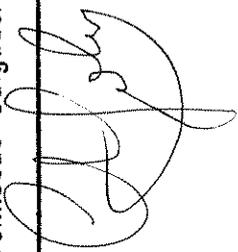


Avaliação de classificadores espectrais de imagens Landsat-TM em áreas rurais densamente ocupadas

- o caso da região de Brotas e Torrinha, São Paulo, Brasil -

Este exemplar corresponde à redação final da tese defendida pelo(a) candidato(a)
Mônica Takako Shimabukuro
Reuso
é aprovada pela Comissão Julgadora.
22/11/96



Mônica Takako Shimabukuro

Dissertação apresentada ao Instituto de Biologia da Universidade Estadual de Campinas como requisito para obtenção do título de Mestre em Biologia Vegetal

Orientador: Prof. Dr. Carlos Alfredo Joly †

Co-orientador: Prof. Dr. Álvaro Penteado Crósta †

Campinas, São Paulo

1996

6260044939

UNIDADE	BC
N.º CHAMADA:	Unicamp
Sh62a	
V.	Ex.
TOMBO BC	29436
PROC.	281/97
C	<input type="checkbox"/>
D	<input checked="" type="checkbox"/>
PREÇO	89,11,00
DATA	16/02/97
N.º CPD	

CM-00097438-0

FICHA CATALOGRÁFICA ELABORADA PELA
BIBLIOTECA CENTRAL DA UNICAMP

Shimabukuro, Mônica Takako
Sh62a Avaliação de classificadores espectrais de imagens Landsat-TM em áreas rurais densamente ocupadas : o caso da região de Brotas e Torrinha, São Paulo, Brasil / Mônica Takako Shimabukuro. -- Campinas, SP : [s.n.], 1996.

Orientador : Carlos Alfredo Joly.

Dissertação (mestrado) - Universidade Estadual de Campinas, Instituto de Biologia.

1. Sensoriamento remoto - Classificação. 2. Processamento de imagens - Técnicas digitais. 3. Agrossilvicultura. I. Joly, Carlos Alfredo. II. Universidade Estadual de Campinas. Instituto de Biologia. III. Título.

O grande Mono fecha os olhos, volta a se coçar e murmura: antes que o sol se ocultasse totalmente - agora corre entre os altos bambus com um animal perseguido pela sombra - consegui reduzir o bosque a um catálogo. Uma página de emaranhada caligrafia vegetal. Matagal de signos: como lê-lo, como abrir caminho por essa espessura? Hanumän sorri com prazer diante da analogia que lhe acaba de ocorrer: caligrafia e vegetação, arvoredo e escritura, leitura e caminho. Caminhar: ler um trecho do terreno, decifrar um pedaço de mundo. A leitura considerada como um caminho para ... O caminho como uma leitura: uma interpretação do mundo natural? Volta a fechar os olhos e vê a si mesmo em outra idade, escrevendo (sobre um papel ou sobre uma rocha? com uma pena ou com um pincel?) o ato de Mahanätaka, onde é descrita sua visita ao bosque do palácio de Rävana. Ele compara sua retórica a uma página de caligrafia indecifrável e pensa: a diferença entre a escritura humana e a divina reside no fato de que o número de signos da primeira é limitado enquanto que o da segunda é infinito: por isso o universo é um texto desprovido de sentido, ilegível até mesmo para os deuses. A crítica do universo (e a dos deuses) chama-se gramática ... Perturbado por esse estranho pensamento, Hanumän salta do muro ao chão, permanece agachado por um instante, ergue-se, espreita os quatro pontos cardeais e, com decisão, penetra no matagal.

Octavio Paz. *O Mono Gramático*. Rio de Janeiro, Guanabara, 1974. cap. 8

A Keiko Shimabukuro

LOCAL E DATA: Campinas, 22/11/96.

BANCA EXAMINADORA

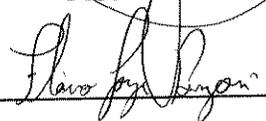
TITULARES

Prof. Dr. Carlos Alfredo Joly - Orientador:



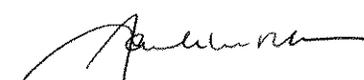
Assinatura

Dr. Flávio Jorge Ponzoni:



Assinatura

Dr. Jansle Vieira Rocha:



Assinatura

SUPLENTES

Dr. Jurandir Zullo Jr.:

Assinatura

APROVADA

Agradecimentos

Ao longo de todos estes anos, desenvolvendo este trabalho em uma área nova no âmbito do Instituto de Biologia da Universidade Estadual de Campinas (IB-UNICAMP), foi imprescindível contar com a colaboração inestimável de muitas professores, funcionários, colegas, profissionais, amigos, familiares e instituições, aos quais agradeço profundamente, destacando, entre outros:

- ao Dr. Carlos Alfredo Joly, Prof. do Departamento de Botânica do IB-UNICAMP, pelo convite para aventurar-me nas aplicações do sensoriamento remoto e dos sistemas de informações geográficas para a ecologia, pelo apoio a consecução da dissertação e a minha carreira profissional e pela amizade;
- ao Dr. Álvaro Penteado Crósta, Prof. do Departamento de Metalogênese e Geoquímica do Instituto de Geologia (IG) da UNICAMP, pela introdução e orientação geral na área de sensoriamento remoto, por franquear a utilização dos equipamentos e aplicativos do laboratório, pelos valiosos incrementos agregados a esse documento, pelo apoio inestimável a consecução da dissertação e a minha carreira profissional e pela amizade;
- ao Dr. Jansle Vieira Rocha, Prof. do Departamento de Planejamento e Produção Agropecuária da Faculdade de Engenharia Agrícola (FEAGRI) da UNICAMP, pela orientação na área de sensoriamento remoto, pelos valiosos incrementos agregados a esse documento e pela amizade;
- ao Dr. Flávio Jorge Ponzoni, Pesquisador Titular da Divisão de Sensoriamento Remoto do Instituto Nacional de Pesquisas Espaciais, pelos valiosos incrementos agregados a esse documento;
- ao Dr. Ardemírio Barros, Prof. do Departamento de Metalogênese e Geoquímica do IG-UNICAMP, pela orientação na área dos sistemas de informações geográficas e pela amizade;
- aos amigos Patrícia Cerqueira Leite do LAPIG-IG-UNICAMP, Marcelo Rossato do IB-UNICAMP, Moacir Américo Cornetti do IG-UNICAMP, Rubens Queiroz e Marcelo Araújo do Centro de Computação (CCUEC-UNICAMP), pelo apoio informático indispensável a consecução de uma dissertação dessa natureza;
- aos colegas de pós-graduação do IG-UNICAMP e amigos Henrique Roig; Heliux de Almeida Guimarães Filho e Edson Ribeiro, sobretudo, pelo companheirismo e pelas sugestões e discussões instigantes sobre o processamento digital de imagens orbitais;
- ao Júlio Botelho, Engenheiro da Prefeitura Municipal de Brotas, pela colaboração valiosa no trabalho de campo;
- ao Jay McCarthy, Eric Augenstein e Daryl Miller, equipe de suporte do ER-Mapper, pela ajuda na solução de dúvidas na utilização do aplicativo;
- ao IG-UNICAMP, pelo acesso aos recursos do Laboratório de Processamento de Informações Geo-referenciadas (LAPIG);

- aos colegas de trabalho e ao próprio Fundo Nacional do Meio Ambiente do Ministério do Meio Ambiente, da Amazônia Legal e Recursos Hídricos (FNMA-MMA), na pessoa de sua Coordenadora Geral Cecília Foloni Ferraz, pela compreensão com as dificuldades inerentes a fase de redação deste documento;
- ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq-MCT), pela concessão de bolsa de estudo e a Fundação de Apoio ao Ensino e à Pesquisa (FAEP-UNICAMP), pelo apoio financeiro;
- a Elza Kiyomi Shimabukuro Garcia, Marco Antônio Garcia Rossi e Filipe Shimabukuro Garcia, pelo apoio logístico e pela amizade;
- ao João Batista Negrão, ex-Secretário do Meio Ambiente da Prefeitura Municipal de Brotas, ao José Carlos de Francisco Júnior e Andréa Baltasar, todos ecologistas atuantes pela conservação do meio ambiente da bela região de Brotas, pelo apoio e pelas informações importantes para o trabalho de campo;
- a Prefeitura Municipal de Brotas, na pessoa de seu prefeito Orlando Pereira Barreto Neto, pelo apoio logístico no trabalho de campo;
- ao Ms. Rubens Lamparelli, funcionário da FEAGRI-UNICAMP, pelo empréstimo das fotografias aéreas da região de Brotas e pela introdução na área de Sensoriamento Remoto;
- ao Eduardo Góes Neves, pesquisador do Museu de Arqueologia e Etnologia da Universidade de São Paulo, pelo empréstimo do aparelho para acesso ao sistema de posicionamento global (GPS)
- a Ellen Cordeiro Perez Pombal e Fátima Maria de Souza funcionárias do Museu de História Natural do IB-UNICAMP, pela amizade;
- ao Dr. George John Sheperd, professor do Departamento de Botânica do IB-UNICAMP, por substituir meu orientador durante seu ano sabático;
- aos colegas do curso de pós-graduação em Biologia Vegetal do Departamento de Botânica do IB-UNICAMP, pela companhia ao longo da jornada;
- Ao Carlos Alberto de Mattos Scaramuzza, por uma miríade de detalhes ao longo da caminhada, mostrando sempre que

*A good world,
dew-drops fall,
by ones, by twos.*
Issa (1763-1827);

- a Valéria Parreira e ao Sérgio Crês; a Carla Barbosa, ao Carlos Alberto Contriciani, a Cibelly Gomes Nunes, por simplesmente e incondicionalmente serem amigos;
- a Laura Hiroko Shimabukuro, a Célia Misako Shimabukuro, a Elza Kiyomi Shimabukuro Garcia e a Lúcia Kazuko Shimabukuro, pelo apoio na minha formação acadêmica, pessoal e sobretudo pela amizade fraternal;
- a vida, por todas as suas cores.

Sumário

1. Resumo	<i>i</i>
2. Abstracts	<i>iii</i>
3. Introdução e Objetivo	<i>1</i>
4. Revisão bibliográfica	<i>5</i>
4.1 Histórico do sensoriamento remoto	<i>5</i>
4.2 Definição de sensoriamento remoto	<i>6</i>
4.3 Conceitos fundamentais do sensoriamento remoto	<i>6</i>
4.4 Plataformas orbitais e sensores - o programa Landsat	<i>9</i>
4.5 Processamento digital de imagens	<i>13</i>
5. Material	<i>58</i>
5.1 Definição da região de estudo	<i>58</i>
5.2 Coleta de dados	<i>59</i>
5.3 Caracterização da área de estudo	<i>61</i>
5.4 Infra-estrutura física	<i>80</i>
6. Métodos	<i>81</i>
6.1 Estrutura e sistemas para o processamento de imagens	<i>81</i>
6.2 Procedimentos metodológicos	<i>82</i>
6.3 Preparação das imagens	<i>82</i>
6.4 Classificação	<i>89</i>
6.5 Coleta dos dados de referência	<i>92</i>
6.6 Análise da exatidão	<i>96</i>
7. Resultados	<i>97</i>
7.1 Preparação das imagens	<i>97</i>
7.2 Classificação	<i>105</i>
7.3 Coleta dos dados de referência	<i>111</i>
7.4 Análise da exatidão	<i>112</i>
8. Discussão	<i>121</i>
8.1 Preparação das imagens	<i>121</i>
8.2 Classificação	<i>127</i>
8.3 Coleta dos dados de referência	<i>132</i>
8.4 Análise da exatidão	<i>134</i>
9. Conclusões e Recomendações	<i>138</i>
10. Bibliografia	<i>141</i>
11. Anexos	<i>167</i>
11.1 Rotina de execução do programa Lestr	<i>167</i>
11.2 Tabelas da análise da exatidão	<i>167</i>

Índice

1. Resumo	i
2. Abstracts	iii
3. Introdução e Objetivo	1
4. Revisão bibliográfica	5
4.1 Histórico do sensoriamento remoto	5
4.2 Definição de sensoriamento remoto	6
4.3 Conceitos fundamentais do sensoriamento remoto	6
4.4 Plataformas orbitais e sensores - o programa Landsat	9
4.5 Processamento digital de imagens	13
4.5.1 Georreferenciamento das imagens	13
4.5.2 Técnicas de realce de imagens	15
4.5.2.1 Seleção de bandas e composições coloridas	17
4.5.3 Classificação de imagens	20
4.5.3.1 Classificação multispectral	24
4.5.3.1.1 Treinamento para a classificação	25
4.5.3.1.2 Classificação não-supervisionada	30
4.5.3.1.3 Classificação supervisionada	34
4.5.3.1.3.1 Classificador máxima verossimilhança gaussiano	35
4.5.3.1.3.2 Classificador distância mínima	38
4.5.3.1.3.3 Classificador Mahalanobis	39
4.5.3.1.3.4 Classificador paralelepípedo	40
4.5.3.2 Avaliação da exatidão da classificação de imagens	41
4.5.3.2.1 Amostragem	44
4.5.3.2.1.1 Número e tamanho de amostra	45
4.5.3.2.1.2 Tipos de amostragem	47
4.5.3.2.2 Matriz de erro	50
4.5.3.2.3 Determinação da exatidão da classificação	51
4.5.3.2.3.1 Técnica descritiva	51
4.5.3.2.3.2 Técnica analítica	52
4.5.3.2.3.2.1 Normalização da matriz de erro	54
4.5.3.2.3.2.2 Modelo log-linear	54
4.5.3.2.3.2.3 Análise Kappa	55
5. Material	58
5.1 Definição da região de estudo	58
5.2 Coleta de dados	59
5.2.1 Aquisição de mapas	60
5.2.2 Aquisição de fotografias aéreas	60
5.2.3 Aquisição de imagens orbitais	61
5.3 Caracterização da área de estudo	61
5.3.1 Hidrografia	61
5.3.2 Geologia	62
5.3.2.1 Mesosóico	64
5.3.2.1.1 Grupo São Bento	64
5.3.2.1.1.1 Formação Pirambóia - TRJp	64
5.3.2.1.1.2 Formação Botucatu - JKb	65
5.3.2.1.1.3 Formação Serra Geral - JKsg e intrusivas básicas	65
5.3.2.1.2 Grupo Bauru	66
5.3.2.1.2.1 Formação Itaqueri - Kti	67
5.3.2.2 Cenozóicas	67
5.3.2.2.1 Aluviões (Qa)	67

5.3.3 Geomorfologia	67
5.3.4 Pedologia	69
5.3.5 Climatologia	70
5.3.6 Paisagens	71
5.4 Infra-estrutura física	80
6. Métodos	81
6.1 Estrutura e sistemas para o processamento de imagens	81
6.2 Procedimentos metodológicos	82
6.3 Preparação das imagens	82
6.3.1 Georreferenciamento das imagens	84
6.3.2 Seleção de bandas	84
6.3.3 Realce radiométrico.	85
6.3.4 Definição da área piloto	85
6.3.5 Reconhecimento da área de estudo	86
6.3.6 Legenda temática do uso das terras	87
6.4 Classificação	89
6.4.1 Seleção dos sítios de treinamento	89
6.4.2 Algoritmos supervisionados	90
6.4.3 Algoritmos não supervisionados	91
6.5 Coleta dos dados de referência	92
6.5.1 Plano de amostragem	93
6.5.2 Processamento dos dados de campo	95
6.6 Análise da exatidão	96
7. Resultados	97
7.1 Preparação das imagens	97
7.1.1 Seleção de bandas	97
7.1.2 Legenda temática do uso das terras	103
7.2 Classificação	105
7.3 Coleta dos dados de referência	111
7.4 Análise da exatidão	112
8. Discussão	121
8.1 Preparação das imagens	121
8.1.1 Seleção de bandas	121
8.1.2 Definição da área piloto	123
8.1.3 Padronização entre legenda temática do uso das terras e dados espectrais.	123
8.2 Classificação	127
8.2.1 Seleção das amostras de treinamento	127
8.2.2 Identificação dos padrões de uso das terras	128
8.2.3 Desempenho dos classificadores	130
8.3 Coleta dos dados de referência	132
8.4 Análise da exatidão	134
9. Conclusões e Recomendações	138
10. Bibliografia	141
11. Anexos	167
11.1 Rotina de execução do programa Lestr	167
11.2 Tabelas da análise da exatidão	167

<i>Figura 7-2. Composição colorida das bandas 5/1/3-RGB da imagem Landsat-TM, de 29/8/92, cobrindo a área de estudo na bacia do rio Jacaré Pepira, Brotas e Torrinha, SP.</i>	100
<i>Figura 7-3. Composição colorida das bandas 5/2/1-RGB da imagem Landsat-TM, de 29/8/92, cobrindo a área de estudo na bacia do rio Jacaré Pepira, Brotas e Torrinha, SP.</i>	101
<i>Figura 7-4. Composição colorida das bandas 4/5/3-RGB da imagem Landsat-TM, de 29/8/92, cobrindo a área de estudo na bacia do rio Jacaré Pepira, Brotas e Torrinha, SP.</i>	102
<i>Figura 7-5. Imagem Landsat-TM da área de estudo na bacia do rio Jacaré Pepira, Brotas e Torrinha, SP, classificada pelo algoritmo de distância mínima euclidiano.</i>	106
<i>Figura 7-6. Imagem Landsat-TM da área de estudo na bacia do rio Jacaré Pepira, Brotas e Torrinha, SP, classificada pelo algoritmo de distância mínima euclidiano baseado nos desvios padrões.</i>	107
<i>Figura 7-7. Imagem Landsat-TM da área de estudo na bacia do rio Jacaré Pepira, Brotas e Torrinha, SP, classificada pelo algoritmo de distância mínima de Mahalanobis.</i>	108
<i>Figura 7-8. Imagem Landsat-TM da área de estudo na bacia do rio Jacaré Pepira, Brotas e Torrinha, SP, classificada pelo algoritmo de máxima verossimilhança.</i>	109
<i>Figura 7-9. Imagem Landsat-TM da área de estudo na bacia do rio Jacaré Pepira, Brotas e Torrinha, SP, classificada pelo algoritmo isodata.</i>	110
<i>Figura 7-10. Mosaico das amostras do uso das terras coletadas na área de estudo na bacia do rio Jacaré Pepira, Brotas e Torrinha, SP.</i>	111
<i>Figura 7-11. Mosaico de células da imagem Landsat-TM classificada pelo algoritmo de distância mínima euclidiano, correspondentes aos pontos amostrais na área de estudo na bacia do rio Jacaré Pepira, Brotas e Torrinha, SP.</i>	112
<i>Figura 7-12. Mosaico de células da imagem Landsat-TM classificada pelo algoritmo de distância mínima euclidiano baseado nos desvios padrões, correspondentes aos pontos amostrais na área de estudo na bacia do rio Jacaré Pepira, Brotas e Torrinha, SP.</i>	113
<i>Figura 7-13 Mosaico de células da imagem Landsat-TM classificada pelo algoritmo de distância mínima de Mahalanobis, correspondentes aos pontos amostrais na área de estudo na bacia do rio Jacaré Pepira, Brotas e Torrinha, SP.</i>	114
<i>Figura 7-14. Mosaico de células da imagem Landsat-TM classificada pelo algoritmo de máxima verossimilhança, correspondentes aos pontos amostrais na área de estudo na bacia do rio Jacaré Pepira, Brotas e Torrinha, SP.</i>	115

Índice de tabelas

<i>Tabela 4-1. Caracterização esquemática dos sensores da série de satélites Landsat.</i>	11
<i>Tabela 4-2. Principais características e aplicações das bandas espectrais do sensor TM para o estudo da vegetação.</i>	12
<i>Tabela 5-1. Coluna litoestratigráfica simplificada da bacia do rio Jacaré-Pepira, São Paulo.</i>	63
<i>Tabela 6-1. Desenvolvimento da legenda temática para a classificação do uso das terras para a área de estudo na bacia do rio Jacaré Pepira, Brotas e Torrinha, SP.</i>	88
<i>Tabela 7-1. Dados estatísticos de bandas da imagem Landsat-TM, 29/8/92, órbita-ponto 220/75 e quadrante C.</i>	97
<i>Tabela 7-2. Matriz de correlação de bandas da imagem Landsat-TM, 29/8/92, órbita-ponto 220/75 e quadrante C.</i>	97
<i>Tabela 7-3. Fatores de índice ótimo (OIF) e a soma dos desvios padrões e dos valores absolutos dos coeficientes de correlação para as possíveis combinações entre as bandas 1, 2, 3, 4 e 5 da imagem Landsat-TM, 29/8/92, órbita-ponto 220/75 e quadrante C.</i>	98
<i>Tabela 7-4. Idade dos plantios de cana de açúcar segundo suas variedades e tipos de solos, nas fazendas cultivadas pela Usina Paraíso na área piloto (região de Brotas e Torrinha, SP) durante o período de aquisição da imagem orbital Landsat-TM em agosto de 1992.</i>	104
<i>Tabela 7-5. Sumário das análises de exatidão executadas para cada algoritmo de classificação.</i>	116
<i>Tabela 7-6. Matriz de erro entre os dados de referência terrestre e da classificação da imagem pelo algoritmo de máxima verossimilhança segundo a legenda temática nº 1.</i>	117
<i>Tabela 7-7. Índice de exatidão global da imagem classificada pelo algoritmo de máxima verossimilhança segundo a legenda temática nº 1.</i>	117
<i>Tabela 7-8. Índice Kappa global e individual entre os dados de referência terrestre e da classificação da imagem pelo algoritmo de máxima verossimilhança segundo a legenda temática nº 1.</i>	118
<i>Tabela 7-9. Matriz de erro entre os dados de referência terrestre e da classificação da imagem pelo algoritmo de máxima verossimilhança segundo a legenda temática nº 2.</i>	118
<i>Tabela 7-10. Índice de exatidão global da imagem classificada pelo algoritmo de máxima verossimilhança segundo a legenda temática nº 2.</i>	118
<i>Tabela 7-11. Índice Kappa global e individual entre os dados de referência terrestre e da classificação da imagem pelo algoritmo de máxima verossimilhança segundo a legenda temática nº 2.</i>	119
<i>Tabela 7-12. Matriz de erro entre os dados de referência terrestre e da classificação da imagem pelo algoritmo de máxima verossimilhança segundo uma redução da legenda temática nº 2.</i>	119
<i>Tabela 7-13. Índice de exatidão global da imagem classificada pelo algoritmo de máxima verossimilhança segundo uma redução da legenda temática nº 2.</i>	119
<i>Tabela 7-14. Índice Kappa global e individual entre os dados de referência terrestre e da classificação da imagem pelo algoritmo de máxima verossimilhança segundo uma redução da legenda temática nº 2.</i>	120
<i>Tabela 8-1. Síntese de aspectos selecionados de alguns trabalhos sobre o cálculo da exatidão da classificação do uso das terras a partir de imagens orbitais.</i>	125
<i>Tabela 8-2. Coeficientes de concordância Kappa global obtidos pelos classificadores Maxver; distância mínima de Mahalanobis, euclidiana simples e baseada em desvios padrões e isodata.</i>	127
<i>Tabela 8-3. Tabela dos coeficientes Kappa global e exatidão global com seus intervalos de confiança para os 5 classificadores, considerando as chaves de classificação de 7, 9 e 13 categorias.</i>	131

Tabela 11-1. Matriz de erro entre os dados de referência terrestre e da classificação da imagem pelo algoritmo de distância mínima euclidiana segundo a legenda temática nº 1. _____	168
Tabela 11-2. Índice de exatidão global da imagem classificada pelo algoritmo de distância mínima euclidiana segundo a legenda temática nº 1. _____	168
Tabela 11-3. Índice Kappa global e individual entre os dados de referência terrestre e da classificação da imagem pelo algoritmo da distância mínima euclidiana segundo a legenda temática nº 1. _____	169
Tabela 11-4. Matriz de erro entre os dados de referência terrestre e da classificação da imagem pelo algoritmo de distância mínima euclidiana segundo a legenda temática nº 2. _____	169
Tabela 11-5. Índice de exatidão global da imagem classificada pelo algoritmo de distância mínima euclidiana a legenda temática nº 2. _____	170
Tabela 11-6. Índice Kappa global e individual entre os dados de referência terrestre e da classificação da imagem pelo algoritmo de distância mínima euclidiana segundo a legenda temática nº 2. _____	170
Tabela 11-7. Matriz de erro entre os dados de referência terrestre e da classificação da imagem pelo algoritmo da distância mínima euclidiana baseada nos desvios padrões segundo a legenda temática nº 1. _____	171
Tabela 11-8. Índice de exatidão global da imagem classificada pelo algoritmo da distância mínima euclidiana baseada nos desvios padrões segundo a legenda temática nº 1. _____	171
Tabela 11-9. Índice Kappa global e individual entre os dados de referência terrestre e da classificação da imagem pelo algoritmo de distância mínima euclidiana baseada nos desvios padrões segundo a legenda temática nº 1. _____	172
Tabela 11-10. Matriz de erro entre os dados de referência terrestre e da classificação da imagem pelo algoritmo de distância mínima euclidiana baseada nos desvios padrões segundo a legenda temática nº 2. _____	172
Tabela 11-11. Índice de exatidão global da imagem classificada pelo algoritmo de distância mínima euclidiana baseada nos desvios padrões segundo a legenda temática nº 2. _____	173
Tabela 11-12. Índice Kappa global e individual entre os dados de referência terrestre e da classificação da imagem pelo algoritmo de distância mínima euclidiana baseada nos desvios padrões segundo a legenda temática nº 2. _____	173
Tabela 11-13. Matriz de erro entre os dados de referência terrestre e da classificação da imagem pelo algoritmo de distância mínima de Mahalanobis segundo a legenda temática nº 1. _____	174
Tabela 11-14. Índice de exatidão global da imagem classificada pelo algoritmo de distância mínima de Mahalanobis segundo a legenda temática nº 1. _____	174
Tabela 11-15. Índice Kappa global e individual entre os dados de referência terrestre e da classificação da imagem pelo algoritmo de distância mínima de Mahalanobis segundo a legenda temática nº 1. _____	175
Tabela 11-16. Matriz de erro entre os dados de referência terrestre e da classificação da imagem pelo algoritmo de distância mínima de Mahalanobis segundo a legenda temática nº 2. _____	175
Tabela 11-17. Índice de exatidão global da imagem classificada pelo algoritmo de distância mínima de Mahalanobis segundo a legenda temática nº 2. _____	175
Tabela 11-18. Índice Kappa global e individual entre os dados de referência terrestre e da classificação da imagem pelo algoritmo de distância mínima de Mahalanobis segundo a legenda temática nº 2. _____	176
Tabela 11-19. Matriz de erro entre os dados de referência terrestre e da classificação da imagem pelo algoritmo isodata. _____	176
Tabela 11-20. Índice de exatidão global da imagem classificada pelo algoritmo isodata. _____	176

1. Resumo

As atividades agrossilvipastoris na bacia do rio Jacaré Pepira, afluente do médio Tietê, tem provocado a destruição da maior parte dos ecossistemas originais. Estudos sobre a dinâmica dos remanescentes da vegetação são primordiais para a sua conservação, e envolvem obrigatoriamente o monitoramento espacial e temporal das mudanças da cobertura terrestre na região. Além dessa aplicação na gestão da biodiversidade, compreender a estrutura, a função e a evolução do uso das terras nas paisagens fornece subsídios fundamentais para o planejamento das atividades primárias e a avaliação das mudanças globais. Tais finalidades requerem ferramentas e métodos para produzir e interpretar dados diacrônicos e sincrônicos, como os oferecidos pelo sensoriamento remoto. Desde que as alterações antrópicas estão se processando com uma gravidade e velocidade bastante intensas é essencial que a aquisição de informações sobre os usos das terras, se torne cada vez mais automatizada e confiável.

