.2. - Neurônio artificial e as redes neurais

O neurofisiologista Warren McCulloch e o matemático Walter Pitts, da Universidade de Illinois em 1943, inspirados pelos avanços na determinação experimental das propriedades eletrofisiológicas do neurônio, e pelos recentes resultados de Alan Turing e John von Neumann que indicavam ser booleana a natureza essencial da inteligência, descontada todas as imprecisões, publicaram o primeiro trabalho sobre “neurônios formais” (KOVÁCS, 1996 p.28).

Segundo KOVÁCS (1996, p. 29), a essência da proposta de McCulloch e Pitts foi: “A inteligência é equivalente ao cálculo de predicados que por sua vez pode ser implementado por funções booleanas. Por outro lado, o sistema nervoso é composto de redes de neurônios, que com as devidas simplificações, tem a capacidade básica de implementar estas funções booleanas. Conclusão: a ligação entre inteligência e atividade nervosa fica estabelecida de forma científica”.

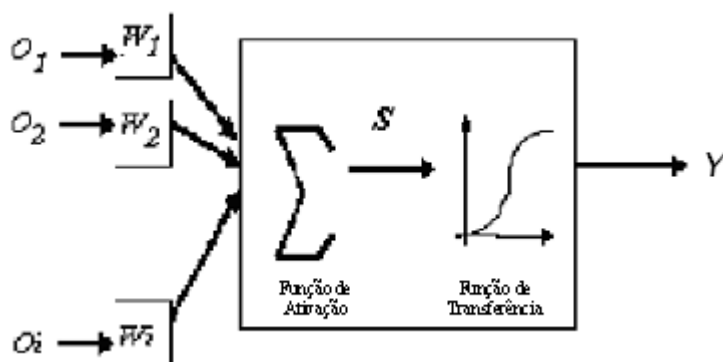
De acordo com TAFNER et al (1996, p.27), o trabalho subsequente de Hebb em 1949 baseou-se em modelos matemáticos que tentavam capturar o conceito de aprendizado por associação. Durante os anos de 1950 e início de 1960 foram desenvolvidos, por Rosenblatt, provas matemáticas mostrando que as máquinas que aprendem (chamadas também de perceptrons) quando treinadas com conjuntos de treinamento linearmente separáveis, convergiriam para uma solução em um número finito de passos iterativos. A solução tomaria a forma de coeficientes de hiperplanos capazes de separar corretamente as classes representadas pelos padrões no conjunto de treinamento.

Segundo GONZALES e WOODS (2000 p.425), resultados tratando do desenvolvimento de novos algoritmos de treinamento para “perceptrons” multicamadas, que resolvem problemas linearmente não separáveis, foram divulgados por Rumelhart, Hinton e Willians em 1986. O sucesso do perceptron com múltiplas

camadas estabeleceu-se como um dos principais modelos de redes neurais atualmente em uso.

Portanto para dissertar sobre redes neurais artificiais antes é necessário redigir sobre o funcionamento de seus neurônios. O modelo matemático de um neurônio artificial pode ser entendido com ajuda da Figura 2.6

FIGURA 2.6 - EXEMPLO DE NEURÔNIO ARTIFICIAL



FONTE: Adaptado de VIEIRA (2000, p.255)

Onde O_i : entrada;
 i : índice da i -ésima entrada;
 W_i : peso;
 S : valor de saída da função de ativação;
 Y : saída

De acordo com ABELÉM (1994, p.18), no neurônio artificial várias entradas (O_i) são possíveis. Cada entrada é multiplicada por um peso correspondente (W_i), gerando entradas ponderadas, igual à força das sinapses. O corpo do neurônio é emulado pela soma ponderada dos valores " $O_i * W_i$ " recebidos pelo neurônio. A partir da soma ponderada das entradas e usando uma função de ativação, um valor de saída é gerado (S). Este valor será comparado com o valor limite para ativação do neurônio. Se este valor alcançar o limite de ativação do neurônio, o valor de saída "Y" é transmitido a outro neurônio simulando o fluxo de impulsos no cérebro.

2.7.2.1 - A função de transferência

A função de transferência do neurônio é quem controla suas saídas, ou seja, gera a saída que será passada a outro neurônio. A função não linear - sigmóide é a

mais utilizada para a classificação de imagens digitais e o principal alvo das pesquisas (ABELÉM, 1994 p.19; VIEIRA, 2000 p.255; TORINO E BARRETO, 2002 p.1437; HANNA E MANDIC, 2003 p.156). Entretanto poderia ser aplicado a outras funções, tais como: logarítmico, linear, binária, etc. Adota-se estas funções não lineares, devido ao fato de simularem mais precisamente os neurônios biológicos (VIEIRA, 2000 p.255).

A função de transferência sigmóide é calculada através da equação 2.4.

$$\alpha = \frac{1}{1 + e^{-s}} \quad (2.4)$$

S – valor de saída da função de ativação;

2.7.3. - Arquitetura das redes neurais artificiais

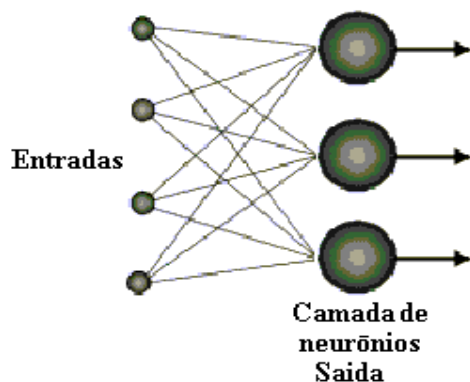
Segundo HAYKIN (2001, p.46-49), existem três classes de arquiteturas de redes fundamentais: Redes alimentadas adiante (“feed-forward”) com camada única, redes alimentadas diretamente com Múltiplas Camadas, Redes Recorrentes.

As redes neurais são organizadas em camadas conectadas, podendo ser a “camada de entrada”, a(s) camada(s) escondida(s) ou oculta(s) e a camada de saída. Na camada de entrada, ou simplesmente nomeada “entrada” devido não ocorrer processamento, os dados são apresentados à rede. A camada de entrada se comunica a uma ou mais camadas ocultas onde ocorre o processamento por meio de um sistema de conexões ponderadas. A camada de saída apresenta o resultado do processamento. O funcionamento da rede é determinado pela conexão entre as unidades de processamento (GERMAN e GAHEGAN, 1996 p.972-973; MURION, 1996a p.996), definida pelos pesos W_i de cada ligação.

Nas redes alimentadas adiante com camada única os neurônios estão organizados na forma de camadas. Uma camada de entrada de nós de fonte que se projeta sobre uma camada de saída de neurônios (nós computacionais), mas não vice-versa. Em outras palavras, esta rede é estritamente do tipo alimentada adiante ou

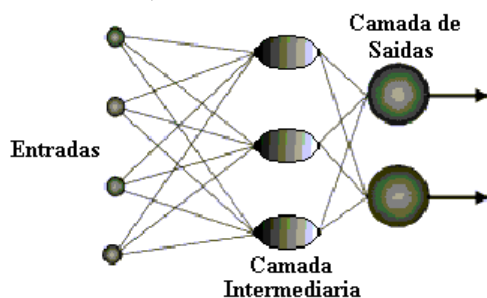
acíclica. Esta é chamada rede de camada única, sendo que a designação “camada única” se refere à camada de saída de nós computacionais. Não é contada a camada de entrada de nós de fonte, porque nesta não é realizada computação (Figura 2.7).

FIGURA 2.7 – ARQUITETURA DA REDE ALIMENTADA ADIANTE COM CAMADA ÚNICA



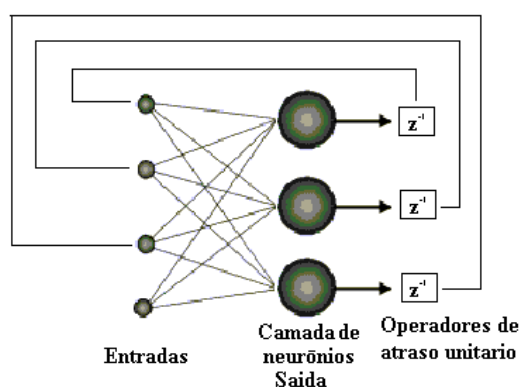
A rede alimentada diretamente com múltiplas camadas é caracterizada pela presença de uma ou mais camadas ocultas, cujos nós computacionais são chamados de neurônios ocultos ou unidades ocultas. A função dos neurônios ocultos é intervir entre a entrada externa e a saída da rede (Figura 2.8). Segundo HAYKIN (2001, p.47), adicionando-se uma ou mais camadas ocultas, tornamos a rede capaz de extrair estatísticas de ordem elevada. Os neurônios da camada de entrada da rede fornecem os respectivos elementos do padrão de ativação (vetor de entrada), que constituem os sinais de entrada aplicados aos neurônios na segunda camada (a primeira camada oculta). Os sinais de saída da segunda camada são utilizados como entradas para a terceira camada, e assim por diante. O conjunto de sinais de saída dos neurônios da camada de saída (final) da rede constitui a resposta global da rede para o padrão de ativação fornecido pelos nós de fonte da camada de entrada (primeira).

FIGURA 2.8 – ARQUITETURA DE UMA REDE NEURAL FEED-FORWARD



As redes recorrentes se caracterizam por terem pelo menos um laço de realimentação. Uma rede recorrente pode consistir, por exemplo, de uma única camada de neurônios, com cada neurônio alimentando seu sinal de saída de volta para as entradas de todos os outros neurônios (Figura 2.9). Os laços de realimentação envolvem o uso de ramos particulares compostos de elementos de atraso unitário (Figura 2.9), o que resulta em um comportamento dinâmico não-linear, admitindo-se que a rede neural contenha unidades não-lineares (HAYKIN, 2001 p.48).

FIGURA 2.9 – ARQUITETURA DA REDE RECORRENTE



2.7.4. - Treinamento da rede neural artificial

Segundo HAYKIN (2001, p.75), uma rede neural aprende acerca do seu ambiente através de um processo iterativo de ajustes aplicados a seus pesos sinápticos. HAYKIN define:

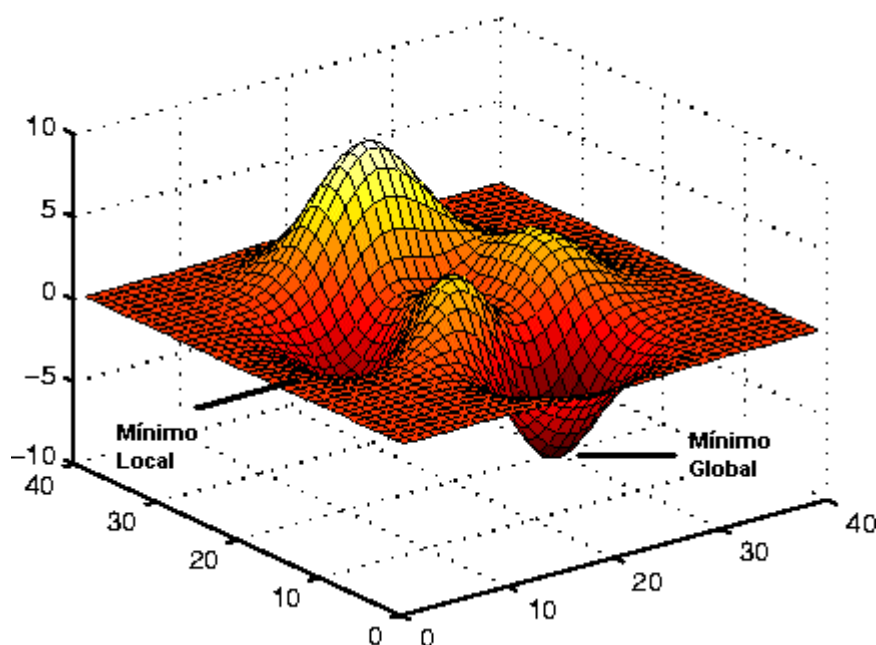
“Aprendizagem é um processo pelo qual os parâmetros livres de uma rede neural são adaptados através de um processo de estimulação pelo ambiente no qual a rede está inserida”.

A rede neural pode ser treinada para resolver problemas de classificação multiespectral. Segundo RIBEIRO e CENTENO (2002, p.57) treinar uma rede neural significa ajustar a sua matriz de pesos, de forma que o vetor de saída coincida com um certo valor desejado para cada vetor de entrada. O treinamento de uma rede pode ser pelo método de aprendizado supervisionado e não-supervisionado.

No aprendizado supervisionado considera-se que o analista possui o conhecimento sobre o ambiente, é capaz de representá-lo por um conjunto de amostras. No entanto o ambiente é desconhecido pela rede neural. Por este motivo o analista fornece à rede uma saída correspondente à entrada. A resposta desejada representa a ação ótima a ser realizada pela rede neural. Os parâmetros da rede são ajustados sob a influência do vetor de treinamento e do sinal de erro. O sinal de erro é definido como a diferença entre a resposta desejada e a resposta real da rede. Este ajuste ocorre iterativamente com o objetivo de fazer a rede neural imitar o analista. Desta forma o conhecimento do ambiente disponível ao analista é transferido a rede neural através do treinamento (HAYKIN, 2000 p.88).

Neste processo iterativo, o erro cometido na classificação, todos os elementos podem ser usados como medida de desempenho da rede na iteração. Este erro é uma função dos parâmetros livres do sistema e pode ser visualizado como superfície de erro onde o mínimo absoluto ou global é procurado (Figura 2.10).

FIGURA 2.10 – SUPERFÍCIE DO ERRO MÉDIO QUADRÁTICO COM MÍNIMO LOCAL E MÍNIMO GLOBAL.



FONTE: Adaptado do MATLAB (1998)

A superfície de erro é obtida a partir do erro cometido em todos os exemplos de treinamento. Qualquer operação do treinamento supervisionado é representada

como um ponto sobre a superfície de erro (Figura 2.10). Para que o sistema melhore seu desempenho, o ponto que está representado sobre a superfície de erro deve ser movido para baixo sucessivamente em direção a um ponto mínimo da superfície de erro. O ponto mínimo pode ter um mínimo local ou um mínimo global (Figura 2.10). O gradiente da superfície de erro em qualquer ponto é um vetor que aponta na direção da descida mais íngreme. O mínimo global é encontrado utilizando-se de gradiente decrescente (KOVÁCS, 1996 p.83-87).

No aprendizado não-supervisionado não há um crítico para supervisionar o processo de aprendizagem e nem exemplo da função a ser aprendida pela rede. São dadas condições para realizar uma medida independente da tarefa da qualidade da representação que a rede deve aprender, e os parâmetros livres da rede são otimizados em relação a esta medida. Uma vez ajustada à rede as regularidades estatísticas dos dados de entrada, ela desenvolve a habilidade de formar representações internas para codificar as características da entrada e assim criar automaticamente novas classes (HAYKIN, 2001 p.90; MURION, 1996b p.1028).

2.7.5. - Algoritmos para o treinamento da rede neural artificial

Existem muitos modelos conexionistas, e uma quantidade de publicações que se dedicam a classificá-los (LIPPMAN, 1987 p. 5-10). Dentre estes modelos, foram selecionados alguns que serão apresentados a seguir com seus respectivos idealizadores:

- Perceptron - proposto por F.Rosenblatt em 1957;
- Perceptron de Multi-Camadas - proposto por P.Werbos, D.Parker, D.Rumelhart em 1974;
- Classificador de Carpenter-Grossberg - proposto por G.Carpenter, S.Grossberg em 1978;
- Rede de Kohonen - proposto por T. Kohonen em 1980;
- Rede de Hopfield - proposto por J.Hopfield em 1982.

No entanto, dentro do escopo desta pesquisa, irá ser exemplificada apenas a característica principal do modelo perceptron multicamadas alimentada adiante utilizando o algoritmo de aprendizado retro-propagação.

Devido a grande maioria dos problemas práticos serem não linearmente separáveis, a rede perceptron multicamadas gerou muito interesse pela habilidade em aprender a reconhecer padrões desta natureza, tendo assim uso generalizado.

De acordo com VIEIRA (2000 p.250), as principais características, desta rede, são:

- Adequado para problemas não linearmente separáveis.
- Rede neural com, no mínimo, três camadas, sendo: camada de entrada, camada(s) intermediária(s) ou escondida(s), camada de saída. Sendo que um neurônio interliga vários outros neurônios de outra camada.
- Vários neurônios podem ser combinados para computar problemas mais complexos.
- Possui um algoritmo de aprendizado, podendo assim melhorar seu desempenho.
- O algoritmo de aprendizado deste modelo funciona por ciclos. Em cada ciclo, um conjunto de padrões é apresentado à rede neural, e calculado o erro em relação ao resultado esperado. O erro é usado para atualizar os pesos, e então um novo ciclo é aplicado.
- Possuem entradas e saídas binárias que assumem os valores +1 ou -1;
- A função de ativação é a função sigmóide;

Diversas pesquisas têm sido realizadas utilizando redes neurais artificiais (RNA) como instrumento de classificação de imagens digitais, como por exemplo, na detecção de focos de incêndio (TORINO E BARRETO, 2002); na classificação de áreas agrícolas, na separação de culturas (VIEIRA, 2000); na medicina para a detecção de câncer em Mamogramas Digitalizados (KINOSHITA, MARQUES E FRERE

2000); no desenvolvimento de métodos para a detecção de movimentos de objetos, através de seqüência de imagens digitais (FERNANDEZ – CABALLERO E AL, 2003) e outras. Com o mesmo propósito, outros pesquisadores têm se empenhado para melhorar a arquitetura ou a classificação na saída da rede, fazendo com que fique mais rápida no seu processamento, aprendizado e treinamento como visto em JI (2000), TONDER E EJIMA (2000), OLIVEIRA E BAUCHSPIESS (2001), MARTINS, GUIMARÃES E FONSECA (2002), YILMAZ E AKIN (2002), BAKKER E HESKES (2003).

VIEIRA (2000 p. 249) reporta que o algoritmo de aprendizado da rede neural mais utilizado na classificação de imagem oriundas do sensoriamento remoto é o retro-propagação, com arquitetura multi-camadas desenvolvido pelo grupo PDP (“Parallel Distributed Processing”) constituído pelos autores RUMMELHART, PARKER, LE CUM e outros em 1980. Pesquisadores partem deste algoritmo, e fazendo as devidas adaptações, emprega-os em outras áreas (RIEDMILLER e BRAUN, 1993 p.586-591; KINOSHITA, MARQUES e FRÈRE, 2000 p.1214-1216; PHILIP e JOSEPH, 2003 p.215-223).

RIEDMILLER e BRAUN (1993 p.586-591), modificaram o algoritmo de aprendizado retro-propagação no intuito de melhorar o tempo de aprendizado obtendo resultados positivos. Como resultado comparativo os autores observaram que enquanto o retro-propagação necessita de 15000 ciclos para convergir o treino, o novo algoritmo (RPROP) necessita de 20 ciclos. Uma característica apresentada pelo algoritmo RPROP é que a entrada dos dados para o treinamento é realizada em grupo enquanto no retro-propagação é realizada tanto em grupo como alternando as classes.

OLIVEIRA e BAUCHSPIESS (2001 p.73-78) utilizaram RNA de base radial para a detecção de bordas empregando cadeias direcionais. Os autores compararam o método de rede neural de base radial com o algoritmo retro-propagação existente na arquitetura multi-camada, verificando uma melhor eficiência na rede neural de base radial. Portanto verifica-se a possibilidade de realizar uma segmentação utilizando

redes neurais para a identificação de regiões similares.

RIBEIRO (2003, 160 p.) utilizou a RNA para integrar variáveis de diferentes fontes no intuito de identificar unidades geológicas (classes de rochas no subsolo terrestre) com a proposta de automatizar a separação das mesmas. O autor encontrou várias unidades geológicas bem definidas no produto final, concluindo assim que as RNA's são eficientes para integrar dados de diferentes fontes.

ANDRADE (2003, 112 p.) utiliza a RNA para identificação de feições na área urbana observando a influência de descritores de forma em imagem de alta resolução espacial (IKONOS II). Conclui que os coeficientes que mais influem na classificação da imagem de alta resolução espacial são: área, perímetro, coeficiente de compacidade e razão entre eixos da elipse. Ainda conclui que a RNA é uma ferramenta útil para a integração de variáveis espaciais e espectrais derivadas de imagens do sensoriamento remoto.

CAPÍTULO 3 - MATERIAL E MÉTODO

3.1 - MATERIAL

O material usado nesta pesquisa, bem como o programa utilizado será descrito a seguir.

3.1.1. - Área de estudo

Uma região, da cidade de Curitiba, próximo as imediações do Centro Politécnico, foi escolhida como alvo do estudo. Esta região está ocupada por uma área urbana típica, com edificações, vias e outros objetos presentes no cotidiano das cidades. A área foi escolhida, devido à disponibilidade de imagens QUICKBIRD II (a imagem foi cedida pela empresa TESE – Tecnologia em Sistemas Espaciais.) e sendo próxima à UFPR possibilita visitas para o levantamento da área de verificação e reconhecimento das classes presente na imagem. A região está situada dentro das coordenadas: 677889.58E, 7184178.02N e 678448.18E, 7183549.42N.

Um programa foi implementado para a classificação mediante RNA's. Este programa possui algumas limitações. Dentre elas, ele suporta apenas recortes de imagens de 500x500 “pixels” e obriga-nos a entrar com imagem compostas por três bandas espectrais. Portanto foram realizados três recortes na imagem QUICK BIRD II.

O primeiro, com 400x400 que está apresentada na Figura 3.1, possuindo as coordenadas 678168.88, 7183828.72 e 678448.18, 7183549.42. Este recorte apresenta vias asfaltadas, casas bem definidas e espaçadas, edificações grandes, como um “Shopping”, árvores e vegetação branda (gramíneas, pequenos arbustos, pastagens etc), este é caracterizado pela vegetação entre as edificações e sua boa distribuição. Um segundo recorte de 500x500 que está apresentado na Figura 3.2, posicionado à esquerda do primeiro, nas coordenadas 677889.58, 7183898.72 e 678238.88,

7183549.42. Este recorte possui as mesmas classes do primeiro recorte, porém destaca-se pela grande aglomeração entre edificações. E finalmente um terceiro recorte, representado pela Figura 3.3, também de 500x500, posicionado acima do segundo recorte, nas coordenadas 677889.58, 7184178.02 e 678238.88, 7183828.72. Este recorte possui as mesmas classes do primeiro, porém caracterizado pelo maior número de vegetação.

FIGURA 3.1 - PRIMEIRO RECORTE DA IMAGEM QUICKBIRD II, DA REGIÃO EM ESTUDO

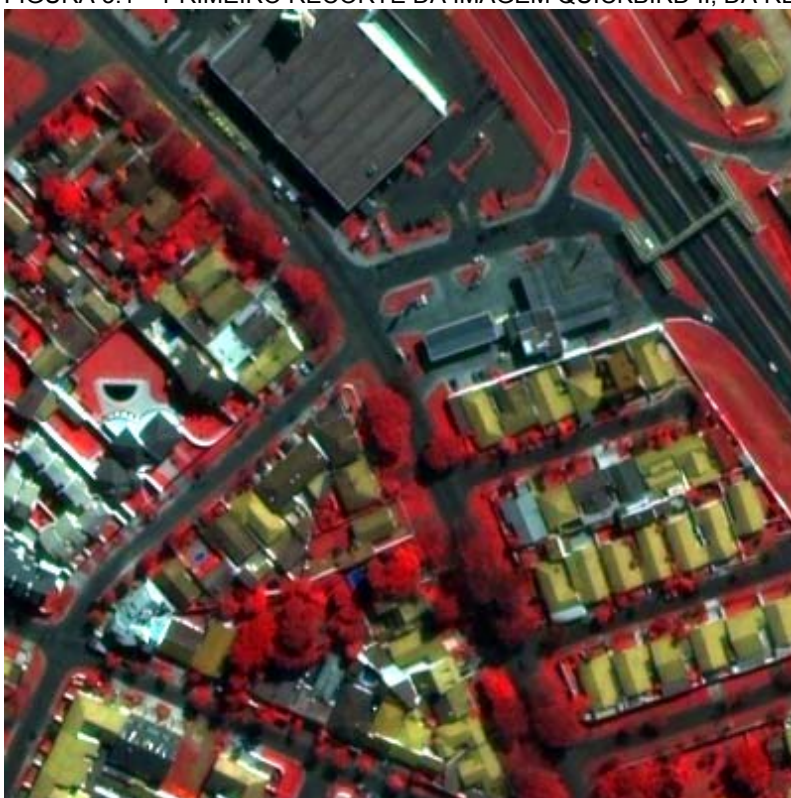


FIGURA 3.2 - SEGUNDO RECORTE DA IMAGEM QUICKBIRD II, DA REGIÃO EM ESTUDO



FIGURA 3.3 - TERCEIRO RECORTE DA IMAGEM QUICKBIRD II, DA REGIÃO EM ESTUDO



3.1.2. - Imagem QUICKBIRD II

Na Tabela 3.1 estão as principais características da imagem de satélite utilizada. Tendo destaque a resolução espacial de 0.61 no nadir para a imagem pancromática e 2.44m na multiespectral. Observa-se que com 25° de inclinação obtém-se uma resolução de 0.72m para a pancromática e 2.88m para a multiespectral. As cenas possuem um imageamento de 16.5km x 16.5km.

A empresa TESE forneceu a imagem utilizada que possui quatro bandas sendo a pancromática com 0.70m de resolução espacial e 2.88m na multiespectral. A mesma executou a fusão das imagens obtendo um produto com 0.70m de resolução espacial, porém com 8 bits de resolução radiométrica, o que impulsionou a realização de uma nova fusão, empregando o método de componentes principais, gerando um produto com 0.7m de resolução espacial e 11 bits de resolução radiométrica. Um recorte deste produto pode ser observado na Figura 3.1.

TABELA 3.1 - PRINCIPAIS CARACTERÍSTICAS DO SENSOR QUICKBIRD II.

Altitude	450 Km
Inclinação	98°
Velocidade	7.1 Km/s
Duração da Órbita	98 min
Tipo de Órbita	Síncrona com o sol
Resolução Espacial	Pancromática: 0.61m / Multiespectral: 2.44m 0.72 m (25°) 2.88m (25°)
Bandas Espectrais	Pan 0.45-0.90 µm Azul 0.45-0.52 µm Verde 0.52-0.60 µm Vermelho 0.63-0.69 µm Infravermelho próximo 0.76-0.90 µm
Resolução Radiométrica	11 bits (2048 níveis de cinza)
Imageamento	Cenas de 16.5km x 16.5km
Frequência de revisita	1 a 3.5, dependendo da latitude (30° off nadir)

Fonte: INTERSAT (2004)

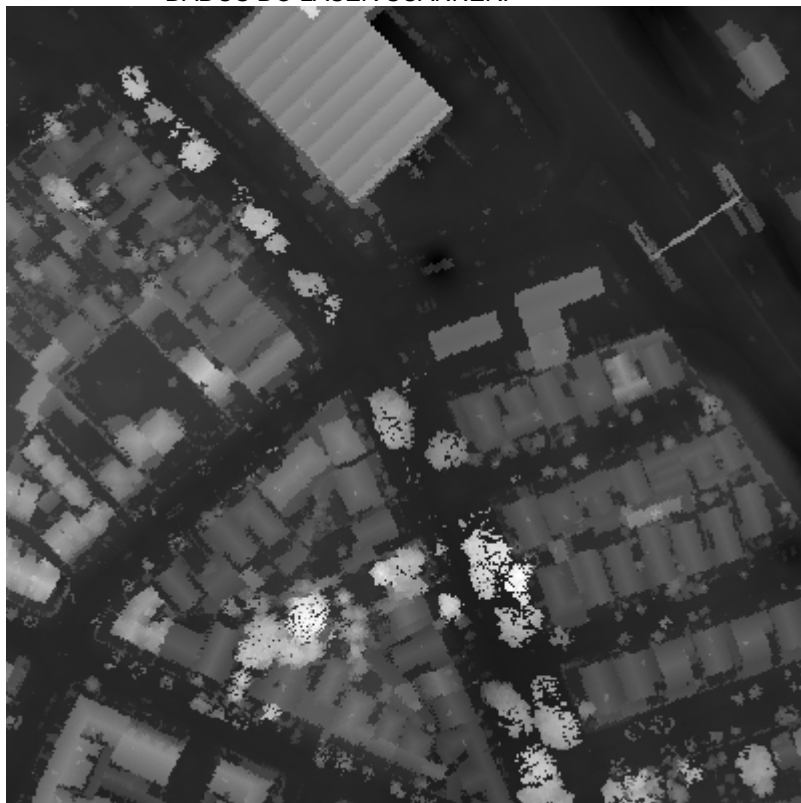
3.1.3. - Grade regular altimétrica

O levantamento laser scanner da região, foi realizado com o sistema OPTECH / INS GPS APPLANIX, pertencente ao Instituto de Tecnologia para o Desenvolvimento (LACTEC - UFPR). O sistema de referência de coordenadas utilizado para o rastreamento foi o WGS84 (World Geodetic System, 1984).

Utilizando os dados disponíveis pelo laser é gerada, através do programa SURFER, uma grade altimétrica regular de 0.7m com as mesmas dimensões da resolução espacial da imagem, bem como as mesmas coordenadas. No intuito de obter apenas as alturas dos objetos em relação ao solo, presentes na cena, foram realizados os seguintes passos:

1. Inicialmente, são coletados pontos (aproximadamente 70 pontos para cada recorte), a partir da grade regular original. Pontos representativos que modelem a superfície do solo, desprezando assim as elevações das edificações, árvores, etc.
2. De posse deste conjunto de pontos, que descrevem a superfície do solo, é gerada uma segunda grade desses pontos.
3. Finalmente, em um terceiro momento, é realizada a subtração das grades. Toma-se a grade original, que possui as elevações com referência geoidal, e subtrai desta, a grade que representa a superfície do solo. Esta operação resulta em uma terceira grade com apenas a altura dos objetos em relação ao solo, contidos na cena. Este produto final pode ser observado através da Figura 3.4.

FIGURA 3.4 - IMAGEM ALTIMÉTRICA DA ALTURA DOS OBJETOS EM RELAÇÃO AO SOLO, ORIUNDA DE DADOS DO LASER SCANNER.



3.1.4. - Programas utilizados

ENVI: Foi utilizado com o objetivo de recortar e registrar a imagem QUICKBIRD II.

ECOGNITION, da Defines: Foi utilizado com o objetivo de segmentar e integrar nesta segmentação dados do laser scanner, considerando também os parâmetros de forma como compacidade e suavidade.

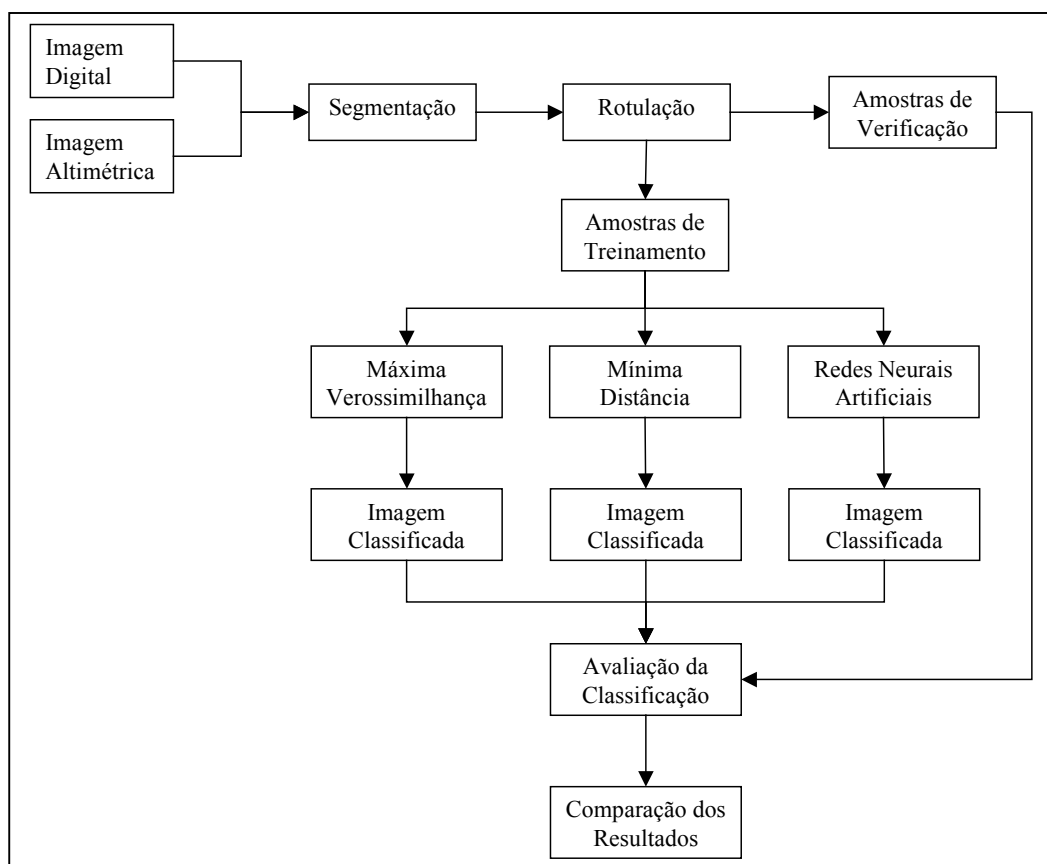
Compilador C++ BUILDER, da Borland: Foi utilizado para implementação do programa, sendo dividido em módulos: operação de dados laser scanner, importação de imagem segmentada, classificadores e verificação.

SURFER, da Golden Software: Foi utilizado para gerar a grade regular a partir dos dados do laser scanner, bem como a subtração das grades gerando uma nova grade com apenas a altura dos objetos presentes na cena.

3.2 - METODOLOGIA

Para descrever a metodologia, inicialmente será introduzido um diagrama de blocos contendo todos os procedimentos necessários para a execução deste trabalho. O diagrama de blocos está representado na Figura 3.5. Nas seções que seguirão estará dissertado detalhadamente cada nível necessário à execução deste projeto.

FIGURA 3.5 - DIAGRAMA DE BLOCOS DA METODOLOGIA A SER EMPREGADA NA DISSERTAÇÃO

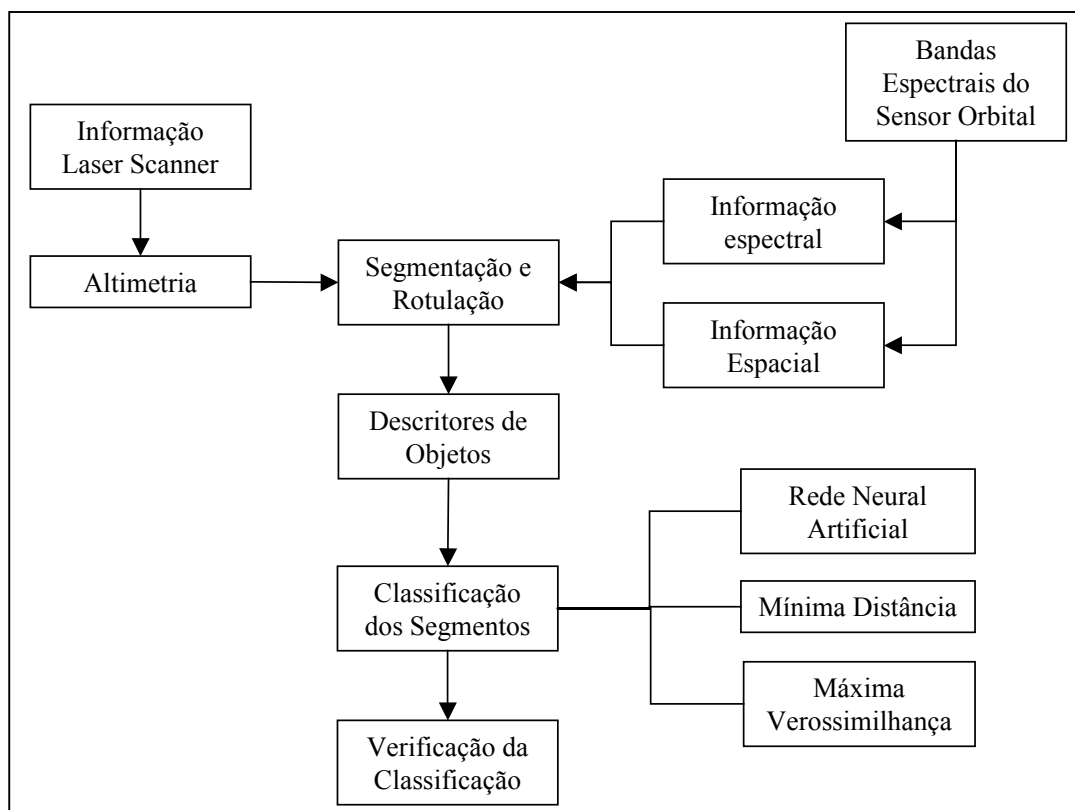


De forma sucinta, inicialmente as imagens espectrais e a grade altimétrica são combinadas e a partir destas é realizada a segmentação. Com a imagem segmentada é realizada a rotulação dos segmentos. Assim torna-se possível a coleta das amostras de treinamento, escolhendo segmentos representativos de cada classe, inicia-se então, o processo de classificação e a geração da imagem temática. Esta nova imagem é comparada com a verdade de campo e assim tem-se uma estimativa da acurácia da classificação.

A metodologia empregada para a implementação do programa em linguagem

C++ pode ser observada pelo diagrama de blocos da Figura 3.6.

FIGURA 3.6 - DIAGRAMA DE BLOCOS DO PROGRAMA IMPLEMENTADO EM LINGUAGEM C++.



3.2.1. - Segmentação da imagem e sua rotulação

3.2.1.1 - Crescimento de Regiões implementado

A implementação de um algoritmo de segmentação é executado mediante a técnica do crescimento de região por agregação de “pixels”, apresentado na seção 2.4.2. Para a área de estudo, o algoritmo implementado gera aproximadamente 70.000 segmentos e os objetos da cena não apresentam suas características de forma, tornando-se difícil a extração destas características dos objetos. Por este motivo, optou-se por utilizar a segmentação espectral-espacial disponível no programa “eCognition”.

3.2.1.2 - Segmentador espacial-espectral do “eCognition”

3.2.1.2.1 - Fundamentos

O agrupamento de “pixels” na segmentação espacial-espectral do “eCognition” é basicamente um algoritmo de crescimento de regiões, que guia a fusão de regiões baseado num critério de semelhança espectral-espacial.

O critério de semelhança é usado para definir a uniformidade dos segmentos. Dois critérios são usados para descrever a separação do objeto da imagem: critério espectral e o critério da forma (suavização e compactação). Estes critérios controlam a homogeneidade dos segmentos ou regiões.

Segundo ECOGNITION (2001, p.3-20), a descrição de heterogeneidade espectral é composta pela soma dos desvios padrão dos valores do contador digital para uma dada banda(c) ponderados pelos pesos w_c , sendo:

$$He = \sum w_c \sigma_c \quad (3.1)$$

Contudo, em muitos casos, a minimização exclusiva da heterogeneidade espectral resulta em segmentos fragmentados, principalmente em imagem muito texturizada. Por esta razão, deve-se considerar além da heterogeneidade espectral o critério de heterogeneidade espacial.

A heterogeneidade espacial é considerada como um desvio de uma forma compacta, descrita pela razão do comprimento da borda do objeto e o número de “pixels” que o compõe. Este cálculo pode ser visto pela equação 3.1. Ainda possibilita descrever a forma suave utilizando ao invés de números de “pixels”, o comprimento de um quadrado ou retângulo que compreenda o objeto (equação 3.2 e 3.3) (ECOGNITION, 2001 p.3-20).

$$hf_{cp} = \frac{l}{\sqrt{n}} \quad (3.2)$$

onde

l = perímetro que compõe o objeto e

n = número de “pixels” que compõe o objeto.

$$hf_{sv} = l/b_A \quad (3.3)$$

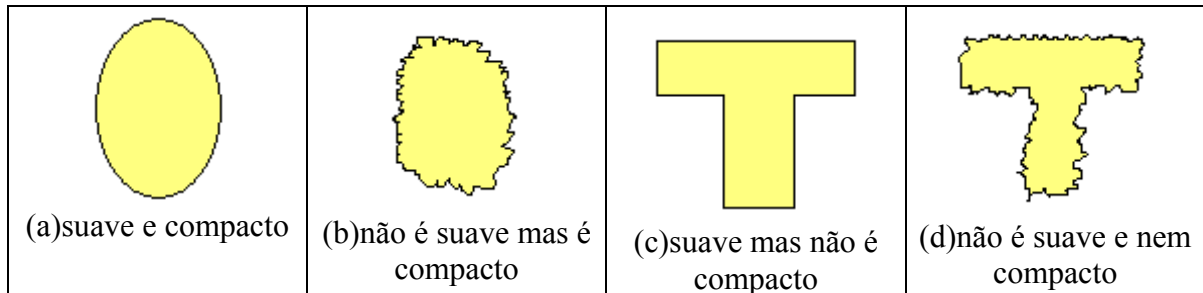
onde

l = perímetro que compõe o objeto e

b_A = perímetro do quadrado que compreende o objeto.

Para esclarecer melhor a diferença entre um objeto compacto e suave, veja os exemplos da Figura 3.7. Na Figura 3.7 (a) observa-se um objeto que é suave e compacto. Em (b) o objeto não é suave porém é compacto. Em (c) o objeto apresenta a característica de suavidade porém não é compacto. Em (d) observa-se que o objeto não é suave e nem compacto.

FIGURA 3.7 - EXEMPLO DE OBJETOS COMPACTOS E SUAVES.



O processo de crescimento de regiões inicia-se com um “pixel” formando um objeto ou região, sendo que, a cada etapa, um par de objetos funde-se para formar um objeto maior. O processo de fusão está baseado no critério de homogeneidade, observando-se a heterogeneidade “h” entre os objetos adjacentes. Na equação 3.4 demonstra-se o cálculo para verificação desta heterogeneidade.

$$h = \sqrt{\sum (f_{1d} - f_{2d})^2} \quad (3.4)$$

onde

h = heterogeneidade entre objetos para um espaço d-dimensional.

f_{1d} e f_{2d} = variâncias espectrais de dois objetos.

A fusão local de objetos é um processo iterativo que objetiva alcançar uma heterogeneidade mínima no processo de agregação. A base é o grau de diferença de heterogeneidade entre duas regiões. Uma vez que a diferença decresce, torna-se maior o critério de fusão. Esta diferença otimiza o processo comparando os atributos das duas regiões. Num dado espaço de atributos, objetos são considerados similares se estiverem espectralmente próximos.

O critério de fusão entre duas regiões adjacentes é descrito pela mudança de heterogeneidade h_{dif} após uma fusão virtual. Dado um determinado critério de heterogeneidade para uma região, o crescimento de heterogeneidade na fusão deve ser minimizado. Existem diferentes possibilidades de descrever a mudança de heterogeneidade antes e depois da fusão. A equação 3.5 corresponde à diferença de heterogeneidade considerando o tamanho dos objetos.

$$h_{dif} = h_m - (h_1n_1 + h_2n_2)/(n_2 + n_1) \quad (3.5)$$

onde

h_1 e h_2 = heterogeneidade para dois objetos;

h_m = heterogeneidade após a fusão virtual; e

n = tamanho dos objetos.

Como já foi citado, o critério de fusão depende da característica espectral e da espacial. De acordo com ECOGNITION (2001, p.3-22), o valor geral do critério de fusão “f” é determinado na equação 3.6, tendo em vista a forma e cor.

$$f = wh_{cor} + (1-w)h_{forma} \quad (3.6)$$

onde

f = critério de fusão;

w = é o peso dado a variável (cor);

h_{cor} = heterogeneidade espectral; e

h_{forma} = heterogeneidade espacial.

As equações 3.7 e 3.8 determinam h_{cor} e h_{forma}

$$h_{cor} = \sum w_c (n_a \sigma_c - (n_1 \sigma_1 - n_2 \sigma_2)) \quad (3.7)$$

onde

c = bandas;

n_a = tamanho do objeto agrupado;

n_1 e n_2 = tamanho dos objetos; e

σ_1 e σ_2 = desvios padrões dos objetos.

$$h_{forma} = w_{cp} h_{cp} + (1 - w_{cp}) h_{sv} \quad (3.8)$$

onde

w_{cp} = peso da compacidade;

h_{cp} = compacidade; e

h_{sv} = suavidade.

A determinação de h_{cp} e h_{sv} , são descritas pelas equações 3.9 e 3.10.

$$h_{sv} = n_a (l_a / b_a) - (n_1 (l_1 / b_1) + n_2 (l_2 / b_2)) \quad (3.9)$$

$$h_{cp} = n_a (l_a / (n_a)^{1/2}) - (n_1 (l_1 / (n_1)^{1/2}) + n_2 (l_2 / (n_2)^{1/2})) \quad (3.10)$$

onde

l_1 e l_2 = são os perímetros dos objetos;

b_1 e b_2 = são os perímetros dos quadrados envolventes;

b_a e l_a = referem-se ao objeto fundido (virtual).

Os pesos são determinados de acordo com os critérios de semelhança definidos pelo usuário. O peso maior é dado para a variável espectral, por ser a imagem multiespectral. Toda a imagem segmentada e os objetos são gerados baseados em critérios ajustáveis de homogeneidade ou heterogeneidade espacial ou espectral.

3.2.1.2.2 - Resultados da segmentação no eCognition

O resultado da segmentação no “eCognition” é uma imagem temática, onde a cada região é atribuído um valor digital, e uma tabela com os descritores de cada região, disponíveis neste sistema. Para a utilização deste resultado nas fases posteriores do processamento, um programa para identificar os segmentos foi implementado.

O programa implementado rotulou os segmentos com um único número para cada segmento. Com estes segmentos foi realizada a amostragem utilizando os critérios apresentados na seção 3.2.5.

3.2.2. - Descrição dos segmentos

No instante da rotulação o algoritmo descreve os segmentos de acordo com sua forma (área, perímetro, coeficiente de compacidade e razão entre eixos da elipse) e a média das alturas presentes na região, ou seja, a média altimétrica a partir da grade regular de elevações.

3.2.3. - Altimetria

Como já mencionado na seção 3.1.3, é gerada uma grade regular com as mesmas dimensões e resolução espacial da imagem a partir das medições pontuais do levantamento do laser scanner. Esta grade contém as alturas de cada “pixel” e possui as mesmas características que a imagem multi-espectral, em termos de tamanho do “pixel”. Outrossim, um conjunto de “pixels” forma a região que descreve o objeto da cena, logo a média das alturas dos “pixels” que compõem uma região é o valor representativo do segmento rotulado.

3.2.4. - Descritores espaciais

Para a descrição da forma dos segmentos neste trabalho, são utilizados a área, o perímetro, o coeficiente de compacidade e a razão entre eixos maior e menor da elipse. Nas próximas seções, eles serão descritos.

3.2.4.1 - Área e Perímetro

Estes são os mais simples descritores utilizados. A área é obtida contando todos os “pixels” que formam a região, e o perímetro é calculado considerando o número de “pixels” conexos (em função do predicado de conectividade escolhido) que constituem o contorno da região.

3.2.4.2 - Coeficiente de compacidade

O coeficiente de compacidade mede o grau de semelhança entre a região e a circunferência de um objeto. O objeto é menos compacto à medida que seu contorno se torna mais irregular. O coeficiente de compacidade (cc), é obtido através da equação 3.11 :

$$cc = \frac{P^2}{4\pi A} \quad (3.11)$$

De acordo com FACON (1993, p.152), o coeficiente de compacidade (cc) é invariante na mudança de escala, já que o mesmo depende dos descritores simples como a área e o perímetro que não dependem da mudança de escala. A compacidade é também invariante na translação e na mudança de orientação.

3.2.4.3 - Razão entre o Eixo Menor e Maior da Elipse

Este descritor emite a informação do contorno que é a razão dos eixos menor e maior da elipse que mais se ajusta ao objeto ou elemento analisado. Estes eixos

podem ser obtidos através da análise por componentes principais (ANDRADE, 2003 p.41).

A partir da determinação da matriz de variância-covariância da posição dos “pixels” pertencentes ao objeto, são calculados os autovetores que correspondem à orientação das 1ª e 2ª componentes. A relação entre os autovalores representa o comprimento relativo dos eixos das componentes principais do segmento (CROSTA, 1992 p.142).

3.2.5. - Coleta de Amostras

Para coletar as amostras que foram utilizadas na classificação da imagem segmentada, foi implementado um programa em linguagem C++ (Figura 3.9). Este algoritmo armazena o valor do rótulo onde é fixado um clique do mouse. Após realizada a coleta de amostras, é possível gerar o arquivo com as amostras e seus descritores de forma e altimetria (Figura 3.8) possibilitando a variação entre classes, intercalando-as. Este princípio produz um melhor aprendizado do conhecimento do projetista pela rede neural artificial.