Esta dissertação teve como objetivo básico avaliar a aplicação de alguns classificadores espectrais de imagens Landsat-TM no levantamento da cobertura e uso das terras em uma área na bacia do rio Jacaré-Pepira, município de Brotas e Torrinha, São Paulo. Foram selecionados 6 tipos de algoritmos de classificação de imagens: isodata, distância mínima euclidiana simples e baseada em desvios padrões, distância mínima de Mahalanobis, máxima verossimilhança e paralelepípedo. Foi adotada uma amostragem aleatória sistemática não alinhada. Os dados de referência terrestres foram adquiridos a partir de 46 amostras de 600 x 600 m, assegurando um mínimo de 2.7% da área total de estudo. Para obter a exatidão dos dados, foram geradas matrizes de erro e calculados os coeficientes de exatidão global e individual (erros de comissão e omissão) e os coeficientes de concordância Kappa global e individual.

Os resultados indicaram o desempenho diferencial dos classificadores, sendo o de máxima verossimilhança e o de distância mínima de Mahalanobis os que apresentaram os melhores resultados. Os processamentos das legendas temáticas com 13 e 9 classes, indicaram os eucaliptais, os corpos d'água e as pastagens como as classes que possuem os menores erros de comissão e maiores coeficientes de concordância Kappa, tendo como referência as imagens classificadas. Quanto aos erros de omissão, as classes canaviais, pastagens e eucaliptais, apresentaram os menores valores para as classificações envolvendo os algoritmos de máxima verossimilhança, distância mínima euclidiana e de Mahalanobis. Contudo, para o algoritmo de distância mínima euclidiana baseado em desvio padrão, a categoria eucaliptais apresentou um valor maior de erro em comparação a classe solo.

Os algoritmos clássicos de classificação da cobertura terrestre, empregados neste trabalho, apresentaram um desempenho insatisfatório, o que provavelmente não atenderia a demanda de usuários finais dos produtos cartográficos.

Como ainda não há procedimentos simples de classificação digital de imagens orbitais e cálculo de sua exatidão, com um desempenho ótimo para todas as classes ou para todas as áreas de uma região, esforços para aperfeiçoar e integrar as diferentes técnicas de classificação, deveriam ser desenvolvidos e documentados detalhadamente, em concomitância com análises de sua exatidão. Os testes comparativos entre as novas técnicas,

tendo as variáveis controladas e a confiabilidade das informações garantida, são fundamentais para determinar precisamente os pontos de estrangulamento dos procedimentos de classificação dos usos das terras, padronizar os métodos e conseqüentemente subsidiar muitos projetos de gestão ambiental.

2. Abstracts

The agriculture, husbandry and forestry activities in Jacaré Pepira river basin, a branch of the middle Tietê river, in the State of São Paulo, in Southeastern Brazil, have brought about increasing environmental impacts in many different natural resources. The mitigation of problems as physical resources contamination and exhaustion, biodiversity loss, climatic changes, etc., is complex and expensive. The different solutions for all these problems, global change questions and agriculture management involve monitoring the spatial and temporal dynamic of land cover and use. This monitoring requires remote sensing tools and methods to produce and to interpret diachronic and synchronic data. Since these environmental consequences of intensive agricultural production systems are serious and growing fast, it is essential that we get more and more automated and reliable land cover information by digital image processing systems.

The aim of this work is to evaluate the application of some spectral classifiers of Landsat-TM images for the land cover and land use survey in Brotas and Torrinha Counties. The image classification algorithms selected were maximum likelihood, Euclidean minimum distance, minimum distance of Mahalanobis, ISOCCLASS and parallelepiped. This study area was chosen because of its great landscape diversity. A grid with 2124 cell of 600 by 600 m was overlaid on the image. The ground truth was checked in 46 of these points (600 by 600 m) to ensure a 2,7% minimal sampling of the whole study area. It was used a not-aligned systematic random sampling. By way of assessing the image classification accuracy, it was used the error matrix and calculated commission error, omission error and Kappa coefficient.

The results show a differential performance of each classifier. Maximum likelihood and Mahalanobis distance shown the best results. The processing involving 13 and 9 land cover types indicated eucalyptus, water bodies and pasture like classes with minor commission error and major Kappa index of agreement. With regard to omission error, the sugar cane, pasture and eucalyptus shown the minor values to maximum likelihood, Euclidean minimum distance and Mahalanobis minimum distance classifiers. However, the eucalyptus showed major error value when compared with soil class in the Euclidean minimum distance algorithm based in standard deviation.

The classical algorithms of digital classification, used in this work, showed unsatisfactory performance. They can not solve the problems of final users of cartographic products, as some others studies indicated.

As there are not a unique procedure of classification and assessment of accuracy yet, with a good performance to all classes and regions, the efforts to develop and integrate different techniques of classification and the accuracy analyses must be stimulated and documented in detail. The comparative tests between new techniques, with controlled variables and reliable land cover information, are fundamental to determine exactly the "bottlenecks" of classification procedures in complex land use, standardize methods and subsidize some environmental management projects.

3. Introdução e Objetivo

“Os satélites de observação da Terra, sobretudo uma nova geração a ser desenvolvida em breve, controlam a saúde do meio ambiente global: o efeito estufa, a erosão da camada superior do solo, a diminuição da camada de ozônio, as correntes dos oceanos, a chuva ácida, os efeitos das enchentes e secas e perigos ainda desconhecidos.”

Carl Sagan, *Pálido ponto azul: o futuro do homem no espaço*, São Paulo, 1994.

A exploração da terra é o sustentáculo da população em boa parte do mundo, especialmente no Brasil, onde muitas vezes tem possibilitado o incremento do padrão de qualidade de vida e da situação sócio-econômica. Porém na maioria das áreas, o manejo dos sistemas de produção agrossilvipastoris apresenta muitas deficiências, caracterizando-se por uma baixa produtividade e impactos negativos no uso dos recursos naturais.

Essas atividades primárias respondem por um componente significativo da dinâmica da cobertura terrestre do planeta, tornando-as uma questão central em um grande número de processos biofísicos associados com as mudanças ambientais globais, formando um sistema complexo e interativo onde as intervenções humanas são conectadas a funções ecológicas em um circuito de retroalimentação (Turner et al., 1995).

Esse contexto explica o fato de que, quaisquer que sejam as idiosincrasias de cada situação ou problema, o diagnóstico sobre a cobertura e o uso das terras é um item fundamental na agenda político-ambiental e técnico-científica. Por exemplo, diversos problemas ecológicos são enfrentados a partir de dados sobre a ocupação dos solos, abrangendo temas como avaliação de impactos ambientais, ciclagem de nutrientes, dinâmica de populações animais, estrutura e evolução de padrões e processos das paisagens, mudança da cobertura e uso das terras; balanço de energia global etc.

Novas tecnologias tem sido utilizadas para auxiliar e superar as dificuldades da coleta dessas informações realizadas por métodos tradicionais, como por exemplo o mapeamento de campo, com sua demanda elevada de tempo, recursos materiais e humanos ou como o estabelecimento de comparações diacrônicas ou sincrônicas com a falta de uniformidade entre os sistemas de classificação empregados (Mather, 1986).

Nesse sentido, o estudo e o monitoramento de questões como a da cobertura e uso das terras e a dos impactos ambientais estão se beneficiando com o uso de técnicas de sensoriamento remoto (SR) e de sistemas de informações geográficas (SIG). Sua disseminação tem tornado cada vez mais exequíveis tarefas como coletar informações,

avaliar e monitorar as condições ambientais de uma área e suas mudanças ao longo de séries temporais, lidar com um grande volume de dados referenciados espacialmente e, principalmente, contribuir para gerenciar eficazmente a exploração da paisagem rural.

Essa ênfase no papel do SR também tem sido verificada na ciência ecológica (Wickland, 1989 e 1991), sendo objeto de destaque na “Sustainable Biosphere Initiative”, uma agenda de pesquisas organizada pela “Ecological Society of America” com o objetivo de direcionar a ciência ecológica no sentido de contemplar melhor o manejo dos recursos terrestres (Lubchenco et al., 1991). Soluções para problemáticas como mudança global, perda de diversidade biológica, conservação de recursos naturais ou desenvolvimento de agroecossistemas sustentáveis implicam na integração dos esforços de pesquisa das subdisciplinas da ecologia acompanhada pela aplicação de novas ferramentas e abordagens como a teledetecção espacial (Matson & Ustin, 1991).

Nos trópicos, o sensoriamento remoto tem sido usado amplamente nas últimas décadas para o manejo de recursos renováveis, com contribuições na classificação, avaliação e mapeamento dos recursos da Terra e no monitoramento da cobertura da vegetação e culturas.

Entretanto, o fato do SR ser uma técnica com origem nos países desenvolvidos, seu sucesso nos trópicos depende em grande parte da eficiência da transferência de tecnologia e do nível de sua adaptação para atender as condições ambientais e culturais da área. Uma solução apropriada não vem geralmente na forma de pacotes fechados, mas sim no desenvolvimento de ferramentas abrangentes e adequadas para os problemas reais e o contexto existente.

Ciente dessa questão, o Brasil tem investido na utilização de imagens aerotransportados e orbitais para o levantamento e monitoramento dos recursos naturais desde a década de 1970 por ocasião do Projeto RADAMBRASIL e do Programa de Sensoriamento Remoto por Satélite (Barbosa & Cunha, 1987). Especificamente as aplicações para fins agrícolas surgem em 1975 (Dallemand et al., 1988). Passados aproximadamente 20 anos, já existe uma comunidade florescente de usuários desses dados e instrumentos que auxiliam no desenvolvimento de inúmeras tecnologias, técnicas, métodos e aplicações, conforme atestam os anais de alguns dos diversos congressos nacionais e internacionais dos últimos anos (Simpósio Brasileiro de Geoprocessamento, 1990; International Symposium on Primary Data Acquisition, 1990; Simpósio Brasileiro de Sensoriamento Remoto, 1990; Simpósio Brasileiro de Sensoriamento Remoto, 1993; Symposium on Resource and Environmental Monitoring, 1994 e Simpósio Brasileiro de Sensoriamento Remoto, 1996).

Quaisquer que sejam as aplicações, o sensoriamento remoto envolve a manipulação e interpretação de imagens direcionadas a coleta de informações no âmbito de uma determinada disciplina da ciência. Há várias técnicas de processamento, analógicas ou digitais, associadas sempre a implantação de um sistema de classificação de um tema específico, relacionando cada elemento da imagem a um objeto real descrito por um rótulo.

Contudo, muitos analistas e planejadores ambientais, frente as potencialidades do sensoriamento remoto, têm utilizado-o de forma indiscriminada, não considerando os limites e a adequabilidade de tais ferramentas para o objeto de interesse e, conseqüentemente, incorrendo em erros.

A maioria dos trabalhos de classificação da cobertura e do uso das terras, implantados por instituições brasileiras, que fazem uso de dados orbitais, os analisam por procedimentos analógicos ou classificadores espectrais padrões (principalmente paralelepípedo, distância mínima e máxima verossimilhança). A grande parte dessas iniciativas não incluem cálculo e avaliação da exatidão dos dados e informações produzidas. Assim como poucas inovações referentes a classificação digital dos usos das terras são implementadas rapidamente nos pacotes de processamento digital de imagem consagrados, a avaliação do erro das informações geradas, apesar de sua importância, está sendo subestimada por muitos administradores nos estudos e manejo de recursos ambientais. Como o custo de programas de controle e monitoramento ambiental são muitas vezes bastante elevados, a falta de precisão acarreta severas inconsistências nas bases de dados, gerando despesas inúteis e dificuldades técnicas muitas vezes responsáveis pelo fracasso desses programas.

Portanto, essa confiabilidade nos dados é fundamental para promover ajustes no desenvolvimento dos projetos de classificação do uso das terras, garantir uma integração do plano de informação em uma base de dados geocodificados e subsidiar adequadamente a gestão dos recursos naturais.

Além disso, a análise de exatidão permite ao gerente de projetos conhecer o potencial e as limitações das técnicas e procedimentos de geoprocessamento e selecioná-las segundo as especificações de cada aplicação particular. Paralelamente, os analistas passam a conhecer a natureza dos erros, possibilitando o aperfeiçoamento do processo de classificação e mapeamento. Esse destaque tem sido amplificado por iniciativas recentes, tratando da precisão de dados espaciais, principalmente no seu aspecto de integração dos dados do SR com os sistemas de informações geográficas (SIG), reavivando a discussão sobre como aperfeiçoar e aferir a acurácia das classificações executadas (Chrisman, 1989; Chrisman, 1991; Goodchild et al., 1994 e Janssen & van der Wel, 1994).

Essa dissertação versa sob a determinação da exatidão das classificações digitais, avaliando as principais técnicas utilizadas nos inventários da cobertura terrestre executados no Brasil.

Diante da pluralidade de situações agroecológicas existentes ainda não há uma série consagrada de procedimentos eficientes e simples para a classificação digital de imagens orbitais e cálculo de sua exatidão. Obter um desempenho ótimo para todas as classes ou para todas as áreas de uma região é um desafio a ser superado com muitos esforços para aperfeiçoar e integrar diferentes técnicas de classificação e as respectivas análises de sua exatidão. Os testes comparativos entre novas técnicas, em um contexto de variáveis controladas e confiabilidade das informações, é um caminho produtivo para determinar precisamente os pontos críticos nas condutas de classificação dos usos das terras, particularmente em áreas complexas; na padronização de métodos e na

incorporação de dados sólidos sobre o uso das terras no cotidiano de projetos sobre avaliação e monitoramento de impactos ambientais, planejamento agrícola e mudança globais.

Nesse contexto, o **OBJETIVO GERAL** dessa dissertação foi avaliar a aplicação de alguns classificadores espectrais de imagens Landsat-TM no levantamento da cobertura e uso das terras em uma área na bacia do rio Jacaré-Pepira, município de Brotas e Torrinha, São Paulo.

Os **OBJETIVOS ESPECÍFICOS** foram:

1. executar todas as etapas do processo de classificação da cobertura e uso das terras, identificando, categorizando, qualificando e mapeando as séries elegidas.
2. verificar a exatidão de alguns classificadores espectrais de imagens digitais na avaliação dos padrões de cobertura terrestre, colaborando para a definição de técnicas e métodos adequados para o reconhecimento das classes de uso das terras.

A consecução destes objetivos foi idealizada e efetivada através das seguintes metas:

1. identificar e caracterizar preliminarmente as principais classes de uso das terras existentes;
2. classificar e descrever o uso das terras através de processamento digital de imagens do sensor Thematic Mapper (TM) do satélite LANDSAT 5;
3. coletar dados de referência terrestre na área de estudo através de uma estratégia amostral;
4. verificar a exatidão dos classificadores espectrais;
5. reavaliar o uso das terras segundo as categorias de uma legenda temática consolidada.

4. Revisão bibliográfica

“There’s no escape from the cartographic paradox: to present a useful and truthful picture, an accurate map must tell white lies.”

Mark Monmonier, *How to lie with maps*, Chicago, 1991.

A necessidade do ser humano de ler o seu entorno, coletando cada vez mais informações com maior precisão sobre a relação de suas atividades sócio-econômicas com a natureza tem motivado o desenvolvimento contínuo do sensoriamento remoto e de suas ferramentas para mapear e monitorar os recursos naturais segundo perspectivas cada vez mais amplas.

Cada avanço ou inovação tecnológica tem confirmado as predições, formuladas por Estes et al. (1983), sobre tendências futuras na área do sensoriamento remoto como: a disseminação dos sistemas sensores radar e das imagens de alta resolução e ampla cobertura; a aplicação de abordagens da inteligência artificial nos processos de reconhecimento de padrões digitais; a expansão do uso de informações suborbitais e orbitais para a modelagem ambiental com o auxílio dos sistemas de informações georreferenciadas e para o planejamento, ordenamento e regulamentação territorial; a implantação de bases de dados geocodificados de escala nacional ou global baseado em dados orbitais e principalmente as previsões sobre as alterações nas percepções, paradigmas, integração interdisciplinar, métodos de análise e modelos científicos nas diferentes disciplinas advindas do uso da teledetecção espacial. Diversas áreas do conhecimento, como a geografia, geomorfologia, pedologia, geologia, meteorologia, oceanografia, hidrologia, ecologia, agricultura, engenharia florestal e civil, planejamento urbano e regional, conservação ambiental e arqueologia, tem sido transformadas pela possibilidade de discernir e monitorar sincrônica e diacronicamente a estrutura e a dinâmica de diferentes fenômenos e sistemas (Curran, 1985).

Face a amplitude do crescimento e estado atual da disciplina, esta revisão é circunstanciada ao escopo de estabelecer uma visão geral e panorâmica dos aspectos associados com o problema focalizado por esta dissertação.

4.1 Histórico do sensoriamento remoto

O sensoriamento remoto (SR), inicia-se com Gaspard Felix Tournachon fotografando, a partir de um balão, Bievre na França em 1858 (Lillesand & Kiefer, 1994). A partir daí o emprego das fotografias aéreas foi gradativamente aumentando, iniciando um processo de expansão técnico-científico da disciplina então conhecida por fotogrametria, um termo proposto em 1867 (Zhizhuo, 1996).

Segundo Colwell (1983), esta primeira fase da história do SR perdura até 1960. A partir daí, em um segundo momento, a interpretação visual de fotografias pancromáticas passa a ser acompanhada pelas pesquisas exploratórias sobre novos sensores a bordo de aviões e satélites. O advento da era espacial, característico daquela década, promoveu uma rápida aceleração do desenvolvimento tecnológico, ampliando a variedade de sistemas sensores remotos multiespectrais e a capacidade de gerar dados digitais além dos analógicos. As inúmeras possibilidades destas novas imagens transformaram tão intensamente a estrutura da disciplina fotogramétrica, a ponto de culminar na necessidade de revisão de sua denominação.

Foi neste contexto que a geógrafa norte-americana Pruitt utilizou pela primeira vez o termo “sensoriamento remoto” no início da década de 1960 em um relatório sobre interpretação de fotografias aéreas para “American Naval Research Office” (Fischer, 1975 e Zhizhuo, 1996).

Nos anos 70, a tendência de incorporação de procedimentos quantitativos, presentes em muitas disciplinas, como é o caso das ciências ambientais, passa a afetar notavelmente o SR. Esta preocupação crescente com a adoção de tratamentos numéricos pôde ser evidenciada pela expansão do uso do processamento quantitativo de imagens baseados em computadores para a análise de produtos de sensores orbitais sob a forma digital (Swain & Davis, 1978 e Campbell, 1996). Exemplifica bem esta mudança de perfil, o aumento da demanda de imagens digitais Landsat em relação às analógicas, passando da razão de 8.000:1 em 1973 (1 no universo de 8.000 imagens vendidas) para 39:1 em 1978 (Curran, 1985).

4.2 Definição de sensoriamento remoto

Em sua acepção original, sensoriamento remoto referia-se simplesmente à observação e à medida de um objeto sem tocá-lo. Com a expansão tecnológica, o conceito passou a abranger toda uma disciplina, sendo aplicado com significados variados pelos pesquisadores da área. Alguns o definem como sendo a técnica de coletar informações a distância (Aronoff, 1989). Outros como uma ciência ou arte de obter informações sobre um objeto, área ou fenômeno através da análise de dados adquiridos por um equipamento que não está em contato com o objeto, área ou fenômeno sobre investigação (Lillesand & Kiefer, 1994). Novo (1989) apresenta uma definição mais detalhada, a qual foi adotada nesta dissertação - “Sensoriamento remoto é a técnica de coleta de informação a distância, com o objetivo de estudar o ambiente terrestre através do registro e da análise das interações entre a radiação eletromagnética e as substâncias componentes da Terra em suas mais diversas manifestações, utilizando de sensores, equipamento para processamento e transmissão de dados, aeronaves etc.”

4.3 Conceitos fundamentais do sensoriamento remoto

Segundo Novo (1989) o sensoriamento remoto opera com ondas eletromagnéticas como a luz visível, infravermelho, calor, microondas, ondas ultravioleta, raio X etc. Essas ondas distribuem-se em um gradiente crescente segundo seus comprimentos de onda e decrescente segundo a frequência de cada radiação, denominado espectro eletromagnético.

A utilização dessas ondas como ferramenta para a compreensão e monitoramento da estrutura e da dinâmica de diversos tipos de fenômenos envolve dois processos básicos: a coleta e análise de dados.

A aquisição dos dados ocorre através de um sensor de ondas eletromagnéticas que pode estar embarcado em uma plataforma aérea ou orbital como no caso dos satélites artificiais. Este processo de aquisição das imagens baseia-se na emissão e reflexão de energia eletromagnética pelos elementos da superfície terrestre e pode ser esquematizado nas seguintes fases: a geração da energia eletromagnética, a propagação da radiação através da atmosfera, a interação energética com os elementos da superfície terrestre, a retransmissão da energia através da atmosfera e a coleta dos dados nos sensores e plataformas (Lillesand & Kiefer, 1994).

O pressuposto para esta detecção de padrões está no fato de que as ondas de cada trecho do espectro eletromagnético interagem distintamente com os materiais da superfície e atmosfera terrestres, dependendo das propriedades elétricas, magnéticas e estruturais de seus elementos, resultando em diferentes proporções de energia refletida, absorvida e transmitida. Este padrão de resposta espectral próprio de cada elemento é denominado assinatura espectral e permite o reconhecimento e identificação de padrões nas imagens (Mather, 1987).

Deste modo, as diferentes intensidades de energia registradas por um sensor são o resultado da integração da energia refletida e/ou emitida pelos diferentes tipos de alvos, em determinadas bandas ou faixas espectrais, definidas como intervalos específicos de comprimento de onda nos quais as características técnicas do sensor possibilitam a coleta da resposta espectral (Silva, 1978).

Imageadores multiespectrais são os responsáveis pela detecção da energia eletromagnética, transformando os sinais luminosos, correspondentes às variações de energia da cena original, primeiro em sinais elétricos e depois em dados numéricos transmitidos para estações de captação espalhadas ao redor do planeta e armazenados em fitas magnéticas. Os sinais registrados, através de detectores sensíveis às diferentes bandas do espectro eletromagnético, produzem diversas imagens multiespectrais para uma mesma área, cada uma delas referente à emissão ou reflexão da cena em uma das bandas (faixas de sensibilidade) do detector.

A análise das informações geradas e o conseqüente diagnóstico dos recursos ou processos sob investigação envolve dois procedimentos básicos: analógico e digital. Em ambos o exame dos dados requer vários aparelhos de interpretação e visualização de dados fotográficos e/ou computadores para processamento de dados digitais.