FIGURA 3.8 - INTERFACE DO ALGORITMO GERADOR DO ARQUIVO DE TREINAMENTO.

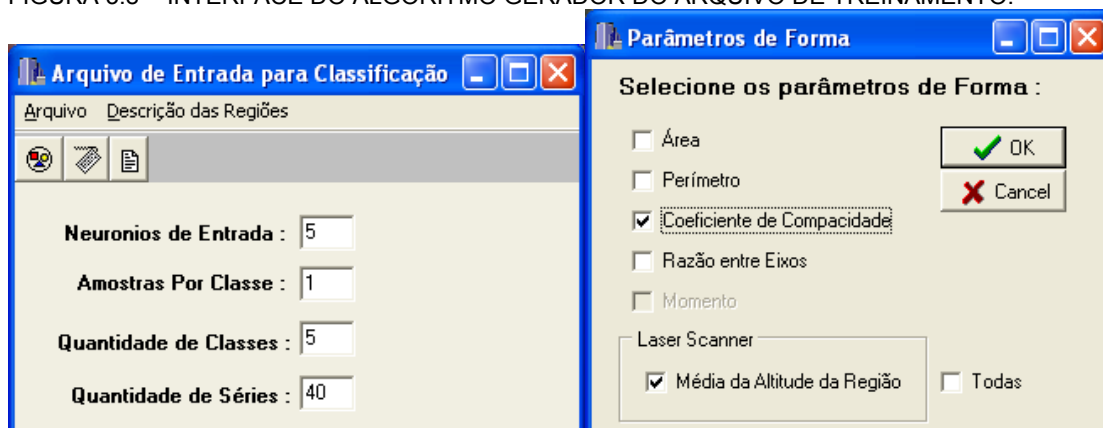
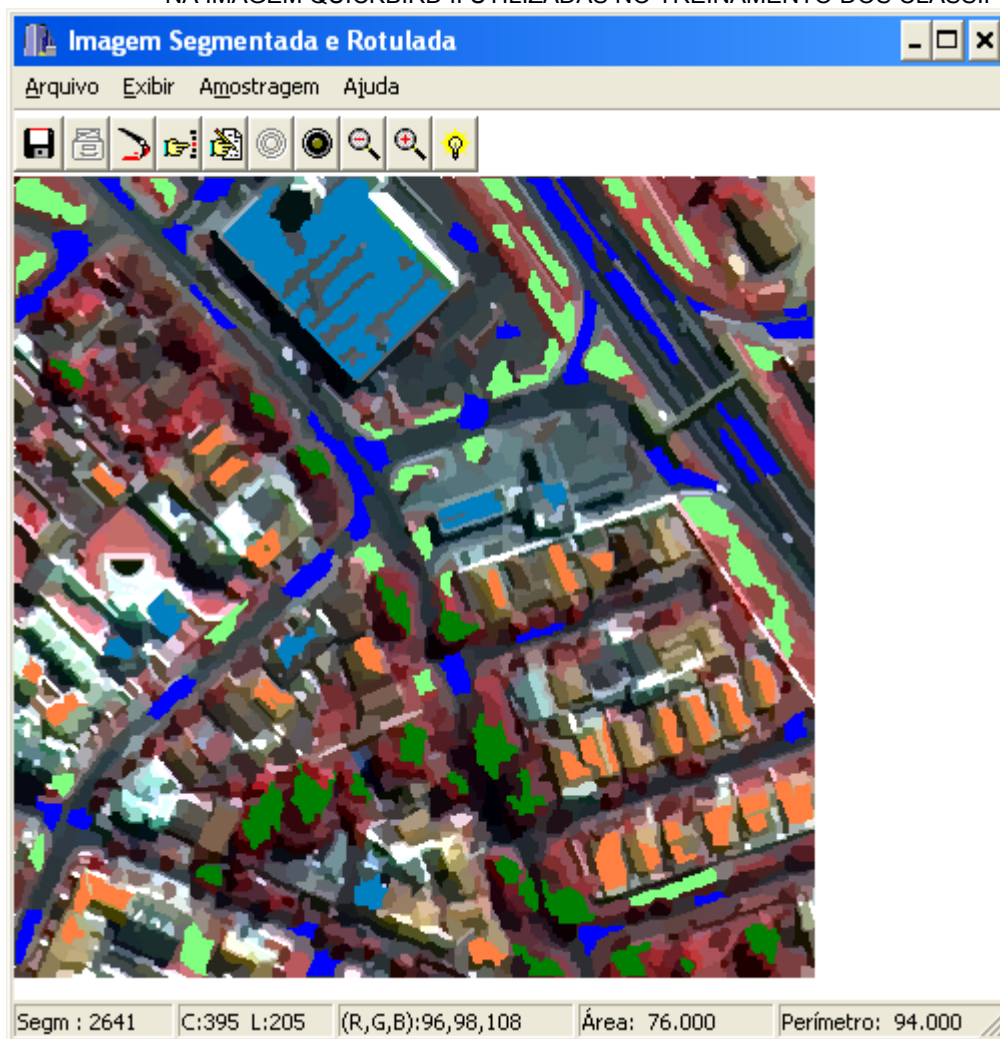
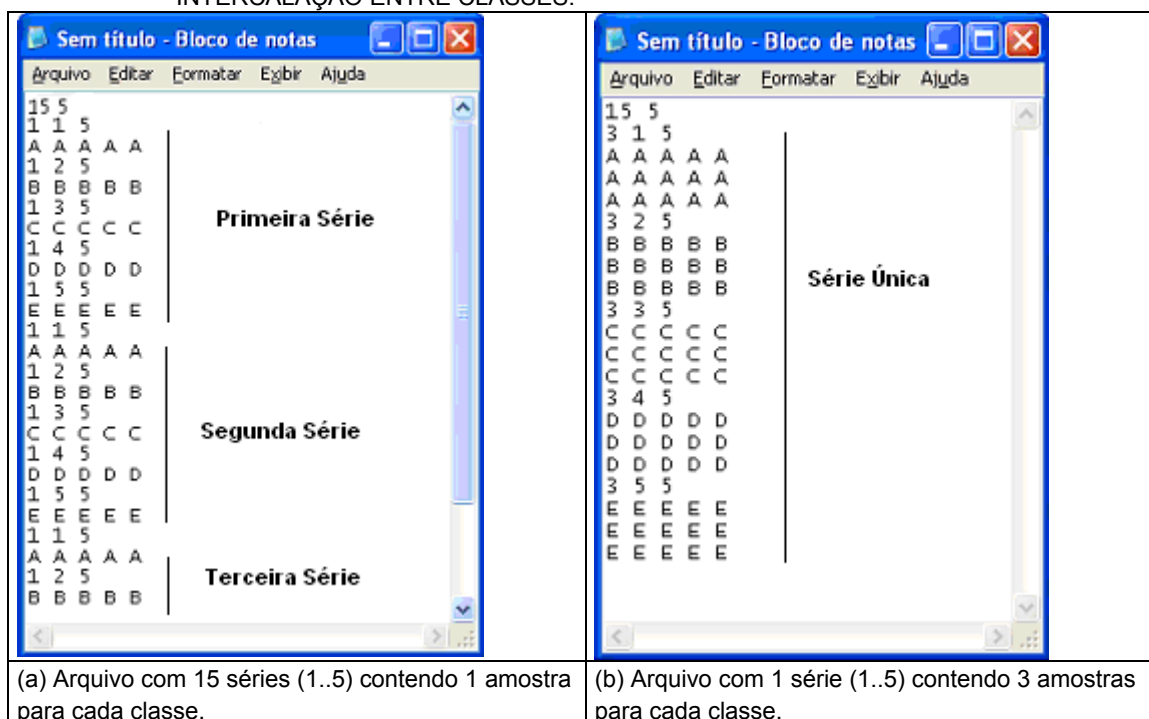


FIGURA 3.9 - INTERFACE DO ALGORITMO COLETOR DE AMOSTRAS COM AS AMOSTRAS COLETADAS NA IMAGEM QUICKBIRD II UTILIZADAS NO TREINAMENTO DOS CLASSIFICADORES.



Analisando usualmente a imagem QUICKBIRD II, percebe-se cinco classes distintas na imagem. Assim foram coletadas 40 amostras para cada classe: Edificação com telhado escuro e Edificação com telhado claro, estrada pavimentada (asfalto), árvore e vegetação rasteira (campo, pastagem, gramíneas, arbustos com baixa altimetria). Com essas classes foram formados os arquivos de amostras com a variação entre classes levando em consideração a alternância das classes no processamento da imagem, ou seja, as amostras foram intercaladas. O arquivo gerado pode ser observado pela Figura 3.10.

FIGURA 3.10 - ARQUIVOS GERADOS PARA TREINAMENTO DOS CLASSIFICADORES CONSIDERANDO A INTERCALAÇÃO ENTRE CLASSES.

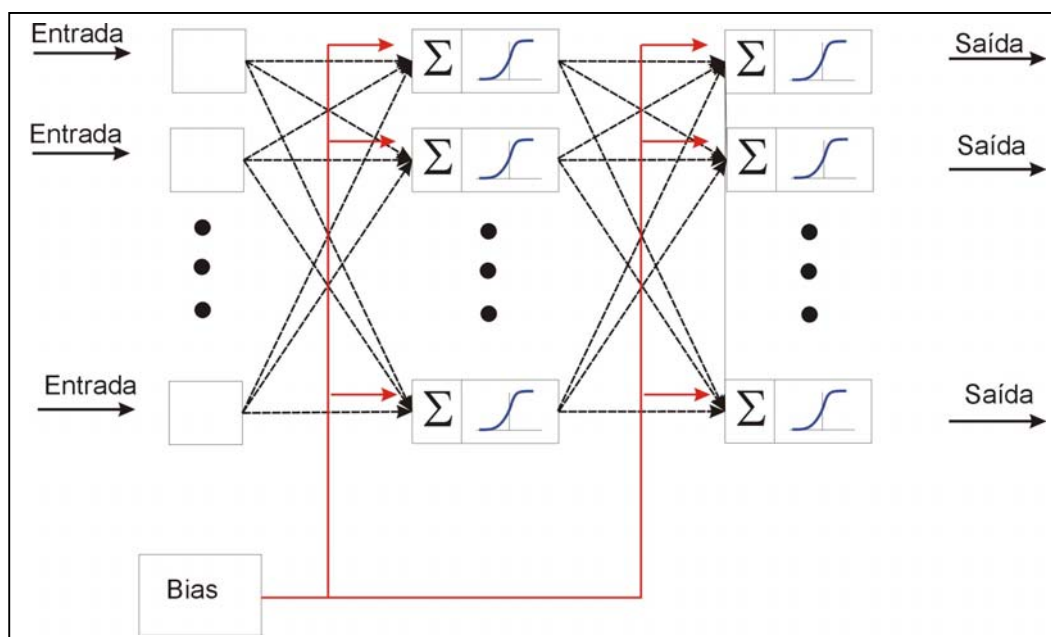


Coletada as amostras da imagem segmentada, inicia-se os processos de classificação. Estes métodos estão apresentados nas seções 3.2.6, 3.2.7 e 3.2.8.

3.2.6. - Implementação da rede neural artificial

Para este trabalho é adotada a arquitetura multicamadas com alimentação adiante (“feed-forward”) com o aprendizado de retro-propagação (“back-propagation”); possui “bias” que tem influência externa ao neurônio, e a função de transferência adotada é a sigmóide como dito anteriormente na seção 2.7.2.1. A Figura 3.11 está representando a arquitetura do modelo multicamada com retro-propagação.

FIGURA 3.11 - ARQUITETURA DO MODELO DE MULTICAMADA COM RETRO-PROPAGAÇÃO.



FONTE: Adaptado de RIBEIRO e CENTENO (2002 p.58).

3.2.6.1 - Algoritmo de aprendizado retro-propagação.

O algoritmo de aprendizado retro-propagação é o algoritmo que ajusta os pesos das conexões intermediárias, sendo isto realizado de trás para frente em função do erro verificado na camada de saída.

Este algoritmo consiste em fornecer variações para alternar os pesos da rede, objetivando minimizar o erro global na saída através do aprendizado. O algoritmo baseia-se no método do gradiente decrescente, que sugere mudanças nos pesos W_{ij} por uma quantidade ΔW_{ij} proporcional ao gradiente da superfície de erro E . (HANNA E MANDIC, 2003 p.156).

$$\Delta W_{ij} = \eta(\sigma E / \sigma W_{ij}) \quad (3.13)$$

onde η é a taxa de aprendizado, i refere-se ao neurônio da camada posterior e j ao neurônio da camada anterior. De acordo com ABELÉM (1994, p.26), executando-se a derivada e fazendo-se a mudança, para cada padrão de entrada individualmente, encontra-se:

$$\Delta W_{ij} = \eta \delta_i \theta_j \quad (3.14)$$

onde

$$\delta_i = (A - \theta_i) * F'(S_i) \quad (3.15)$$

$$F'(S_i) = F(S_i) * (1 - F(S_i)) \quad (3.16)$$

Onde δ_i representa a diferença entre a saída esperada (A), e a saída da rede (θ_i), ou seja, o erro local da unidade i para o padrão procurado, isto multiplicado pela primeira derivada da função de ativação (F) do valor S_i . A Equação 3.15 é conhecida como regra delta (VIEIRA, 2000 p. 259). E finalmente o termo θ_j é a saída produzida pela rede para o padrão procurado na conexão j .

As funções mais utilizadas, que ajudam no controle da ativação do neurônio e que fazem uma simulação aproximada do modelo de um neurônio biológico, são as funções não lineares (ABELÉM, 1994 p.19; VIEIRA, 2000 p.260; BECERIKLI, KONAR E SAMAD, 2003 p.252) já dito anteriormente, e são muito utilizadas no modelo de retro-propagação. Será adotada, para este trabalho a função sigmóide, onde os pontos de saturação estão entre 0 e 1 (Equação 3.17).

$$\alpha_i = \frac{1}{1 + e^{-S_i}} \quad (3.17)$$

Já para se fazer o ajuste dos pesos para as camadas intermediárias, substitui-se o erro local apresentado na Equação 3.15, pelo somatório dos produtos entre os pesos que ligam o neurônio em questão com os neurônios da camada de saída (i) e os deltas (δ) dos respectivos neurônios da mesma. Este cálculo é realizado utilizando a Equação 3.18.

$$\delta_i = \sum (\delta_j * W_{ij}) * F'(S_i) \quad (3.18)$$

Portanto se consegue fazer o ajuste tanto para os pesos da camada de saída com as equações 3.14 e 3.15, como para as camadas intermediárias com as equações 3.14 e 3.18. Se tiver mais camadas escondidas calcula-se o delta (δ) tendo como

referência a camada escondida anterior no cálculo do delta.

Para iniciar o processo de treinamento desta rede, antes é necessário que todos os pesos sejam inicializados randomicamente, garantindo desta forma que a rede não fique saturada com grandes valores de pesos. Após são realizados os seguintes passos:

1. Selecionar um conjunto de treinamento e aplicar esta entrada à rede;
2. Calcular a saída da rede;
3. Calcular o erro entre a saída da rede e a saída desejada (alvo); e
4. Ajustar os pesos da rede de maneira a minimizar o erro.

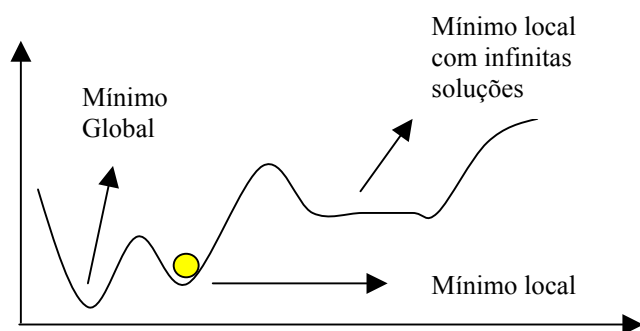
Repetir o passo 1 até o passo 4 para cada vetor do conjunto de treinamento, até o erro se tornar aceitável, baixo para o conjunto inteiro.

Observa-se que os passos 1 e 2 constituem a etapa de propagação, onde a entrada é propagada através da rede, da entrada até a saída. Os passos 3 e 4, constituem a etapa de retro-propagação, onde o erro calculado é propagado de volta através da rede para ajustar os pesos.

De acordo com VIEIRA (2000, p.264-265), apesar do grande sucesso do modelo de retro-propagação, existem alguns problemas que se deve citar, como o longo período de treinamento, principalmente para problemas complexos e mínimos locais.

Sabendo que a superfície do erro geralmente é cheia de vales e desníveis, e o algoritmo empregado é um tipo de gradiente descendente, existindo uma possibilidade de ficar preso em um mínimo local como mostrado na Figura 3.12.

FIGURA 3.12 - SUPERFÍCIE DE ERRO, CORTE TRANSVERSAL.



FONTE: Adaptado de BOTELHO (2002 p.12)

Durante o treinamento os pesos podem ser ajustados para valores muito grandes, levando a paralisar a rede.

Devido os problemas citados acima no modelo de retro-propagação, inúmeras propostas tem sido feitas nos últimos anos (VIEIRA, 2000; ABELÉM, 1994; HERTZ, 1991; WASSERMAN, 1989). Foi proposta por RUMELHART, HINTON E WILLIAMS em 1986 uma alternativa chamada de “momento”, para diminuir o tempo de treinamento, assegurando um maior grau de estabilidade para a rede, mostrando-se eficiente em alguns casos.

O termo momento especifica a quantidade que será acrescentada da variação do velho peso ($\Delta W_{ij}(t)$) para o novo peso ($\Delta W_{ij}(t+1)$). Esta quantidade é computada usando a equação 3.19.

$$\Delta W_{ij}(t+1) = \eta \delta_i \theta_j + \mu \Delta W_{ij}(t) \quad (3.19)$$

Onde μ é o momento.

Segundo VIEIRA (2000, p.264), o termo “bias” representa uma influência externa no neurônio e possui uma grande influência na convergência da rede. A “bias” determina onde a função de ativação tem sua ascensão mais íngreme (Equação 3.20).

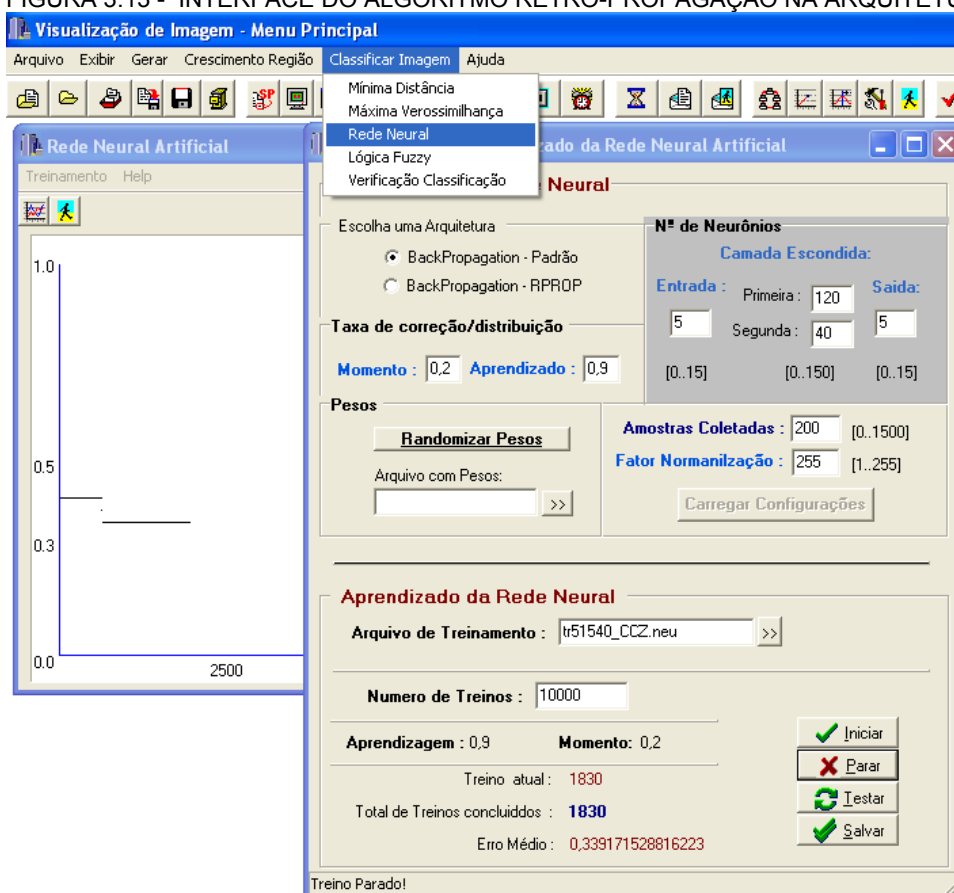
$$S_i = \sum(\theta_j * W_{ij}) + \theta \quad (3.20)$$

onde

θ é o termo “bias”.

Na Figura 3.13 observa-se a interface do algoritmo de retro-propagação na arquitetura multicamada implementado em linguagem C++. Também é exibido um gráfico onde no eixo das ordenadas está o erro médio do treinamento e no eixo das abscissas está o total de treinos a serem executados.

FIGURA 3.13 - INTERFACE DO ALGORITMO RETRO-PROPAGAÇÃO NA ARQUITETURA MULTICAMADA.



3.2.6.2 - Treinamento da rede neural artificial para a classificação.

Para que o treinamento da rede neural não se torne tendencioso é necessário que se intercale as amostras de treinamento. Como exemplo prático desta situação pode-se citar o aprendizado de línguas estrangeiras. Quando o cérebro humano é condicionado a apenas uma língua estrangeira, este aprende satisfatoriamente aquela língua. Porém, quando é condicionado a outro país onde é exposto a outra língua a tendência é esquecer a língua anterior e se adequar apenas a língua que foi exposta mais recentemente. Porém quando o cérebro é exposto a várias línguas de uma única vez, este se adapta e aprende todas de forma satisfatória.

Para verificação do treinamento foi adotado o coeficiente “Kappa” descrito na seção 3.2.10.

3.2.6.3 - Metodologia para a definição da arquitetura

Segundo GERMAN e GAHEGAN (1996, p.973), a escolha da arquitetura da rede é totalmente experimental. Porém para iniciar o processo, em um “perceptron” de multicamadas, pode-se adotar o critério da segunda camada intermediária ser 1/3 da primeira camada intermediária. Portanto foram executados testes que ajudaram na escolha da arquitetura para a melhor classificação desta imagem. Testes com 2 camadas escondidas foram executados.

No primeiro experimento adotou-se 120 e 40 neurônios na primeira e na segunda camada escondida, respectivamente. Nos experimentos seguintes o número de neurônios foi reduzido, com o intuito de encontrar a arquitetura da rede mais apropriada para este problema. O arquivo de entrada para treinamento possui 5 classes com 5 descritores (3 informações espectrais, 2 descritores de forma). Quando se introduziu apenas 1 descritor de forma, a arquitetura que melhor se enquadrou foi a que possuía 90 neurônios para a primeira camada e 30 para a segunda.

No segundo experimento, após várias tentativas, adotou-se 60 e 20 neurônios na primeira e na segunda camada escondida, respectivamente. O arquivo de entrada para treinamento possui 5 classes com 5 descritores (3 informações espectrais, 1 descritores de forma e a diferença de altitude).

O melhor resultado obtido foi o que adotou uma arquitetura de 60 neurônios para a primeira camada e 20 para a segunda, utilizando no arquivo de treinamento, a variável “diferença de altitude” (segundo experimento). Justifica-se esta escolha devido a separabilidade entre classes utilizando a variável “diferença de altitude” oriunda dos dados do laser scanner ser melhor para separar as classes que os descritores de forma. Esta abordagem, será melhor indagada na seção 4.2.5.

3.2.7. - Implementação do Método distância mínima

O método de classificação da distância mínima atribui cada elemento desconhecido à classe cuja média é mais próxima a ele, ou seja, a atribuição de uma classe ao elemento depende da distância em relação à classe a qual este foi atribuído. A equação 3.21 descreve o critério de decisão deste método.

$$d(x, w_j) = \sqrt{(X - M_j)^T (X - M_j)} \quad (3.21)$$

onde

X = elemento representado por um vetor contendo informação espectral, descritor de forma e altimetria;

M = vetor de média da classe, contendo informação espectral, descritor de forma e altimetria;

j = classe em análise.

Como neste estudo consideram-se regiões como elemento da imagem, as classes foram descritas a partir das médias das variáveis (área, forma, coeficiente de compactidade, razão entre eixos e altimetria), calculadas a partir das amostras coletadas para cada classe.

3.2.8. - Implementação do Método Máxima Verossimilhança

O problema da classificação de imagens de satélite também pode ser encarado do ponto de vista estatístico, atribuindo o pixel para a classe mais provável. Para isto são necessários conceitos básicos de probabilidade. A probabilidade condicional descrita pelo Teorema de Bayes (RICHARDS e JIA, 1999 p.396) serve para estimar a probabilidade de uma classe ocorrer. O elemento x pertence à classe w_i se

$$P(x \in w_i) > P(x \in w_j) \quad i \neq j$$

ou

$$P(w_i | x) > P(w_j | x) \quad \text{com} \quad P(w_i | x) = \frac{P(x | w_i)P(w_i)}{P(x)} \quad (3.22)$$

Assumindo uma distribuição gaussiana dos “pixels” dentro de cada classe tem-se para M-variáveis

$$P(x | w_i) = (2\pi)^{-m/2} |\Sigma_j|^{-1/2} \exp\left(-0.5 * (x - \mu_j)^t \Sigma^{-1} (x - \mu_j)\right) \quad (3.23)$$

Para fins de facilitar os cálculos, na equação 3.22, os termos $P(x)$ e $P(w_j)$ podem ser considerados constantes, do que resulta:

$$x \in w_i \text{ se } P(x | w_i) > P(x | w_j)$$

Substituindo (3.23) e fazendo os devidos cálculos e simplificações obtêm-se a equação 3.24 que expressa o critério de decisão do método

$$G(x) = \log\left(|\Sigma_j|\right) + (x - \mu_j)^t \Sigma^{-1} (x - \mu_j) \quad (3.24)$$

$$x \in w_j \text{ se } G_j(x) < G_k(x) \quad \forall j \neq k$$

onde

$|\Sigma_j|$ = Determinante da matriz de variância-covariância;

Σ^{-1} = Inversa da matriz de variância-covariância;

μ_j = Vetor médio da classe j.

3.2.9. - Área de verificação

Para avaliar se as informações que compõem o mapa resultante de uma classificação possuem confiabilidade temática, faz-se necessário avaliar a precisão do treinamento da classificação. O método tradicional de verificação da precisão de uma classificação utiliza uma matriz de erros ou matriz de confusão. Para analisar o desempenho de métodos de classificação utiliza-se o coeficiente “Kappa”, sendo este o mais usual na determinação da acurácia da classificação.

Porém para avaliar o mapa temático resultante é necessária uma área de verificação (verdade de campo), onde todos os elementos da cena serão identificados. Nesta pesquisa, como áreas de verificação foram vetorizadas, os objetos da cena foram identificados através de uma análise visual para uma posterior avaliação. A

vetorização pode ser observada na Figura 3.14. A partir desta imagem vetorizada, utiliza-se o ENVI para transformar os dados vetoriais ao formato raster, e assim produzir uma imagem temática com a verdade do campo, com a mesma resolução que a imagem original (Figura 3.15) para uma posterior comparação e análise da acurácia.

FIGURA 3.14 - VERDADE DE CAMPO VETORIZADA.

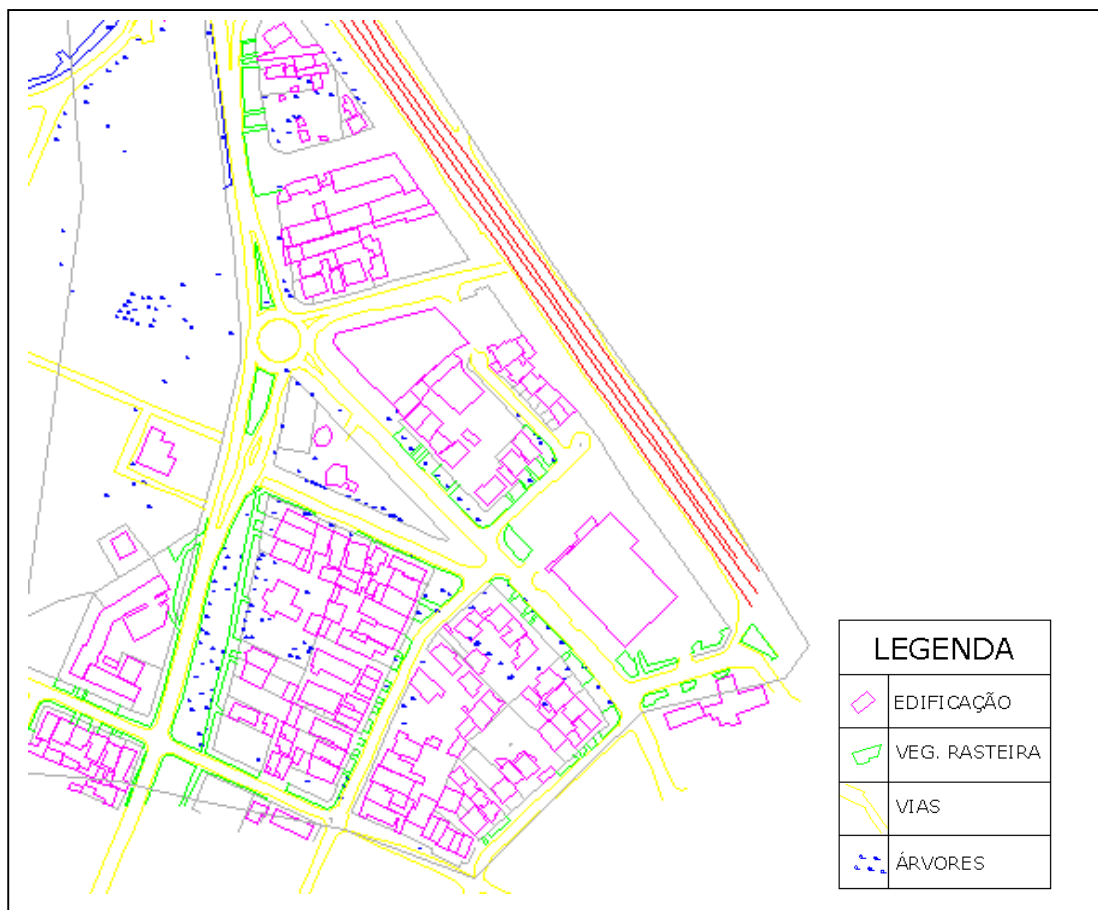
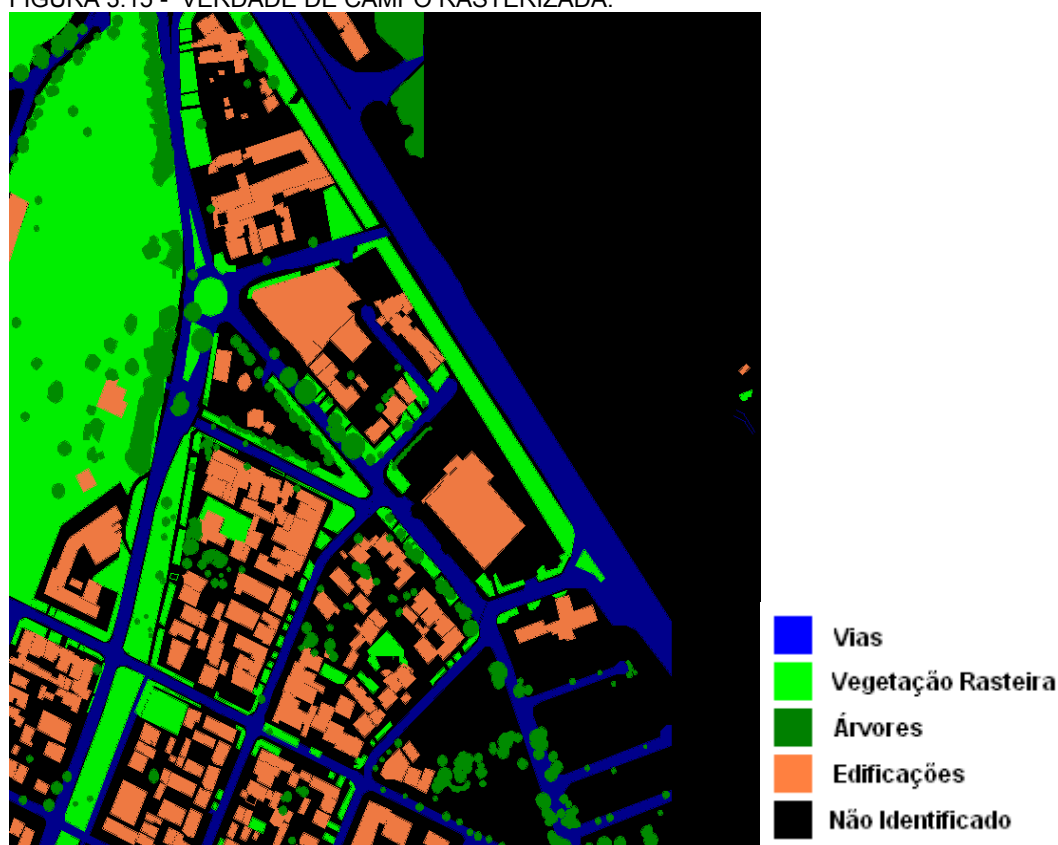


FIGURA 3.15 - VERDADE DE CAMPO RASTERIZADA.



3.2.10. - Avaliação da imagem classificada

Para a comparação dos resultados são utilizadas tabelas que possuem índices de acurácia global, adotado para este estudo o índice “kappa”. Nas mesmas, estão exibidos os coeficientes que expressam a exatidão das classes, tanto do ponto de vista do usuário como a exatidão do ponto de vista do produto. Todos estes parâmetros são derivados da matriz de confusão.

O coeficiente “kappa” tem sido utilizado tanto para avaliar a eficiência da classificação em métodos convencionais “pixel” a “pixel”, como em métodos de classificação por regiões dando relevância aos objetos da cena, como exemplo cita-se os autores LIRA e MALETTI (2002 p.958) e SOHN e REBELLO (2002 p.1276).

A matriz de confusão é um resumo do sucesso da classificação pois nela são representados o número de “pixels” corretamente classificados e o número de erros por classe. Esta possui o tamanho $m \times m$, sendo m o número de classes já definidas. Sua

diagonal compreende o número de elementos corretamente classificados, suas linhas representam as classes da verdade de campo e suas colunas as classes do mapa obtido através da classificação (JENSEN, 1996 p.249). A partir desta matriz é possível calcular um índice de precisão da classificação total (índice “Kappa” - k) dado pela equação 3.25:

$$k = \frac{Po - Pc}{1 - Pc} \quad (3.25)$$

onde:

$$Po = \sum_{i=1}^M X_{ii} \quad (3.26)$$

$$Pc = \sum_{i=1}^M X_{i+} X_{+i} \quad (3.27)$$

sendo:

P_0 = concordância observada.

P_C = concordância esperada, sendo X_{i+} os elementos pertencentes a uma determinada classe e X_{+i} os elementos classificados em uma determinada classe de objeto.

M = total de classes.

O coeficiente “Kappa” pode apresentar valores entre -1 e +1, e o mesmo pode mostrar o desempenho da classificação de acordo com a seguinte Tabela3.2:

TABELA 3.2 - COEFICIENTE “KAPPA” E O DESEMPENHO DA CLASSIFICAÇÃO

“Kappa”	Desempenho da Classificação
< 0	Péssimo
$0 < \kappa < 0,2$	Mau
$0,2 < \kappa < 0,4$	Razoável
$0,4 < \kappa < 0,6$	Bom
$0,6 < \kappa < 0,8$	Muito Bom
$0,8 < \kappa < 1,0$	Excelente

FONTE: Moreira, 2001 p.196.

No resultado da classificação, isto é, na avaliação da acurácia por meio da matriz de confusão se distinguem dois tipos principais de erros: erros de omissão e

erros de inclusão.

Os erros de omissão são “pixels” pertencentes a uma determinada classe que foram erroneamente classificados como sendo de outra. Estes erros servem para descrever com que precisão uma determinada classe foi definida. Isto leva ao conceito de exatidão do ponto de vista do produtor (criador) do mapa temático. Quanto maior o erro de omissão maior será a exatidão do ponto de vista do produtor. A exatidão do ponto de vista do produtor pode ser calculada pela equação 3.28.

$$EP = \frac{M(i,i)}{SL(i)} * 100 \quad (3.28)$$

onde EP : exatidão do ponto de vista do produtor;
M(i,i) : elemento da diagonal;
SL(i) : soma parcial da linha i da matriz de confusão.

Os erros de inclusão correspondem aos “pixels” que, sendo na realidade de outra classe, foram incluídos na classe considerada. O usuário estaria interessado em saber se pode confiar no produto. Estes erros evidenciam se o total de “pixels” representados como uma determinada classe na imagem temática, realmente corresponde a esta classe. Quanto menor o erro de inclusão maior será a exatidão do ponto de vista do usuário. A exatidão do ponto de vista do usuário pode ser calculada pela equação 3.29.

$$EU = \frac{M(i,i)}{SC(i)} * 100 \quad (3.29)$$

onde EU : exatidão do ponto de vista do usuário;
M(i,i) : elemento da diagonal;
SC(i) : soma parcial da coluna i da matriz de confusão.

3.2.10.1 - Algoritmo de avaliação implementado

Pela Figura 3.16, observa-se a interface do algoritmo implementado em linguagem C++ que exerce a função de avaliação da classificação. O algoritmo exige uma imagem temática com a verdade de campo onde executa a interseção entre esta e

a imagem classificada, resultando o percentual corretamente classificado, bem como os erros de omissão e inclusão associados a esta classe (Figura 3.17).

FIGURA 3.16 - INTERFACE DO MÓDULO DE AVALIAÇÃO DA CLASSIFICAÇÃO.

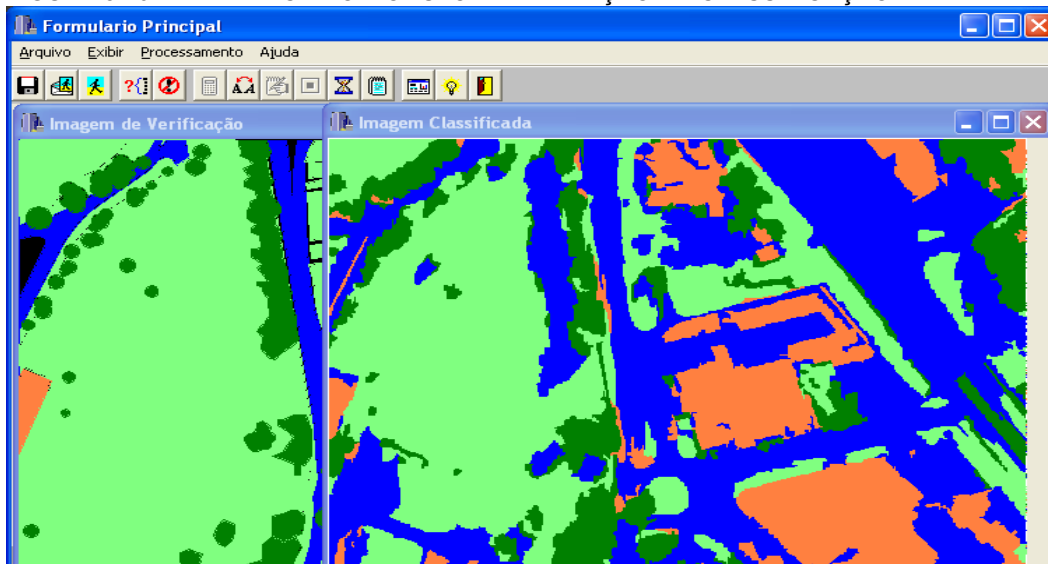


FIGURA 3.17 - INTERFACE COM A EXATIDÃO DO PONTO DE VISTA DO USUÁRIO E DO PRODUTO NA CLASSIFICAÇÃO DO OBJETO.

Faltou	1	2	3	4	5	Total	ErroOmissão	Exat.Produto
1	45378	2180	4110	0	2554	54222	0,163	0,837
2	28741	60864	2125	0	11473	103203	0,41	0,59
3	3595	360	25389	0	528	29872	0,15	0,85
4	18010	5808	4960	0	4200	32978	1	0
5	4926	2525	2430	0	19727	29608	0,334	0,666
Total	82640	65929	34054	0	34282			
ErroInclusão	0,451	0,077	0,254	0	0,425			
Exat.Usuário	0,549	0,923	0,746	0	0,575			
Kappa	0,578							
Extra	1	2	3	4	5			
1	55125	0	0	0	0			
2	0	10615	0	0	0			
3	0	0	13470	0	0			
4	0	0	0	0	0			
5	0	0	0	0	18734			

CAPÍTULO 4 - RESULTADOS E DISCUSSÕES

4.1 - SEGMENTAÇÃO DA IMAGEM

Os três recortes, mencionados na seção 3.1.1, foram segmentados utilizando os mesmos parâmetros apresentados na seção 3.2.1.2. Eles foram executados desta forma para possibilitar a constituição de um mosaico ao final da classificação. Na Figura 4.1 apresenta-se os polígonos dos segmentos gerados a partir do algoritmo de segmentação do “eCognition”. Na Figura 4.3(a) e 4.3(b) são apresentadas as imagens segmentadas a partir do “eCognition”.

Para a realizar a segmentação no “eCognition”, foram introduzidas as 4 bandas espectrais disponíveis e a imagem altimétrica, todas com o mesmo peso. A imagem da Figura 4.1(b) é o resultado da segmentação com os parâmetros de forma e escala conforme Tabela 4.1. Observando a Figura 4.1 nota-se que os objetos da cena estão bem caracterizados pela sua forma, portanto é possível usar parâmetros de forma para separar os objetos. O “eCognition” leva em consideração a forma e a topologia dos objetos na cena (descrito na seção 3.2.1.2.1).

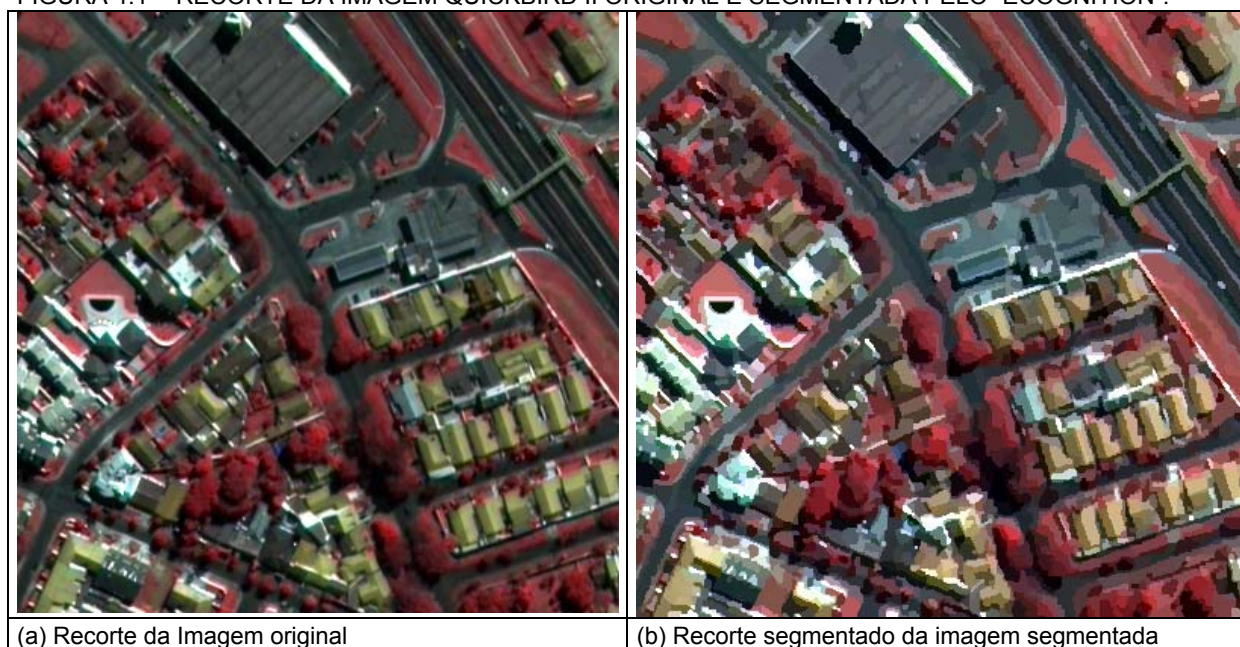
TABELA 4.1 - TABELA QUE FORNECE OS PARÂMETROS UTILIZADOS EM CADA NÍVEL DE SEGMENTAÇÃO NO ECOGNITION.

<i>Nivel</i>	<i>Escala</i>	<i>Escala</i>		<i>Forma</i>	
		Cor	Forma	Suavidade	Compacidade
1	7	0.7	0.3	0.7	0.3
2	5	0.3	0.7	0.7	0.3

Como pode ser observado na Tabela 4.1, foram utilizados 2 níveis de segmentação, caracterizando assim uma segmentação espacial-espectral (descrito na seção 2.4.3). No primeiro nível foi utilizada uma escala de valor igual a 7, parâmetro de cor igual a 0.7 e forma igual a 0.3 isto devido ao fato dos objetos da cena estarem caracterizados mais pela cor (informação espectral). À suavidade e compacidade foram dados os mesmos pesos de 0.7 e 0.3 valores “default” do “eCognition”. No

segundo nível foi acrescida a imagem altimétrica gerada a partir da grade regular (Figura 4.1). O “eCognition” gera os segmentos do segundo nível a partir dos segmentos criados pelo primeiro nível. Portanto, com base no primeiro nível, a configuração do segundo nível utilizou uma escala com valor igual a 5, na tentativa de aglomerar mais os segmentos de um objeto, para representá-lo melhor, parâmetro de cor igual a 0.3 e forma 0.7 isto devido, neste nível, ter sido incluído a imagem altimétrica produzida a partir dos dados laser scanner. Foi mantida a mesma configuração do nível 1 para os parâmetros de suavidade e compacidade.

FIGURA 4.1 - RECORTE DA IMAGEM QUICKBIRD II ORIGINAL E SEGMENTADA PELO “ECOGNITION”.



Já para a escolha das três bandas espectrais, foram realizados testes estatísticos, verificando a relevância de cada banda na separabilidade entre os objetos na cena. Estes testes podem ser observados nas próximas seções.