No analógico, a imagem é fornecida sob a forma de um produto análogo à uma fotografia convencional. Embora seu custo seja semelhante ao formato digital, sua aquisição só pode ser feita de acordo com uma escala cartográfica e uma composição de bandas e coloração predefinidas. As informações são adquiridas através de interpretação visual da imagem e delineamento de áreas em uma transparência sobreposta. Após a extração, as linhas geradas podem ser editadas por métodos cartográficos tradicionais em um mapa ou digitalizadas via mesa digitalizadora ou leitor óptico ("scanner") e manipuladas por um

sistema de informações geográficas para gerar um plano de informação digital e um mapa temático.

Na forma digital, uma imagem é um conjunto de elementos discretos de informação, representada numericamente sob a forma de uma matriz de células denominadas pixels (um acrônimo para "picture element"). Deste modo, a estrutura da imagem é definida pelos pixels e seus atributos:

- posicional, a localização expressa por um sistema de coordenadas e
- numérico, a média da intensidade de energia refletida ou emitida pelos diferentes materiais presentes no pixel e expressa por um número digital ou ND.

Este parâmetro numérico também pode ser compreendido como uma terceira coordenada Z e nos sensores de última geração varia entre 0 a 255. Esses NDs representam a resposta espectral dos diferentes objetos de uma região, podendo ser visualizados através de uma associação com diferentes tonalidades de cinza e analisados por fórmulas matemáticas implantadas em algoritmos computacionais (Crosta, 1992).

As informações contidas nos pixels têm natureza distinta de acordo com os diferentes satélites e sistemas sensores e são circunstanciadas pelos parâmetros de:

- resolução espacial - capacidade do sistema sensor em visualizar e distinguir um objeto na superfície terrestre e pode ser entendido de forma simplificada como o tamanho do pixel;
- resolução espectral - capacidade de discriminação de diferenças espectrais definida pela largura do intervalo de comprimento de onda detectado em cada banda espectral do sistema sensor e pelo número de bandas coberto pelo sensor e
- resolução radiométrica - capacidade de discriminação de diferenças de radiância, representada pelo número de níveis de cinza adotados pelo sensor para exprimir a radiância refletida ou emitida pelos alvos (Crósta, 1992 e Lillesand & Kiefer, 1994)

A imagem composta pelos pixels pode ser visualizada através de uma única banda em modo monocromático ou através da composição de 3 bandas em uma representação falsa cor, onde para cada uma delas é atribuída uma das três cores primárias (vermelho, verde e azul).

A análise digital ou analógica das imagens requer a discriminação de padrões na resposta espectral dos diferentes alvos. Como os elementos respondem variadamente ao longo dos diferentes comprimentos de ondas, o comportamento da assinatura espectral de dois objetos distintos pode ser semelhante em uma determinada faixa ou completamente dispare em outro trecho do espectro eletromagnético. Deste modo, a interpretação de uma imagem envolve a análise das curvas de reflectância espectral, ou seja um gráfico relacionando a reflectância espectral de diferentes alvos, como vegetação, solo, água etc., com um gradiente de comprimentos de onda. A configuração dessa curva aponta as características espectrais do objeto e fornece indicações para a identificação, qualificação e quantificação dos elementos constituintes (Landgrebe, 1978).

Dados de referência sobre os recursos ou processos em estudo como localização, extensão e caracterização também são utilizados quando disponíveis para auxiliar na análise dos dados. Estas informações são geralmente compiladas em formas de tabelas ou mapas, e/ou articuladas a outros planos de informações georreferenciadas através de sistemas de informações geográficas, constituindo elementos decisivos para apoiar os processos de decisão dos usuários (Lillesand & Kiefer, 1994).

4.4 Plataformas orbitais e sensores - o programa Landsat

Na década de 60, o leque de possibilidades oferecido pelas plataformas espaciais estimulou pesquisadores e principalmente, militares a se direcionarem no desenvolvimento das áreas da óptica, eletrônica, telecomunicações e tratamento de informações. Esta articulação entre os domínios do sensoriamento remoto e o da pesquisa espacial tem suscitado até hoje o florescimento de toda uma nova área do conhecimento e da tecnologia com uma gama de aplicações teóricas e práticas sempre crescente. Atualmente existem satélites e sensores para múltiplas finalidades dentre as quais destacam-se:

- as oceanográficas, como por exemplo: o Nimbus 7, a série “Marine Observation Satellite” (MOS-1 e MOS-1b) (Osawa et al., 1995); o SEASAT-1; o “Indian Remote Sensing Satellite” (IRS-P4), cujo lançamento está previsto para o final de 1996 (Chandrasekhar, 1995);
- as meteorológicas, como por exemplo: a série “National Oceanic and Atmospheric Administration” (NOAA), a série “Geostationary Operational Environmental Satellites” (GOES), a série “Defense Meteorological Satellite Program” (DMSP), METEOSAT etc.;
- as de análise de recursos naturais, como por exemplo: o “Sattelite Probatoire d'Observation de la Terre” (SPOT), o “Indian Remote Sensing Satellite” (IRS-1A; IRS-1B; IRS-P2, IRC-1C e IRC-P3); o Almaz-1 e o Resours-02 do governo russo, o “European Remote Sensing Satellite” (ERS-1 e 2), o RADARSAT, o “Japanese Earth Resources Satellite” (JERS-1), o “Advanced Earth Observing Satellite” (ADEOS), cujo lançamento está previsto para este ano (Higashi, 1995); o “China Brazil Earth Resource Satellite (CBERS), cujo lançamento está previsto para 1997, o “Advanced Land Observing Satellite” (ALOS) (Osawa et al., 1995) e pelo menos 16 outras plataformas com início de operação planejado entre 1996 e 2004 (Thibault, 1995).

Dentre todos os sistemas orbitais de sensoriamento remoto hoje em operação, o sistema Landsat da NASA, tem sido até o momento o mais importante e o mais utilizado para as condições brasileiras, principalmente devido as facilidades físicas e orçamentárias para a aquisição dos produtos proporcionadas pelo sistema de recepção de imagens de Cuiabá, MT, operacional desde 1973, de processamento eletrônico e fotográfico de imagens de Cachoeira Paulista, SP, e de distribuição nacional de imagens (Novo, 1989).

Estimulado pelo potencial das imagens orbitais, antevisto pelos produtos gerados nos primeiros satélites meteorológicos e nas primeiras missões tripuladas, a NASA com a colaboração do Departamento do Interior iniciou um programa de satélites de recursos naturais em 1967, denominado inicialmente “Earth Resources Technology Satellites” - ERTS e posteriormente renomeado retroativamente como série Landsat em 1975. O

primeiro satélite foi lançado em 23 de julho de 1972, à título probatório, para testar a viabilidade de gerar imagens sobre recursos naturais a partir de plataformas orbitais automáticas. Seguindo o sucesso dos experimentos com os dados do Landsat-1 devidos principalmente a adoção do princípio de “open skies”, *id est* acesso mundial indiscriminado, foram sendo lançados mais 4 satélites da série, deflagrando o grande impulso para a aplicação da teledetecção espacial para o planejamento e gerenciamento dos recursos naturais. Os Landsat-2 (lançamento em 22/1/75), Landsat-3 (5/5/78), Landsat-4 (16/7/82) e Landsat-5 (1/3/84) foram suprindo a perda natural de eficiência dos sistemas sensores dos satélites precedentes da série e incorporando novos avanços tecnológicos tanto na plataforma como nos sensores (Lillesand & Kiefer, 1994).

A partir do LANDSAT-4, foi adicionado ao principal sensor, o “Multispectral Scanner” (MSS), um novo e aperfeiçoado sistema de coleta de informação espectral, o “Thematic Mapper” (TM), ambos descritos na Tabela 4-1. Tanto o MSS como o TM são sensores passivos pois sua operação requer uma fonte de radiação externa para viabilizar o registro da reflexão ou emissão de radiação eletromagnética por alvos da superfície terrestre. Ambos também trabalham com trechos do espectro eletromagnéticos cuja transmissão atmosférica é totalmente comprometida pela presença de nuvens.

O Landsat-5, atualmente em funcionamento, está posicionado em uma altitude de 705 km, mais baixa do que a de 900 km seus predecessores 1, 2 e 3. Sua órbita é repetitiva, circular, heliossincrônica (ângulo sol-terra-satélite constante em 37,5°), quase polar e inclinada de 98, 2° em relação ao Equador. Deslocando de norte para sul, na parte iluminada da Terra, a plataforma cruza o Equador a cada aproximadamente 99 minutos, em torno das 9:45 h no horário solar local, completando 14,5 órbitas por dia. Em períodos de 16 dias, todo o globo terrestre é coberto, em faixa contínua de 185 km de largura, posicionadas 2752 km uma das outras.

O sensor TM foi criado inicialmente para a discriminação da vegetação, com bandas espectrais selecionadas para distinguir características das suas respostas espectrais (Salomonson et al. 1980), mas com outras aplicações nas áreas hidrológica, geológica, mineralógica, pedológica etc. (vide Tabela 4-2).

No caso específico da utilização das imagens TM para a fitoecologia, fenômenos físicos discretos presentes nas interações de radiação eletromagnética com o complexo solo-planta conferem às curvas de reflectância padrões típicos com picos e vales associados as características do alvo imageado. A variação da magnitude das formas das curvas fornece as bases para a discriminação de dados sobre a cobertura vegetal. Enquanto as respostas espectrais nas bandas do visível (1, 2 e 3) são primariamente controladas pelos pigmentos da planta (vide Tabela 4-2), a reflectância é regida: no infravermelho próximo principalmente pela estrutura física do mesófilo foliar (banda 4); no infravermelho médio pelo conteúdo hídrico do alvo (banda 5 e 7) e no infravermelho termal quase completamente pela emissão associada a alterações térmicas (banda 6) (Townshend, 1984).

Tabela 4-1. Caracterização esquemática dos sensores da série de satélites Landsat.

Características	MSS	TM
Missões	1 a 5	4 e 5
Resolução espectral (banda / μm)		
azul		1 / 0.45-0.52
verde	4 (1) ¹ / 0.5-0.6	2 / 0.52-0.60
vermelho	5 (2) ¹ / 0.6-0.7	3 / 0.63-0.69
infravermelho próximo	6 (3) ¹ / 0.7-0.8	4 / 0.76-0.90
infravermelho próximo	7 (4) ¹ / 0.8-1.1	
infravermelho médio		5 / 1.55-1.75
infravermelho médio		7 / 2.08-2.35
termal		6 / 10.4-12.5
Resolução espacial (m)	79/82 ² x 79/82 ²	30/120 ³ x 30/120 ³
Resol. radiométrica (níveis de cinza)	128	256
Resolução temporal (dias)	16	16
Área coberta (km)	185 x 185	185 x 185
nº de pixels/cena	7.722.000	38.020.000

¹ nova designação das bandas no Landsat 4 e 5² 79 m no Landsat 1 a 3; 82 m no Landsat 4 e 5³ 30 m nas bandas 1 a 5 e 7; 120 m na banda 6

Embora a imagem produzida pelos instrumentos do LANDSAT não substitua a precisão em termos de resolução espacial das foto aéreas, suas características a tornam adequada para levantamentos de informações e análise de problemas nas escalas cartográficas de 1:50.000 e em certos casos até mesmo de 1:25.000.

Tabela 4-2. Principais características e aplicações das bandas espectrais do sensor TM para o estudo da vegetação.

<p>1 - azul (comprimento de onda 0,45-0,52µm)</p> <ul style="list-style-type: none"> • A capacidade de penetração na água das ondas com este comprimento permite a avaliação de sedimentos em suspensão nos corpos d'água e o mapeamento de regiões costeiras e batimétrico. • Como o pigmento clorofila presente nas folhas verdes geralmente absorve 70 a 90 % da luz na região azul do espectro eletromagnético, com um pico de absorção situado ao redor do 0,45 (Mather, 1987), os dados dessa banda são úteis para a discriminação e a cartografia de tipos florestais, a diferenciação entre solo e vegetação e a identificação de características das culturas como perturbações fisiológicas ou deciduidade (Lillesand & Kiefer, 1994).
<p>2 - verde (comprimento de onda 0,52 - 0,60µm)</p> <ul style="list-style-type: none"> • Como os pigmentos presentes nas plantas (clorofila, carotenóides e antocianina) absorvem muito pouco na região verde (Hoffer, 1987), a especificação desta banda foi elaborada para medir este pico de reflectância da vegetação, habilitando-a para a identificação de fisionomias vegetais, de seu vigor fenológico e de características das culturas (Lillesand & Kiefer, 1994).
<p>3 - vermelho (0,63-0,69µm)</p> <ul style="list-style-type: none"> • A clorofila usualmente absorve de 70 a 90% da luz da região vermelha do espectro, em torno de 0,65µm (Mather, 1987). • O processo de senescência foliar ou mesmo qualquer distúrbio fisiológico resulta em um aumento na reflectância da luz visível, principalmente na região do vermelho, devido a degradação da clorofila ou perda de sua eficiência fotossintética (Hoffer, 1978). • O fato do intervalo desta banda do TM estar dirigido para esta região de elevada absorção do pigmento clorofila, permite sua utilização para diferenciação de espécies vegetais, bem como a identificação de características das culturas (Lillesand & Kiefer, 1994).
<p>4- infravermelho próximo (0,76 - 0,90 µm)</p> <ul style="list-style-type: none"> • A região do infravermelho é o melhor segmento do espectro para distinguir plantas e suas condições fisiológicas. O coeficiente de reflectância eleva-se muito no intervalo de 0,70 a 1,35µm em razão da falta de absorção destes comprimentos de onda tanto por pigmentos como pela água líquida. As mudanças na reflectância neste tipo de radiação são controladas pela estrutura interna das folhas, associando-se à variação no tamanho e forma das células e dos espaços intercelulares. • A reflectância do infravermelho também é sensível a alterações metabólicas no mesófilo foliar, apresentando valores maiores em folhas saudáveis e menores nas senescentes ou acometidas por uma fitopatologia (Hoffer, 1987). • Dentro da amplitude desta banda, os solos refletem bem menos que a vegetação verde. Porém, entre 1,3 e 2,2µm, esta relação inverte-se e a cobertura pedológica passa a ter uma reflectância maior. A combinação deste comportamento diferencial torna estes comprimentos de onda valiosos para distinguir a vegetação do solo nu e estimar a fitodensidade ou fitocobertura. • Estes atributos tomam esta banda adequada para a classificação da vegetação, para avaliação do vigor fenológica, da biomassa e da umidade edáfica e para a delimitação de corpos d'água (Lillesand & Kiefer, 1994).
<p>5 - infravermelho médio (1,55 - 1,75µm)</p> <ul style="list-style-type: none"> • No trecho entre 1,35 e 2,5 µm, onde esta situada esta banda, a reflectância ainda está submetida a algum efeito da estrutura interna foliar, porém o principal fator modulador é a água presente no tecido foliar (Mather, 1987). Nestes comprimentos de onda, a existência de bandas onde a água apresenta elevada absorção (aproximadamente 1,4; 1,95 e 2,7 µm) explica a resposta espectral da vegetação verde verificada por Hoffer & Johannsen (1969) na qual a reflectância nas ondas infravermelho diminui com o aumento da quantidade de água, possibilitando distinguir plantas com deficiência hídrica. • As características dos comprimentos de onda desta banda habilitam-na para a aplicação na determinação da umidade da vegetação e dos solos e na discriminação entre neve e nuvens (Lillesand & Kiefer, 1994).
<p>6 - infravermelho termal (10,40 -12,50µm)</p> <ul style="list-style-type: none"> • Esta banda confere aos dados do sensor TM uma utilidade potencial para avaliação da estricção térmica da vegetação, da umidade edáfica e para o mapeamento termal (Lillesand & Kiefer, 1994).
<p>7 - infravermelho médio (2,08 - 2,35µm)</p> <ul style="list-style-type: none"> • Também posicionada no segmento do espectro eletromagnético onde a reflectância é função do conteúdo hídrico da folha, diminuindo a medida que a hidratação no tecido foliar aumenta, esta banda, além de aplicações para discriminação mineralógica e geológica, é útil para estimar o estado hídrico da vegetação (Lillesand & Kiefer, 1994).

4.5 Processamento digital de imagens

O olho humano não é capaz de extrair todas as características presentes nas imagens. A sensibilidade do olho humano está restrita a faixa visível do espectro eletromagnético e o poder de discriminação de tonalidades de cinza raramente ultrapassa 30 tons. Além dessa limitação, o homem é incapaz de processar os grandes volumes de informações presente nas imagens. Por exemplo, uma cena gerada pelo sensor TM cobre 185 km² da terra, contendo mais de 9 milhões de pixels, cada um deles com um valor de brilho variável de 0 a 225 tonalidades de cinza, em cada uma das 7 bandas, equivalendo a mais de 16 milhões de números. A capacidade humana de análise da resposta espectral é ainda mais restringida pelos vários tipos de degradação e distorções dos processos de aquisição, transmissão e visualização de imagens.

O objetivo primordial do processamento digital de imagens (PDI) é superar estas limitações naturais do sentido da visão dos seres humanos e estruturar sistemas semi-automatizados para a extração de informações a partir da manipulação e interpretação de imagens digitais via computador (Crosta, 1992 e Lillesand & Kiefer, 1994).

A idéia básica do PDI é inserir uma imagem digital, píxel a píxel, em um computador e extrair rapidamente as informações de interesse. O computador é programado para alimentar uma equação ou uma série de equações com esses dados, comparando e agrupando-os de acordo com suas semelhanças. Com os resultados do cálculo para cada píxel é gerada uma nova imagem digital passível de visualização ou de manipulações adicionais. O número de processamentos é muito grande, incluindo realces, filtragens, visualização de composições coloridas, bandas individuais e escalas diversas, integração com bases de dados geocodificadas no ambiente dos sistemas de informação geográfica, incorporação de informações de caráter auxiliar, diversos métodos de classificação numérica, textural, "fuzzy", etc. Estas diversas formas de manipulação das imagens digitais podem ser resumidas em 5 tipos: retificação e registro, melhoramento, classificação, integração e modelagem biofísica (Lillesand & Kiefer, 1994). Em seguida, serão melhor desenvolvidas os tipos de técnicas relacionados com o propósito dessa dissertação.

4.5.1 Georreferenciamento das imagens

As imagens orbitais contêm freqüentemente distorções associadas às características do sensor, à curvatura da Terra e ao movimento da plataforma. Deste modo, geralmente é necessário aplicar procedimentos de correção geométrica sobre as imagens. Esse processo de georreferenciamento tem como objetivo corrigir as imagens distorcidas para criar uma representação mais fiel da cena original. Estabelecendo-se uma relação entre os pontos da imagem e seus correspondentes na Terra, toda a imagem pode ser retificada adequadamente, conferindo-lhe os atributos cartográficos de escala e projeção.

Conforme muito bem descrito por Janssen & van der Wel (1994), no processo de registro da imagem, estabelece-se uma relação entre os dados linha-coluna da imagem e um sistema de coordenadas dos dados de uma outra imagem ou de um sistema de projeção particular de um mapa. Esta associação pode ser feita de duas formas: via parâmetros orbitais do satélite ou pela localização do pontos de controle terrestre ("ground control

points - GCPs”) na imagem e no sistema de coordenadas de referência. A última abordagem não-paramétrica é considerada geralmente como mais realista porque o modelo de geometria orbital adotado na descrição dos erros é incompleto e causa distorções geométricas (Mather, 1987).

Esses pontos de controle podem ser identificados tanto no sistema de coordenadas da imagem (i, j - linhas e colunas) como no sistema de coordenadas de referência (x, y - coordenadas cartográficas). Esses sistemas podem ser relacionados através de equações polinomiais. Apesar de uma transformação afim de primeiro grau (“first-degree affine”) frequentemente atender os requisitos de uma imagem de satélite, a complexidade desses polinômios está vinculada à geometria da imagem e ao tipo de projeção empregado pela base cartográfica selecionada.

O cálculo do polinômio é feito por uma minimização da soma dos quadrados. Determinada a solução ótima do polinômio, a computação dos resíduos nas direções x e y (δ_x e δ_y) permite seu emprego para determinar a acuidade através da raiz quadrada do desvio quadrático médio (RMS) ou do desvio padrão. O erro RMS na direção x é obtido por:

$$RMS_x = \left[\frac{1}{n} \sum_{i=1}^n (\delta_{xi})^2 \right]^{\frac{1}{2}},$$

onde δ_{xi} é o resíduo do ponto de controle i e n , o número de GCPs.

O RMS_y é gerado analogamente, possibilitando a sua subsequente combinação com RMS_x para produção de um RMS planimétrico (RMS_{xy}):

$$RMS_{xy} = \left[RMS_x^2 + RMS_y^2 \right]^{\frac{1}{2}}.$$

Uma estimativa da acuidade de caráter mais estatístico envolveria o desvio padrão. A soma dos resíduos é dividida pela redundância (r) definida pelo grau de liberdade determinado pelo polinômio aplicado:

$$S_x = \left[\frac{1}{r} \sum_{i=1}^n (\delta_{xi})^2 \right]^{\frac{1}{2}}.$$

O uso de um grande número de GCPs faz com que o erro RMS e o desvio padrão aproximem-se.

As acuidades obtidas podem ser interpretadas através da construção de limites de confiança, empregando-se uma abordagem Gaussiana caso os GCPs sejam independentemente identificados.

A obtenção de um registro da imagem bem feito está vinculada à uma identificação precisa dos GCPs. Selecionar esses pontos a partir de mapas pode agregar incertezas inerentes ao fato de uma representação cartográfica ser uma idealização e generalização da realidade. Uma alternativa para superar essa dificuldade tem sido o crescente incorporação do GPS (“Global Positioning Systems”), limitada apenas pela determinação dos GCPs na

imagem orbital onde a identificação é fortemente circunstanciada pela resolução do sensor adotado (Janssen & van der Wel, 1994).

O segundo passo no georreferenciamento é a aplicação de um algoritmo de amostragem na imagem para gerar uma nova grade correspondente aos eixos euclidianos do sistema de referência escolhido. Em um dos métodos consagrados, o do “vizinho mais próximo”, o valor dos pixels da nova imagem é semelhante àquele do ponto mais próximo correspondente na imagem bruta. Essa amostragem aparentemente causa um menor impacto sobre a imagem, uma vez que “copia” o valor do pixel original para a imagem resultante. É um procedimento indicado para tratar imagens multispectrais pois preserva a relação entre as bandas, permitindo operações de cruzamento de bandas (Taylor, 1989).

As informações necessárias para a implementação do georreferenciamento são o *datum* geodético, a projeção do mapa e o tipo de coordenada, os quais podem ser escolhidos a partir de uma lista fornecida pelo aplicativo, tal qual o método de amostragem.

4.5.2 Técnicas de realce de imagens

O melhoramento de imagens reúne técnicas de processamento para aperfeiçoar a visualização das informações contidas em uma imagem, de acordo com os objetivos e interesses do usuário. Esses recursos podem ser agrupados em três grandes categorias segundo Lillesand & Kiefer (1994): manipulação de contraste, manipulação de características espaciais e manipulação de múltiplas imagens.

As **manipulações de contraste** englobam técnicas para a extração de informações das imagens como a separação por nível de cinza (“gray-level thresholding”), segmentação de níveis de cinza (“level slicing”) e o aumento de contraste (“contrast stretching”) (Lillesand & Kiefer, 1994). Dentre estas, a última certamente é a mais complexa e importante a ponto de merecer um tratamento mais detalhado.

Como raramente há uma variação de níveis de cinza com amplitude de 0 a 255 em uma única cena, o aumento de contraste é um procedimento para expandir o intervalo dos níveis de intensidade para seu valor máximo. Os principais motivos desta concentração de NDs em uma pequena faixa são: a inexistência de dois extremos de reflectância concomitantemente em um mesmo local da superfície terrestre e o efeito da bruma atmosférica, iluminação solar da cena insuficiente, desempenho do sensor inadequado ou características específicas dos alvos representados na cena (Crósta, 1992). A vantagem principal dessa técnica é a manutenção da estrutura e textura dos dados, pois não há modificações na relação entre eles. Os tipos de aumento de contraste comumente utilizados são: aumento linear de contraste, aumento bi-linear de contraste; aumento não linear de contraste (com as seguintes subdivisões em equalização de histograma, normalização de histograma, aumento logarítmico de contraste e aumento exponencial de contraste) (Ahern & Sirois, 1989; Crósta, 1992 e Schowengerdt, 1983) e aumento “quase-linear” de contraste (Guo, 1991).

A **manipulação de características espaciais**, o segundo grande grupo de técnicas de melhoramento, abarca recursos conhecidos como filtragem espacial através dos quais

pode-se realçar ou suavizar seletivamente as frequências espaciais dos NDs. Estas frequências representam a textura da variação tonal presente em uma imagem. A textura, em função da distribuição espacial das tonalidades de cinza, pode assumir em algumas áreas, um aspecto rugoso, correspondendo a feições de alta frequência espacial onde os níveis de cinza são muito contrastantes em um número de pixels relativamente pequeno, ou um aspecto suave, correspondendo a feições de baixa frequência onde a variação dos níveis de cinza é gradual em um número de pixels relativamente grande (Lillesand & Kiefer, 1994).