FIGURA 4.2 - IMAGEM DO PRIMEIRO RECORTE SEGMENTADO NO ECOGNITION E CODIFICADO PELO PROGRAMA

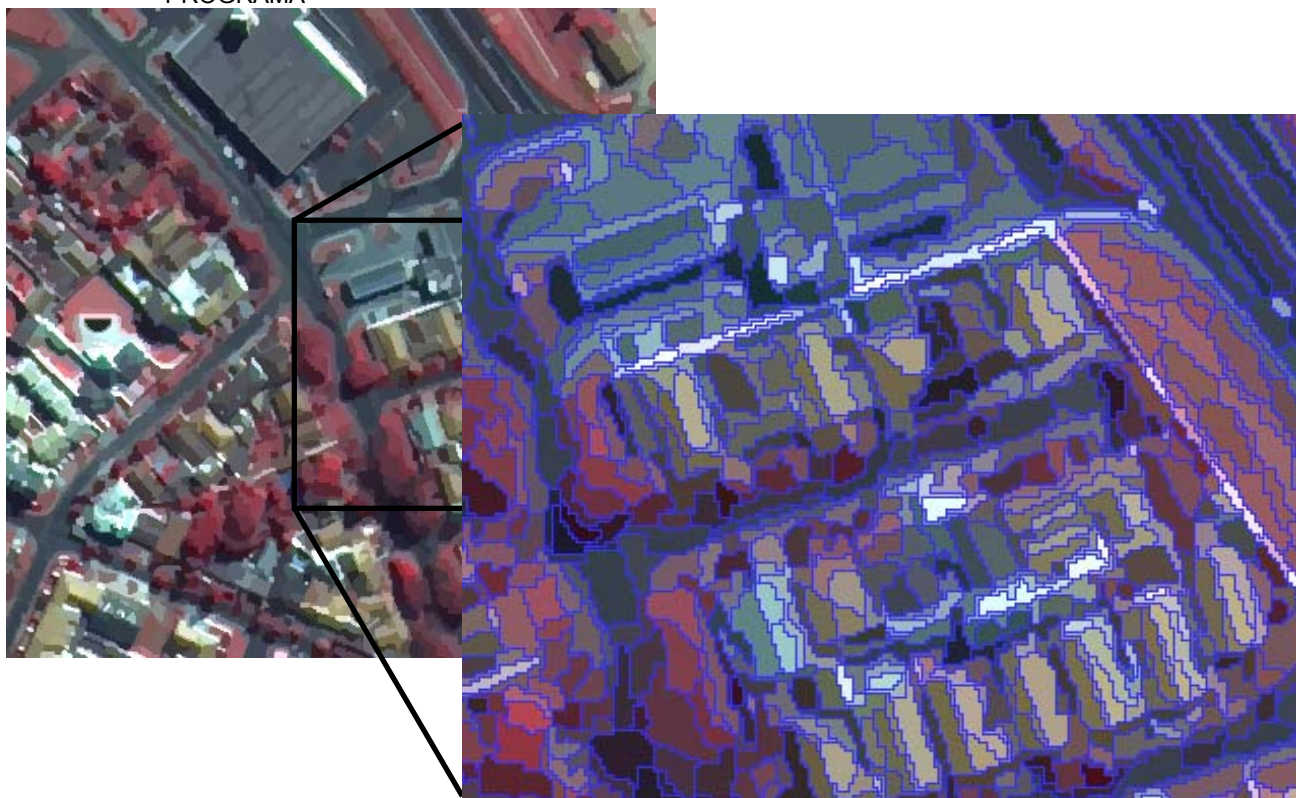
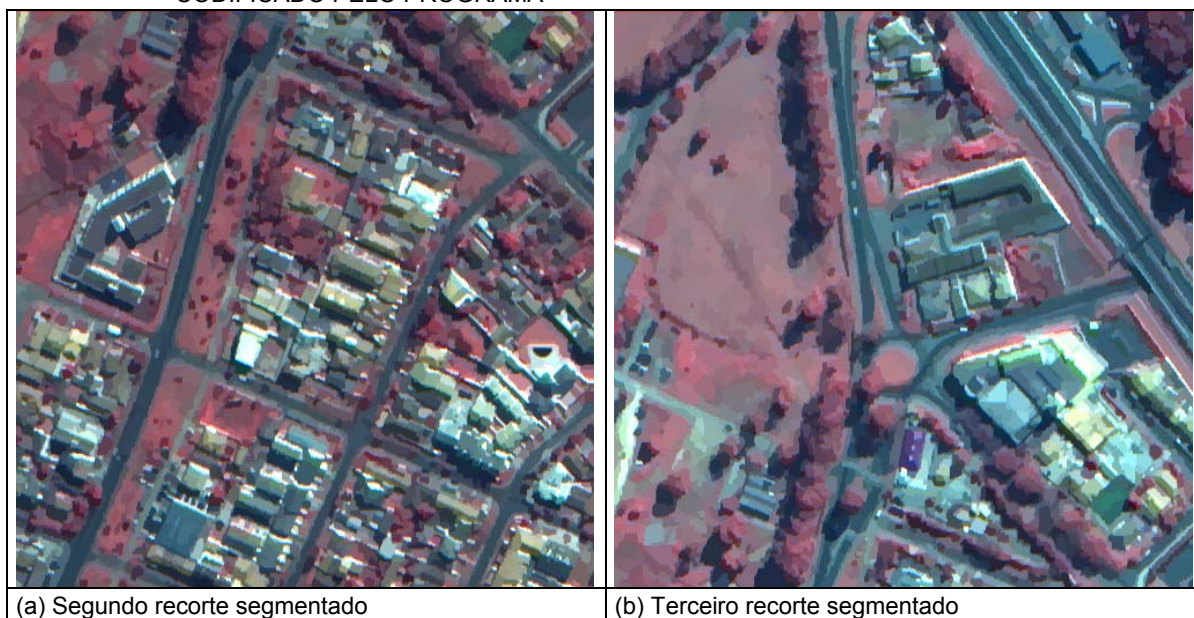


FIGURA 4.3 - IMAGENS DO SEGUNDO E TERCEIRO RECORTE SEGMENTADO NO ECOGNITION E CODIFICADO PELO PROGRAMA



(a) Segundo recorte segmentado

(b) Terceiro recorte segmentado

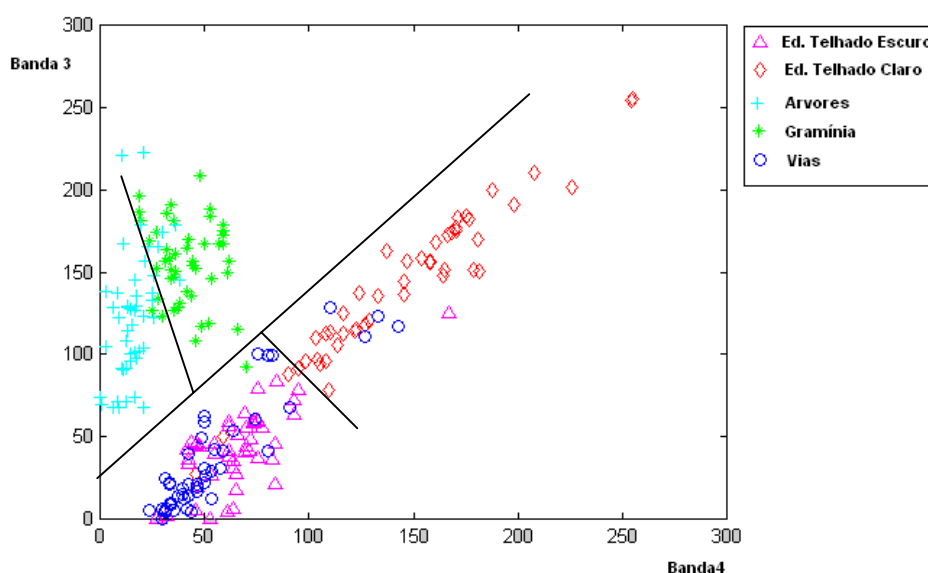
4.2 - VERIFICAÇÃO DA RELAÇÃO ENTRE CLASSES SEGUNDO OS PARÂMETROS ESPECTRAIS E ESPACIAIS.

Com o intuito de encontrar uma melhor separabilidade entre classes, foram realizados estudos comparativos entre as bandas espectrais e sua relação com as informações espaciais. Para tal, foram utilizados gráficos comparando estas informações, apresentados nas Figuras 4.4 a 4.16. Todas as variáveis foram normalizadas entre 0 e 255.

4.2.1. - Separabilidade entre informações espectrais.

A Figura 4.4 evidencia a relação espectral entre classes nas bandas 4 e 3. Nota-se a presença de três grupos que podem ser separados com retas, este são representados pelas classes: “edificações de telhado claro”, “gramínea” e “árvores”. Um quarto grupo é evidenciado, sendo este composto por duas classes confusas: “Edificação de telhado escuro” e “vias”.

FIGURA 4.4 - RELAÇÃO ENTRE BANDA 4 (INFRA-VERMELHO) E BANDA 3 (VISÍVEL).



A Figura 4.5 evidencia a relação espectral entre classes nas bandas 1 e 2. Nota-se que com o traçar de uma reta é possível separar as edificações de telhado claro das demais. Analisando a Figura 4.6, percebe-se que a mesma situação da Figura

4.5 ocorre entre as bandas 2 e 3, porém a classe “edificação de telhado claro” apresenta-se mais separada das demais classes.

FIGURA 4.5 - RELAÇÃO ENTRE BANDA 1 (VISÍVEL) E BANDA 2 (VISÍVEL).

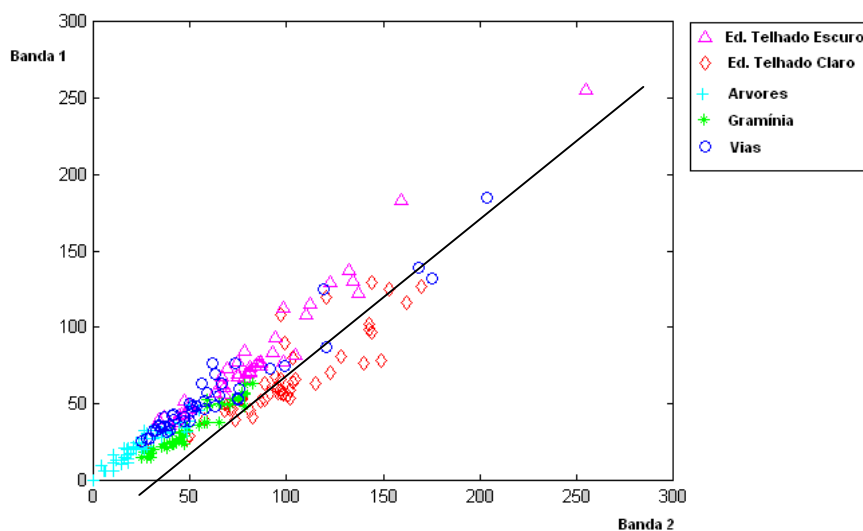
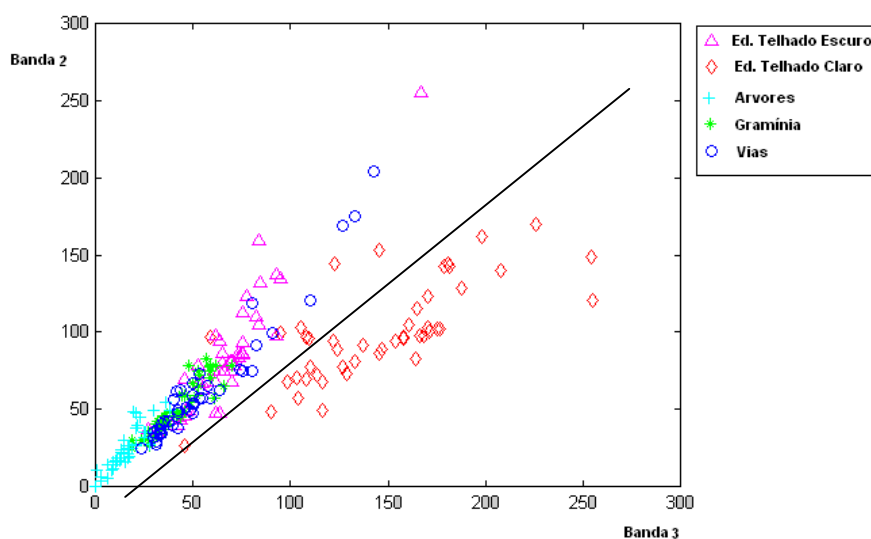


FIGURA 4.6 - RELAÇÃO ENTRE BANDA 3 (VISÍVEL) E BANDA 2 (VISÍVEL).



A Figura 4.7 mostra a relação espectral entre classes nas bandas 2 e 4. Nota-se que é possível separar 4 grupos distintos com 3 retas. Um grupo com “gramíneas”, outro com “árvores”, e outro com “edificação de telhado claro” e um quarto grupo com “vias” e “edificação de telhado escuro”. Neste último grupo, as classes apresentam-se misturadas, dificultando a separabilidade. Observando a Figura 4.8, nota-se a mesma situação, entre as bandas 1 e 4, dos mesmos grupos da Figura 4.7,

porém mais misturados.

FIGURA 4.7 - RELAÇÃO ENTRE BANDA 4 E BANDA 2.

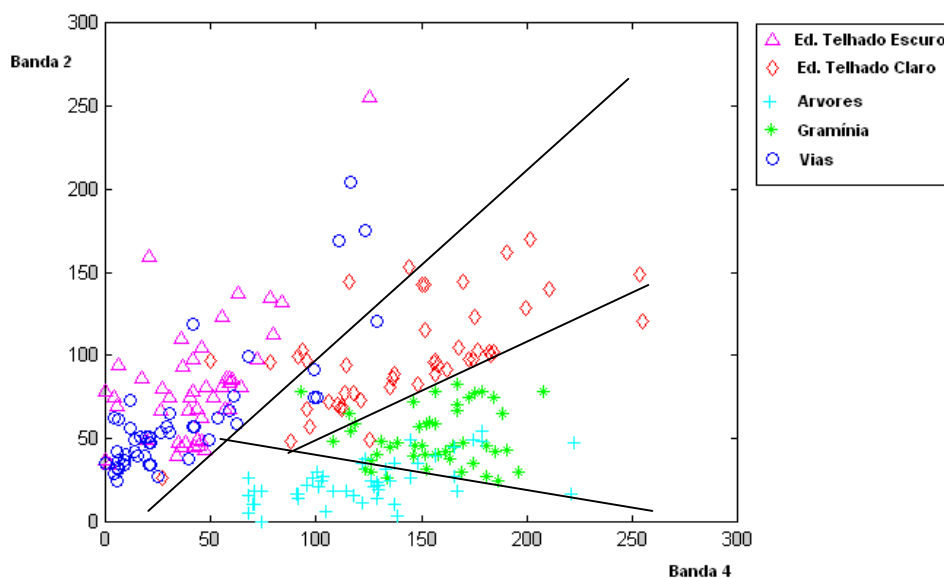
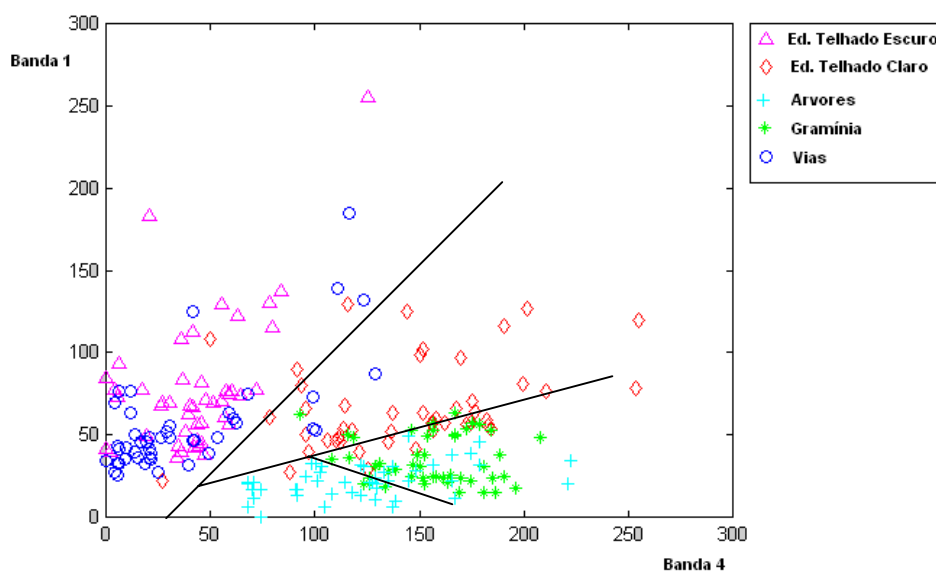


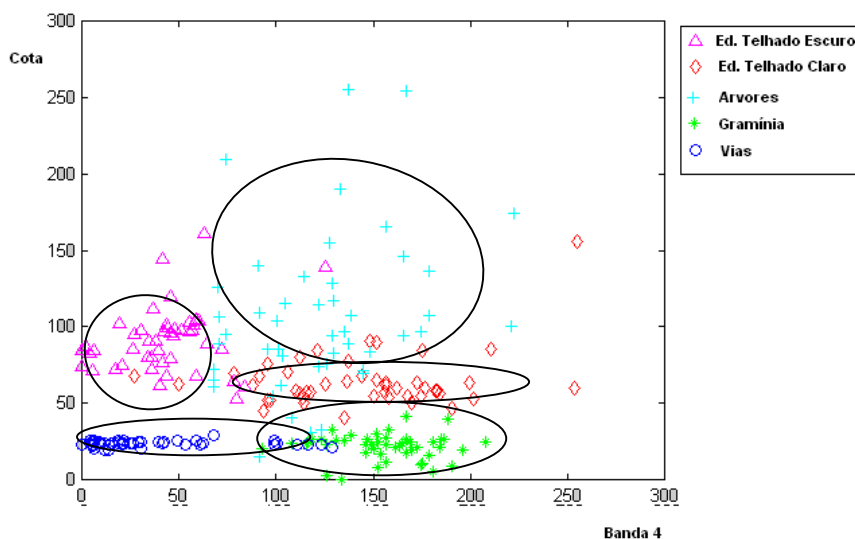
FIGURA 4.8 - RELAÇÃO ENTRE BANDA 4 (INFRA-VERMELHO) E BANDA 1 (VISÍVEL).



4.2.2. - Separabilidade entre informações espectrais e dados do laser scanner.

A Figura 4.9 exibe a relação espectral e espacial entre classes representadas pela banda 4 e dados do laser scanner (altura) respectivamente. Nota-se que todas as classes ficaram bem caracterizadas, com exceção da classe “árvore” que possui um espalhamento (desvio padrão) maior que as demais.

FIGURA 4.9 - RELAÇÃO ENTRE ALTITUDE E BANDA 4.



A Figura 4.10 evidencia a variação das classes representadas pela banda 2 e dados do laser scanner (altitude) respectivamente. Nota-se quatro grupos distintos, “árvore”, “edificação de telhado escuro”, “edificação de telhado claro” e um quarto grupo com mistura entre “vias” e “gramíneas”. Comparando a Figura 4.10 com a Figura 4.9, nota-se a classe “árvore” melhor caracterizada. O mesmo ocorre para a Figura 4.11 e Figura 4.12 onde é representada a relação entre classes na banda 2 e altitude, e banda 1 e altitude, respectivamente. Porém, comparando estas duas últimas figuras com a Figura 4.10, percebe-se que as classes da Figura 4.11 e 4.12 estão mais misturadas.

FIGURA 4.10 - RELAÇÃO ENTRE ALTITUDE E BANDA 3.

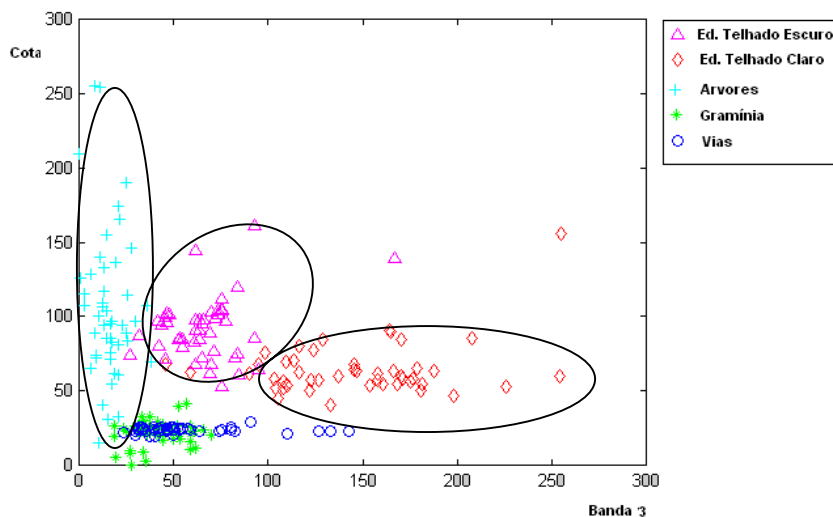


FIGURA 4.11 - RELAÇÃO ENTRE ALTITUDE E BANDA 2.

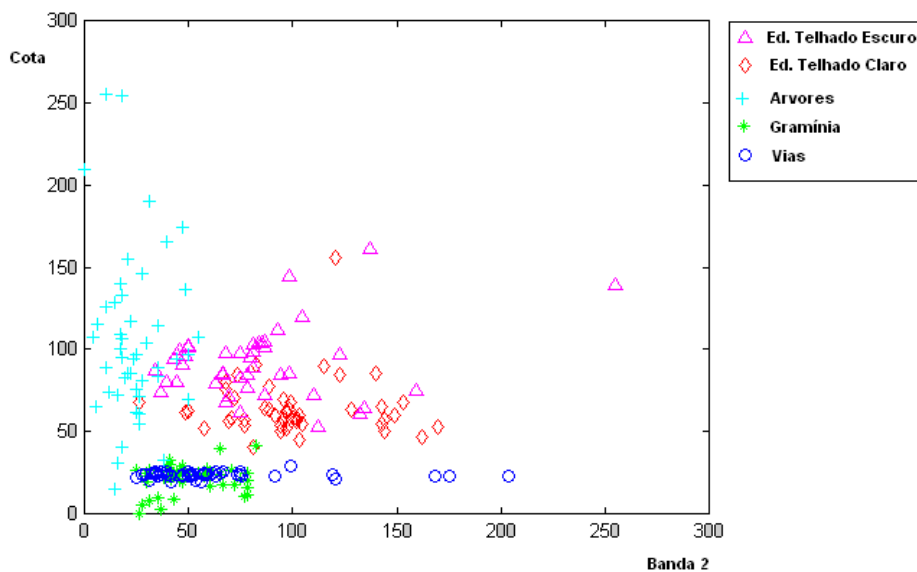
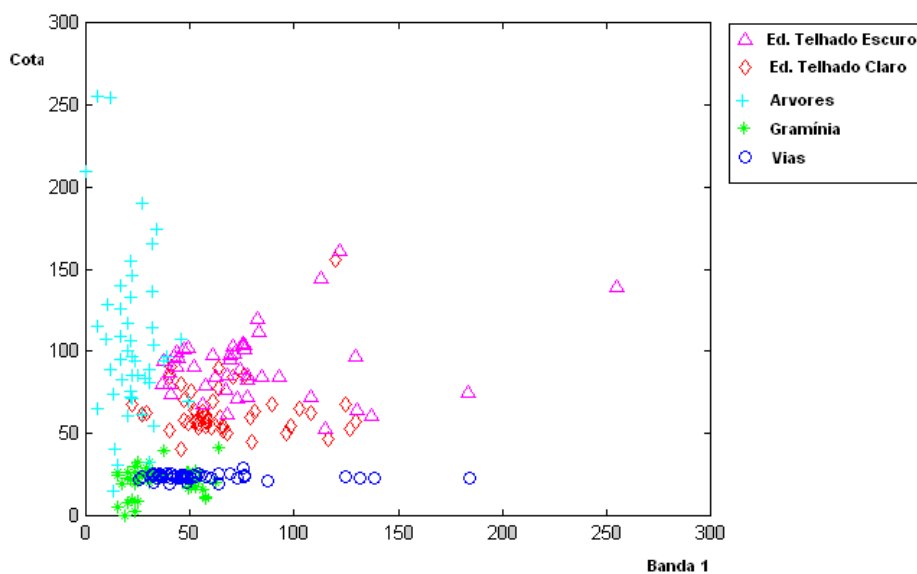


FIGURA 4.12 - RELAÇÃO ENTRE ALTITUDE E BANDA 1.



4.2.3. - Separabilidade entre dados do laser scanner e descritores de forma.

A Figura 4.13 evidencia a relação espacial entre classes representadas pela área do segmento e sua altitude média respectivamente. Nota-se que todas as classes não foram bem caracterizadas existindo uma confusão, principalmente, entre “árvores” e “edificação de telhado escuro”, “vias” e “gramínea”. O mesmo ocorre para as combinações da diferença de altitude com o perímetro, diferença de altitude com razão

entre eixos da elipse e diferença de altitude com coeficiente de compacidade apresentadas nas Figuras 4.14, 4.15, 4.16, respectivamente.

FIGURA 4.13 - RELAÇÃO ENTRE ALTITUDE E ÁREA

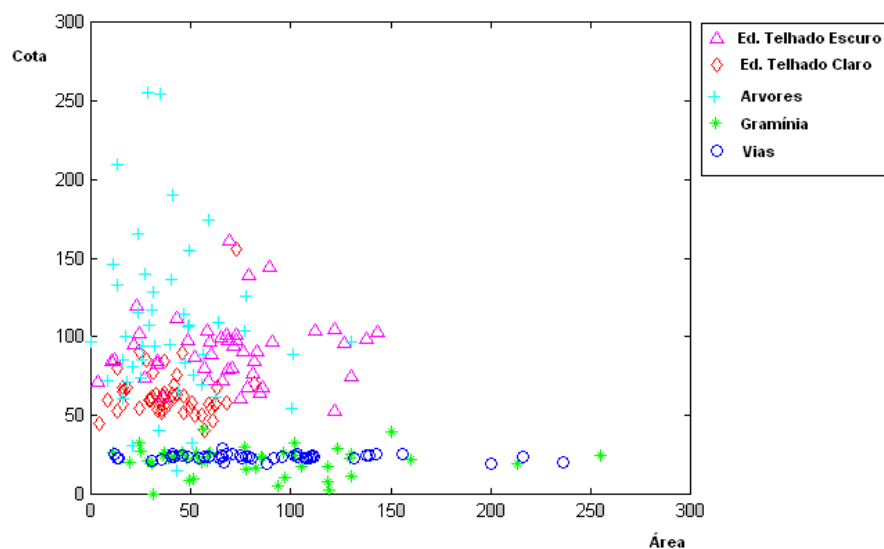


FIGURA 4.14 - RELAÇÃO ENTRE ALTITUDE E PERÍMETRO.

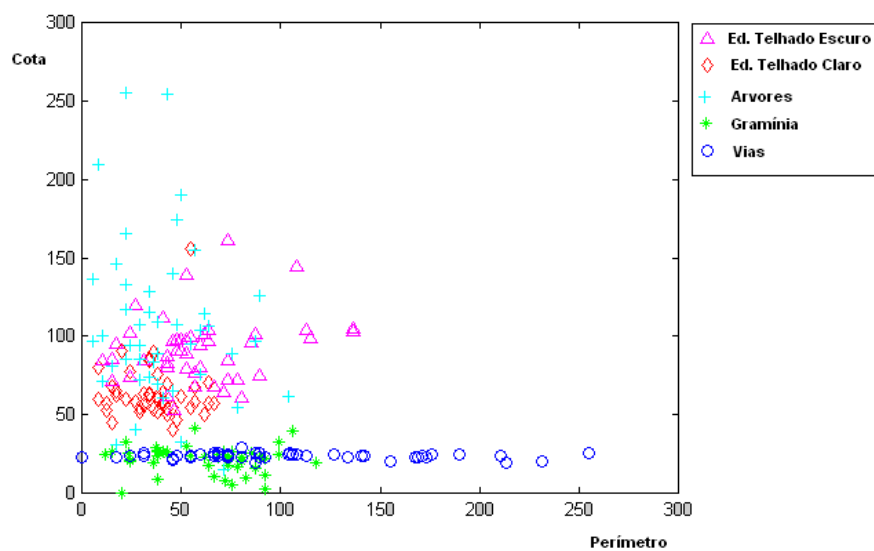


FIGURA 4.15 - RELAÇÃO ENTRE ALTITUDE E RAZÃO ENTRE EIXOS DA ELIPSE.

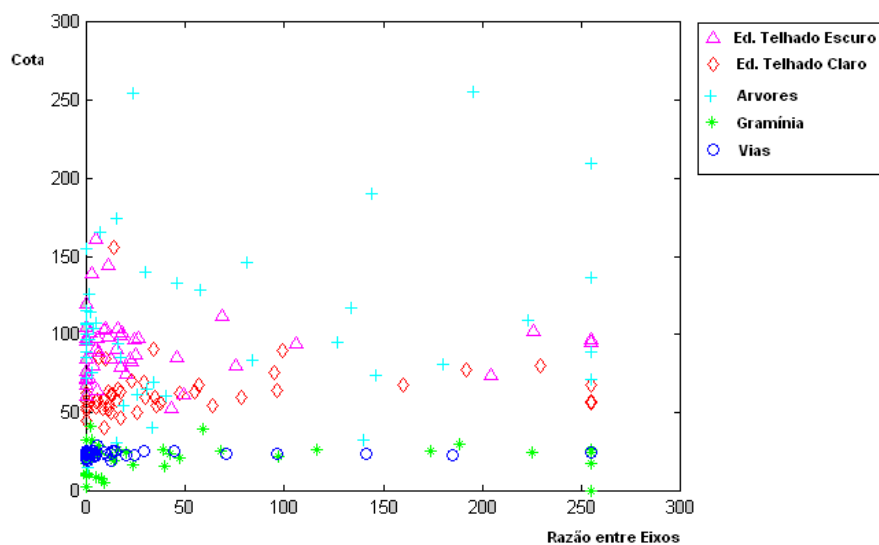
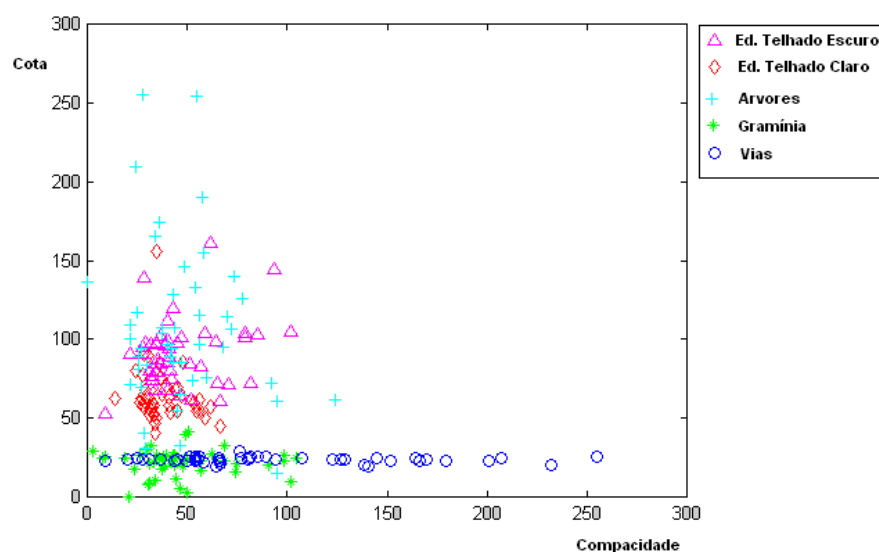


FIGURA 4.16 - RELAÇÃO ENTRE ALTITUDE E COEFICIENTE DE COMPACIDADE.



4.2.4. - Conclusão da análise dos gráficos.

Conclui-se com a análise dos gráficos apresentados, que a maior separabilidade encontra-se na relação entre as bandas espectrais, e na relação entre as bandas espectrais e a altitude fornecida pelo laser scanner.

Na relação entre bandas espectrais, as que se mostraram mais eficientes na separabilidade entre as classes apresentadas, foram as bandas 2, 3 e 4. Já na relação

entre bandas espectrais e altitude do laser scanner a combinação que mostrou maior eficiência na separabilidade foi a composição entre a banda 4 e a altitude do laser scanner.

No intuito de verificar os resultados obtidos por meio de análise dos gráficos foram realizados cálculos estatísticos. Estes cálculos serão apresentados nas próximas seções.

4.2.5. - Separabilidade estatística entre informações espectrais e espaciais.

Para caracterização entre classes de uma forma estatística foi realizado o cálculo da distância entre as médias das classes de acordo com suas informações espectrais e espaciais. Os valores da média das classes podem ser observados nas Tabelas 4.2.

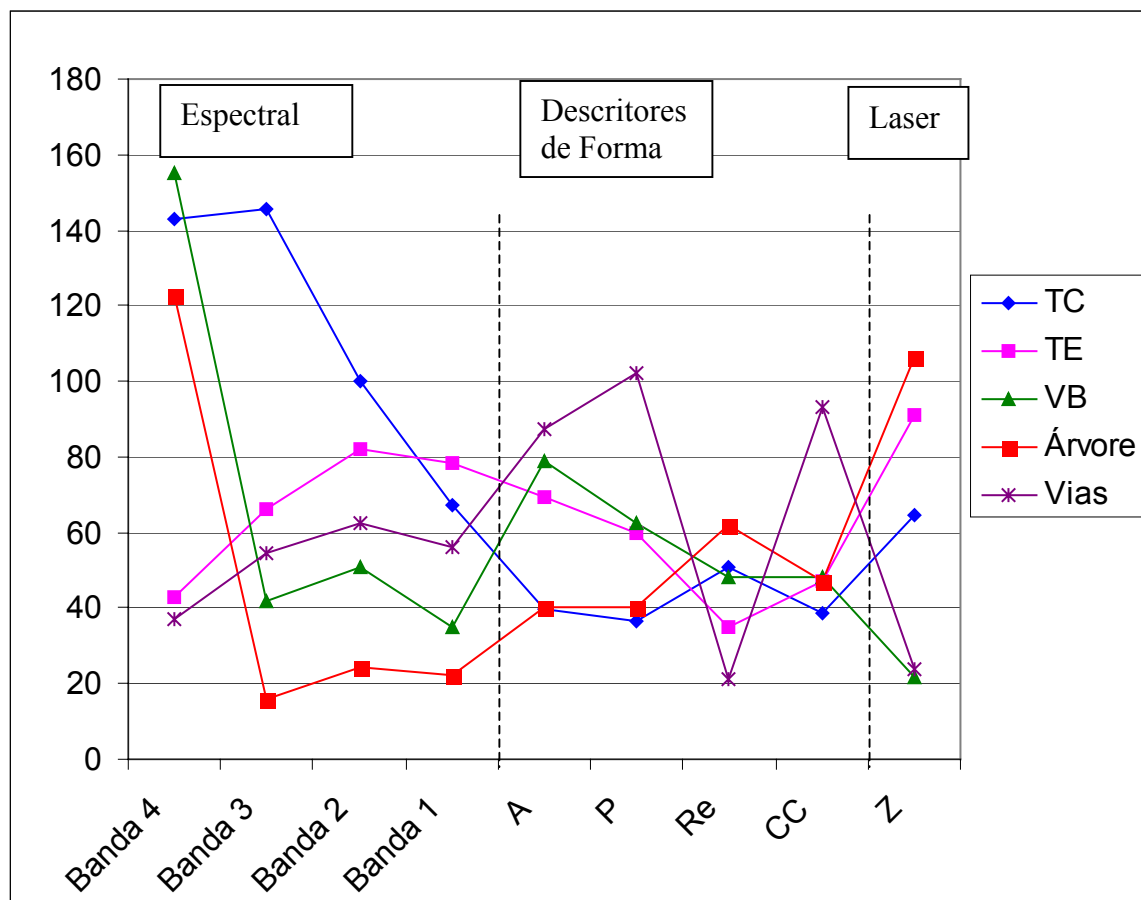
TABELA 4.2 - MÉDIA NORMALIZADA DAS CLASSES

	Telhado Claro	Telhado Escuro	Vegetação Branda	Árvore	Vias
Banda 4	142.69	43.08	155.37	122.75	36.935
Banda 3	145.63	66.007	41.603	15.991	54.365
Banda 2	99.802	82.051	51.056	24.206	62.408
Banda 1	67.282	78.383	34.703	22.202	56.005
Área	39.879	69.188	78.627	40.452	87.176
Perímetro	36.344	60.052	62.555	40.307	102.14
Re	50.951	34.936	48.086	61.738	20.988
CC	38.634	47.199	47.97	47.166	93.321
Altitude	64.47	91.125	21.599	106.15	23.573

* Vegetação Branda (Gramínea, pastagem, pequenos arbustos e etc).

No intuito de analisar a separabilidade das classes utilizando as melhores informações espectrais e espaciais, foi gerado um gráfico (Figura 4.17) partindo da Tabela 4.2 em que os dados estão normalizados.

FIGURA 4.17 - MÉDIAS DAS CLASSES (TC – TELHADO CLARO; TE – TELHADO ESCURO; VB – VEGETAÇÃO BAIXA; ARVORE, VIAS) DE ACORDO COM AS INFORMAÇÕES ESPACIAIS E ESPECTRAIS.



Na Figura 4.17, observa-se que as informações espectrais (banda 1, banda 2, banda 3 e banda 4) estão melhor espaçadas entre si que as informações espaciais, existindo uma confusão entre as classes telhado escuro e vias na banda 4. Para os dados do laser scanner (Z) nota-se uma melhor separabilidade entre classes que os parâmetros de forma (Área, Perímetro, Razão entre eixos e coeficiente de compacidade). Percebe-se, portanto, a baixa separabilidade das classes utilizando estes descritores de forma.

De acordo com CHUVIECO (1990, p.346), existem alguns critérios quantitativos que permitem evidenciar numericamente categorias distintas. Este procedimento é denominado separabilidade estatística. Um dos métodos pode ser descrito pela equação 4.1.

$$D_i(A, B) = \frac{|\overline{ND_A} - \overline{ND_B}|}{dp_A + dp_B} \quad (4.1)$$

onde:

ND_A = Média do nível digital da classe A;

ND_B = Média do nível digital da classe B;

dp_A = Desvio Padrão classe A;

dp_B = Desvio Padrão da classe B;

i = variável.

Portanto, os desvios padrão em relação às informações espaciais e espectrais, podem ser observados pela Tabela 4.3, respectivamente.

TABELA 4.3 - DESVIO PADRÃO DAS CLASSES NORMALIZADAS DE ACORDO COM AS INFORMAÇÕES ESPECTRAIS E ESPACIAIS

	Telhado Claro	Telhado Escuro	Vegetação Branda	Árvore	Vias
Banda 4	46.401	24.207	25.327	38.124	35.992
Banda 3	44.049	21.904	13.165	8.2896	27.857
Banda 2	31.354	38.549	16.919	13	38.453
Banda 1	26.819	40.253	14.533	10.445	31.671
Área	17.689	33.968	50.799	26.052	49.267
Perímetro	15.597	29.418	26.731	23.327	60.418
Re	73.414	66.291	82.401	84.423	51.106
CC	11.492	19.951	24.117	23.703	59.695
Altitude	17.836	21.173	8.520	51.383	1.757

* Vegetação Branda (Gramínea, pastagem, pequenos arbustos e etc).

Assim foi calculada a distância entre cada classe utilizando a equação 4.1, para cada informação espacial e espectral. Logo, calculou-se as distâncias geométricas entre informações espaciais e espectrais, utilizando a equação 4.2.

$$D(A, B) = \sqrt{D_1(A, B) + D_2(A, B) + \dots + D_n(A, B)} \quad (4.2)$$

onde

$D_n(A, B)$ = Distância entre as classes A e B considerando a variável “n”;

$D(A, B)$ = Distância total normalizada entre classes A e B.

Na Tabela 4.4 está sendo apresentado o valor do cálculo da distância entre classes, considerando todas as variáveis disponíveis, utilizando a equação 4.2. Na

última linha da tabela, observa-se o total da distância de cada classe. Para obter-se a estimativa da distância global que descreve a separabilidade entre as classes considerando a variável em questão, soma-se o total do valor de cada coluna, obtendo-se 108.31.

TABELA 4.4 - MATRIZ DAS DISTÂNCIAS ENTRE CLASSES CONSIDERENDO TODAS AS VARIÁVEIS

	Telhado Claro	Telhado Escuro	Vegetação Branda	Árvore	Vias	Distância Global
Telhado Clara	0	2.1587	2.7873	3.0883	3.1257	
Telhado Escuro	2.1587	0	3.3993	2.4671	3.0995	
Vegetação Branda	2.7873	3.3993	0	2.232	2.1093	
Árvore	3.0883	2.4671	2.232	0	2.6107	
Vias	3.1257	3.0995	2.1093	2.6107	0	
SOMA	11.16	11.125	10.528	10.398	10.945	

* Vegetação Branda (Gramínea, pastagem, pequenos arbustos e etc).

Na Tabela 4.5 é mostrado o resultado da distância global considerando apenas uma variável. Para cada variável espacial (A – área, P – perímetro, Re – razão entre eixos da elipse, CC – coeficiente de compacidade e Z – altura dos objetos da cena) e espectral (B1 – Banda1, B2 – Banda2, B3 – Banda3 e B4 – Banda4) foi realizado o cálculo da distância global considerando as seguintes classes: TC – telhado claro, TE – telhado escuro, VB = Vegetação Branda (Gramínea, pastagem, pequenos arbustos e etc), Árvore, Vias. Na mesma tabela, é apresentada a combinação entre as bandas espectrais (B4, B3, B2) em estudo como modelo de referência para uso nas análises puramente espectrais.

TABELA 4.5 - DISTÂNCIA GLOBAL PARA CADA VARIÁVEL ESPECTRAL E ESPACIAL CONSIDERANDO AS CLASSES EM ESTUDO, (N=1)

Nº Variáveis	Variáveis	TOTAL
N=1	B4	41.43
	B3	47.699
	B2	29.139
	B1	25.346
	A	15.879
	P	18.888
	Re	5.7896
	CC	13.225
	Z	54.692
N=3	B4, B3, B2	77.263

Observando a Tabela 4.6, nota-se que os maiores valores de distância acumulada estão nas informações espectrais da banda 4(B4), com 41.43, e banda 3(B3), com 47.699, e na informação espacial fornecida pelo laser scanner (Z), com 54.692. Observa-se que a separabilidade fornecida pelos dados do laser foram melhores que todas as demais, já mencionado na Figura 4.17. O valor da distância associado aos dados do laser scanner (Z) é o mais próximo ao valor da distância combinando as três melhores bandas espectrais (B4, B3, B2). Portanto, como a melhor variável de separabilidade entre classes é a altitude, foi realizada a combinação entre pares de variáveis (N=2), sendo que os dados do laser scanner foram combinados com outras variáveis (Tabela 4.6).

TABELA 4.6 - DISTÂNCIA GLOBAL PARA DIFERENTES COMBINAÇÕES DE 2 VARIÁVEIS CONSIDERANDO AS CLASSES EM ESTUDO, (N=2)

Nº variáveis	Variáveis	TOTAL
N = 2	Z, B4	77.028
	Z, B3	80.232
	Z, B2	67.467
	Z, A	58.803
	Z, P	60.010
	Z, Re	55.829
	Z, CC	58.098

Observando a Tabela 4.6 e analisando as distâncias associadas às informações espectrais, nota-se que os melhores resultados correspondem a combinação entre a altitude do laser e a banda 3 (Z, B3), com 80.232, e banda 4 (Z, B4), com 77.028. Já para as informações espaciais, os melhores resultados estão associados a composição

entre a altitude do laser e ao perímetro (Z, P), com 60.010.

A mesma análise foi realizada considerando três variáveis (N=3). Para isto foram considerados os dados do laser scanner, as bandas espectrais e o descritor de forma que apresentaram maior distância na Tabela 4.6. Estas distâncias são mostradas na Tabela 4.7. Observa-se que as melhores composições espectrais são as que envolvem as bandas 4 e banda 3 e a diferença de altitude, com 97.187. Nota-se, também, que entre os melhores não há descritores de forma. Quando substitui uma banda espectral pela diferença de altitude (Z) aumenta a qualidade.

TABELA 4.7 - DISTÂNCIA GLOBAL PARA DIFERENTES COMBINAÇÕES DE 3 VARIÁVEIS CONSIDERANDO AS CLASSES EM ESTUDO, (N=3).

Nº Variáveis	Variáveis	TOTAL
N=3	Z, B4, B3	97.187
	Z, B4, B2	86.664
	Z, B3, B2	86.686
	B4, B3, B2	77.263

No intuito de verificar a contribuição das melhores variáveis apresentadas pela Tabela 4.7, foi realizado uma análise considerando quatro e cinco variáveis (Tabela 4.8). Observando as combinações, utilizando cinco variáveis (N=5), com melhores descritores de forma, que são área e perímetro, as melhores bandas espectrais (bandas 4 e 3) e a diferença de altitude, nota-se que o valor da distância global obtido, 101.2, é muito próximo a distância global da combinação envolvendo as três melhores bandas espectrais e a altitude (Z, B4, B3, B2), com 102.99, que utiliza quatro variáveis. Isto significa que a contribuição da composição entre estas três bandas espectrais torna-se indispensável para uma classificação que busca uma melhor eficiência. Ainda, conclui-se que se o usuário não tiver acesso a dados oriundos do laser scanner (altitudes), os descritores de forma podem ser significativos, porém como encontra-se disponível, para este estudo, dados do laser scanner, não vale a pena considerar estes descritores de forma.

TABELA 4.8 - DISTÂNCIA GLOBAL PARA DIFERENTES COMBINAÇÕES DE 4 E 5 VARIÁVEIS CONSIDERANDO AS CLASSES EM ESTUDO, (N=4 E N=5).

Nº Variáveis	Variáveis	TOTAL
N=4	B4, B3, B2, A	79.596
	B4, B3, B2, P	80.768
	Z, B4, B3, B2	102.99
N=5	Z, B3, B2, A, P	91.606
	Z, B4, B3, A, P	101.2

Da Tabela 4.8, conclui-se que a melhor combinação é composta pelas bandas espectrais 2, 3, 4 e com a altitude do laser. Portanto, na análise de conjuntos de cinco variáveis (N=5), foram consideradas as variáveis laser scanner, banda 4, banda 3 e banda 2 combinadas com os descritores de forma (área, perímetro, coeficiente de compacidade e razão entre eixos da elipse) (Tabela 4.9).

TABELA 4.9 - DISTÂNCIA GLOBAL PARA DIFERENTES COMBINAÇÕES DE 5 VARIÁVEIS CONSIDERANDO AS CLASSES EM ESTUDO, (N=5).

Nº Variáveis	Variáveis	TOTAL
N=5	Z, B4, B3, B2, A	104.68
	Z, B4, B3, B2, P	105.18
	Z, B4, B3, B2, Re	103.22
	Z, B4, B3, B2, CC	104.34

Observando a Tabela 4.9, nota-se que as melhores combinações entre os descritores são as que envolvem o perímetro e a área. Porém todos os valores das distâncias globais estão muito próximos, demonstrando que os descritores de forma têm uma menor influência na separabilidade das classes, que é dado pelo aumento significativo do valor da distância global (TOTAL). Esta variabilidade do valor da distância global pode ser observada, comparando os valores das combinações com 5 variáveis desta tabela com a distância obtida para a combinação das informações espectrais e a altitude do laser, mostrada na Tabela 4.8.

Na Tabela 4.10, é apresentada a combinação entre todas as variáveis disponíveis (o laser, bandas espectrais 2, 3, 4, área, perímetro, razão entre eixos e coeficiente de compacidade), oito variáveis, como modelo de referência para avaliação da influência dos descritores. Na mesma tabela, são apresentados valores das distâncias das combinações entre o laser, as bandas espectrais e a combinação entre pares de descritores de forma, assim formando combinações com seis e sete variáveis.

TABELA 4.10 - DISTÂNCIA GLOBAL PARA DIFERENTES COMBINAÇÕES DE 6, 7 E 8 VARIÁVEIS CONSIDERANDO AS CLASSES EM ESTUDO, (N=6, N=7 E N=8)..

Nº variáveis	Variáveis	TOTAL
N=6	Z, B4, B3, B2, Re, CC	104.56
	Z, B4, B3, B2, A, P	106.82
N=7	Z, B4, B3, B2, A, P, CC	108.1
	Z, B4, B3, B2, A, P, Re	107.04
N=8	Z, B4, B3, B2, A, P, Re, CC	108.31

Observando a Tabela 4.10 e comparando o valor da distância da combinação que possui o maior número de variáveis (N=8), com os demais valores das distâncias das combinações calculadas, confirma-se a pouca contribuição dos descritores de forma na separabilidade das classes, devido seus valores estarem muito próximos.

Para avaliar a influência da inclusão de cada variável na separabilidade entre classes, foi realizado o cálculo do percentual que cada variável contribui. Este cálculo consiste na razão entre o valor da distância global em questão e a máxima separabilidade, dada pela distância considerando todas as variáveis (Tabela 4.10). Estes valores podem ser observados na Tabela 4.11.

TABELA 4.11 - RESUMO COM AS MELHORES DISTÂNCIAS GLOBAIS PARA CADA NÍVEL DE COMBINAÇÕES ENTRE VARIÁVEIS

Nº variáveis	Variáveis	% em relação ao máximo	Diferença entre níveis (%)	Diferença em relação ao máximo (%)
N=1	Z	50,74	-	49,26
N=2	Z, B3	74,07	23,33	25,93
N=3	Z, B4, B3	89,73	15,66	10,27
N=4	Z, B4, B3, B2	95,09	5,36	4,91
N=5	Z, B4, B3, B2, P	97,11	2,02	2,89
N=6	Z, B4, B2, B1, A, P	98,62	1,51	1,38
N=7	Z, B4, B2, B1, A, P, CC	99,81	1,19	0,19
N=8	Z, B4, B2, B1, A, P, Re, CC	100	0,19	0

Observando a Tabela 4.11, nota-se que a altimetria e os dados espectrais são responsáveis por 95.09% da maior separabilidade possível, ou seja, os descritores de

forma estão decidindo 4.91% da separabilidade total das classes. Percebe-se que a inclusão de um descritor de forma reduz esta decisão para 2.89%. Observando que a diferença entre níveis, a partir deste número de variáveis (N=5), torna-se muito próximos, abaixo de 1.51%, decidindo 1.38% do máximo de separabilidade, levando a concluir que a inclusão de novas variáveis de forma não serão pertinentes, comparado com a separabilidade máxima.