A filtragem de frequência espacial pode ser executada por alterações nos atributos espaciais da imagem via sistema de coordenadas x , y . Operacionalizados através da aplicação do processo de convolução, estes filtros podem ser distinguidos em três tipos básicos (Crosta, 1992): passa-baixa ou suavização (de média, de média ponderada, de moda e de mediana), passa-alta (laplacianos e de realce de borda) e direcionais (de Roberts, de Sobel e de Prewitt). Outra forma de implementar a filtragem é através da análise de Fourier, uma transformação matemática complexa efetuada no domínio das frequências espaciais, individualizadas em seus diversos componentes.

A terceira categoria de técnicas de realce das imagens, a **manipulação de múltiplas imagens**, inclui as operações aritméticas entre bandas; extração de componentes principais, canônicos e vegetacionais; transformação do espaço de cores matiz, saturação e intensidade ("Hue, Saturation, Intensity" - HSI) e o realce por decorrelação.

Como as bandas espectrais geradas pelos imageadores são contíguas, características da imagem capturada, como por exemplo o efeito de sombra criado pelos condicionantes topográficos locais e pelos ângulos de iluminação e visada do sensor, podem originar um problema comum dos dados de sensoriamento remoto, a elevada redundância de informações nas diferentes bandas. Esta correlação inter-banda tem sido ampliada pelo constante desenvolvimento da sensibilidade dos sensores e conseqüentemente ampliando o tempo de cálculo computacional para gerar novas imagens sem necessariamente melhorar sua interpretação.

As técnicas para remover essa correlação sem perder substancialmente informações incluem, além da manipulação de caracteres espaciais através de filtros de frequência apresentada nos parágrafos anteriores, a seleção de bandas e respectivas cores para gerar composições coloridas, os métodos de compressão de dados, tais como:

- manipulações matemáticas entre bandas (por exemplo, razões de bandas e índices de vegetação);
- transformação HSI e o realce por decorrelação, através dos quais pode-se aperfeiçoar a atribuição de cores nas composições falsa-cor e
- técnicas para extração de componentes vegetacionais, principais e canônicos, para ampliar a utilização e visualização da informação de mais de três bandas simultaneamente, reduzindo a dimensionalidade das n -bandas originais em novas bandas ou componentes com informações concentradas e não correlatas (Haack & Jampoler, 1995; Lillesand & Kiefer, 1994 e Crosta, 1992).

A compressão da dimensionalidade dos dados originais para obter realces diversos nas imagens também pode ser atingida através das quatro operações aritméticas básicas entre bandas multiespectrais e/ou multitemporais. De um modo geral, a adição e multiplicação atuam realçando semelhanças espectrais, contrapondo-se ao destaque das diferenças conferido pela subtração e divisão ou razão de bandas. Todas estas técnicas operam sobre os NDs de duas ou mais bandas segundo a regra matemática adotada, gerando como arquivo de saída uma única banda. Contudo, essa redução da dimensionalidade sempre implicará em uma perda parcial de informações (Crosta, 1992).

A importância do monitoramento da vegetação suscitou o desenvolvimento de uma série de transformações lineares dos dados espectrais, com características específicas em função do imageador selecionado e das características da formação vegetal enfocada. São exemplos consagrados dessas técnicas, os recursos para extração de componentes vegetacionais como a transformação "tasseled cap" (Crist & Cicone, 1984) e os diversos índices de vegetação. Essas combinações entre bandas espectrais reúnem um grande número de variantes, das quais destacam-se, por exemplo, o índice de vegetação transformado (TVI - "transformed vegetation index") e principalmente o índice de diferença vegetal normalizado (NDVI - "normalized difference vegetation index"), largamente empregado com os dados do sensor AVHRR do satélite meteorológico NOAA (Lillesand & Kiefer, 1994).

4.5.2.1 Seleção de bandas e composições coloridas

No contexto dos procedimentos para redução da dimensionalidade apresentados no item anterior, a seleção de bandas e a formulação de composições coloridas é um processo de extrema importância para a interpretação visual e a classificação digital das imagens. Sua complexidade está relacionada à dependência das peculiaridades dos dados analisados (Conese & Maselli, 1993).

Como o olho humano emprega três cores primárias e o sensor TM do satélite LANDSAT imageia em sete bandas espectrais é necessário selecionar as três melhores bandas para criar uma composição colorida no espaço de cores vermelho-verde-azul ("Red-Green-Blue"- RGB), o mais amplamente utilizado nos sistemas de tratamento digital de imagens devido a sua congruência com o sistema visual dos seres humanos (Sheffield, 1985).

A técnica é conceitualmente bastante simples, consistindo em utilizar os valores dos NDs de três bandas de uma mesma imagem como pesos em uma combinação linear das cores vermelho, verde e azul. Sendo essas as cores primárias, uma infinidade de cores derivadas podem ser geradas e relacionadas a cada alvo, auxiliando a interpretação da imagem. A escolha das bandas a serem reunidas, também denominado de triplete, e a definição da melhor composição colorida é função de fatores como: o tema considerado, os padrões ambientais da área imageada, a qualidade das bandas da imagem e a influência diária e sazonal na aquisição da imagem.

Apesar da seleção de bandas ser importante devido a composição colorida ser normalmente o primeiro contato entre o usuário e as informações contidas nas imagens e a

interpretação analógica continuar a ser um componente corriqueiro das aplicações de sensoriamento remoto, existe pouca literatura sobre o assunto (Haack & Jampoler, 1995).

Os tripletos TM 3-2-1 (cor verdadeira) e TM 4-3-2 (falsa cor) parecerem ser os mais usuais historicamente, contudo a boa diferenciação de alvos consagrou a composição 5-4-3 para a classificação de imagens Landsat-TM orientadas para a execução de levantamentos: da cobertura terrestre (Fuller et al., 1989; Fuller & Parsell, 1990; Janssen & Middelkoop, 1992; Fuller et al., 1994 e Sharma et al., 1995); da citricultura, cafeeicultura e canavicultura (Tardin et al., 1992); florestais (Karteris, 1990 e Congalton et al., 1993); de florestas inundáveis (Sader et al., 1995); de desmatamentos (Jusoff et al., 1995) e de campos naturais (Fuller et al., 1989). Também são citadas como adequadas para efeito de processamento digital as seguintes combinações: 4-3-2 para finalidades florestais (Karteris, 1990), 7-5-1 e 5-4-1 para inventário da cobertura vegetal (respectivamente, Booth & Oldfield, 1989 e Sharma et al., 1995) e 7-4-2 para tipificação de florestas tropicais úmidas (Hill & Foody, 1994).

A 5-4-3 também possibilita bons resultados na visualização de imagem Landsat-TM nos levantamentos das seguintes situações: agroecossistemas (Townshend et al., 1983); cultivo de arroz inundado (Tennakoon et al., 1992); desmatamento (Jusoff et al., 1995) e floresta tropical (Nelson et al., 1994 e Tuomisto et al., 1994). Bouchardy & Rodriguez (1992) optaram pela composição 4-3-1 para uma análise visual prévia de uma área agrícola muito heterogênea antes da aplicação de procedimentos de classificação digital.

A composição TM 1-2-3, apesar de ser mais familiar ao usuário pois recria as cores naturais, freqüentemente não fornece um máximo conteúdo informacional, pois envolve somente bandas da região eletromagnética do visível, onde as ondas curtas sofrem mais os efeitos da interferência atmosférica e da nebulosidade. Muitos estudos indicam que imagens criadas a partir de diferentes partes do espectro eletromagnético fornecem mais informações (Haack & Jampoler, 1995).

A seleção das bandas Landsat-TM para alimentar os classificadores pode envolver um número maior de bandas, embora haja um custo maior em termos de processamento de dados. Foram obtidos bons resultados com as seguintes composições: 5-4-3-2-1 para avaliação de campos naturais (Fuller et al., 1989); 7-5-4-3-2 para classificação de comunidades florestais temperadas e de sub-bosque (Stenback & Congalton, 1990); 5-4-3-1 para a estimativa da orizicultura inundável (Tennakoon et al., 1992); 7-5-4-3-2-1 para mapeamento do uso das terras (Viñas & Baulies, 1995), para classificação das formações vegetais da tundra (Joria & Jorgenson, 1996) e para propósitos florestais em áreas temperadas (Karteris, 1990; Fiorella & Ripple, 1993 e Bauer et al., 1994) e em áreas tropicais (Garcia & Alvarez, 1994).

Há inúmeras técnicas de seleção de bandas ou separabilidade. Mausel et al. (1990) compararam quatro medidas de separabilidade (análise de divergência, divergência transformada, distância de Bhattacharya e distância de Jeffries-Matusita) para identificar o melhor grupo de bandas para classificar uma área agrícola através de um procedimento paramétrico, concluindo que a divergência transformada e a distância de Jeffries-Matusita são os melhores métodos. Haack et al. (1986) através destes métodos selecionaram as

bandas 3-4-5 do sensor TM. Outros pesquisadores também consideram este triplete como o mais adequado para a classificação de formações vegetacionais (Lauer & Waltz, 1983; Haas & Waltz, 1983; Dottavio & Dottavio, 1984 e Haas, 1992). Os índices de separabilidade são baseados nas medidas de distância estatística entre classes espectrais de interesse (Swain, 1978 e Thomas et al. 1987). Chavez et al. (1982) desenvolveram o fator de índice ótimo ("optimum index factor"- OIF) para selecionar a melhor combinação de três bandas baseado na quantidade de variância e correlação total dentro e entre várias combinações de bandas. Dwividi & Rao (1992), aplicando o OIF nas imagens de uma planície de inundação, determinaram o triplete TM 1-3-5 como detentor das melhores informações, contudo não conseguiram estabelecer uma correlação consistente com a exatidão dos dados de classificação.

Sheffield (1985) sugeriu um método baseado no volume do elipsóide formado pela distribuição dos dados no espaço n-dimensional associado a matriz de variância-covariância de um triplete. Isto reduz a seleção de bandas com alta correlação. É possível selecionar o triplete de bandas com o elipsóide de maior volume através do cálculo e ordenação dos determinantes da submatriz de covariância (3 x 3) da matriz original (7 x 7). Através desse método concluiu que as bandas TM 1, 4 e 5 era o melhor triplete. Kenk et al. (1988) chegou a mesma conclusão.

Ma & Olson (1989) mostraram que o valor do índice de sobreposição de valor de brilho ("brightness value overlapping index" - BVOI) é a melhor medida do grau de sobreposição no valor do brilho entre tipos de cobertura.

Sharma et al. (1995) ao testar diferentes métodos como o determinante da matriz de covariância de Sheffield, o OIF, a distância de Bhattacharya, a análise de divergência, mostraram que a inclusão de bandas do infravermelho médio em uma combinação de bandas do visível e infravermelho próximo aumentava a separabilidade entre as classes. Portanto, a combinação das bandas 2-3-4-5 confere tanto a máxima separabilidade como a melhor precisão na classificação de categorias vegetais.

Os procedimentos estatísticos de compressão de dados, já citados como análise de variáveis canônicas (Conese et al., 1987) e componentes principais (Taylor, 1974; Merembeck et al., 1977 e Richards, 1984), têm sido considerados efetivos na identificação de subgrupos de bandas espectrais apropriados para classificação da cobertura terrestre (Dean & Hoffer, 1982). Esses métodos são eficientes e simples, mas transformam as variáveis originais em um sistema de referência diferente, dificultando a interpretação dos resultados finais, a qual só pode ser bem feita com o auxílio de um conhecimento metuculoso das características da área de estudo (Fung & LeDrew, 1987 e Chavez & Kwarteng, 1989).

Em 1993, Conese & Maselli propõem a utilização da análise de informação mútua, baseada na teoria de entropia, para a seleção de bandas Landsat-TM no estudo de vegetação. Na análise de informação mútua, as diferenças entre a entropia máxima e a observada são utilizadas como uma medida de correspondência entre os diferentes atributos. Se esses são probabilisticamente independentes, então a informação probabilística global é zero. Com base nisto, a análise de informação mútua pode ser aplicada para determinar

quais bandas da cena são mais informativas em relação a variável ecológica escolhida, ou seja, qual a porcentagem de informação sobre um certo tema expressa por cada banda.

Haack & Jampoler (1995), utilizando imagens do Landsat-TM, analisaram o desempenho de três composições coloridas na classificação de culturas agrícolas em uma região da Califórnia, USA: a composição falsa-cor padrão 4-3-2, simulacro dos filmes infravermelho; o tripleto 4-5-3 identificado a partir da aplicação da técnica de divergência transformada e os três primeiros componentes principais resultantes da redução dos dados das sete bandas da imagem. Dentre as combinações testadas, a 4-5-3 apresentou os melhores resultados.

Crósta (1992), baseado na sensibilidade diferenciada do olho humano em relação às cores, sugere a otimização da combinação colorida através da atribuição da cor verde para a banda de maior informação textural e do azul e vermelho para as bandas com feições espectrais sutis. Baseado no mesmo princípio, Sheffield (1985) aconselha designar a cor verde para a banda de maior variância dentro da imagem, vermelho para a banda com a segunda maior variância e azul para a remanescente.

4.5.3 Classificação de imagens

Durante a recente história dos satélites de exploração dos recursos terrestres, as atividades de cartografia da cobertura da terra têm gradativamente ampliado sua automatização através da adoção de métodos digitais de classificação dos dados multiespectrais gerados através de teledetecção (Manselli et al., 1992).

A classificação digital de uma imagem orbital tem como objetivo principal a categorização automática de todos os pixels de uma imagem em classes de cobertura terrestres ou temas. Normalmente envolve a associação dos dados multiespectrais, representados por cada píxel na imagem, a um rótulo descritivo de um objeto na superfície. Esta relação é estabelecida com base na comparação entre cada píxel e as **assinaturas espectrais** do objeto de interesse, ou seja suas características de reflectância determinadas em locais da área de estudo onde sua presença seja indubitável (Mather, 1987).

Nos procedimentos de classificação costuma-se utilizar o termo **padrão** para designar um grupo de valores de radiância para cada píxel, obtidos em várias bandas espectrais (Mather, 1987 e Lillesand & Kiefer, 1994). Assim o processo de classificação também pode ser descrito como o reconhecimento de padrões, ou identificação do padrão associado a cada píxel na imagem, no que se refere as características dos objetos em um ponto da superfície terrestre correspondente. Há basicamente três tipos de reconhecimento de padrões. O reconhecimento de padrões espectrais se refere a família de procedimentos de classificação onde as informações espectrais, píxel a píxel, são utilizadas como base para a classificação dos temas. O reconhecimento de padrões espaciais envolve a categorização dos pixels da imagem com base em suas relações espectrais com os pixels ao redor dele. Classificadores espaciais podem considerar aspectos como textura, proximidade do píxel, tamanho, forma, direção, repetição e contexto. Estes classificadores tentam reproduzir um tipo de síntese espacial feita pelo olho humano durante o processo de interpretação visual e apresentam um procedimento mais complexo e computacionalmente mais intenso em relação aos de reconhecimento de padrões espectrais. O reconhecimento de padrões temporais usam

a diacronia para auxiliar na identificação de atributos. Muitas vezes determinados temas apresentam distintos padrões espectrais e espaciais ao longo do tempo, por exemplo uma cultura em desenvolvimento, facilitando ou permitindo sua identificação em imagens multitemporais (Lillesand & Kiefer, 1994).

As características ou variáveis, tais como as bandas 1, 2, 3, 4, 5 e 7 do sensor TM-Landsat, adotadas para definir as bases dos padrões, são denominadas **atributos**. Neste contexto, um padrão é simplesmente um conjunto de medidas bem definidas, de natureza espectral ou espacial, no universo composto pelos atributos escolhidos (Swain, 1978 e Mather, 1987). Para reconhecimento desses padrões, o roteiro de abordagem segue, primeiro, da seleção do atributo com a melhor capacidade de caracterizar os padrões para, segundo, a escolha de um método adequado para comparação de padrões descritivos dos objetos a serem classificados com os padrões não identificados presentes na imagem. O processo finaliza-se com a verificação do grau de exatidão do processo de alocação das categorias selecionadas (Mather, 1987).

As frequências de distribuição dos NDs (níveis de cinza) dos atributos, ou seja cada par de bandas componente de uma imagem multiespectral, podem ser visualizadas graficamente através de um sistema de coordenadas cartesianas em um **diagrama de dispersão** ("scatter diagram"), onde o espaço euclidiano recebe a denominação específica de **espaço de atributos**. A distribuição dos pontos nessas representações podem formar áreas isoladas de alta concentração de pixels, os agrupamentos ("clusters"), normalmente relacionadas a uma classe, *i.e.* um tipo ou um conjunto de fenômenos integrantes da cobertura da superfície terrestre. Como o número de bandas dos sensores usualmente é muito superior a dois, os mesmos conceitos de associação de um atributo a um dos eixos do espaço euclidiano e do uso da distância inter-pontos para a delimitação dos grupamentos, estendidos para n-dimensões, definem um espaço de atributos n-dimensional. Esse modelo geométrico do processo de reconhecimento de padrões permite uma percepção da distribuição dos valores dos atributos e expressa esses dois conceitos fundamentais do procedimento da classificação digital: a representação do alvo de interesse no atributo selecionado através dos eixos do espaço de atributos e a possibilidade de estabelecer nesse espaço critérios de decisão para a classificação dos elementos da cobertura terrestre. Entretanto, seu equivalente algébrico tem um desempenho melhor nas situações reais pois é extensível a mais de 3 dimensões e aplicável em programas de computadores (Mather, 1987 e Crosta, 1992).

Quando essa operação de rotulação abrange todos pixels presentes na área focalizada, a distribuição espacial de um elemento terrestre tal como, formações vegetais, cobertura pedológica, qualidade de água etc., expressa-se sob a forma de um mapa temático.

Uma das formas adotadas para melhorar o desempenho das classificações tem sido o uso de imagens multitemporais, devido principalmente a introdução de informações relativas as mudanças nos estágios fenológicos dos diferentes tipos de cobertura vegetal (Dye & Wood, 1989; Fuller & Parsell, 1990; Conese & Maselli, 1991; Wickland, 1991; Townshend, 1992; Fuller et al., 1994 e Viñas & Baulies, 1995).

Até recentemente a classificação era baseada apenas nas informações espectrais das imagens (classificação por pixels ou por pontos) (Kartikayan et al. 1994 e Fung & Chan, 1994). Nas duas últimas décadas houve um aumento no estudo da estrutura espacial da imagem permitindo sua incorporação nos classificadores digitais. O princípio é tentar definir a textura do pixel sob o processo de classificação, *i.e.* aferir a homogeneidade dos vizinhos mais próximos do pixel. Este conceito é amplamente utilizado de forma intuitiva na interpretação analógica das fotografias aéreas.

Em 1979, Haralick fez uma revisão das medidas texturais. Conners & Harlow (1980) examinaram e compararam os diferentes algoritmos texturais. Curran (1988), Woodcock et al. (1988a e b) e Ribeiro (1994) fizeram uso de semi-variogramas para análise das informações texturais das imagens. A análise textural tem sido utilizado em mapeamentos e estimativas de coberturas terrestres (Skidmore, 1989b; Gordon & Phillipson, 1986; Franklin & Peddle, 1990 e St-Onge & Cavayas, 1995), monitoramento ambiental (Cross et al., 1988) e em estudos ecológicos (Saxon, 1984).

Woodcock & Strahler (1987) usaram a variância local para medir a estrutura espacial da imagem como uma função da resolução espacial. Segundo os resultados obtidos por este trabalho, a variância local de imagens com a resolução espacial dos sensores HRV e Thematic Mapper é relativamente baixa em ambientes agrícolas, confirmando a adequabilidade do uso de classificadores mais tradicionais, ao contrário dos ecossistemas florestais, conurbações e áreas urbanas onde a alta variância local sugere a utilização de classificadores baseados na análise da textura, da mistura (a contribuição dos elementos menores que as resolução espacial para a radiância registrada por cada pixel) e do contexto.

Uma outra abordagem recente para incrementar a precisão do processo de classificação é o uso de redes neurais artificiais, cujo desenvolvimento remonta aos últimos sete anos (Paola & Schowengerdt, 1995). Neste caso, a classificação ou agrupamento de dados é executado através de redes de unidades de processamento simples, dispostas em um conjunto de camadas interconectadas por canais, aos quais são atribuído pesos (Foody, 1995a). Dentre os muitos tipos de redes neurais diferentes como por exemplo o mapa de atributos auto-organizáveis (Schaale & Furrer, 1995), o modelo de percepção multicamadas tem sido um dos mais utilizados (Wilkinson, 1996). Algumas das vantagens destes procedimentos são: a independência de suposições estatísticas sobre a distribuição dos dados espectrais (Foody, 1995b e Paola & Schowengerdt, 1995), as dimensões reduzidas das áreas de treinamento (Hepner et al., 1990 e Zhuang et al., 1994), o processamento de imagens mistas combinando dados de sensores diferentes como o Landat-TM e o ERS--SAR ("Synthetic Aperture Radar") (Wilkinson et al., 1995), o potencial para manipular bases de dados complexas dos computadores de processamento paralelo (Wilkinson, 1996) e a precisão global satisfatória atingida por estes classificadores, comparável as das técnicas paramétricas convencionais e em alguns casos superior (Wilkinson, 1996).

Além da consideração dos tons e textura, a análise visual da imagem frequentemente envolve a avaliação do contexto de um objeto, *i.e.* a relação de um objeto com outros circunvizinhos. Alguns objetos ocorrem, outros não, em determinadas regiões ou em determinadas associações espaciais ou seqüências temporais com outros elementos. Enquanto textura refere-se a relação espacial entre os pixels originais da imagem em uma

janela de um determinado tamanho, contexto é uma função da posição entre pixels, classificados ou não, em qualquer lugar da imagem. Normalmente a aplicação dos algoritmos contextuais é feita sobre o resultado de um processamento de classificação preliminar, rerotulando as categorias dos pixels segundo a informação contextual. Vários tipos de formas de contexto podem ser utilizados na identificação de um padrão na imagem através das classes de outros padrões associados, como por exemplo, a distância, direção, contiguidade e inclusão (Campbell, 1996).

Um outro grupo de procedimentos explora essas relações na atribuição do rótulo ao pixel, utilizando informações auxiliares ("ancillary data"). São dados extra imagem capazes de refinar a classificação dos padrões, articulando dados ambientais georreferenciados com as imagens orbitais (Hutchinson, 1982; Kenk et al, 1988; Wu et al., 1988; Janssen et al. 1990; Zhuang et al., 1991 e Congalton et al., 1993). Uma das formas de desenvolvimento desta abordagem tem sido executada através de modelos de regras de decisão ("Knowledge-based classification"), cuja implantação pode ser feita em um sistema de informações geográficas ("GIS rule-based model"). O conhecimento das relações entre as características dos diversos tipos de cobertura terrestre, representados pelos planos de informação georreferenciados, e os padrões espectrais é codificado em um sistema de classificação das imagens. Este conhecimento necessário para identificar corretamente a classe de cobertura é descrito e armazenado por uma série de fatos independentes e regras de decisão. O mecanismo para geração de soluções para o problema ("problem-solving mechanism") decide como ordenar os fatos e heurísticamente formular uma classificação da imagem. A complexidade deste processo de tecer inferências a partir das relações entre fatos reside em analisar e decidir sob as incertezas provenientes de evidências díspares ou inconfiáveis ou de imprecisões na tradução das informações/experiência em uma base de conhecimento (Marble & Pequet, 1983; Skidmore, 1989a; Janssen & Middelkoop, 1990; Marble, 1990; Srinivasan & Richards, 1990; Bolstad & Lillesand, 1992 e Sader et al., 1995).

Esta técnica é apenas um dos aspectos da crescente sinergia entre o uso dos sistemas de informação geográficas e do sensoriamento remoto, cuja associação pode ser abordada segundo três formas: a teledetecção como instrumento para a geração de bases de dados geocodificadas; a recíproca deste primeiro uso, ou seja as bases de dados geocodificadas como informação auxiliar para o aperfeiçoamento do processamento de dados remotos e a conjugação de dados remotos com geocodificados como suporte para a modelagem e análise ambiental (Wilkinson, 1996).

Os métodos de classificação de imagens podem ser combinados formando classificadores híbridos (Chuvieco & Congalton, 1988b; Fiorella & Ripple, 1993; Lauver & Whistler, 1993; Garcia & Alvarez, 1994; Hardin, 1994; Lillesand & Kiefer, 1994, Wilkinson et al. 1995 e Jensen, 1996). Não existe uma única forma correta de se abordar o problema da classificação. Abordagens particulares devem ser selecionadas dependendo da natureza dos dados analisados, dos recursos computacionais disponíveis e da finalidade da classificação (Lillesand & Kiefer, 1994). Nesta mesma linha de articulação de potencialidades distintas, um caminho promissor é a complementaridade dos produtos de sensores diferentes, como por exemplo a fusão multitemporal de imagens de alta resolução

radar, ERS-SAR, com ópticas, Landsat-TM, para a classificação supervisionada de culturas agrícolas (Mangolini & Arino, 1996).

4.5.3.1 Classificação multispectral

Um grande número de investigações tem explorado as aplicações de diversos recursos estatísticos para discriminação dos tipos de cobertura de um território com base em suas assinaturas espectrais (entre os quais, sem a pretensão de formular uma lista exaustiva, pode-se citar: Hixon et al, 1980; Conners & Harlow, 1980; Wharton et al., 1981; Justice & Townshend, 1981; Justice & Townshend, 1982; Dutra & Mascarenhas, 1984; Richards, 1984; Townshend, 1984; Atkinson et al., 1985; Yool et al., 1986; Fung & LeDrew, 1987; Kenk et al., 1988; Booth & Oldfield, 1989; Franklin & Peddle, 1990; Gong & Howarth, 1990a; Gong & Howarth, 1990b; Townshend et al. 1991, Gong & Howarth, 1992; Gong et al., 1992; Maselli et al. 1992; Green et al., 1993; Kartikeyan et al., 1994; Kushwaha et al., 1994 e Lark, 1994).

O problema de rotulação dos pixels é abordado por essas técnicas seguindo dois critérios básicos: a existência de conhecimentos prévios sobre a identidade dos padrões de cobertura terrestre distingue os **supervisionados dos não-supervisionados** e a suposição sobre a forma da distribuição dos dados em um espaço de atributos N-dimensional distingue os **paramétricos dos não-paramétricos** (Sabins, 1977).

Deste modo, os métodos supervisionados caracterizam-se pela disponibilidade do número de padrões distintos existentes na área levantada pela imagem e pela possibilidade de estimar as propriedades estatísticas dos atributos elegidos para descrever cada um desses padrões. Com esses dados, são elaborados **modelos**, representando padrões ideais ligados aos tipos de cobertura da superfície terrestre, denominados **classes de informação** (categorias de cobertura terrestre), com os quais os grupos de pixels presentes na imagem são comparados. Após a determinação da combinação mais similar, cada pixel recebe um rótulo relacionando-o a uma determinada classe (Mather, 1987).

Por sua vez, os métodos não-supervisionados podem ser tratados como uma autoclassificação, aplicável quando assume-se não haver informações preliminares sobre o número ou características dos padrões verificáveis na imagem. Neste caso, seleciona-se arbitrariamente um número inicial de padrões, denominados **padrões básicos**, nos quais cada pixel é realocado progressivamente. O método simplesmente utiliza os valores espectrais dos atributos para determinar grupos de pixels sem sobreposição, conhecidos por **classes espectrais**, os quais são identificados através de informações auxiliares provenientes de observações de campo e de documentos fotográficos e cartográficos (Mather, 1987).

Os classificadores **paramétricos** trabalham com a suposição de que os valores de reflectância, com os quais determina-se os parâmetros de distribuição necessários ao processo de classificação, são amostras de variáveis multidimensionais e com uma distribuição normal. Estes requisitos devem ser observados em todos os atributos selecionados para cada classe nominal (Jensen, 1996). Por sua vez, os métodos **não paramétricos** ou robustos não exigem nenhum pressuposto quanto a forma da distribuição multivariada das classes (Schowengerdt, 1983 e Hardin, 1994).

4.5.3.1.1 Treinamento para a classificação

Os procedimentos de classificação iniciam-se com o treinamento do classificador espectral selecionado para reconhecer as classes de interesse. Desse modo, o objetivo do processo de treinamento é reunir dados espectrais para a definição de um conjunto de estatísticas capaz de descrever os padrões de respostas espectrais de cada classe de cobertura terrestre presente na imagem (Swain, 1978). A determinação da localização, forma, tamanho e orientação das nuvens de pixels, associadas a cada tema, no espaço de atributos viabiliza a obtenção de regras de decisão para a classificação de cada pixel (Merchant, 1982).

Devido as dificuldades e alto custo para a obtenção das informações para todo o universo de atributos, trabalha-se com estimativas da população obtidas através de amostras de pixels, as **amostras de treinamento**. Há dois critérios básicos para a escolha dessas áreas: representatividade de toda a variabilidade existente para cada classe e concordância em cada série com os pressupostos sobre a forma da distribuição dos dados relacionados com o tipo de classificador elegido (Haralick, 1976).

Como os vetores médios e matrizes de variância-covariância (um análogo não-padronizado de matriz de correlação) para determinação das assinaturas espectrais das classes são extraídos através dessas amostras, a execução do treinamento responde por boa parte do acerto ou comprometimento de toda a classificação (Schowengerdt, 1983), a ponto de Hixson et al. (1980) considerar o desenvolvimento de estatísticas de treinamento relativamente mais importante para a precisão de uma classificação que a seleção da técnica de classificação e Büttner et al. (1989) atribuir o sucesso da classificação e conseqüentemente a confiabilidade das informações geradas, à qualidade do processo de treinamento.

Estas técnicas de amostragem também podem ser distinguidas em não-supervisionadas ou supervisionadas (Schowengerdt, 1983).

A primeira etapa do treinamento para os métodos de classificação não-supervisionados é a aplicação de um algoritmo de agrupamento (um "clustering method") como k-médias ou isodata (Duda & Hart, 1973) em uma área de treinamento selecionada pela sua heterogeneidade como forma de assegurar uma boa amostragem das classes possíveis e de suas variabilidades. Este programa identifica as concentrações naturais dos dados no espaço k-dimensional dos atributos, derivando de cada grupo as assinaturas espectrais correspondentes a uma classe desconhecida. O esforço de identificação destas classes pode ser efetuado pelo analista neste momento ou após a classificação de toda a imagem e pode culminar na ausência de uma correspondência explícita com as categorias do tema selecionado. Deste modo, o treinamento para o método de classificação não-supervisionada explora a estrutura interna dos dados, sem incorrer na limitação provida pela necessidade de um conhecimento prévio sobre a área focalizada. Esta característica de independência torna este tipo de treinamento útil para o desenvolvimento de esquemas híbridos de classificação multiespectral ao apontar áreas homogêneas com potencial para proverem amostras de treinamento para as técnicas supervisionadas (Schowengerdt, 1983).

Por sua vez, os métodos de classificação de imagens digitais supervisionados constroem seus cálculos a partir dos dados preliminares das classes temáticas da imagem e dos descritores estatísticos dos atributos selecionados, cuja configuração para caracterizar cada classe depende do tipo de classificação elegido. Por exemplo, o método do paralelepípedo precisa de estimativas dos valores extremos em cada atributo para cada classe, enquanto métodos de análise discriminante utilizam o vetor médio e a matriz de variância-covariância das classes (Mather, 1987 e Lark, 1994).

A fase de treinamento para a classificação supervisionada não é um processo automático, exigindo uma interação entre o analista e os dados da imagem, onde muitas vezes ciência e arte conjugam-se (Lillesand & Kiefer, 1994). Um completo conhecimento da área de estudo e a posse de dados de referência são necessários. As amostras de treinamento são normalmente obtidas através de trabalhos de campo, fotografias aéreas ou interpretação de mapas, sendo que suas posições na imagem são definidas por análise visual e delimitadas pelo operador com auxílio do "mouse".

A validade das estimativas estatísticas depende basicamente de dois fatores: o tamanho e a representatividade da amostra.

O tamanho da amostra está relacionado ao: número de atributos, ou seja bandas espectrais a serem avaliadas; número das propriedades estatísticas a serem estimadas (Swain, 1978 e Mather, 1987); tipo de classificador utilizado e o grau de heterogeneidade dos elementos das classes (Hixson et al., 1980 e Lark, 1994). Geralmente a heterogeneidade é determinada pela análise visual da composição colorida no monitor do computador. Essa análise pode incorrer em enganos causados pela equalização do histograma (Pratt, 1975) ou pela limitação de três canais para o uso das bandas nas composições coloridas (RGB), suscitando uma eventual exclusão da heterogeneidade inclusa em uma das outras bandas não utilizadas no triplete (Büttner et al., 1989).

Freqüentemente é difícil determinar o número de pixels exigido. Em extensas áreas homogêneas, como grandes monoculturas agrícolas, esta tarefa se torna relativamente fácil, ao contrário de áreas com grande variabilidade espacial e espectral, como as diferentes formações de vegetação natural (Büttner et al., 1989) ou áreas onde o parcelar agrícola ou de outro elemento da paisagem é muito reduzido, comum na estrutura fundiária dominada por minifúndios.

Quando se usa qualquer classificador baseado em informações estatísticas, o limite teórico do menor número de pixels nos quais uma área de treinamento deve estar incluída é $n+1$, onde n é o número de bandas espectrais. Na prática, um mínimo de $10n$ a $100n$ pixels como tamanho da amostra é indicado como adequado por Swain (1978), mesmo considerando a melhora das estimativas de média vetorial e matriz de covariância com o aumento do número de pixels nos grupos de treinamento (Lillesand & Kiefer, 1994). Para Mather (1987) o número de pixels por classe deveria ser no mínimo 30 vezes o número de atributos (bandas).

Além da questão da dimensão da amostra, durante a captura das áreas de treinamento é imprescindível considerar a sua representatividade. Nem sempre uma classe de cobertura terrestre é representada por uma única classe espectral. Por exemplo, a

agropecuária de uma região pode apresentar vários tipos de culturas e cada tipo de cultura por sua vez pode apresentar várias classes espectrais, originadas por diferenças nos sistemas de produção, sistemas de cultivo, datas de plantio, umidade edáfica, variedades cultivadas, condições fisiográficas e atmosféricas, ou nas combinações desses fatores e de outras variáveis. Assim sendo, todas as classes espectrais constituintes de cada classe de informação devem ser adequadamente representadas nos conjuntos de estatísticas usadas para classificar uma imagem.

Outro ponto importante a observar é a dispersão geográfica das amostras. É interessante capturar pequenas áreas de amostragem de várias partes da imagem ao invés de amostrar uma grande e única mancha, pois aumenta-se a chance das amostras de treinamento serem representativas de todas as variações do tipo de cobertura presentes na cena sob análise.

Assim o analista, seguindo uma abordagem de custo-benefício, deve fornecer um tamanho suficiente de amostras para assegurar uma determinação exata dos parâmetros estatísticos e para representar a variabilidade espectral total da cena. Após a escolha das áreas de amostragem, uma verificação é feita para analisar a qualidade dos dados, observando-se se a homogeneidade espectral e a normalidade da distribuição. Nesta fase, as áreas de treinamento com mais de uma classe espectral são recoletadas e pixels estranhos às características da amostra são eliminados. A exclusão de classes com discriminação problemática e ocorrência restrita na imagem é recomendável, evitando-se confusões com classes distribuídas extensivamente. Desse modo, admite-se um erro de identificação de uma classe ocasional na cena como recurso para preservar a exatidão da classificação de uma classe espectral análoga, porém comum na imagem (Lillesand & Kiefer, 1994).

Deste modo, o treinamento para uma classificação adequada implica em um processo de refinamento para superar omissões de classes espectrais importantes e inclusões de classes redundantes do ponto de vista computacional. Este refinamento das amostras de treinamento é geralmente um ponto crucial para melhorar a exatidão de uma classificação. Contudo uma similaridade elevada entre padrões de resposta espectral de duas categorias não pode ser superada por nenhuma técnica (Lillesand & Kiefer, 1994).

Estes aperfeiçoamentos nas amostras de treinamento são efetivados por uma ou mais das técnicas de refinamento apresentadas em seguida:

- contaminação da amostra por pixels aberrantes

A contaminação da amostra por valores atípicos também é um dos fatores geradores de erros na extração de amostras de treinamento, pois compromete a média da amostra e a matriz de variância-covariância para aquela classe (Mather, 1987). Campbell (1980) analisa as formas de detecção desses valores exóticos e propõem estimativas da média e matriz de variância-covariância mais insensíveis a presença de pixels aberrantes. Estas estimativas conferem um peso maior para observações cuja origem assume-se ser do corpo principal dos dados e menor para observações avaliadas como desvios.

A distância de Mahalanobis é empregada para identificar os elementos atípicos em relação aos membros padrões da amostra. A determinação de seu quadrado pode ser feita por:

$$d_m^2 = (\mathbf{x}_m - \bar{\mathbf{x}})' \mathbf{S}^{-1} (\mathbf{x}_m - \bar{\mathbf{x}})$$

onde m representa o índice de contagem dos elementos da amostra; \mathbf{x}_m o m -ésimo valor da amostra (vetor de pixel); $\bar{\mathbf{x}}$ o vetor médio, \mathbf{S} a matriz de variância-covariância de uma dada amostra (classe) e $\bar{\mathbf{x}}'$ a transposição do vetor $\bar{\mathbf{x}}$ (Mather, 1987).

Os elementos díspares da amostra podem ser identificados visualmente através da representação gráfica da distância de Mahalanobis, ou algumas funções derivadas, contra a distribuição das probabilidades normais (Healy, 1968). Funções dessa distância são utilizadas como ponderações para o cálculo de estimativas robustas da média e da matriz de variância-covariância (Campbell, 1980). Os pixels com valores elevados para a distância de Mahalanobis, ou seja os relativamente dissimilares da média da classe de treinamento, são atenuados considerando a forma da distribuição de probabilidade dos elementos da classe de treinamento (Mather, 1987).

- autocorrelação espacial entre pixels

A especificação do tamanho mínimo da amostra de treinamento é válida somente se pixels selecionados são independentes, situação dificilmente constatada nos casos de adjacência. Por exemplo, se a determinação de um pixel aponta-o como pertence a classe floresta, a probabilidade do elemento adjacente também integrar a mesma categoria é razoavelmente elevada. Esta correlação entre pontos próximos recebe a denominação de autocorrelação espacial.

A verificação dessa autocorrelação espacial na amostra de treinamento traduz uma super estimativa do número de elementos de informação independentes presentes na amostra. Esta correlação é subestimada pelas fórmulas estatísticas padrões, reduzindo a exatidão da classificação em execução (Mather, 1987).

Portanto um ponto importante a ser observado é o grau de autocorrelação, cujo cálculo pode ser efetuado através de blocos de pixels cujo espaçamento é definido pela seqüência 1, 2, 3 ... unidades de distância e exprimindo a correlação entre o grupo de pixels original e seus primeiros, segundos, terceiros e subsequentes vizinhos mais próximos na forma de um correlograma. Seu valor depende: da associação natural entre pixels vizinhos; da dimensão do pixel onde o aumento do tamanho do pixel é acompanhado pela diminuição da distância de autocorrelação e dos efeitos dos pré-processamento dos dados (Mather, 1987).

Segundo Campbell (1981), a matriz de variância-covariância é consideravelmente maior quando determinada de pixels randomicamente escolhidos dentro de uma classe que de um polígono contínuo de pixels de uma mesma classe. A representação gráfica dessa mesma relação através de um elipsóide definido pelo vetor médio e a matriz de variância-

covariância de pixels de contíguos e randômicos para as bandas MSS 6 e 7 do Landsat 2 resultou em uma localização dos centros das 2 elipses próxima, porém com orientação, tamanho e forma distintas. Diante desse quadro, a amostragem de pixels randômica em uma área de treinamento seria mais indicada em relação ao uso de blocos contínuos.

Para lidar com o mesmo problema, Labovitz & Masuoka (1984) utiliza uma estratégia de amostragem sistemática onde o espaçamento inter-amostras é definido pelo grau de autocorrelação espacial positiva verificado nos dados.

- representação gráfica dos padrões de resposta espectral

Diversos formatos permitem a visualização gráfica da distribuição da resposta espectral da área de treinamento. Um deles, o histograma, ao permitir uma verificação visual da normalidade da distribuição, é particularmente pertinente nos casos de aplicação do classificador de máxima verossimilhança.

Apesar da utilidade dos histogramas para ilustrar a distribuição das categorias individuais, esses gráficos não possibilitam uma comparação espectral adequada entre diferentes categorias. Nesses casos, é necessário uma forma de expressão gráfica da coincidência espectral para permitir a distinção espectral das classes. Essa representação é obtida através do já mencionado diagrama de dispersão ou coincidência espectral (“scatter diagram”) onde ilustra-se, para cada banda, a média da resposta espectral de cada tema e a variância da distribuição, indicando as sobreposições entre os padrões das respostas das categorias. Sua análise aponta quais combinações de bandas podem discriminar melhor as categorias. Quando há alta sobreposição em todos os atributos considerados, a exatidão da classificação de uma determinada classe pode ser prejudicada (Lillesand & Kiefer, 1994).

Embora os diagramas de coincidência espectral sejam uma visualização simultânea do comportamento espectral das categorias ao longo de todas as bandas, os diagramas de dispersão tridimensionais fornecem uma representação mais precisa da sobreposição entre categorias ao representar a distribuição multivariada dos dados.

- expressão quantitativa da separação das categorias

A disjunção estatística entre as características espectrais de cada classe podem ser aferidas para cada par de classes a serem classificadas. Essa distinção é expressa em formato matricial através de um parâmetro estatístico, como por exemplo, uma das muitas expressões matemáticas de divergência. Essa distância de covariância ponderada entre as médias das categorias relaciona-se geralmente de modo positivo com a separação dos padrões de treinamento, de modo que quanto maior a divergência, maior a probabilidade de identificar corretamente as classes (Lillesand & Kiefer, 1994).

- auto-classificação dos dados de treinamento

A classificação das amostras de treinamento consiste em uma outra técnica para avaliar a distinção espectral dos temas. Somente as áreas de treinamento são submetidas a uma classificação preliminar com a finalidade de determinar a exatidão da identificação das categorias dos pixels de treinamento. Essas porcentagens são geralmente apresentadas em forma de matriz de erro e são indicativas apenas do comportamento do classificador em

situações mais controladas onde as amostras de treinamento representam áreas homogêneas de cada classe da temática classificada (Lillesand & Kiefer, 1994).

- classificação interativa preliminar

Algumas funções interativas, incorporadas a maioria dos pacotes de processamento de imagens modernos, possibilitam a visualização de como as amostras de treinamento são aplicáveis para a classificação de toda a imagem. Normalmente a execução preliminar de um algoritmo computacionalmente eficiente, como a distância mínima para médias, pode ser acompanhada visualmente, apontando aproximadamente as áreas classificadas com os parâmetros estatísticos derivados das áreas de treinamento (Lillesand & Kiefer, 1994).

- classificação representativa da subcena

O resultado de uma classificação preliminar de um subconjunto representativo da variabilidade de toda imagem pode ser sobreposto interativamente sobre a imagem original, permitindo a verificação da relação entre classes selecionadas e os dados brutos (Lillesand & Kiefer, 1994).

4.5.3.1.2 Classificação não-supervisionada

No sentido estrito, a classificação não-supervisionada não requer o fornecimento pelo usuário de nenhuma informação prévia sobre as características do tema a ser avaliado. Sua utilização é apropriada nos casos da disponibilidade de pouca ou nenhuma informação detalhada sobre a distribuição dos tipos de classes. Baseado na premissa de que uma cobertura terrestre específica gera pixels com valores digitais próximos, este tipo de classificador reúne os dados da imagem em número de classes definido por esses grupos naturais existentes na imagem. Em outras palavras, esses métodos procuram localizar grupos na distribuição de pixels no espaço de atributos selecionados (Lillesand & Kiefer, 1994 e Niblack, 1985).

Avaliar a estrutura inerente dos dados, identificando os centros e limites destes grupos em um espaço p-dimensional implica em procedimentos conhecidos como **análise de agrupamento** com amplas aplicações em diversos campos (Richards, 1986). Tais agrupamentos podem ser obtidos via diferentes métodos, dos quais um de utilização comum é o algoritmo iterativo K-médias. Basicamente após a locação do número desejado de centros de agrupamentos pelo analista, os conjuntos são movimentados até atingirem a máxima separabilidade.

O algoritmo classifica a imagem pixel a pixel, agrupando-os, somente segundo os seus números digitais, em classes espectrais. O passo seguinte é submeter esses agrupamentos a um processo de reconhecimento *a posteriori*, através da comparação com algum tipo de informação de referência, para determinar a identidade e o valor significativo do grupo em termos das classes de informação.

As técnicas não supervisionadas também podem subsidiar o emprego de classificadores supervisionados, explicando uma outra denominação para a conduta não supervisionada, **classificação exploratória**. Uma das formas é através da avaliação do comportamento de uma classe espectral em um espaço de atributos p-dimensional, uma vez

que um dos pressupostos *sine qua non* da aplicação de alguns métodos é a unimodalidade da distribuição dos pontos. Na outra situação, a classificação não-supervisionada é processada como um passo preliminar para refinar a base de conhecimento sobre as classes espectrais presentes na imagem, esclarecendo se as categorias a serem adotadas pela subsequente classificação supervisionada são homogêneas espectralmente e discerníveis entre si. Por exemplo, as classes espectrais definidas por uma análise não-supervisionada podem servir de amostras de treinamento para as rotinas de classificação supervisionada, pois asseguram que todas as classes espectrais presentes na cena sejam representadas no processo de treinamento para a abordagem supervisionada (Mather, 1987 e Lillesand & Kiefer, 1994). Essa abordagem híbrida tem tido ampla aceitação devido a sua contribuição para conferir maior eficiência ao processo de classificação temática (Mayer & Fox, 1981; Wharton & Turner, 1981; Justice & Townshend, 1982 e Bouchardy & Rodriguez, 1992).

Os algoritmos baseados na análise de agrupamento, como o citado k-médias, criam arbitrariamente "sementes" no espaço de atributos multidimensional, representando cada uma um dos centros de agrupamentos definidos pelo analista. Cada píxel desconhecido na imagem é então rotulado como pertencente ao agrupamento cujo vetor médio é mais semelhante ao seu. Esta alocação de cada ponto é efetuada segundo a regra de decisão "distância mais curta para o centro" aplicada sobre a distância euclidiana relativa entre o píxel sob classificação e os centros das "sementes". Assim uma determinada classe espectral é atribuída para cada píxel se sua distância euclidiana elevada ao quadrado para o centro desta classe for menor que sua distância aos centros de outras classes. Nos casos em que as distâncias de um píxel para dois centros forem iguais, a designação é feita arbitrariamente para uma das classes. No final dessa primeira rodada do processo de classificação, as coordenadas dos centróides de cada grupo no espaço de atributo são calculadas através da obtenção da média dos valores digitais de todos os pixels em cada eixo do espaço de atributos.

Substituindo as sementes iniciais por estes novos centros, os pontos são novamente submetidos a rotulação pela mesma regra de decisão, a "distância mais curta para o centro". Após a distância de todos os pontos ter sido recalculada em relação a estes novos centros, um terceiro grupo de centróides é determinado. Estas iterações prosseguem até a posição dos vetores médios dos centros se estabilizar sem serem observadas alterações significativas. Esse processo de recálculo e realocação dos pixels fazem com que os centróides aleatoriamente escolhidos migrem em direção aos verdadeiros centróides das classes (Mather, 1987 e Lillesand & Kiefer, 1994).

Como esse processo de alocação dos pixels consiste basicamente em um seqüência ordenada das distâncias euclidianas entre os pontos e os centróides, a potência quadrada da distância pode ser usada para conservar a relação original das medidas. O cálculo algébrico da distância euclidiana elevada ao quadrado computadas também implica na multidimensionalidade do espaço de atributos (Mather, 1987).

Quando não se conhece o número de agrupamentos de pontos do conjunto, esquemas mais elaborados são necessários, tais como o algoritmo isodata (Duda & Hart, 1973). É um tipo de k-média onde é possível incluir os seguintes parâmetros adicionais para direcionar as divisões e as fusões das classes:

- um limite da distância mínima entre dois centros de agrupamento para que dois centros muito próximos sejam reunidos em um mesmo grupo;
- um limite no desvio padrão de cada banda para cada classe de modo que um agrupamento com uma variabilidade muito alta seja dividido em dois;
- um número mínimo de pixels em cada agrupamento para que os com um número de pixels inferior a esse limite tenham seus elementos alocados para outros agrupamentos;
- um número máximo de agrupamentos passíveis de fusão em qualquer iteração, evitando uma excessiva aglutinação (Niblak, 1985).

A coesão dos pontos de um agrupamento em uma região esférica do espaço de atributos é um pressuposto básico no qual esses algoritmos estão escorados. Essa concentração pode ser aferida através do conjunto de seus desvios padrões considerados um a um em cada eixo do espaço p-dimensional da imagem. Uma dimensão acentuada desses desvios padrões torna o agrupamento alongado no sentido do eixo relacionado. Uma separação adequada entre os agrupamentos acompanhada de um posicionamento dos seus centros segundo uma distância superior a um limite pré-estabelecido é uma segunda suposição desse tipo de classificador.

Sendo iterativos, os algoritmos da família k-médias fazem uso intensivo dos recursos computacionais, especialmente se os dados não possuem uma estrutura clara com agrupamento bem separados. Nessas situações, é comum, a menos que certas precauções sejam adotadas, a ocorrência de iterações em falso onde agrupamentos divididos em um ciclo são fundidos no subsequente para serem novamente dividido na terceira iteração. Neste contexto, explica-se a prática de, antes de selecionar-se a cena toda, aplicar as técnicas de classificação não supervisionadas em subáreas heterogêneas da imagem. Deste modo, assegurar-se uma boa representatividade das classes espectrais existentes para orientar a escolha do número de centros iniciais de agrupamentos e dos limites de alongamento e proximidade, cujos valores são muito variáveis de caso para caso para permitir a formulação de recomendações de emprego geral (Mather, 1987).