Portanto, conclui-se que as bandas espectrais e a altitude fornecida pelo laser scanner são indispensáveis para a separabilidade das classes. Analisando todas as tabelas apresentadas, conclui-se também que os descritores de forma em estudo contribuem muito pouco em relação a altimetria fornecida pelo laser scanner. Para a verificação destes cálculos serão apresentados, nas próximas seções, resultados empregando classificadores que levam em consideração as variáveis de altitude, fornecida pelo laser e informações espectrais. Devido os resultados apresentados, serão realizadas combinações com estas variáveis (laser e bandas espectrais), empregando um descritor de forma, ou seja, composições entre cinco variáveis (N=5).

4.3 - CLASSIFICAÇÃO ORIENTADA A REGIÕES UTILIZANDO REDES NEURAIS ARTIFICIAIS COM O PROGRAMA IMPLEMENTADO EM C++.

Foram realizados vários testes para a definição da arquitetura e quantidade de seus neurônios propícios para a resolução do problema apresentado neste estudo.

Para a utilização da rede neural artificial, o primeiro passo consiste na escolha da arquitetura a ser empregada. Para este trabalho, foi empregado o método de rede neural multicamada com o aprendizado de retro-propagação. O segundo passo consiste na definição da quantidade de seus neurônios que serão utilizados. Foram utilizadas duas camadas intermediárias, sendo 60 neurônios na primeira camada, e 20 neurônios na segunda camada. Foram coletados 46 amostras para cada classe, distribuídas pelos três recortes realizados, sendo as classes: Edificações de telhado claro, Edificações de telhado escuro, Gramíneas, Árvores e Vias.

4.3.1. - Testes usando diferentes combinações de variáveis.

Devido o programa implementado possuir a restrição de utilizar apenas três bandas espectrais, foram considerados as três melhores informações espectrais (banda 4, banda3 e banda 2), representado pela letra “I” nas tabelas apresentadas.

Inicialmente foram realizados testes puramente espectrais (I), com o objetivo de referenciar os demais testes, e posteriormente combinado com a altimetria fornecida pelos dados do laser scanner, obtendo-se assim uma combinação com quatro variáveis (três espectrais e uma espacial). Em um terceiro momento foram adicionadas às quatro variáveis, uma quinta variável, o descritor de forma. Os testes foram organizados desta maneira com o intuito de verificar a influência dos dados altimétricos e os descritores de forma na rede neural artificial.

Depois da imagem ser classificada, esta imagem temática produzida foi comparada com a área de verificação, mencionado na seção 3.2.9, do segundo e terceiro recorte.

Portanto, nas Tabela 4.12 e Tabela 4.13 estão sendo apresentados resultados através do cálculo do índice “kappa” e a exatidão das classes em relação ao produtor e ao usuário, os quais irá apontar o melhor resultado. Observando-se estas tabelas nota-se que o melhor resultado é o que apresenta as informações espectrais (I) e a altitude fornecida pelo laser (Z), com 65.2.

TABELA 4.12 - RESULTADOS DA REDE NEURAL PARA O SEGUNDO RECORTE REALIZADO NA IMAGEM.

Composição	Kappa	Exatidão das Classes				
		Telhados	Vegetação Branda	Árvores	Vias	
I	41.1	61.8	49.9	63.6	57.9	Produtor
		70.3	72.3	62.2	45.8	Usuário
I+Z	65.2	85.8	73.3	48.4	74.2	Produtor
		84.1	78.2	64.3	65.0	Usuário
I+Z+A	33.7	38.1	57.7	37.5	64.0	Produtor
		69.7	69.3	21.8	41.0	Usuário
I+Z+P	37.7	45.3	71.0	1.9	66.9	Produtor
		67.8	69.2	7.7	41.3	Usuário
I+Z+Re	38.5	37.5	71.3	29.3	65.7	Produtor
		79.8	67.4	25.1	43.2	Usuário
I+Z+CC	36.5	44.8	70.8	5.8	62.9	Produtor
		64.4	69.4	16.4	40.4	Usuário

TABELA 4.13 - RESULTADOS DA REDE NEURAL PARA O TERCEIRO RECORTE REALIZADO NA IMAGEM.

Composição	Kappa	Exatidão das Classes				
		Telhados	Vegetação Branda	Árvores	Vias	
I	41.6	51.2	61.5	70.9	58.0	Produtor
		84.7	44.3	37.4	55.3	Usuário
I+Z	51.1	71.4	57.9	72.5	68.6	Produtor
		51.9	89.9	43.8	61.6	Usuário
I+Z+A	39.9	67.1	61.5	29.4	60.5	Produtor
		34.6	78.0	46.2	55.8	Usuário
I+Z+P	41.3	66.8	60.8	31.0	65.0	Produtor
		36.2	78.6	47.5	56.2	Usuário
I+Z+Re	41.0	65.0	60.0	34.0	64.7	Produtor
		36.4	78.7	46.8	55.3	Usuário
I+Z+CC	40.3	72.0	61.2	30.2	58.7	Produtor
		35.7	78.0	48.4	55.7	Usuário

* Vegetação Branda (Gramínea, pastagem, pequenos arbustos e etc).

Observando-se as Tabelas 4.12 e 4.13, nota-se que quando são combinados os valores dos descritores de forma à altitude e os dados espectrais, percebe-se que a rede neural confunde-se. Os resultados apresentados por estas combinações (Informação espectral + diferença de altitude + descritores de forma) estão próximos à que utilizou apenas a informação espectral, em ambas tabelas. Portanto conclui-se que a rede neural não localizou, na superfície de erros, o mínimo global, mas sim um mínimo local, ou um mínimo local situado em um plano dentro da superfície de erro, do qual não conseguiu sair. O valor do erro neste mínimo local é muito próximo do erro associado ao resultado final da rede neural com informação espectral. Portanto os descritores de forma combinados com a altitude do laser e os dados espectrais das bandas 4, 3 e 2 não são apropriados para melhorar a discriminação das classes.

4.4 - CLASSIFICAÇÃO ORIENTADA A REGIÕES UTILIZANDO DISTÂNCIA MÍNIMA.

Para a execução deste método foi utilizada a metodologia já mencionada na seção 3.2.7. Segue, na próxima seção, a comparação do laser scanner com os descritores de forma e os dados espectrais disponíveis nas bandas 4, 3 e 2.

4.4.1. - Testes usando diferentes combinações de variáveis.

Da mesma forma como expostos na seção 4.3.1, os passos e critérios foram os mesmos na execução dos testes. Os resultados podem ser observados nas Tabelas 4.14 e 4.15.

Nota-se que o melhor resultado, para a Tabela 4.14, é o que considerou as informações espectrais (I), a altitude do laser (Z) e o descritor de forma área (A), com um “kappa” de 55.3. Ainda observando esta tabela, nota-se que todos os valores obtidos estão muito próximos ao valor da combinação das informações espectrais (I) e a altitude do laser (Z), com 53.0. Para a Tabela 4.15, o melhor resultado está na composição que envolve o descritor de forma perímetro (P), com 48.7. Observando o valor obtido para a combinação da informação espectral (I) e altitude do laser (Z), 47.6, nota-se a mesma proximidade mostrada na Tabela 4.14. Estes resultados concretizam as estatísticas apresentadas na seção anterior (seção 4.2), que os descritores de forma pouco contribuem para a classificação dos objetos na cena.

TABELA 4.14 - RESULTADOS DO MÉTODO DE MÍNIMA DISTÂNCIA PARA O SEGUNDO RECORTE REALIZADO NA IMAGEM.

Composição	Kappa	Exatidão das Classes				
		Telhados	Vegetação Branda	Árvores	Vias	
I	35.1	53.2	69.9	35.3	63.2	Produtor
		66.7	23.4	77.6	46.4	Usuário
I+Z	53.0	65.5	39.1	65.8	77.9	Produtor
		83.7	66.0	78.6	47.8	Usuário
I+Z+A	55.3	69.5	41.1	67.8	76.2	Produtor
		85.6	63.5	77.4	50.1	Usuário
I+Z+P	54.6	67.5	39.7	67.2	78.0	Produtor
		86.1	66.3	77.6	49.2	Usuário
I+Z+Re	52.7	64.8	38.7	65.9	77.8	Produtor
		83.7	65.2	78.1	47.8	Usuário
I+Z+CC	54.1	66.6	38.1	64.9	81.1	Produtor
		86.8	65.8	79.1	48.4	Usuário

* Vegetação Branda (Gramínea, pastagem, pequenos arbustos e etc).

TABELA 4.15 - RESULTADOS DO MÉTODO DE MÍNIMA DISTÂNCIA PARA O TERCEIRO RECORTE REALIZADO NA IMAGEM.

Composição	Kappa	Exatidão das classes				
		Telhados	Vegetação Branda	Árvores	Vias	
I	29.7	77.5	25.7	54.9	35.0	Produtor
		29.6	60.7	81.6	33.5	Usuário
I+Z	47.6	73.6	37.1	59.2	77.7	Produtor
		48.7	77.2	89.7	47.1	Usuário
I+Z+A	47.9	74.4	37.6	59.8	77.5	Produtor
		50.9	76.3	87.4	47.7	Usuário
I+Z+P	48.7	74.6	37.5	58.3	81.6	Produtor
		51.6	76.5	90.8	47.6	Usuário
I+Z+Re	47.4	73.5	36.2	59.1	77.8	Produtor
		48.6	77.1	89.6	47.0	Usuário
I+Z+CC	48.3	74.0	37.5	57.0	83.2	Produtor
		53.9	77.2	91.2	46.5	Usuário

* Vegetação Branda (Gramínea, pastagem, pequenos arbustos e etc).

Portanto, das Tabelas 4.14 e 4.15, conclui-se que as composições que envolveram os descritores de forma, informação espectral e diferença de altitude obtiveram valores muito próximos à composição que envolve apenas a diferença de altitude e informação espectral, levando assim a irrelevância da utilização dos descritores de forma na classificação de regiões, através do método de mínima distância, uma vez tendo dados de altitude oriundas do laser scanner.

4.5 - CLASSIFICAÇÃO ORIENTADA A REGIÕES UTILIZANDO MÁXIMA VEROSSIMILHANÇA.

Para a execução deste método foi utilizada a metodologia já mencionada na seção 3.2.8. Segue, na próxima seção, a comparação do laser scanner com os descritores de forma.

Os resultados obtidos estão apresentados nas Tabelas 4.16 e 4.17, correspondendo ao segundo e terceiro recorte respectivamente.

Observando-se a Tabela 4.16, nota-se que o melhor resultado está na composição que envolve informações espectrais, altimetria do laser e o descritor de forma área (I+Z+A), com 41.0. Observando os demais valores do “kappa”, nota-se que

todos os valores que envolvem descritores de forma, possuem um valor muito próximo do valor obtido por meio da composição informação espectral (I) e altimetria do laser (Z), com 39.6. Para a Tabela 4.17, observa-se que a melhor composição está no que envolve o descritor de perímetro (I+Z+P). Observando-se os demais valores que envolvem descritores de forma, nota-se a mesma proximidade da Tabela 4.16. Portanto a contribuição das variáveis de forma é muito baixa, o que não justifica seu uso.

TABELA 4.16 - RESULTADOS DO MÉTODO DE MÁXIMA VEROSSIMILHANÇA PARA O SEGUNDO RECORTE REALIZADO NA IMAGEM.

Composição	Kappa	Exatidão das Classes				
		Telhados	Vegetação Branda	Árvores	Vias	
I	32.0	51.3	55.2	49.5	46.1	Produtor
		65.6	20.0	57.1	52.1	Usuário
I+Z	39.6	40.6	6.1	88.9	55.5	Produtor
		87.0	53.2	57.8	46.4	Usuário
I+Z+A	41.0	45.7	7.4	86.6	55.5	Produtor
		85.8	53.3	58.8	46.3	Usuário
I+Z+P	38.5	42.5	6.1	90.5	48.9	Produtor
		82.5	49.5	56.5	46.7	Usuário
I+Z+Re	40.3	43.1	5.5	84.7	59.6	Produtor
		82.1	63.9	60.5	44.9	Usuário
I+Z+CC	38.2	42.9	6.1	88.9	49.5	Produtor
		81.2	52.0	56.4	46.3	Usuário

* Vegetação Branda (Gramínea, pastagem, pequenos arbustos e etc).

TABELA 4.17 - RESULTADOS DO MÉTODO DE MÁXIMA VEROSSIMILHANÇA PARA O TERCEIRO RECORTE REALIZADO NA IMAGEM.

Composição	Kappa	Exatidão das Classes				
		Telhados	Vegetação Branda	Árvores	Vias	
I	30.1	70.2	22.3	62.5	31.5	Produtor
		27.8	34.5	75.6	46.2	Usuário
I+Z	45.0	62.3	5.8	80.4	64.3	Produtor
		49.8	45.3	77.3	50.9	Usuário
I+Z+A	43.6	63.6	6.5	78.5	62.3	Produtor
		47.7	46.9	77.3	49.5	Usuário
I+Z+P	45.3	63.2	6.4	80.4	64.4	Produtor
		50.4	47.7	77.1	51.5	Usuário
I+Z+Re	42.6	63.3	3.4	75.6	65.4	Produtor
		44.9	34.9	79.3	48.2	Usuário
I+Z+CC	44.9	63.8	6.6	81.1	61.5	Produtor
		47.3	48.2	77.3	51.9	Usuário

* Vegetação Branda (Gramínea, pastagem, pequenos arbustos e etc).

4.6 - VERIFICAÇÃO DA INFLUÊNCIA DO LASER SCANNER PERANTE OS DESCRITORES DE FORMA UTILIZANDO OS CLASSIFICADORES IMPLEMENTADOS

Estatisticamente observa-se que os descritores de forma estão decidindo apenas 4.19% da separabilidade das classes, comparado com a máxima separabilidade possível, mencionada na seção 4.2. Observando os resultados, pelas Tabelas 4.12 a 4.17, obtidos por meio dos classificadores implementados, concretiza-se a estatística da seção 4.2, apontando que estes descritores de forma pouco contribuem para o produto final, se comparados com a altimetria fornecida pelo laser scanner. Devido o motivo apresentado acima, foi descartado, para este estudo, os descritores de forma, utilizados como variáveis de integração.

4.7 - COMPARAÇÃO ENTRE CLASSIFICADORES

Para comparar os classificadores, inicialmente foram selecionados os melhores valores de “kappa” encontrados para as combinações exibidas nas Tabelas 4.12 a 4.17. Estes valores e combinações podem ser observados pela Tabela 4.18.

TABELA 4.18 - MELHORES RESULTADOS DO INDICE KAPPA PARA OS MÉTODOS DE CLASSIFICAÇÃO EMPREGADOS NO ESTUDO.

Variável	Método		
	Redes Neurais Artificiais	Distancia Mínima	Máxima Verossimilhança
I	41.6	35.1	32.0
I + Z	65.2	53.0	45.0

I = imagem QUICK BIRD II com as bandas 4,3 e 2;

Z = Diferença de altimetria, dados oriundos do laser scanner.

Observando a Tabela 4.18, nota-se que a rede neural exhibe os melhores valores do “kappa” encontrado pelos diferentes métodos e recortes, tanto na utilização de apenas informações espectrais, quanto na composição de informação espectral e diferença de altitude, com 41.6 e 65.2 respectivamente. Com estes resultados conclui-se que a rede neural é o melhor método para a classificação de regiões utilizando

dados oriundos do laser scanner.

Nas Figuras 4.18, 4.19 e 4.20 estão apresentadas as imagens que obtiveram o melhor produtor final.

FIGURA 4.18 - IMAGEM DO PRIMEIRO RECORTE CLASSIFICADA PELA REDE NEURAL ARTIFICIAL UTILIZANDO A COMPOSIÇÃO I+Z (INFORMAÇÃO ESPECTRAL E ALTIMETRIA DO LASER)

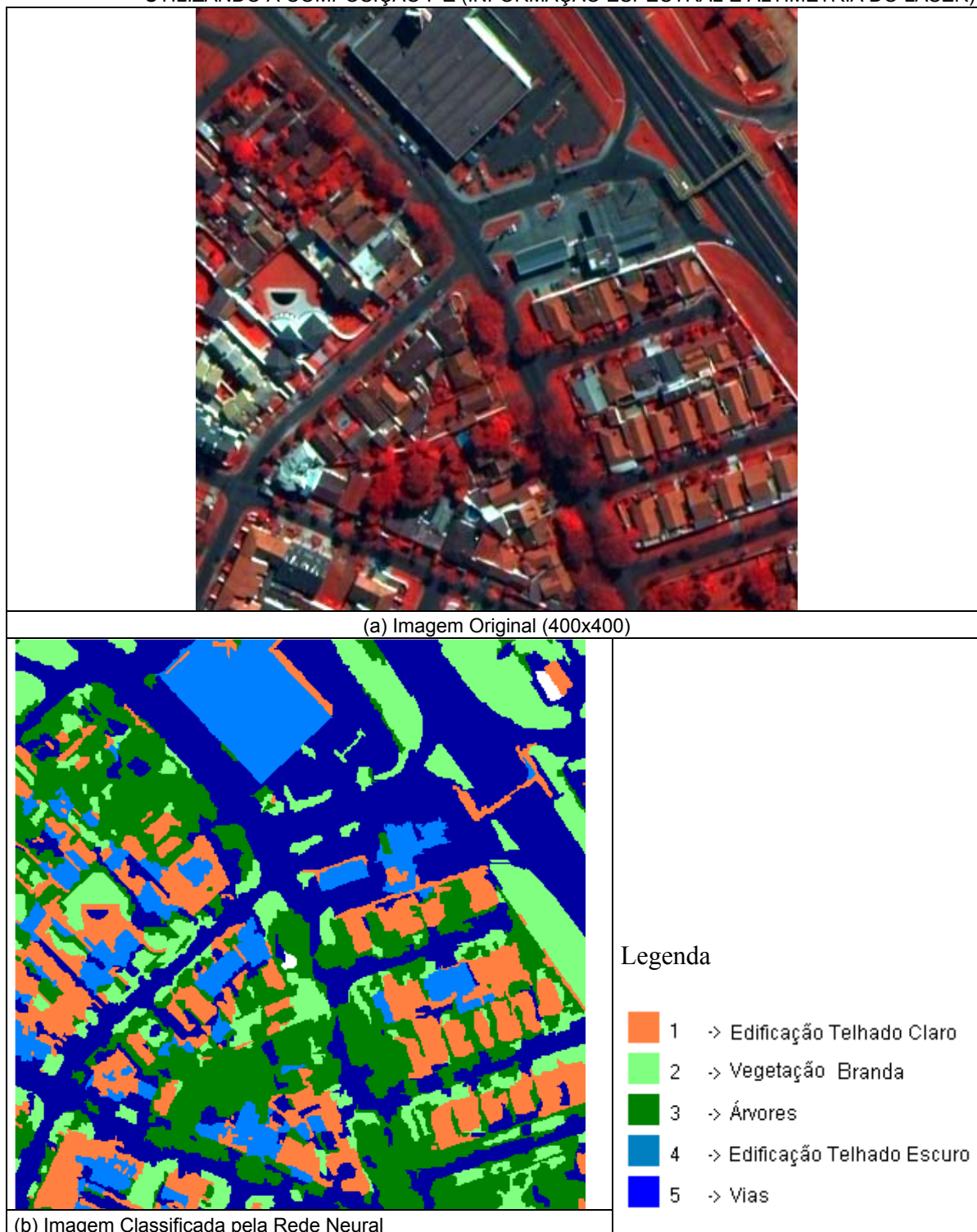


FIGURA 4.19 - IMAGEM DO SEGUNDO RECORTE CLASSIFICADA PELA REDE NEURAL ARTIFICIAL UTILIZANDO A COMPOSIÇÃO I+Z (INFORMAÇÃO ESPECTRAL E ALTIMETRIA DO LASER)

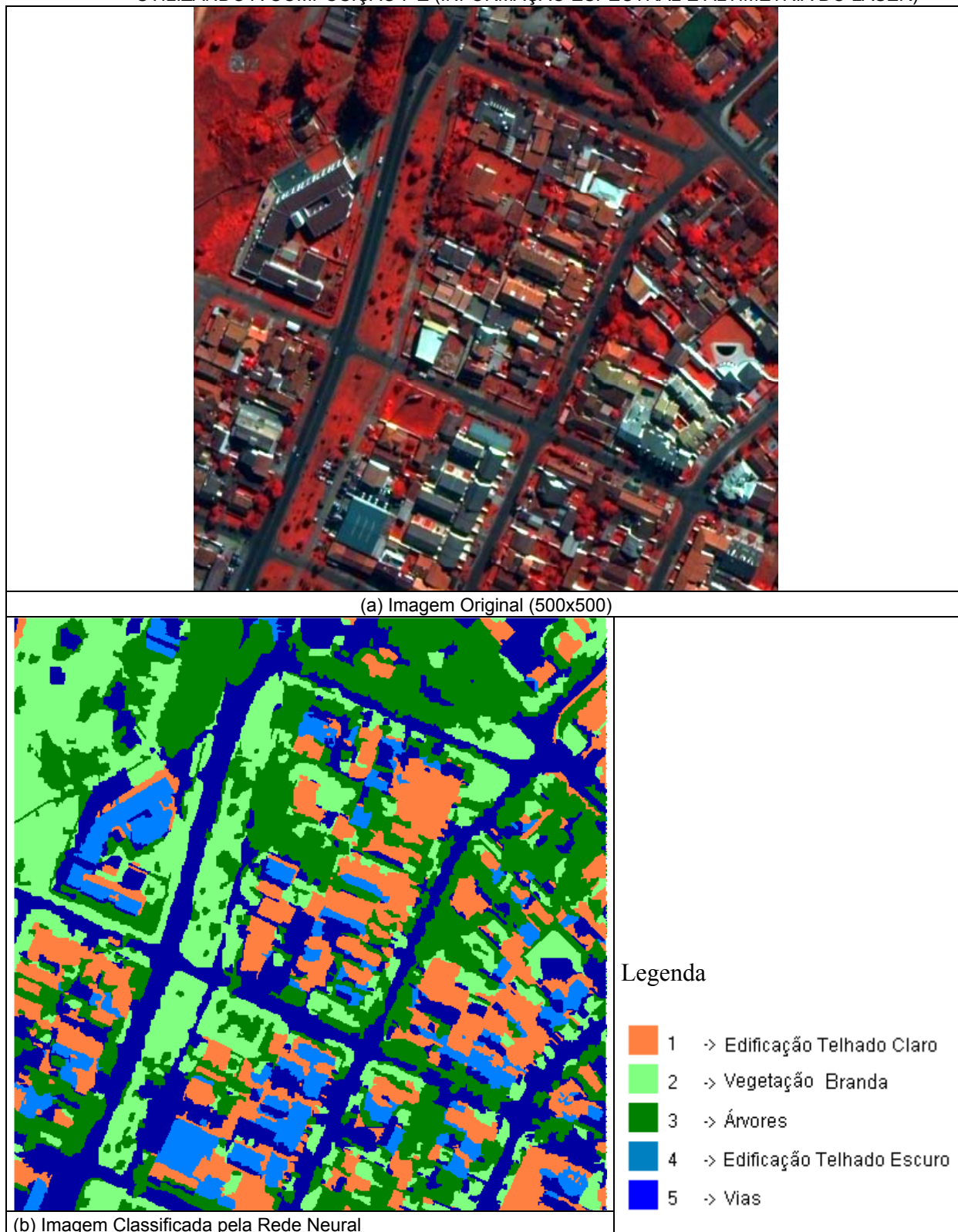
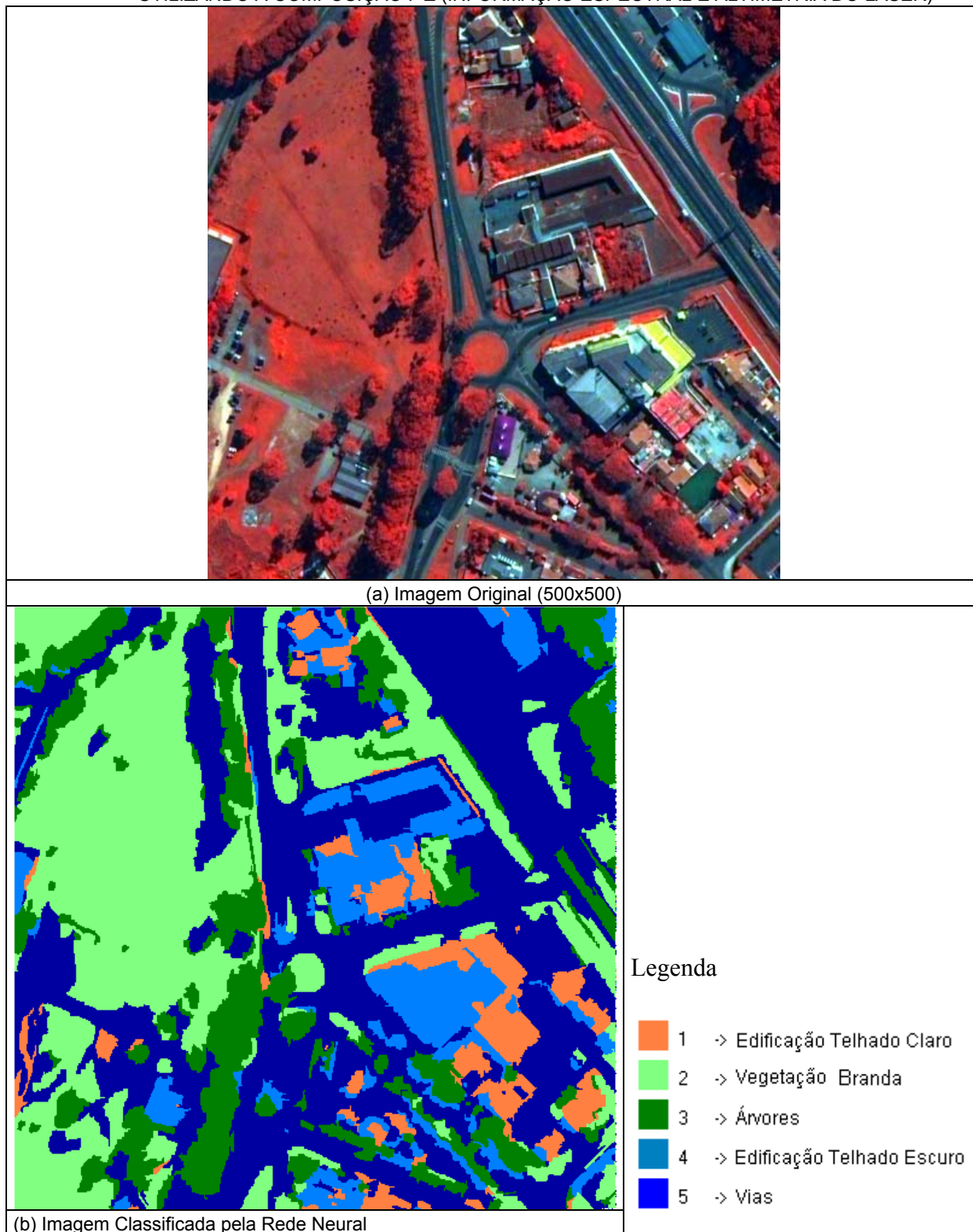