Uma alternativa, para simplificar, é utilizar somente a fusão de agrupamentos. O analista superestima o número esperado de centróides de agrupamentos em um processo de classificação dos pixels segundo a distância mais curta para o centro. Analogamente as técnicas anteriores, posições estáveis para os centróides são definidas. Em seguida, uma medida da coesão de cada agrupamento é determinada pela soma das potências quadradas das distâncias euclidianas dos pixels aos centros de seus respectivos agrupamentos. A raiz quadrada dessa soma dividida pelo número de pontos do agrupamento fornece a raiz quadrada do desvio quadrático médio (“root mean square deviation”) para o agrupamento. A localização dos pares de agrupamentos passíveis de uma combinação capaz de diminuir uma unidade do número de centróides e concomitantemente ampliar o menos possível o desvio dos mínimos quadrados global, é feita através do cálculo do seguinte parâmetro:

$$\frac{n_i n_j}{n_i + n_j} \sum_{k=1}^p (y_{ik} - y_{jk})^2 \quad (i = 2, k; j = 1, i - 1)$$

para cada par de centróides dos agrupamentos (y_i e y_j) onde p é o número de dimensões no espaço de atributos. Os centróides dos dois agrupamentos para os quais o parâmetro computado tem o menor valor são reunidos segundo uma média ponderada, cujos pesos são o número de pixels em cada agrupamento. Enquanto o número de agrupamentos for superior ou igual a um limite mínimo fornecido pelo analista, a realocação dos pixels e o cálculo do desvio dos mínimos quadrados global são processados para reduzir o número de centróides. O procedimento prossegue para cada número de agrupamentos entre os limites máximo e mínimo ou até a interrupção final promovida pelo operador em função da análise analógica dos resultados obtidos na classificação da imagem (Mather, 1987).

Outra opção em termos de métodos de classificação não-supervisionada adota, ao invés do conceito de coesão no espaço de atributos como uma medida de similaridade ou semelhança, os histogramas multidimensionais dos pixels sob classificação. Neste caso, considerando um exemplo tri-dimensional devido a inviabilidade da visualização de um arranjo multi-dimensional, os eixos x e y representam os atributos e o z o número de pontos relacionado com cada combinação (x,y) dos valores dos pixels. Os picos presentes na superfície resultante representam as áreas do espaço de atributos relativamente condensadas, contrapondo-se aos vales associados a regiões despovoadas na imagem. A localização dos agrupamentos é definida pelos centros desses picos (Goldberg & Shlien, 1977 e 1978; Narendra & Goldberg, 1977 e Wharton, 1980).

Outra estratégia comum para abordar a classificação não-supervisionada são algoritmos estruturados sob o conceito de sensibilidade à textura ou a rugosidade da imagem como base para a definição dos centróides dos agrupamentos. Entendendo-se a textura como a variância multidimensional presente em uma janela de observação deslocada sobre a imagem, o classificador emprega um limite de variância determinado pelo usuário, abaixo do qual a janela é avaliada como homogênea e acima como heterogênea. Os centros de agrupamento são gerados seqüencialmente a partir das médias das janelas homogêneas encontradas na imagem. Alcançando-se o teto de centróides especificado pelo operador, os dois agrupamentos com a distância no espaço de atributos mais próxima são reunidos em um novo com estatísticas combinadas, em um processo iterativo até completar a análise da imagem toda. Os agrupamentos resultantes e seus descritores cuja distância estatística é inferior a especificada pelo analista são novamente unidos, até a definição dos conjuntos finais (Lillesand & Kiefer, 1994).

Em uma outra conduta híbrida, os dados de áreas de treinamento para a classificação supervisionados podem ser mobilizados para incrementar os resultados de análise de agrupamentos nas situações em que a cobertura de certas classes é insuficiente para uma análise exclusivamente não supervisionada (Lillesand & Kiefer, 1994).

Independente qual o método de classificação não supervisionado adotado, os resultados são a identificação de classes espectrais distintas na imagem, as quais podem ser identificadas como tipos de cobertura mediante a associação com dados de referência terrestre. Entretanto, as classes espectrais podem se relacionar com mais de uma categoria de informações ou várias classes espectrais podem ser atribuídas a uma categoria de

informação. Nesses casos, essas categorias de informação são espectralmente similares e não podem ser diferenciadas no grupo de dados analisado.

4.5.3.1.3 Classificação supervisionada

Comparativamente, a abordagem supervisionada é formulada sobre a definição de categorias úteis para a análise do tema classificado e a posterior verificação de sua distinção espectral, contrapondo-se a abordagem não supervisionada, onde primeiro determina-se as classes espectralmente separáveis para depois aferir a sua utilidade para o processo de classificação. Desse modo, essa família de métodos supervisionados baseia seu processamento em uma base de conhecimentos *a priori* sobre a região imageada. Essas informações são fornecidas pelo analista no processo de treinamento já descrito no item Treinamento para a classificação. Através de observações de campo, análise de fotografias aéreas ou da base cartográfica disponível, são especificados o número de classes e amostras representativas para a definição das suas características estatísticas. A variabilidade da resposta espectral de cada classe de cobertura terrestre é descrita multidimensionalmente por uma nuvem de pontos no espaço de atributos selecionado. Os principais algoritmos desse grupo empregam as características estatísticas dessas nuvens, ou seja, o padrão de resposta espectral dos tipos de cobertura como chaves de interpretação para a alocação dos pixels desconhecidos para suas classes apropriadas no processo de classificação (Mather, 1987 e Lillesand & Kiefer, 1994). Os programas mais usuais têm uma natureza paramétrica, ou seja trabalham com a suposição de uma forma gaussiana para os vetores médios de cada classe em cada atributo elegido para o treinamento.

A natureza dos dados brutos disponíveis e do produto final a ser gerado são os parâmetros para a seleção de um classificador particular ou regra de decisão (Friedman, 1980). Alguns dos algoritmos de classificação supervisionada mais empregados ou citados pelos livros básicos são os métodos: do paralelepípedo, de distância mínima (euclidiana e de Mahalanobis) e, principalmente, o de máxima verossimilhança (Niblack, 1985; Richards, 1986; Mather, 1987; Lillesand & Kiefer, 1994 e Jensen, 1996).

A consecução de uma classificação supervisionada proveitosa e confiável dos dados digitais pode ser orientada pelo seguinte roteiro:

- elaborar uma legenda temática de classificação, compondo o conjunto de tipos de cobertura terrestre ou classes de informação para identificação da imagem;
- selecionar áreas de treinamento representativas para cada classe, de modo a garantir a inclusão da sua variabilidade espectral, através de consultas a base cartográfica disponível, produtos suborbitais e levantamentos de campo;
- selecionar o classificador adequado, cujo processamento é alimentado por estimativas de parâmetros específicos;
- alocar todos os pixels da imagem em n classes de informação através do algoritmo treinado;
- avaliar estatisticamente a exatidão da classificação e

- gerar tabelas sintéticas ou mapas temáticos para expressão da classificação produzida (Richards, 1986 e Jensen, 1996).

4.5.3.1.3.1 Classificador máxima verossimilhança gaussiano

O classificador de máxima verossimilhança (MaxVer - “maximum likelihood”) permite a inclusão de uma ponderação das distâncias às médias de cada classe no processo de alocação dos pixels (Crósta, 1992). Adotando a premissa de distribuição normal ou gaussiana dos dados, parâmetros estatísticos para os padrões de resposta espectral de cada categoria (Castleman, 1979) são calculados a partir de vetores \mathbf{x} representativos de cada pixel nas áreas de treinamento:

$$\mathbf{x} = \begin{Bmatrix} B_{xy1} \\ B_{xy2} \\ \vdots \\ B_{xyk} \end{Bmatrix}$$

onde B_{xyk} é o ND para (x,y) -ésimo pixel na banda k .

Esses dados possibilitam a descrição da distribuição da assinatura espectral de cada classe por meio de um vetor médio \mathbf{m}_i e da matriz de covariância Σ_i :

$$\mathbf{m}_i = \begin{Bmatrix} \mu_{i1} \\ \mu_{i2} \\ \vdots \\ \mu_{ik} \end{Bmatrix} \quad \Sigma_i = \begin{Bmatrix} \sigma_{i11} & \sigma_{i12} & \dots & \sigma_{i1l} \\ \sigma_{i21} & \sigma_{i22} & \dots & \sigma_{i2l} \\ \vdots & \vdots & \dots & \vdots \\ \sigma_{ik1} & \sigma_{ik2} & \dots & \sigma_{ikl} \end{Bmatrix}$$

onde μ_{ik} representa o valor médio dos NDs coletados para a classe i nas k bandas das amostras de treinamento e σ_{ikl} a covariância da classe i entre as bandas de k à l , definida por

$$\sigma_{ikl} = \sum_{p=1}^N \frac{[x_k(p) - \mu_{ki}][x_l(p) - \mu_{li}]}{(N - 1)}$$

onde N é o número de pixels na classe i da amostra de treinamento p e $x_k(p)$ e $x_l(p)$ são os valores digitais do pixel em duas bandas k e l .

Alimentando-se desses parâmetros, o algoritmo, após calcular a probabilidade estatística do vetor médio relacionado a um pixel desconhecido pertencer a cada classe de cobertura terrestre definida no processo de treinamento, associa o pixel à categoria com a probabilidade mais elevada ou o mantém sem identificação caso os valores sejam inferiores a um limite estipulado pelo analista.

Essa regra de decisão da máxima verossimilhança, cuja aplicação a um vetor \mathbf{x} correspondente a um pixel permite alocá-lo a uma determinada classe, é elaborada matematicamente através da teoria de Bayes e pode ser apresentada sinteticamente da seguinte forma:

$$\mathbf{x} \in \omega_i \text{ se } g_i(\mathbf{x}) \geq g_j(\mathbf{x}) \text{ para } j \neq i \text{ e } i = 1, \dots, M$$

onde ω_i representa uma possível classe da imagem, M o número total de classes e $g_i(\mathbf{x})$ a função discriminante

$$g_i(\mathbf{x}) = -\ln|\Sigma_i| - (\mathbf{x} - \mathbf{m}_i)' \Sigma_i^{-1} (\mathbf{x} - \mathbf{m}_i).$$

A concentração dos pixels nas amostras de treinamento de uma determinada classe pode ser expressa por uma função de densidade de probabilidade. Essa denominação é atribuída a representação gráfica da distribuição dos pixels nas áreas de treinamento em um diagrama de dispersão através de isolinhas de probabilidade. A interpretação desses contornos ao redor da média de cada classe está relacionada a probabilidade da existência de um pixel na população definida pelo processo de treinamento com um certo arranjo de NDs. Assumindo-se uma amostragem adequada da variabilidade da classe de cobertura terrestre focalizada, esses elipsóides n -dimensionais também descreveriam a probabilidade de um pixel desconhecido e localizado em um ponto do espaço de atributos estar associado a uma determinada categoria. Além disso, a forma dessas figuras equiprobabilísticas indica a sensibilidade do classificador a covariância verificada nas áreas de treinamento (Crósta, 1992 e Lillesand & Kiefer, 1994).

Além dessa abordagem de máxima verossimilhança, nas situações de disponibilidade de um conhecimento prévio sobre diferenças na probabilidade de ocorrência das séries de cobertura terrestre é recomendável o emprego de uma regra de decisão Bayesiana, baseada na seguinte função discriminante (Hord, 1982 e Richards, 1986):

$$g_i(\mathbf{x}) = \ln p(\omega_i) - \frac{1}{2} \ln|\Sigma_i| - \frac{1}{2} (\mathbf{x} - \mathbf{m}_i)' \Sigma_i^{-1} (\mathbf{x} - \mathbf{m}_i)$$

Apesar de alguns autores (Jensen, 1986 e Lillesand & Kiefer, 1994), apresentarem esse classificador como uma extensão da técnica MaxVer, Schowengerdt (1983) e Richards (1986) demonstram matematicamente que a regra de decisão de máxima verossimilhança é uma simplificação da regra Bayesiana onde, devido a inexistência de estimativas da probabilidade de ocorrência *a priori* de cada classe, assume-se uma igualdade de seus valores.

De fato, a maioria das classificações MaxVer assume probabilidades de ocorrência iguais e o mesmo custo de classificação para todas as classes. Em situações particulares, a acurácia pode ser incrementada iterativamente conferindo-se um peso a cada classe, conforme demonstrou Strahler (1980) em uma aplicação bem sucedida dessas probabilidades da ocorrência *a priori* de cada classe como recurso para aperfeiçoar a classificação, incorporando os efeitos do relevo e outras características do terreno. De um modo geral, a utilização de probabilidades *a priori* distintas é conveniente quando as

amostras de treinamento não são proporcionais à expressão da série na imagem total ou é necessário identificar todos os pixels pertencentes a uma classe, mesmo sob a pena da incorporação de um grande número de erros de comissão ou inclusão (Curran, 1985).

Para uma abordagem mais profunda sobre a teoria matemática e de probabilidades das regras de decisão empregadas no classificador de máxima verossimilhança e em sua extensão Bayesiana, fontes de referência adequadas são Haralick & Fu (1983); Schowengerdt (1983) e Richards (1986).

Devido a complexidade em termos de processamento, esse algoritmo é o mais dispendioso tanto na etapa de treinamento como na demanda de recurso da CPU, porém seus resultados são geralmente os mais precisos (Tomlins, 1981). O problema do grande número de cálculos exigidos para a classificação de cada pixel, suscitou o desenvolvimento de muitas abordagens para aumentar a sua eficiência. Por exemplo, a área de busca para cada classe pode ser restringida a partir de informações obtidas durante o treinamento e, em seguida, incorporada em um glossário ("look up table") para implementação em grupos de dados similares. Desse modo, a classificação de um pixel passa a ser um procedimento muito simplificado de indexar seu arranjo de NDs na tabela do glossário (Moik, 1980; Bolstad & Lillesand, 1991 e Ahearn & Wee, 1991). Outros autores utilizam métodos de redução de dimensionalidade dos dados, tais como a extração de componentes principais ou canônicos, como um meio de otimizar a MaxVer (Lillesand & Kiefer, 1994). Um terceira alternativa para conciliar a simplificação dos cálculos e conservação da precisão, é o emprego de classificadores do tipo árvores de decisão, estratificados ou fatiados. Sua execução é processada em uma série de etapas, separando certas classes da forma mais simplificada em cada passo. Por exemplo, a identificação de certas classes pode ser efetuada com somente 2 ou 3 bandas, tornando conveniente o uso de um classificador do tipo paralelepípedo. A inclusão de um número maior de bandas ou a aplicação de um MaxVer fica restrita àquelas categorias de cobertura terrestre com ambigüidades residuais entre as classes espectrais devido a sobreposição de características no espaço de atributos (Lillesand & Kiefer, 1994 e Jensen, 1996). Venkateswarlu & Sing (1995a) exploram uma forma híbrida de conferir maior velocidade ao processamento do MaxVer, articulando uma série de lógicas utilizadas por outros autores em iniciativas precursoras de aperfeiçoamento do algoritmo, obtendo um incremento de aproximadamente cinco vezes no desempenho do classificador.

Apesar desse inconveniente em termos de complexidade de pré-processamento e consecução do algoritmo, ao longo dos anos, o MaxVer consolidou-se como um método seguro e estável, em situações de validade da suposição de multinormalidade dos dados no espaço de atributos n-dimensional, devido a exequibilidade de seus requisitos em termos de áreas de treinamento e sua relativa solidez diante de distúrbios nos padrões de distribuição dos dados. Contudo, quando aplicado em paisagens de elevada heterogeneidade espacial, onde a forma da distribuição espectral das categorias de cobertura terrestre não obedece o critério da normalidade, esse classificador ou qualquer outro de natureza paramétrica não apresenta bons resultados na classificação e na estimativa de áreas, especialmente quando há intensas discrepâncias na extensão das classes (Maselli et al., 1990 e 1992). Com a intenção de superar essas dificuldades geradas pela correspondência reduzida entre classes de informação e espectrais, o emprego de classificadores não-paramétricos, *id est* sem

premissas sobre a natureza dos dados, foi proposto por Skidmore & Turner (1988). Mas, concomitantemente ao bom desempenho tanto na classificação como na aferição de áreas comparativamente aos métodos paramétricos convencionais, esses algoritmos são extremamente sensíveis a anomalias de distribuição e ao tamanho das amostras de treinamento, uma vez que a ausência de suposições sobre a forma dos dados permite a inclusão de distribuições sem representatividade estatística da população como um todo. Para superar essas dificuldades dos classificadores paramétricos e não-paramétricos, Maselli et al. (1992) formula um classificador híbrido baseado na inserção de probabilidades *a priori*, geradas pela técnica de Skidmore & Turner (1988), no procedimento do MaxVer, descrita por Swain & Davis (1978) e Strahler (1980), articulando as vantagens do MaxVer com a flexibilidade e vocação para estimativa de áreas dos classificadores não-paramétricos. Em 1995, Maselli et al. retomam esta abordagem e obtêm resultados na classificação espectral ainda mais precisos com a inclusão de informações auxiliares.

4.5.3.1.3.2 Classificador distância mínima

Embora esse bom desempenho do MaxVer no processo de reconhecimento de padrões e a grande disponibilidade de variações no seu emprego isoladamente ou articulado com outras técnicas explique o fato desse algoritmo ser um dos classificadores espectrais mais utilizados, sua eficiência depende inteiramente da exatidão da estimativa do vetor médio e da matriz de covariância, ou seja, exige um grande número de amostras de treinamento por cada classe. Se o número de amostras for reduzido é recomendável o uso de classificadores de distância mínima, tais como os baseados na distância euclidiana ou de Mahalanobis (Merembeck & Turner, 1980 e Venkateswarlu, 1988). Sua utilização também é sugerida como mais adequada à análise dos dados nas classificações baseadas em classes de escala nominais como floresta, pastagem, cana de açúcar etc (Jensen, 1996). Além disso, em geral, esses classificadores requerem menos tempo para a execução de seus cálculos comparado ao MaxVer, suscitando o desenvolvimento de técnicas baseadas na distância euclidiana para reduzir o tempo necessário de alocação da CPU (Hodgson, 1988; Bryant, 1988 e Venkateswarlu & Raju, 1992). Como a distância euclidiana não faz uso das informações de covariância das classes sob análise, seus resultados são insatisfatórios com classes cuja distribuição dos pixels é alongada. Já a distância de Mahalanobis, por usar informações de covariância, apresenta alguma sensibilidade ao caso de classes com distribuição alongada quando comparado com a distância euclidiana, além de ter gerado bons resultados com pequenas amostras de treinamento (Merembeck & Turner, 1980).

Esse grupo de classificadores emprega a regra de decisão da distância mínima para as médias das classes, também denominada centróide (Mather, 1987). Computacionalmente é uma das mais simples, tornando esses algoritmos pouco dispendiosos e comumente utilizados (Curran, 1985 e Jensen, 1996). Seu resultado como classificação pode ter um desempenho comparável a outros algoritmos com demandas maiores de processamento da CPU, tais como o de máxima verossimilhança (Hixson et al., 1980). Sua desvantagem está na insensibilidade à variância espectral de cada classe (Curran, 1985).

Esse método identifica os pixels desconhecidos em poucos passos. Primeiro, os vetores médios das classes propostas são calculados através dos NDs médios nas amostras

de treinamento em todas as bandas selecionadas. Na seqüência, a distância de cada pixel para cada vetor médio é determinada. Depois, os pixels de toda a imagem recebem a designação da categoria cujo vetor médio é mais próximo, separando o espaço n -dimensional em regiões de decisão, com divisas retilíneas posicionadas em equidistância de dois ou mais centróides. Por último, em muitos algoritmos de distância mínima, um delimitador é fixado ao redor dos vetores médios, tornando a região de decisão uma hipersfera, além da qual o pixel é denominado desconhecido. Esse limite assume valores variáveis em função da coesão de cada série, estimada através do desvio padrão dos pixels constituintes das amostras de treinamento nas bandas selecionadas para uma determinada classe (Mather, 1987).

Em síntese, a exploração da informação da covariância da amostra pelo classificador MaxVer é o principal ponto de distinção com o algoritmo da distância mínima euclidiano. Comparativamente, pela distância mínima euclidiana, a alocação de um pixel a uma determinada série é executada considerando-se apenas a distância da média mais próxima, não importa qual a sua direção em relação a média. Por sua vez, no MaxVer, o processo de decisão incorpora informações sobre a direção a partir da matriz de covariância. Entretanto nas situações em que a covariância da classe é dominada por algum tipo de ruído, ao invés da dispersão natural das informações de uma classe espectral, o desempenho superior do classificador MaxVer é anulado (Richards, 1986).

Existem vários meios de se calcular essa distância mínima. O mais usualmente aplicado (Richards, 1986) é a distância euclidiana baseada no teorema de Pitágoras,

$$D_{AB} = \sum_{i=1}^n |a_i - b_i|.$$

Uma outra alternativa por exemplo é a distância “round the block” (Swain & Davis, 1978),

$$D_{AB} = \sqrt{\sum_{i=1}^n (a_i - b_i)^2}.$$

4.5.3.1.3.3 Classificador Mahalanobis.

É possível decompor a função discriminante do classificador de máxima verossimilhança (vide item 4.5.3.1.3.1), adotada nas situações em que a probabilidade de ocorrência de classes *a priori* é igual, na fórmula empregada pela distância mínima euclidiana, de onde também pode ser derivada a de Mahalanobis. Primeiro, a reversão do sinal da função permite considerá-la como uma medida de distância ao quadrado, desde que o componente quadrático tenha as mesmas dimensões e o outro termo seja constante:

$$d(\mathbf{x}, \mathbf{m}_i)^2 = \ln|\Sigma_i| + (\mathbf{x} - \mathbf{m}_i)' \Sigma_i^{-1} (\mathbf{x} - \mathbf{m}_i).$$

Deste modo a classificação pela máxima verossimilhança pode ser considerada como um classificador de distância mínima, uma vez que a alocação do pixel é processada na base

da menor distância $d(\mathbf{x}, \mathbf{m}_i)$ tal qual uma distância euclidiana, com a diferença de que a medida é sensível a direção e modificada segundo a categoria analisada. (Richards, 1986).

Nos casos em que todas as covariâncias das classes são iguais, o logaritmo neperiano da expressão anterior deixa de ser diferencial, reduzindo a distância para:

$$d(\mathbf{x}, \mathbf{m}_i)^2 = (\mathbf{x} - \mathbf{m}_i)' \Sigma^{-1} (\mathbf{x} - \mathbf{m}_i)$$

A raiz quadrada desta expressão é conhecida como a distância de Mahalanobis. As vantagens do classificador baseado nessa distância são a sua maior rapidez e a retenção de sensibilidade a direção através da matriz de covariância Σ , a qual pode ser uma média de classe ou uma variância ponderada (Richards, 1986).

4.5.3.1.3.4 Classificador paralelepípedo

A popularidade do classificador paralelepípedo para aplicações de sensoriamento remoto pode ser atribuída a sua rapidez, eficiência computacional e necessidade reduzida de informações fornecidas pelo analista (Rose & Rosendahl, 1983; Mather, 1987 e Lillesand & Kiefer, 1994).

Este método agrega sensibilidade à variância de cada categoria ao estimar a faixa de ocorrência dos NDs fornecidos pelo conjunto de treinamento, para cada uma das classes temáticas em cada um dos atributos selecionados. Esses valores extremos possibilitam a delimitação de paralelepípedos no espaço de atributos n-dimensional, correlacionados cada um a uma das classes de informação. Os pixels são identificados, como pertencentes a uma determinada classe, segundo a coincidência de sua posição com uma região de decisão definida por cada paralelepípedo. Os pontos situados fora de todos os volumes definidos são considerados como desconhecidos ou como bem definiu Mather (1987), *terra incognita*. Portanto, para cada classe específica, uma estimativa dos valores mínimos e máximos dos pixels em cada uma das bandas envolvidas na classificação é obtida através das áreas de treinamento fornecidas pelo usuário.

Sua operação é análoga a um fatiamento por níveis (“density slice”), com a diferença que a formação de regiões na imagem, segundo a similaridade dos NDs, ocorre simultaneamente em todas os atributos (Curran, 1985).

A eventual sobreposição dos paralelepípedos, devido a semelhança espectral das classes escolhidas ou a alta correlação espectral dos alvos nos atributos, acarreta problemas na alocação dos pixels nas séries propostas. Uma forma simples, porém imprecisa de solucionar esse problema, é designar o ponto para o primeira região de decisão com a qual sua localização combine, tornando a ordem de análise um fator muito importante, mas de difícil sistematização. Desse modo, essa alternativa somente se torna conveniente nos casos em que os dados tem uma estruturação boa e o número de pixels em conflito de identificação é muito baixo (Mather, 1987). Uma outra abordagem é a intervenção do analista sob a conformação dos paralelepípedos sobrepostos, definindo limites sob medida para criar regiões de decisão adjacentes porém independentes (Curran, 1985). Uma forma mais elaborada para a identificação desses pixels conflitantes é aplicar uma regra de distância

mínima para a escolha da melhor classificação sobre a distância euclidiana entre o ponto sob juízo e o centróide do paralelepípedo (Mather, 1987).

Contraoendo-se a facilidade e rapidez de processamento dessa técnica, o levantamento, a partir dos dados de treinamento, apenas dos limites máximo e mínimo de cada tema em cada atributo pode implicar em: uma representação equivocada ou insuficiente das classes espectrais reais; um desperdício da informação contida nos dados intermediários das áreas de treinamento e uma suposição nem sempre verificável de que a forma da distribuição dos dados no espaço n-dimensional possa ser englobada por uma caixa. Nesse contexto, esse classificador não permite a obtenção de resultados especialmente precisos (Mather, 1987).

4.5.3.2 Avaliação da exatidão da classificação de imagens

O cuidado com a utilização de mapas nunca deve ser negligenciado. Os mapas trazem consigo erros inerentes e inevitáveis a sua geração. Um mundo tridimensional é obrigatoriamente distorcido para que possa ser registrado em uma folha de papel ou visualizado em uma tela de vídeo. Além disso um mapa deve oferecer uma visão seletiva da realidade, a fim de não esconder do usuário informações de interesse em um “mar” de detalhes (Monmonier, 1991).

A precisão dos produtos cartográficos ou resultados oriundos de sua manipulação pode ser comprometida pela inserção de erros no próprio procedimento de criação, manipulação e expressão dos mapas através de sistemas de informações geográficas ou outros aplicativos de cartografia. Os erros cartográficos, como delimitação de áreas, transferência de escala, generalização etc., podem ser distinguidos, quanto a sua origem, em intrínsecos, como os provenientes da base cartográfica utilizada para a geração dos planos de informação digitais, ou operacionais, oriundos da entrada e processamento dos mapas pelo usuário e pelos algoritmos (Keifer et al., 1991).

A geração de mapas de cobertura terrestre a partir de dados de sensoriamento remoto não foge a essa regra. Congalton & Green (1993) incluíram nas fontes geradoras de confusão entre os dados de classificação e de referência, além dos erros de classificação e de delineamento das classes, as diferenças de registro geométrico entre os dados de referência e os classificados; erros de delineamento proveniente da digitalização dos locais escolhidos para a verificação da precisão; erros de entrada de dados de referência na geração da base de dados para verificação de precisão; erro de interpretação dos dados de referência tais como, equívocos de fotointerpretação e inconsistências na identificação da heterogeneidade intra e inter-unidades de paisagem (Forman & Godron, 1986) e diferenças temporais na cobertura terrestre entre a tomada de informações dos sensores e dos dados de campo.

Corves & Place (1994) dividem em dois, os fatores que afetam a precisão e a confiabilidade das informações obtidas pela classificação de dados de sensoriamento remoto: o erro de classificação e o erro de rotulação. O primeiro é decorrente da sobreposição espectral de duas ou mais classes definidas pelo usuário e o segundo ocorre em situações onde há um conhecimento incompleto dos tipos de cobertura presentes na região por parte do analista.

Skidmore & Turner (1992) citaram dois tipos de erros que podem ser encontrados em mapas: o de localização, referente ao erro na posição dos elementos, tais como, pontos, linhas e grades; e o de atributo, no qual o elemento foi designado erroneamente a uma categoria (Veregin, 1989 e Chrisman, 1989).

Na verdade ambos os erros ocorrem conjuntamente contribuindo para o erro global, sendo difícil isolá-los (Chrisman, 1989). As técnicas para o estabelecimento da exatidão geralmente não diferenciam esses dois tipos de erros. Porém alguns autores propuseram métodos distintos para analisá-los (Hord & Brooner, 1976 e Skidmore & Turner, 1992).

Muitas vezes a avaliação do reconhecimento e da delimitação de áreas da cobertura terrestre a partir de dados de SR é subjetiva e a qualidade dos produtos gerados é desconhecida. Face a necessidade de se estabelecer critérios mínimos de qualidade para os produtos obtidos a partir do uso de técnicas de sensoriamento remoto, alguns autores sugeriram níveis mínimos para a aceitação desses produtos: Anderson (1976) propôs níveis de aceitação entre 85% e 90% e Aronoff (1982b) citou a necessidade de se estabelecer um nível mínimo de exatidão com um nível máximo de confiança, exemplificando 85% de exatidão a um nível de confiança de 95%. Watson et al. (1992) sugeriram um nível mínimo de exatidão de 95%.

Pesquisas recentes, na área de exatidão de dados espaciais e integração de dados de sensoriamento remoto em sistemas de informação geográfica tem revivido a discussão da avaliação de exatidão (Chrisman, 1989; 1991 e Janssen & van der Wel, 1994).

O uso adequado dos dados georreferenciados somente é possível se a qualidade desses dados é conhecida. No caso das classificações digitais, dada a sua complexidade, há uma maior necessidade de melhorar a confiabilidade nos resultados (Congalton, 1991). Hord & Brooner (1976) utilizam o termo exatidão da classificação para referirem-se a correspondência entre os rótulos designados para os pixels e as classes “verdadeiras”. Aronoff (1989) enfatiza o significado estatístico da exatidão, definindo a exatidão de classificação como “... the probability that the class assigned to a location on the map is the class that would be found at that location in the field...”. Story & Congalton (1986) fazem alusão à essa definição de Aronoff como exatidão do usuário ou confiabilidade, uma vez que mede o valor de um mapa para um usuário particular (Janssen & van der Wel, 1994).

A avaliação da exatidão da classificação das coberturas terrestres derivada de dados de SR tem sido reconhecida como uma ferramenta valiosa no julgamento da adequabilidade desses dados para várias aplicações (Story & Congalton, 1986; Lillesand & Kiefer, 1994; Congalton & Green, 1993 e Janssen & van der Wel, 1994), uma vez que determina o valor da informação.

É essencial que pesquisadores e usuários de dados de sensoriamento remoto tenham um forte conhecimento tanto dos fatores envolvidos como das técnicas usadas na execução de qualquer avaliação da exatidão (Congalton, 1991 e Janssen & van der Wel, 1994).

Os estudos sobre exatidão de classificação de imagens e mapas de cobertura terrestre, derivados de dados de sensoriamento remoto tem abordado erros causados por pré-processamento (Smith & Kovalick, 1985), por técnicas interpretativas manuais

(Congalton & Mead, 1983) e digitais (Story et al., 1984 e Congalton & Rekas, 1985), por sistemas imageadores (Williams et al., 1983), por técnicas de amostragem, cálculo de exatidão e comparação de resultados (Hord & Brooner, 1976; van Genderen & Lock, 1977; Ginevan, 1979; Hay, 1979; Aronoff, 1982a e Congalton et al., 1983).

Uma importante observação é que apesar de haver um limitado número de métodos para avaliar a exatidão dos dados, as definições utilizadas diferem muito de autor para autor, especialmente na avaliação de exatidão temática (Janssen & van der Wel, 1994).

Historicamente, a exatidão foi tratada mais como uma observação e frequentemente era efetuada de uma maneira ocasional. Segundo Congalton (1991), a exatidão da classificação dos dados de SR não era considerada como uma parte integrante dos projetos. Muitas vezes a avaliação foi somente qualitativa e obtida pela visitação de poucas áreas, sendo a qualidade da classificação avaliada arbitrariamente (Congalton & Green, 1993). Até os anos 80, muitos estudos registravam simplesmente um número para expressar a exatidão da classificação. Em muitos desses casos obtinha-se a exatidão genérica (“non-site-specific accuracy”) (Congalton, 1991). Segundo Mead & Szajgin (1982), a exatidão genérica é usualmente expressa como similaridade entre a área total de cada categoria classificada e a área da classe correspondente, determinada por medidas de campo ou resultante da fotointerpretação. Essa medida compara somente a área total sem considerar a sua localização. Esse tipo de estimativa, segundo Congalton (1991), pode resultar em valores super estimados de exatidão, principalmente quando ocorre compensação de erros entre categorias no balanço geral. Para superar tais problemas passaram a analisar a chamada exatidão específica, a qual considera a natureza espacial dos dados. Esse método compara dois grupos de dados (classificados e de referência) gerando uma matriz de erro a partir da qual podem ser quantificados os erros de omissão e comissão (Mead & Szajgin, 1982).

Nos anos 70, foram propostos alguns critérios e técnicas para testar a exatidão dos mapas, tais como os de Hord & Brooner (1976) e os de van Genderen & Lock (1977), os quais analisaram o problema da exatidão como a determinação de um número ótimo de amostras da verdade terrestre e de um número permissível de erros de classificação dessas amostras, assumindo que o erro de classificação de uma dada área possa ser determinado de forma dicotômica (ou seja, ou é certo ou errado). Segundo Ginevan (1979) a metodologia sugerida por ambos os trabalhos é falha do ponto de vista operacional e estatístico, por considerar apenas a probabilidade de aceitação de uma mapa de baixa exatidão (“consumer’s risk”) e o número mínimo de amostras da verdade terrestre, esquecendo de avaliar a probabilidade de rejeitar um mapa de alta exatidão (“producer’s risk”).

No início da década de 80, novas técnicas foram propostas (Rosenfield et al., 1982 e Congalton et al., 1983). Como resultado desses trabalhos e de outros, a matriz de erro (Story & Congalton, 1986) tornou-se o padrão para registrar a exatidão de classificação de dados de sensoriamento remoto. Em adição, o uso da estatística Kappa (Cohen, 1960) foi recomendada por muitos pesquisadores como uma medida aceitável da exatidão. No período transcorrido entre os anos 80 e 90 foram incorporados outros fatores no estudo da exatidão dos dados espaciais, incluindo esquemas de amostragem e tamanho de amostra (Hixson et al., 1981; Card, 1982; Rosenfield, 1982; Rosenfield et al., 1982 e Curran & Williamson, 1986), esquema de classificação e autocorrelação espacial (Hixson et al., 1980; Campbell,

1981; Congalton, 1988a e Congalton, 1991). As técnicas estabeleceram-se e a obtenção da exatidão foi incorporada como um componente fundamental de qualquer projeto de mapeamento. Outras importantes considerações tornaram-se evidentes tais como as técnicas de verificação de campo, incorporação da avaliação da exatidão em projetos globais e regionais, avaliando todas as fontes de erro nos dados espaciais, não apenas na exatidão da classificação (Congalton & Green, 1993). Por exemplo, Lunetta et al. (1991) observou o efeito do erro associado com aquisição de dados de sensoriamento remoto, tais como processamento, análise, conversão e apresentação final, no processo de tomada de decisão.

Smith et al. (1991) explorou o uso de diferentes algoritmos e seus efeitos nos mesmos grupos de dados e Fenstermaker (1991) discutiu os temas de melhoramento global em termos do programa nacional de avaliação e monitoramento ambiental (EMAP) da Agência de Proteção Ambiental (EPA) dos EUA.

4.5.3.2.1 Amostragem

Uma das formas de se verificar a exatidão de uma imagem seria observar se todos os pixels foram rotulados corretamente. A realização de uma verificação de todas partes da região em estudo torna-se inviável, devido ao tempo e alto custo (van Genderen & Lock, 1977). Assim, muitos usuários e pesquisadores utilizam de procedimentos de amostragem ou planejamento de amostragem, para estimar a exatidão da classificação (Benson et al., 1971 e Rosenfield, 1982).

Rosenfield (1982) e Dicks & Lo (1989) definem o planejamento da amostragem como o campo da estatística aplicada que envolve a determinação do tamanho da amostra, da estratégia de amostragem e do teste estatístico para avaliação das amostras. Ginevan (1979) estabeleceu alguns critérios necessários para elaborar um esquema de amostragem adequado:

- probabilidade de aceitação de uma mapa de baixa exatidão (“consumer’s risk”) baixa;
- probabilidade de rejeição de um mapa de alta exatidão (“producer’s risk”) baixa;
- requisição de um número mínimo de amostras da verdade terrestres.

Muitos pesquisadores tem expressado opiniões sobre os esquemas de amostragem apropriados para aferir a exatidão dos mapas temáticos (Hord & Brooner, 1976; Ginevan, 1979; Rhode, 1978; Fitzpatrick-Lins, 1981 e Congalton, 1988b).

Corves & Place (1994) sugerem um método alternativo para determinar a exatidão através de mapas de confiabilidade, baseados em arquivos de probabilidade. Os autores também apontam as falhas dos métodos de obtenção da exatidão elaborados a partir de amostras para áreas com poucas informações atualizadas da cobertura e/ou em áreas de difícil acesso, onde a verificação das amostras é frequentemente inviável. Além disso, indicaram a ineficiência dos métodos baseados em amostras para fornecer informações da distribuição espacial da confiabilidade do mapa temático. Outro problema observado na avaliação da exatidão é o uso de amostras de treinamento para testar o resultado da classificação, onde ocorrem situações em que a exatidão pode ser super estimada (Congalton, 1991).

4.5.3.2.1.1 Número e tamanho de amostra

O número adequado de amostras necessário para a avaliação da exatidão depende do nível de erro aceitável nos dados e do nível de confiança na estimativa, ambos estipulados pelo analista. Outros fatores significativos no estabelecimento do tamanho da amostra são o tempo e o custo do procedimento de avaliação da exatidão (Baile et al., 1994).

Vários pesquisadores estudaram a determinação do número de amostras para avaliação da exatidão de dados de SR (Hord & Brooner, 1976; van Genderen et al., 1978; Hay, 1979; Ginevan, 1979; Rosenfield et al. 1982, Aronoff, 1982 e Aronoff, 1985).

O problema do número de amostras pode ser solucionado usando a função de densidade de probabilidade (Ginevan, 1979). A distribuição binomial tem sido reconhecida como o modelo matemático apropriado para ser utilizado no processo de avaliação da exatidão (Ginevan 1979, Rosenfield & Melley 1980; Aronoff 1982a, 1982b e 1985 e Congalton, 1988b) e tem sido abordada em vários trabalhos (por exemplo, van Genderen & Lock, 1977; Ginevan, 1979; Aronoff, 1985 e Goodchild et al., 1994).

O número de amostras necessário para um determinado nível de confiança e para um determinado erro aceitável pode ser aferido a partir da distribuição binomial através da seguinte fórmula:

$$n = \frac{pq}{\left(\frac{E}{z_\alpha}\right)^2}$$

onde p é a exatidão requerida; q é o resultado de $1 - p$; E é o erro permitido ou grau de precisão desejado e z_α é a abcissa da curva normal onde define-se o início da área relacionada ao nível de significância α (Cochran, 1977 e Goodchild et al., 1994).

Outros pesquisadores tais como Rosenfield & Melley (1980) e Rosenfield (1982), usaram a aproximação da distribuição binomial para a normal.

Uma vez que a avaliação da exatidão se baseia em amostras, as estimativas obtidas requerem um intervalo de confiança para serem realmente úteis. Esse intervalo é o segmento dentro do qual há um nível de confiança pré-determinando que o valor da estimativa estará correto. O comprimento desse intervalo é afetado pelo número de amostras usado para derivar a estimativa e pelo nível de confiança próprio. Quanto maior o número, menor será o intervalo de confiança (Bailey et al., 1994).

Um procedimento para determinar um número de amostras com um número de erros tolerável capaz de satisfazer os três critérios de Ginevan (1979) pode ser desenvolvido em um ramo da estatística conhecido como amostragem de aceitação, adotado por vários pesquisadores como Vaughn (1974), Guttman et al. (1982), Aronoff (1982a) e Janssen & van der Wel (1994). Como o procedimento da amostragem de aceitação permite o cálculo do número ótimo de amostras antes da amostragem real, minimiza as observações de campo. Ginevan (1979) usou a função de densidade de probabilidade binomial segundo esta abordagem.

A amostragem de aceitação faz uso de testes de hipótese, utilizados para determinar se uma função de estimativa suporta a hipótese que uma amostra seja proveniente de uma população com um valor paramétrico específico, tal como uma distribuição normal com um determinado desvio padrão (Mikhail & Ackerman, 1976). Portanto a hipótese estatística é uma afirmação sobre a distribuição de probabilidade de uma variável aleatória. Em geral, qualquer teste estatístico requer a formulação de uma hipótese nula (H_0), a expectativa teórica de uma variável randômica, e uma hipótese alternativa (H_a). A hipótese nula sempre especifica um grupo de parâmetros teóricos sobre a distribuição de uma variável randômica contra o qual os parâmetros estimados da amostra são comparados. O resultado dessa comparação determina se o parâmetro da amostra está em concordância suficiente com a hipótese nula. Se este é o caso, há evidências suficiente para aceitar a hipótese nula como verdadeira. Caso contrário, a hipótese nula é rejeitada como falsa e a hipótese alternativa é aceita como verdadeira (Bailey et al., 1994).

Apesar destas técnicas serem utilizadas para obtenção do número de amostras necessário para obter a exatidão global e/ou das categorias, Congalton (1991) questiona a validade destas técnicas para a construção da matriz de erro. Para Congalton (1991), o pensamento tradicional sobre técnicas de amostragem não é adequado para tratar o grande volume de dados das imagens provenientes de sensoriamento remoto, sugerindo que um mínimo de 50 amostras seja coletado para cada categoria. Hay (1979) utilizando um plano de amostragem estratificada e testes estatísticos baseados na distribuição normal e em sua aproximação de Poisson, chega a mesma conclusão, recomendando o número de amostras 50.

Para Justice & Townshend (1981) a área de uma amostra deveria ser relativa a resolução espacial e a precisão geométrica da imagem gerada:

$$A = [PD[1 + 2PG]]^2$$

A fórmula indica a área mínima de um sítio de amostragem A em relação ao diâmetro do pixel no terreno PD em metros e a precisão geométrica de um pixel PG em unidades de pixels.

Outra abordagem para a determinação do tamanho ideal das amostras é baseada na porcentagem da área total de estudo (Taylor, 1989) ou da imagem digital (Congalton, 1988b). A porcentagem utilizada para definir o tamanho adequado das amostras tem variado de pesquisador para pesquisador. Tais como 0.48% a 1.4% da áreas de estudo (Congalton, 1988b); 1% (Hill & Mégier, 1988), 2.2% e 2.7% (Taylor, 1989) e 0.7% a 1.5% (Gallego & Delincé, 1991).

Apesar desses estudos, ainda não foi encontrado um método consensual. Atualmente, a seleção do número e tamanho das amostras é feita através de considerações práticas. Portanto um balanço entre a exequibilidade e a adequação estatística ainda consiste em um tema de pesquisa (Rocha, 1992).

4.5.3.2.1.2 Tipos de amostragem

Há vários métodos de amostragem que podem ser aplicados em estudos de cobertura terrestre utilizando dados de sensoriamento remoto. Os mais comuns são: o aleatório simples, o estratificado aleatório, o sistemático e o por aglomerados (também denominado por dois estágios ou aglomerados agrupado ou aninhado). Outras abordagens existem, mas são essencialmente variações desses quatro métodos.

- Amostragem aleatória simples

É um método que seleciona n unidades amostrais de N unidades na população, de tal modo que todas as amostras possuem uma chance igual de serem selecionadas (Cochran, 1977). Cada unidade amostral é identificada por um par de coordenadas, selecionada através de um gerador de números randômicos. No caso de exatidão de mapas, a amostragem aleatória é executada sem reposição.

- Amostragem estratificada aleatória

É um método de amostragem no qual um conhecimento *a priori* sobre uma população é usado para dividir ou estratificar a população em subpopulações não sobrepostas e mais similares que a população original (Barrett & Nutt, 1979). Essa abordagem é usada quando alguém quer conhecer uma informação específica sobre certa subpopulação e para aumentar a precisão das estimativas para a população inteira (Cochran, 1977). A amostragem estratificada permite um controle sobre o número de amostras observadas em cada categoria do mapa, garantindo a seleção de um número mínimo de amostras em cada estrato ou categoria (Stehman, 1996).

- Amostragem por aglomerados

A amostragem por aglomerados é um método no qual as unidades amostrais não são simples pixels mas, grupos de pixels onde cada elemento (pixel) deve ser único para somente uma unidade amostral ou grupo (Kish, 1965). A amostragem por agrupamento é executada por várias razões, incluindo conveniência e economia de custos. É mais fácil e barato, pois se visita poucas e grandes áreas ao invés de muitas e pequenas. A desvantagem deste método, quando comparado com o aleatório, é o fato da variância para um dado esforço amostral ser maior, devido a homogeneidade dos elementos e dos agrupamentos. Além da análise estatística ser mais complexa

- Amostragem sistemática

É o método no qual as unidades amostrais são selecionadas segundo algum intervalo de tempo ou espaço (Barrett & Nutt, 1979). O primeiro ponto é alocado aleatoriamente e cada unidade sucessiva é coletada em intervalos previamente especificados. A vantagem deste método é a facilidade e a conveniência na obtenção das amostras e a distribuição uniforme das amostras sobre a população (Cochran, 1977). A maior desvantagem é que o procedimento de seleção não atribui chances iguais de inclusão na amostra para cada unidade na população, implicando em uma estimativa de erro de amostragem não totalmente confiável (*id est*, a variância). Além disso, se a população contém alguma periodicidade ou

gradiente, então o espaçamento regular das unidades amostrais pode resultar em uma amostra não-representativa (Barry & Baker, 1968).

- Amostragem estratificada sistemática não-alinhada

Este método combina as vantagens da aleatoriedade e da estratificação com aspectos úteis da amostragem sistemática, enquanto evita a possibilidade da incorporação de tendências devido a presença de periodicidade (Berry & Baker, 1968). Este método está bem ilustrado em Congalton (1988) e é muitas vezes chamado de amostragem aleatória sistemática, quando amostras são aleatoriamente selecionados dentro do estrato.

Apesar da disponibilidade de vários tipos de métodos de amostragem, atualmente há poucas concordâncias a respeito da escolha adequada para aplicação em projetos de mapeamento de cobertura e uso das terras.

A amostragem aleatória simples, por exemplo, produz estimativas com propriedades estatísticas desejáveis, tais como a consistência e não tendenciosidade da estimativa da média. Uma estimativa é considerada consistente se é igual ao parâmetro populacional quando a população inteira é amostrada (*id est*, n igual a N). Uma estimativa é não tendenciosa se o valor da média da estimativa para qualquer amostra de mesmo tamanho for igual ao parâmetro populacional. Hord & Brooner (1976) usaram amostragem aleatória simples com reposição. Contudo este tipo de amostragem apresenta alguns inconvenientes, tais como seu alto custo, seu elevado dispêndio de tempo, principalmente em áreas de difícil acesso (Zonneveld, 1974), e as deficiências em amostrar adequadamente categorias de áreas reduzidas (van Genderen et al., 1978).

O uso de amostragem estratificada aleatória, no qual os estratos são definidos por tipos de cobertura das terras, é defendido por Rudd (1971), Zonneveld (1974), van Genderen et al. (1978), Ginevan (1979) e Bailey et al. (1994). Van Genderen et al. (1978) relataram a tentativa de pesquisadores em suprir as deficiências dos métodos randômicos, inserindo alguma forma de amostragem estratificada, concluindo que a amostragem estratificada aleatória parece ser o melhor método de amostragem em estudos ambientais utilizando imagens de SR. Vários outros autores utilizaram esta técnica como aponta Hay (1979). Berry & Baker (1968) utilizaram a estratificação geométrica ao invés de tipos de cobertura. Berry (1962) e Berry & Baker (1968) usaram o procedimento de amostragem estratificada sistemática não-alinhada para selecionar amostras em estudos similares e recomendaram esse procedimento para uso em testes de exatidão dos mapas de cobertura e uso das terras produzidos pelo Serviço Geológico do Estados Unidos da América (USGS).

A consistência e a inexistência de tendenciosidade das amostras obtidas pelas técnicas de amostragem aleatória simples ou estratificada aleatória explicam certa preferência por essas abordagens (Congalton, 1988b). Contudo, possuem sérios problemas em termos de verificação de campo dos pixels localizados em regiões de difícil acesso ou sem informações auxiliares.

Sob tais condições, a amostragem por aglomerado ou agrupamento tem sido freqüentemente selecionada como uma estratégia de amostragem alternativa. Entretanto, a amostragem por aglomerados tende a selecionar pixels de áreas de cobertura terrestre

previamente conhecida, desviando a amostragem e superestimando a exatidão da imagem classificada. Os agrupamento de pixels selecionados para avaliar a exatidão neste tipo de técnica representam áreas de cobertura homogênea. Isto direciona o problema para as áreas complexas, *id est* os pixels correspondentes a misturas de diferentes tipos de cobertura não são levantados pelas amostras. Em consequência, os valores de exatidão baseados em tais amostras devem ser significativamente mais altos em relação ao verdadeiro valor para todo o mapa de cobertura. Mesmo quando as amostras são representativas do mapa total, os valores numéricos de exatidão falham em fornecer indicações de como e para onde os pixels identificados como não confiáveis são alocados na imagem classificada (Corves & Place, 1994).

Outro problema relativo a amostragem por aglomerados é causado pela autocorrelação espacial dos pixels vizinhos cuja tendência é desviar as estimativas amostrais com o aumento do tamanho da unidade amostral. Da análise de autocorrelação, Congalton (1988) concluiu que a coleta de dados deve ser feita através de pequenos grupos de pixels, usando não mais que 10, ou no máximo 25 pixels por grupo.

Quanto a amostragem por aglomerados, Congalton (1988) sugere sua utilização para tratar pequenos problemas de registro nos dados ou para treinar áreas muito grandes. Mas é preciso conferir a devida atenção para a questão da independência dos pixels em uma imagem, uma vez que grupos muito grandes podem ter pouca utilidade por adicionarem poucas informações para a amostra. Todd et al. (1980) e Rhode (1978) utilizaram a amostragem por aglomerados devido a fatores econômicos.