FIGURA 4.20 - IMAGEM DO TERCEIRO RECORTE CLASSIFICADA PELA REDE NEURAL ARTIFICIAL UTILIZANDO A COMPOSIÇÃO I+Z (INFORMAÇÃO ESPECTRAL E ALTIMETRIA DO LASER)



4.8 - DESEMPENHO DAS REDES NEURAIAS NA IDENTIFICAÇÃO DE CASAS

A seguir foi analisado o desempenho do método redes neurais na identificação de casas, representadas pela letra “T”, nas Tabelas 4.12 e 4.13, que representa as edificações de telhado claro e escuro. Para isto, foram considerados os recortes 2 e 3, por existir fotos aéreas destas regiões que permitiram a restituição dos prédios. Para o 2º recorte, nota-se que a exatidão, da melhor composição, do ponto de vista do produtor é de 85.8. Já do ponto de vista do usuário, a exatidão é de 84.1. Para o 3º recorte, a exatidão da melhor composição, do ponto de vista do produtor é de 71.4. Já do ponto de vista do usuário, a exatidão é de 51.9. Portanto, o produto, apresenta uma confiabilidade boa para o 2º recorte, tanto do ponto de vista do produtor como do ponto de vista do usuário para a classe edificação. Já para o 3º recorte, o produto, apresenta uma confiabilidade boa do ponto de vista do usuário e uma confiabilidade baixa do ponto de vista do produtor.

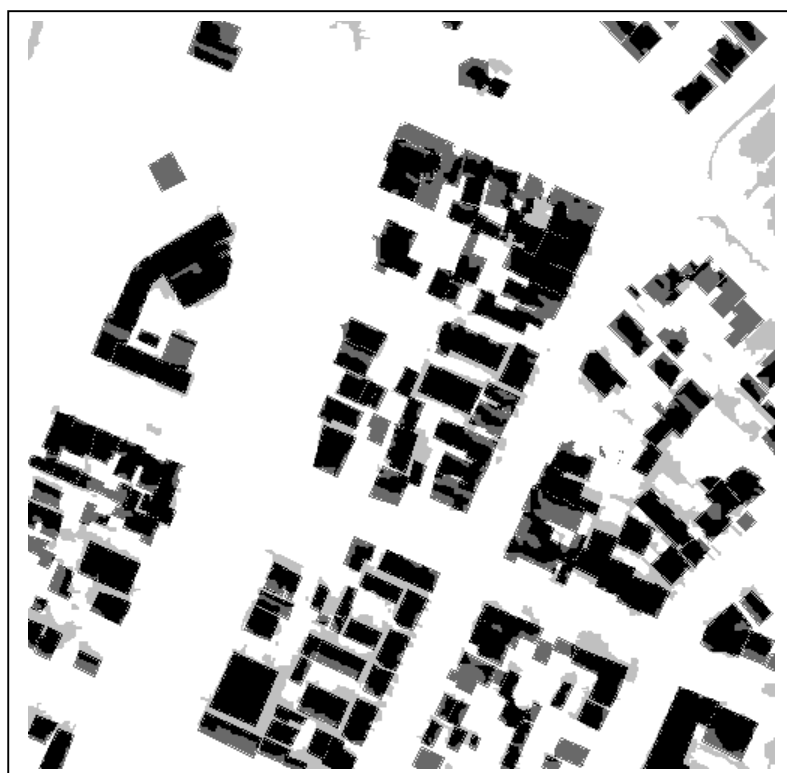
Para verificar o sucesso da detecção de casas nas imagens, foram contadas todas as edificações da imagem original apresentada pelas Figuras 4.18, 4.19 e 4.20. Nas mesmas figuras foi realizada a contagem de todas as edificações não identificadas pelo classificador rede neural. Os resultados estão sendo apresentados na Tabela 4.19, onde se observa que a maioria das casas foi identificada, em torno de 96 – 100%, dependendo do recorte. Considerando todos os recortes, a taxa de sucesso é de 93.6%, um valor alto, ou seja, de cada 100 casas, apenas 8 não foram identificadas. Porém observando-se as imagens nota-se que a maioria das edificações identificadas foram extrapoladas em sua área, hora pelo muro que separa as edificações, hora pela sombra da edificação e hora pela mistura de objetos nos “pixels” de bordas. Isto é também constatado pelo baixo valor da acurácia “do usuário” citado acima.

TABELA 4.19 - IDENTIFICAÇÃO DAS EDIFICAÇÕES NA CLASSIFICAÇÃO DAS IMAGENS.

	Primeiro Recorte	Segundo Recorte	Terceiro Recorte	Total
Edificações identificadas na imagem original	78	117	25	220
Edificações não identificadas pelo algoritmo de redes neurais	3	12	-	15
TOTAL	81	129	25	-
TOTAL %	96%	90.7%	100%	93.6%

Pela Figura 4.21, nota-se a eficiência do produto na identificação de edificações do meio urbano. Percebe-se que o programa superdimensiona a identificação das edificações. Justifica-se este fato devido aos muros existentes na área que separa uma edificação da outra, a transição entre o valor digital de uma classe e outra e as sombras existentes na cena.

FIGURA 4.21 - IMAGEM COM A PERSPECTIVA DA EFICIENCIA DO PRODUTO.



- Fundo
- Edificação encontrada com sucesso
- Edificação encontrada, mas inexistente na realidade
- Edificação não encontrada

CAPÍTULO 5 - CONCLUSÕES E RECOMENDAÇÕES

No desenvolvimento desta pesquisa, foram realizados diferentes experimentos, no intuito de avaliar a eficiência do uso da altimetria derivada do laser como variável auxiliar na classificação de imagem de alta resolução espacial. Os dados do laser scanner melhoram a qualidade da classificação comparado com uma classificação puramente espectral. Através da análise dos resultados obtidos por meio dos três métodos de classificação empregados, pode-se concluir que a altimetria fornecida pelo laser scanner possibilita a obtenção de uma imagem temática de melhor qualidade do que os descritores de forma, geralmente utilizados como variável auxiliar.

É possível, através dos três métodos empregados nesta pesquisa (rede neural, distância mínima e máxima verossimilhança), integrar dados de diferentes fontes, no intuito de se obter uma imagem temática da cobertura do solo em regiões urbanas.

As três metodologias de classificação empregadas tiveram um bom desempenho na classificação de imagem de alta resolução, sendo a melhor delas, a que empregou as redes neurais artificiais.

Em seu treinamento, o algoritmo considera implicitamente a escala e dispersão das classes e exige poucas amostras para seu treinamento, apesar de ser muito trabalhosa em seu treinamento devido a escolha da quantidade de neurônios a ser utilizado. O método de mínima distância, com sua simplicidade, alcança índices próximos ao da rede neural, possui boa capacidade de integrar dados de diferentes fontes, mas exigir um maior número de amostras para seu treinamento do que o método de redes neurais.

Os índices encontrados do “kappa” são baixos. Como a avaliação da classificação a ser realizada compara o resultado com a digitalização dos contornos dos objetos, obtidos por restituição fotogramétrica, uma parte dos erros apontados pelo índice “kappa” corresponde a “pixels” localizados nas fronteiras dos objetos. Tais

fronteiras dependem do sucesso da segmentação. A avaliação da imagem temática, por tanto inclui erros na segmentação e na classificação e não pode diferenciá-los.

Nos experimentos, o potencial das redes neurais para identificação de casas foi avaliado e comprovou-se que esta metodologia permite detectar com sucesso a maioria das construções. Analisando o número de casas identificadas, pode-se notar que o sucesso do classificador na detecção de casas é bom, mas que ele não é eficiente em sua delimitação. Este problema também é oriundo da segmentação.

Como a classificação orientada a regiões depende de uma prévia segmentação, a mesma deve ser a mais rigorosa possível. Percebe-se que os resultados das classificações ficaram limitados ao resultado obtido pela segmentação. Não é possível corrigir os erros vindos da segmentação na classificação.

Sugerem-se estudos que aprimorem a segmentação de imagem de alta resolução, que possibilitem uma melhor determinação dos objetos, o que irá aumentar a qualidade do mapa temático resultante, ou mesmo a construção de uma nova metodologia específica para imagem de alta resolução espacial, onde se considere o tamanho, a forma e textura, ou seja, que extraia diferentes objetos em diferentes escalas, simultaneamente.

CAPÍTULO 6 - REFERÊNCIA BIBLIOGRAFICA

- ABELÉM, A. J. G., **Redes Neurais Artificiais Na Previsão De Séries Temporais**, publicação Dissertação de Mestrado pela Universidade Católica do Rio de Janeiro, RJ, Brasil, 197 p., 1994.
- ANDRADE, F. A. **Integração de variáveis espectrais e forma na classificação de imagens de alta resolução utilizando redes neurais artificiais**. Dissertação de mestrado pela Universidade Federal do Paraná, Curitiba - PR, 112 p., 2003.
- ANTUNES, A. F. B., **Classificação de Ambiente Ciliar Baseada em Orientação a Objeto em Imagens de Alta Resolução Espacial**. Tese de Doutorado pela Universidade Federal do Paraná, Curitiba – PR, 145 p., 2003.
- BAKKER, B. e HESKES, T. **Clustering ensembles of neural network models**. *Neural Networks*, v.16, p.261-269, 2003.
- BECERIKLI Y., KONAR A. F. e SAMAD T. **Intelligent optimal control with dynamic neural networks**. *Neural Networks*. V. 16. p. 251 – 259. 2003.
- BOTELHO, M. F. **Aplicação De Redes Neurais Para Identificação De Feições Espaciais**. Monografia apresentada ao Curso de Graduação em Engenharia de Agrimensura da Universidade Federal de Viçosa, Viçosa – MG, 35 p., 2002
- BOTELHO, M. F. e CENTENO, J. S. **A Influência Da Alternância Entre Classes E Amostras Na Entrada Da Rede Neural Na Classificação De Imagem Digital**. Anais do III COLOQUIO de Ciências Geodésicas, p. CD, Curitiba – PR, 14 p., 2003.
- BRONSVELD, K.; CHUTIRATTANAPAN, S.; PATTANAKANOK, B.; SUWANWEAKAMTORN, R. ; TRAKOOLDIT, P. **The use of local knowledge in land use/land cover mapping from satellite images**. *ITC Journal*. v. 4, p. 349 – 348. 1994.
- CENTENO, J. S.; KISHI, R. T. E BÄHR. **Recognition of Buildings Using Scanned Maps and Laser Scanner Altitude Data**. *Photogrammetrie Fernerkundung Geoinformation*, p. 19-28, 1999.
- CHENG, T. **Fuzzy Objects: Their Changes and Uncertainties**. *Photogrammetric Engineering & Remote Sensing*, v.68, n.1, p. 41-49, 2002.
- CHUVIECO, EMILIO. **Fundamentos de Teledetección espacial**. Ediciones Rialp, S.A. Madrid. 452 P. 1996
- CLAUSI, D. A.; ZHAO, Y. **Rapid extraction of image texture by co-ocurrence using a hybrid data structure**. *Computers & Geosciences*, v.28, p.763-774, 2002.
- CROSTA, A P. **Processamento Digital de Imagens de Sensoriamento Remoto**. Gráfica da UNICAMP. Campinas, S. P., 1992.

- ECOGNITION. **User Guide**. Definiens Imaging. München, Germany, 2001.
- FACON, J. **Processamento e Análise de Imagens**. Universidad Nacional de Córdoba, Argentina, 253 p. 1993.
- FERNÁNDEZ – CABALLERO et al. **On Motion Detection Through a Multi-Layer Neural Network Architecture**. Neural Networks, v.13, p.291-303, 2000.
- FERRO, C. J. S.; WARNER, T. A. **Scale and Texture in digital image classification**. Photogrammetric Engineering & Remote Sensing, v.68, n.1, p. 51-63, 2002.
- FHILIP, N. S.; JOSEPH, K. B. **A neural network tool for analyzing trends in rainfall**. Computers & Geosciences, v.29, n.02, p.215-223, 2003.
- FILHO, O. M.; NETO H. V. **Processamento Digital de Imagens**. Editora BRASPORT. Rio de Janeiro, 327 p. 1999.
- FRANKLIN, S. E.; MAUDIE, A. J.; LAVIGNE, M. B.. **Using spatial Co-ocurrence texture to increase forest structure and species composition classification accuracy**. . Photogrammetric Engineering & Remote Sensing, v.67, n.7, p. 849-855, 2001.
- FUNG, T.; CHAN, K. C. **Spatial composition of spectral classes: A structural approach for image analysis of heterogeneous land-use and land-over types**. Photogrammetric Engineering & Remote Sensing, v.60, n.2, p. 173-180, 1994.
- GERMAN, G. W. H. e GAHEGAN, M. N. **Neural Network Architectures for the Classification of Temporal Image Sequences**. Computers & Geosciences, v.22, n.09, p.969-979, 1996.
- GONZALES, R.C.; WOODS, R.E. **Processamento de Imagens Digitais**. Edgard Blücher Ltda. São Paulo, 527 p., 2000.
- HAALA, N.; Brenner, C. **Extraction of buildings and trees in urban environments**. Photogrammetric Engineering & Remote Sensing, v.54, n.2, p. 130-137, 1999.
- HANNA A. I., MANDIC, D. P. **A complex-valued nonlinear neural adaptive filter with a gradient adaptive amplitude of the activation function**. Neural Networks. V. 16. p. 155 – 159. 2003.
- HARRIS, R. J.; MURRAY, R.. **IHS transform for the integration of RADAR imagery with other remotely sensing data**. Photogrammetric Engineering & Remote Sensing. V. 61, n. 12, p. 1631- 1641, 1990.
- HAYKIN, S. **Redes Neurais. princípios e prática**. Porto Alegre: Bookman, 900 p., 2001.
- HINTON, G.. **Learning in papallel networks**. BYTE: Artificial Inteligence, USA, New York, Vol. 10 n° 4, p.265-273, 1985.
- HUISING E. J., PEREIRA L. M. G. **Commercial Implications of Topographic Terrain Mapping Using Scanning Airbone Laser Radar**. Photogrammetric Enginnering and Remote Sensing, Vol. LXIII, n° 4, Abril 1997
- INTERSAT. **Nossos satélites**. www.intersat.com.br. Acesso em Janeiro 2004.

- JAIN, R. KASTURI, R. SCHUNCK, B. G. **Machine Vision**. McGraw Hill, New York, 457 p., 1995.
- JENSEN J. R. **Introductory digital image processing a remote sensing perspective**. 2 Ed. New Jersey, 279p. 1996.
- Jl,C.Y. **Land-Use Classification of Remotely Sensed Data Using Kohonen Self-Organizing Feature Map Neural Networks**. Photogrametric Engineering & Remote Sensing, v.66, n.12, p.1451– 1460, 2000.
- KIEMA, J. B. K.; BÄHR, H. P.. **Fusion of colour infrared imagery and airborne laser scanning data in the automatic classification of urban environments**. Revista Brasileira de Cartografia, Nº 52. P.29-36, 2000.
- KINOSHITA, S.K.; MARQUES, P.M.A. e FRÈRE, A.F. **Utilização da Análise de Componentes Principais e Redes Neurais Artificiais para a Classificação de Nódulos em Mamogramas Digitalizados**. Anais do CBEB'2000, p.1214-1216, 2000.
- KOK, R., SCHNEIDER, T., AMMER, U., 1999. **Object-based classification and application in alpine Forest environment**. International Archives of Photogrammetry and Remote Sensing. Valladolid. Espanha. V. 32, p. 110-155.
- KOVÁCS, Z. L. **Redes Neurais Artificiais – Fundamentos e Aplicações**. Collegium cognitio, 2ª edição, São Paulo, 279 p., 1996.
- LIPPMANN, R. **An Introduction to Computing with Neural Works**. IEEE Computer Society. Vol. 3, nº 4, p. 4-22, 1987.
- LIRA, J. e MALETTI, G. **A supervised contextual classifier based on a region-growth algorithm**. Computers & Geosciences, v.28, n.08, p.951– 959, 2002.
- MARTINS, M. P.; GUIMARÃES, L. N. F. e FONSECA, L. M. G. **Classificador de Texturas por Redes Neurais**. II Congresso Brasileiro de Computação, p.6, 2002.
- MASAHARU, H; HASEGAWA, H.. **Three-dimensional city modeling from laser scanner data by extracting building polygons using region segmentation method**. International Archives fo Photogrammetry and Remote Sensing. Vol. XXXIII, Part B3. P.556-562, Amsterdam 2000.
- MATLAB. **User`s Guide – Neural Network Toolbox**, version 3.0. Mth Works. 1998.
- MOREIRA, M. A. **Fundamentos do sensoriamento remoto e metodologias de aplicação**. Instituto Nacional de Pesquisas Espaciais- INPE. São José dos Campos, 1 Ed. 310 p, 2001.
- MURNION, S. D. **Comparison of Back Propagation and Binary Diamond Neural Networks in the Classification of a Landsat TM Image**. Computers & Geosciences, v.22, n.09, p.995-1001, 1996a.
- MURNION, S. D. **Spatial Analysis Using Unsupervised Neural Networks**. Computers & Geosciences, v.22, n.09, p.1027– 1031, 1996b.
- NISHIDA, W. **Uma Rede Neural Artificial para Classificação de Imagens Multiespectrais de Sensoriamento Remoto**. Dissertação (Mestrado em

- Engenharia de Produção) - Curso de Pós-Graduação em Engenharia de Produção, Universidade Federal de Santa Catarina. Florianópolis, 1998.
- OLIVEIRA, K. C. e BAUCHSPIESS A. **Classificação de Imagens Codificadas por Cadeias Direcionais Utilizando Redes Neurais Artificiais**. V Congresso Brasileiro de Redes Neurais, PUC, Rio de Janeiro – RJ –Brasil, p. 73-78, 2001.
- PEREIRA, S. V..**Classificação e avaliação de sistemas de recuperação de imagens por conteúdo**. Dissertação de mestrado pela Universidade Federal do Paraná, Curitiba - PR, 91 p., 2001.
- RIBEIRO, S. R. A.. **Integração de Imagens Landsat TM e dados auxiliares para a delimitação automática de unidades fisionômicas usando Redes Neurais Artificiais**. Tese de Doutorado pela Universidade Federal do Paraná, Curitiba – PR, 160 p., 2003.
- RIBEIRO, S. R. A.; CENTENO J. A. S. **Clasificación del uso del suelo utilizando redes neurales y el algoritmo maxver**. Innovación, Año 14 n.1, p.55-60. Antofagasta – Chile, 2002.
- RICHARDS, B. **Remote sensing and digital image analysis**. Springer –Verlag, Berlin. 1993
- RIEDMILLER, M.; BRAUN, H. **A direct adaptive method for faster backpropagation learning: The RPROP algorithm**. Proceedings of the IEEE, Int. Conference on Neural Network, San Francisco, USA, p. 589-591, 1993.
- SALI E.e WOLFSON H. **Texture classification in aerial photographs and satellite data**. Int. J. Remote Sensing, V. 13 n.18, p. 3395 – 3408, 1992.
- SCHNEIDER, W.; STEINWENDNER, J. **Landcover Mapping by Interrelated Segmentation and Classification of Satellite Images**. Photogrammetric and Remote Sensing. V. 32, part 7-4-3 w6, june, 1999.
- SCHRÖDER, M.; VALESSA, M., REHRAUER, H.; SEIDEL, K.; DATEU, M.. **Gibbs random field models: a toolbox for spatial information extraction**. Computers & Geosciences, v.26, p.423-432, 2000.
- SHERSTHA W. E., THOMPSON P. Y., DEAN R. G., HARRELL H., **Coastal & Highway Mapping by Airborne Laser Swath Mapping Tecnology**. Third International Airborne Remote Sensing Conference and Exhibition, 7-10 July 1997, Copenhagen, Denmark
- SMITH, M. J.; PRIESTNALL, G.; ASAL, F..**Combining LIDAR and photogrammetry for urban and rural landscape studies**. International Archives fo Photogrammetry and Remote Sensing. Vol. XXXIII, Part B3. P. 44-50, Amsterdam 2000.
- SOHN, Y. e REBELLO, S. **Supervised and Unsupervised Spectral Angle Classifiers**. Photogrammetric Engineering & Remote Sensing, v.68, n.12, p.1271–1280, 2002.

- SONKA, M; HLAVAC, V; BOYLE, R.. **Image Processing, analysis, and Machine Vision**. Brooks/Cole Publishing Company. USA, 1998.
- TAFNER, M. A.; XEREZ, M.; RODRIGUES FILHO, I. W.. **Introdução e princípios de neurocomputação redes neurais artificiais**. P.193. Blumenau, 1996.
- TONDER, G. J. e EJIMA, Y. **The patchwork engine: imagem segmentation from shape symmetries**. Computers & Geosciences, v.13, p.291-303, 2000.
- TORINO, C. V. e BARRETO, D. T. **Detecção de Fumaça Através do Processamento de Imagem Utilizando Redes Neurais em um Sistema Híbrido**. XIV Congresso Brasileiro de Automática, Natal – RN, p.1434-1439, 2002.
- TOUTIN, T. **Error Tracking in Ikonos Geometric Processing Using a 3D Parametric Model**. Photogrammetric Engineering & Remote Sensing, v.69, n.1, p.43-51, 2003.
- VIEIRA, C. A. O. **Accuracy Of Remotely Sensing Classification Of Agricultural Crops: A Comparative Study**, Thesis Submitted to the University of Nottingham for the Degree of Doctor of Philosophy, p.396, 2000.
- WARNER, T. A. ; SHANK, M.. **An evaluation of the potential fuzzy classification of multispectral data using artificial neural networks**. Photogrammetric Engineering & Remote Sensing. v.63, n.11, p.1285-1294, 1997.
- YILMAZ, S.; DEMIRCIOGLU, C. E AKIN, S. **Application de artificial neural networks to optimum bit selection**. Computers & Geosciences, v.28, p.261-269, 2002.