Fitzpatrick-Lins (1978b e 1978a) e Ling et al. (1979) foram favoráveis ao uso da amostragem sistemática alinhada. As abordagens sistemáticas facilitam a coleta de dados de referência tornando mais ameno o esforço de localização dos pontos no terreno e permitindo a coleta de dados simultânea para o treinamento e para a exatidão (Congalton, 1991). Essa técnica distribui as unidades amostrais equitativamente sobre a totalidade da região de interesse e pode ser tratada como se fosse randômica caso eventuais efeitos eventuais de gradientes ou padrões periódicos presentes na população sejam anulados pela amostragem (Freund & Williams, 1972). Essas periodicidades nos erros prejudicam o uso de amostragem sistemática (Congalton, 1988a). Fitzpatrick-Lins (1981) demonstrou que essa técnica também é influenciada pela área das categorias.

Rosenfield et al. (1982) utilizou um método de amostragem estratificada sistemática não-alinhada e uma amostragem aleatória adicional dos pontos das categorias sub-representadas. Na mesma linha de abordagem, Congalton (1991) sugere a utilização de combinações de amostragem aleatória e sistemática, a fim de obter um balanço entre a validade estatística e a aplicação prática.

Congalton (1988b) executou simulações de 5 tipos de amostragens (aleatória simples, aleatória estratificada, de aglomerados, sistemática e estratificada sistemática alinhada) em três áreas distintas e concluiu que em todos os casos a amostragem aleatória simples sem reposição, seguida da amostragem aleatória estratificada estimaram as estatísticas populacionais adequadamente. Para as áreas agrícolas, a aleatória estratificada teve resultado semelhantes a aleatória simples. Ao contrário, a sistemática estratificada

alinhada e a sistemática não tiveram bons resultados para avaliar os erros de dados de sensoriamento remoto. Essas abordagens, apesar de diminuírem o esforço amostral em campo, uma vez que facilitam a localização dos pontos no terreno, superestimaram os parâmetros populacionais de áreas de agricultura e pastagem menos complexas. Ambas as técnicas não apresentaram bons resultados devido a periodicidade dos dados causados pela autocorrelação entre os erros.

Congalton (1991) e Janssen & van der Wel (1994) recomendam a amostragem estratificada para a determinação da exatidão. Stehman (1995) questiona a falta de fórmulas de estimativas dos parâmetros de interesse na obtenção da exatidão.

4.5.3.2.2 Matriz de erro

A forma mais comum para representar a exatidão da classificação de um mapa é através do estabelecimento da porcentagem da área do mapa corretamente classificada em comparação com dados de referência (também denominado verdade terrestre) (Story & Congalton, 1986). Geralmente utiliza-se uma matriz de erro para obter esta porcentagem (Hoffer, 1975; Mead & Meyer, 1977 e Card, 1982). Esse recurso permite ao analista determinar o desempenho da classificação para categorias individuais como para a classificação total (Hoffer & Fleming, 1978 e Congalton et al., 1993), favorecendo a visualização dos resultados e auxiliando na verificação da aleatoriedade dos erros.

O uso da matriz de erro para representar a exatidão tem sido recomendado por muitos pesquisadores e deveria ser adotado como padrão convencional de registro ("standard reporting convention") (Congalton, 1991). Também denominada de matriz de confusão, tabela de contingência (Story & Congalton, 1986), matriz de avaliação (Aronoff, 1985) ou matriz de erros de classificação (Chrisman, 1991), a matriz consiste de um vetor quadrado agrupado em linhas e colunas que expressam o número de unidades amostrais (pixels, agrupamentos ou polígonos) designado para uma categoria particular em relação a uma categoria real. As colunas representam geralmente os dados de referência enquanto as linhas indicam a classificação gerada com os dados de SR. A diagonal principal representará a concordância entre ambas.

Embora a matriz seja simples por si só, há algumas confusões a respeito de sua configuração ("lay-out"). Pode parecer trivial, mas as conseqüências podem ser consideráveis. Diferentes designações das linhas e das colunas na matriz indubitavelmente obstruem uma interpretação rápida. As denominações "dados classificados" (Story & Congalton, 1986 e Mather, 1987), "interpretados" (van Genderen et al., 1978), ou "observados" (Aronoff, 1982b) são utilizadas para indicar o mesmo conceito. Similarmente, "classes" ou "dados de referência" (Congalton et al., 1983 e Story & Congalton, 1986) são referidos também como "verificado" (Aronoff, 1982b), "identificado" (Hay, 1979), "conhecido" (Lillesand & Kiefer, 1994) ou "verdadeiro" (Card, 1982). Por outro lado, Curran (1985) e Aronoff, 1982b usam termos idênticos - "observado" para indicar tipos diferentes de dados, respectivamente, de "referência" e "classificados" (Janssen & van der Wel, 1994). A configuração apresentada por Story & Congalton (1986) em sua descrição de matriz de erro será adotada nesta dissertação.

A matriz de erro pode ser usada como o ponto inicial para uma série de técnicas estatísticas analíticas e descritivas (Congalton, 1991 e Janssen & van der Wel, 1994).

4.5.3.2.3 Determinação da exatidão da classificação

4.5.3.2.3.1 Técnica descritiva

A estatística descritiva mais simples é a exatidão global também denominada de proporção de pixels corretamente classificados - PCC (Veregin, 1989), que é um valor da exatidão da classificação total. A exatidão global é calculada pela divisão da soma dos números de classificações corretas (diagonal principal) pelo número total de amostras da matriz (Congalton, 1991).

A partir da matriz de erro é possível também avaliar o desempenho da classificação para cada categoria individual, um requisito para aplicações específicas (Story & Congalton, 1986).

As medidas da classificação de classes individuais podem ser calculadas usando-se duas abordagens: a exatidão do produtor e do usuário e os erros de omissão e comissão (Story & Congalton, 1986).

A exatidão do usuário expressa a probabilidade de uma amostra de uma imagem classificada estar realmente representando aquela categoria no campo e a exatidão do produtor é a probabilidade das amostras de referência de uma certa classe (verdade terrestre) serem corretamente classificadas (Story & Congalton, 1986). A exatidão do usuário é às vezes denominada de confiabilidade (Congalton & Rekas, 1985), uma vez que o usuário está interessado na confiabilidade do mapa, ou seja, quanto realmente o mapa representa o que está no terreno. É calculado dividindo-se o número de amostras corretamente classificadas de uma dada categoria pelo número total de amostras classificadas como a categoria em estudo (total da linha) (Story & Congalton, 1986 e Congalton, 1991). A exatidão do produtor (Story & Congalton, 1986 e Congalton, 1991) é obtida dividindo-se o número de amostras de uma certa categoria classificadas corretamente pelo total de amostras de dados de referência da categoria em estudo (valor da soma de toda a coluna).

Como a exatidão do usuário e do produtor podem ser de interesse para o mesmo analista, esses termos podem ser enganosos. Portanto, Janssen & van der Wel (1994) expressam a exatidão temática em termos de erro de omissão e erro de comissão.

O erro de omissão refere-se a amostra de uma certa classe de dados de referência que não são classificados como tais, também denominado erro de exclusão.

O erro de comissão refere-se as amostras de uma certa classe de dados classificados que foram erroneamente rotulados, ou seja, atribuí-se a uma categoria uma unidade pertencente a outra, também denominado erro de inclusão.

Janssen & van der Wel (1994) explica visualmente o modo como os erros de omissão e comissão são calculados. Ambos os tipos de erros são muito importantes no

estágio de treinamento interativo porque indicam qual categoria não apresenta uma distribuição representativa.

A exatidão da classificação de um mapa quando calculada a partir do número de unidades amostrais, conjunto de pixels ou segmentos, não considera a área de cada categoria. Assim, Fitzpatrick-Lins (1981) e Rosenfield et al. (1982), sugerem para o cálculo da exatidão do mapa, a introdução de um peso para cada categoria em função da sua área. Esse peso equivale a razão entre a área ocupada por cada categoria e a área total do mapa. Prisley & Smith (1987) propõem uma ponderação em função da importância das categorias face ao objeto do mapeamento.

4.5.3.2.3.2 Técnica analítica

Em adição às técnicas descritivas, a matriz de confusão pode ser utilizada para gerar técnicas estatísticas analíticas (Congalton, 1991), tais como: intervalo de confiança (Hord & Brooner, 1976; Fitzpatrick-Lins, 1981 e Rosenfield et al., 1982); teste de hipótese para decidir a aceitabilidade da exatidão de um mapa ou classe (Ginevan, 1979; Rosenfield et al. 1982 e Aronoff, 1982a), análise de variância (Rosenfield, 1982) e análise multivariada discreta (Congalton et al., 1983; Rosenfield & Fitzpatrick-Lins, 1986 e Hudson & Ramm, 1987).

Hord & Brooner (1976) vinculam a estimativa da exatidão da classificação obtida a partir de amostras ao estabelecimento de intervalos de confiança, baseados na distribuição binomial dos dados. Esta trabalho descreve um procedimento pelo qual 95% de intervalo de confiança pode ser derivado dado o tamanho da amostra e a PCC, sugerindo o uso somente do limite inferior.

Outros autores acompanhando esta abordagem, propuseram o limite inferior de um intervalo de confiança de 95% como valor de exatidão da classificação como indica Rosenfield et al. (1982). Os intervalos de confiança para PCC podem ser lidos de um monograma binomial fornecido por manuais estatísticos ou calculados usando a distribuição binomial exata. A probabilidade binomial pode ser calculada para a classificação como um todo ou para classes individuais.

Em adição à distribuição binomial discreta, a aproximação da binomial para a normal também tem sido usada para calcular os limites de confiança da exatidão estimada para a população (Aronoff, 1982a). No caso de uma aproximação da distribuição binomial à normal para grandes amostras, Rosenfield et al. (1982) sugerem a correção da continuidade. Rosenfield & Melley (1980) descrevem a correção requerida para esse ajuste. Fitzpatrick-Lins (1981) obteve a exatidão mínima de um mapa a partir do limite inferior do intervalo de confiança para a exatidão da imagem e através da introdução de um termo referente a correção de continuidade expresso em porcentagem. Essa aproximação é válida se o tamanho da amostra for grande. Este pré-requisito não é sempre satisfeito, como Ginevan (1979) notou, referindo-se a Hord & Brooner (1976). Aronoff (1982a) também comenta a inviabilidade do uso desta aproximação para teste de exatidão, devido ao alto custo decorrente do tamanho da amostra. Rosenfield et al. (1982) e Fitzpatrick-Lins (1981) salientaram que o uso dessa equação de continuidade somente se aplica quando a

amostragem é representativa de cada categoria no mapa, em função de sua área. Caso contrário, os autores sugerem a ponderação da área dessas categorias.

Para algumas aplicações, os resultados da classificação deveriam ter um PCC mínimo. Nestes casos a técnica de amostragem de aceitação deveria ser utilizada (Ginevan, 1979; Rosenfield et al., 1982 e Aronoff, 1982a). A vantagem desta análise é que o menor tamanho de amostra pode ser determinado se os riscos e a exatidão esperada são definidos previamente.

Após a execução de uma amostragem apropriada, limites de confiança e testes de hipótese podem ser aplicados tanto para categorias individuais como para a classificação de toda a imagem.

Rosenfield et al. (1982) descreveram um método baseado no teste de hipótese e na distribuição binomial acumulada para determinar se uma dada categoria apresenta o valor de exatidão esperado para o número de pontos amostrais utilizados. Para a aplicação desse teste é definido um nível crítico inferior ao número mínimo de pontos que devem ser interpretados na amostra, de forma a aceitar a hipótese que a categoria possui a exatidão específica em um dado nível de significância (ver item 4.5.3.2.1.1).

Neste teste estatístico podem ocorrer dois tipos de erros: o erro tipo I e o erro tipo II. O nível de significância (α) define a possibilidade de rejeitar H_0 erroneamente (erro tipo I). E o poder do teste ($1-\beta$) pode ser definido como a possibilidade de erroneamente aceitar H_0 para situações em que prevalece H_1 (erro tipo II).

Em geral, os termos risco do consumidor e do produtor são usualmente usados para indicar α e β , respectivamente (Aronoff, 1982a). Obviamente, o risco do consumidor define o risco de erroneamente aceitar H_0 , cuja maior consequência atinge o consumidor. Note que o termo risco do produtor e do usuário tem uma completa diferença de significado da exatidão do usuário e do produtor.

Aronoff (1982a) considerou o procedimento testado por Ginevan (1979) como estatisticamente válido e enfatizou a importância de incluir tanto o risco do consumidor e do produtor na avaliação da exatidão. De fato, tanto o risco do produtor como do consumidor deveriam ser mínimos, o que é difícil por causa de sua interdependência. Um pequeno risco do produtor pode ser obtido pelo aumento do risco do consumidor ou um aumento do tamanho da amostra.

Aronoff (1985) introduziu o valor mínimo de exatidão. Durante um teste de exatidão, ao invés de rejeitar uma classificação como insuficientemente exata, dever-se-ia avaliar o número de erros de classificação aceitáveis para uma exatidão previamente definida. Essa exatidão mínima é um índice muito útil para comparar testes de exatidão com diferentes tamanhos de amostra.

Além das técnicas mencionadas, alguns autores sugerem a utilização de técnicas de análise de variância ou análise multivariada, para a obtenção da exatidão de uma ou mais categorias separadamente, em substituição ao teste de hipótese (Aronoff, 1982a).

Freqüentemente denominadas de análise de tabela de contingência, essas duas técnicas foram sugeridas por Card (1982) como “a estrutura mais natural para avaliar a exatidão, tanto para mostrar os resultados empíricos como para uma análise estatística fácil”.

As técnicas de análise multivariada discreta introduzidas por Congalton et al. (1983) e endossadas por Rosenfield e Fitzpatrick-Lins (1986), Hudson e Ramm (1987) e outros, têm sido utilizadas amplamente como ferramentas para obter a exatidão de mapas derivados de dados de SR, mostrando-se mais adequadas ao estudo da exatidão de dados discretos das imagens digitais.

Congalton (1983) apresentou alguns inconvenientes na aplicação da análise de variância como fazer uso somente da diagonal principal da matriz de erro e requerer uma distribuição normal dos dados. Os elementos da diagonal principal da matriz de erro podem ser convertidos a uma distribuição normal através de várias transformações (Snedecor & Cochran, 1989). Outra premissa da análise de variância é a independência dos dados, cuja verificação é rara nos dados de sensoriamento remoto devido a confusão entre as categorias (Rosenfield, 1982).

Ao contrário, as técnicas de análise multivariada discreta não exigem a independência dos dados e nem requerem transformações quanto ao tipo de distribuição. Essas técnicas são desenhadas especialmente para lidar com dados categóricos (“categorical data”), justamente o caso dos dados digitais das imagens orbitais. Além disso esse método também usa a matriz de erro inteira e não apenas os elementos da diagonal.

Congalton, et al. (1983) analisou três métodos para comparação de matrizes de erro utilizando análise multivariada discreta: normalização da matriz de erro, modelo “log-linear” descrito por Fienberg (1980) e Bishop et al. (1975) e coeficientes de concordância.

4.5.3.2.3.2.1 Normalização da matriz de erro

A normalização da matriz de erro é um processo onde cada valor das células é influenciado pelos valores das outras células das linhas e colunas correspondentes, obtidas através do procedimento de ajuste proporcional iterativo. Um peso é sucessivamente acrescido as linhas e colunas de uma matriz, balanceado-as. Portanto cada célula é uma combinação de dados de referência e de classificação, sendo representativas dos erros de comissão e omissão para dada categoria. A normalização permite que os valores das células correspondentes de duas ou mais matrizes de erro possam ser comparados diretamente sem considerar diferenças entre tamanhos de amostras e inclusão de erros de omissão e comissão. Como este procedimento considera que todas as células são de igual peso ou importância, Bishop et al. (1975) propõe um ajuste para lidar com categorias de diferentes importâncias, eliminando os efeitos do tamanho da amostra enquanto incorpora erros de omissão e comissão (Congalton et al., 1983).

4.5.3.2.3.2.2 Modelo log-linear

O modelo log-linear descrito por Fienberg (1980) e Bishop et al. (1975) é um método de comparação que permite simultaneamente analisar os vários fatores com

influência sobre a exatidão da classificação. Neste método, o modelo de combinação de variáveis e suas interações que fornece um bom ajuste para a matriz de erro é escolhido usando um procedimento de seleção de modelo, semelhante ao utilizado em regressão (por exemplo, “forward selection”). O usuário pesquisa sistematicamente todas as possíveis combinações das variáveis e suas interações e opta pela combinação mais simples capaz de fornecer um bom ajuste para os dados (Congalton et al., 1983). Assim, a matriz de erro formulada com diferentes datas, algoritmos e analistas pode ser avaliada determinando os fatores necessários para explicar a exatidão da classificação encontrada.

4.5.3.2.3.2.3 Análise Kappa

Outro procedimento multivariado discreto usado na avaliação da exatidão é a análise Kappa (Cohen, 1960). É uma técnica utilizada para calcular uma medida de concordância global, baseada na diferença entre a concordância real de classificação e a probabilidade de concordância (*id est*, a concordância por puro acaso). A concordância real corresponde à concordância entre os dados da classificação e os de referência, presentes na diagonal principal. A concordância por puro acaso é dada pelo produto dos valores das linhas e colunas marginais incorporando indiretamente os elementos fora da diagonal principal (Congalton et al., 1983). Esta técnica envolve a matriz de erro completa, incluindo informações a respeito de erro de comissão e omissão (Janssen & van der Wel, 1994).

A estatística KHAT (estimativa de Kappa) é utilizada para testar a concordância entre os dados observados e os classificados contidos na matriz de erro usando a máxima verossimilhança da distribuição multinomial (Congalton & Mead, 1983). Um teste de significância de KHAT pode ser aplicado à matriz para determinar se a concordância entre a classificação e os dados de referência é significativa.

A estatística Kappa foi originalmente desenvolvido por Cohen (1960), apesar da maioria dos artigos citarem Bishop et al. (1975) como uma fonte da fórmula (Hudson & Ramm, 1987).

Essa medida de concordância, denominada de KHAT (\hat{K}), é calculado por:

$$\hat{K} = \frac{N \sum_{i=1}^r x_{ii} - \sum_{i=1}^r (x_{i+} x_{+i})}{N^2 - \sum_{i=1}^r (x_{i+} x_{+i})}$$

onde r é o número de linhas de uma matriz; x_{ii} , o número de observações na linha i e coluna i ; x_{i+} e x_{+i} , respectivamente os totais marginais das linha i e coluna i e N , o número total de observações (Bishop et al., 1975).

Congalton et al. (1983) propuseram pela primeira vez o uso do coeficiente de concordância Kappa como uma medida de exatidão para classificação de imagens orbitais. Posteriormente sua aplicação foi amplamente difundida na comunidade de SR (Congalton & Mead, 1983, Rosenfield & Fitzpatrick-Lins, 1986; Hudson & Ramm, 1987, Rocha, 1992, Janssen & van der Wel, 1994 e Fidalgo, 1994).

Dois de seus pontos positivos consistem no fato de promover uma remoção da chance casual de concordância e de todos os elementos da matriz de erro, e não apenas a diagonal principal, contribuir para o seu cálculo (Rosenfield & Fitzpatrick-Lins, 1986 e Campbell, 1987).

Apesar de sua importância, o coeficiente Kappa inicialmente proposto apresenta algumas limitações. Uma delas é considerar toda a matriz de erro, tratando igualmente os tipos de erro, sem ponderações sobre a área ou a importância de cada categoria no valor global. Cohen (1968) desenvolveu um Kappa ponderado para atender esses casos onde algumas discordâncias são mais graves que outras.

Rosenfield & Fitzpatrick-Lins (1986) propuseram um coeficiente Kappa condicional para classes individuais.

Em 1975, Bishop et al. já haviam descrito uma medida de concordância condicional K_i e sua estimativa \hat{K}_i , para a determinação da concordância entre as categorias, sendo:

$$\hat{K}_i = \frac{Nx_{ii} - x_{i+}x_{+i}}{Nx_{i+} - x_{i+}x_{+i}}$$

Rosenfield & Fitzpatrick-Lins (1986) compararam alguns métodos para o estabelecimento da exatidão de classificação como o Kappa e o Kappa condicional, o método de Turk (1979); o método de Hellden (1980) e o método de Short (1982). Sua recomendação recaiu sobre o uso dos coeficientes Kappa e Kappa condicional para a avaliação da classificação global e de categorias individuais, respectivamente. Congalton et al. (1983) e Stehman (1996) fornecem detalhes adicionais na aplicação de Kappa em sensoriamento remoto.

Hudson & Ramm (1987) alerta para a necessidade de cautela no uso do coeficientes de concordância Kappa, citando alguns erros numéricos e de formulação em alguns trabalhos, como por exemplo os erros na estatística Kappa e sua variância presentes em Congalton et al. (1983), cujo *erratum* foi publicado em Congalton et al. (1984). Rosenfield & Fitzpatrick-Lins (1986) indicam algumas versões errôneas da variância de Kappa, tais como o de Cohen (1960), de Spitzer et al. (1967), Cohen (1968) e Everitt (1968). A formulação feita por Fleiss et al. (1969) é considerada correta.

O Kappa tem sido considerado geralmente como representação da proporção de concordância obtida depois da remoção da proporção de concordância cuja ocorrência pode casual (Pe). Entretanto, Foody (1992) levanta o problema da super estimação decorrente do cálculo da proporção de acaso de concordância (Pe) no cálculo do coeficiente presente na fórmula propostos por Cohen (1960) e Rosenfield & Fitzpatrick-Lins (1986). Sua exposição indica que este elemento da equação não reflete a proporção de concordância presente na classificação menos a chance de concordância casual, uma vez que inclui a concordância real (Brennan & Prediger, 1981) ou a concordância “por causa” (Aickin, 1990), em adição a chance de concordância. O efeito desse problema é reduzir a magnitude do coeficiente Kappa e assim a exatidão aparente da classificação. Nesse mesmo trabalho, Foody (1992) sugere abordagens alternativas para o coeficiente Kappa, tais como as discutidas por Brennan & Prediger (1981) e Aickin (1990).

A estimativa KHAT padrão utilizado por Bishop et al. (1975) supõem que a amostragem seja multinomial, requisito basicamente satisfeito pela amostragem aleatória simples (Stehman, 1996). Congalton (1991) aponta a necessidade de valorizar o efeito do uso dessas fórmulas padrão em estudos cuja amostragem não é aleatória simples. Stehman (1996) apresenta um estimava alternativa de Kappa para ser empregada com uma amostragem estratificada aleatória, sob a qual é derivada a variância dessa estimativa e uma estimativa da variância baseada na amostra. Esse estimativa, denominada de KS, também é apropriada para amostragem sistemática dentro de estratos, porém a variância e a estimativa de variância apresentadas não são adequadas para uma amostragem estratificada sistemática.

Segundo Bishop et al. (1975), nas situações onde \hat{K} é assintoticamente normal, um intervalo de confiança para \hat{K} pode ser formulado a partir da sua variância ($\hat{\sigma}^2$):

$$\hat{\sigma}^2(\hat{K}) = \frac{1}{N} \left\{ \frac{\theta_1(1-\theta_1)}{(1-\theta_1)^2} + \frac{2(1-\theta_1)(2\theta_1\theta_2 - \theta_3)}{(1-\theta_2)^3} + \frac{(1-\theta_1)^2(\theta_4 - 4\theta_2)^2}{(1-\theta_2)^4} \right\}$$

sendo,

$$\theta_1 = \sum_{i=1}^r \frac{x_{ii}}{N}; \theta_2 = \sum_{i=1}^r \frac{x_{i+}x_{+i}}{N^2}; \theta_3 = \sum_{i=1}^r \frac{x_{ii}}{N} \left(\frac{x_{i+}}{N} + \frac{x_{+i}}{N} \right) \text{ e } \theta_4 = \sum_{i=1}^r \frac{x_{ij}}{N} \left(\frac{x_{i+}}{N} + \frac{x_{+j}}{N} \right)^2.$$

Conseqüentemente, o coeficiente Kappa pode fornecer informações tanto de uma simples matriz como comparar estatisticamente matrizes (Cohen, 1960).

O teste entre dois KHATs independentes permite que quaisquer duas matrizes possam ser comparadas a fim de determinar se há diferenças significativas. O teste de significância para determinar a significância da diferença entre matrizes (Cohen, 1960), é construído a partir da variância de KHAT $\hat{\sigma}^2$:

$$\frac{\hat{K}_1 - \hat{K}_2}{\sqrt{\hat{\sigma}_1^2 + \hat{\sigma}_2^2}} \sim Z$$

Desta forma, matrizes de erro gerados por muitos algoritmos de pré-classificação podem ser agora comparados dois a dois, para determinar que classificação é significativamente melhor que a outra. Pesquisadores podem também usar esses procedimentos para testar os efeitos de fatores individuais na exatidão da classificação. Contudo, esse procedimento deveria ser limitado a apenas uma variação de um fator na classificação, por vez.

e Salis et al., 1995), autoecologia de espécies vegetais (Lieberg, 1990; Aidar, 1992; Lobo, 1993; Marques, 1994 e Lobo & Joly, 1996), recuperação e implantação de matas ciliares (CEPAM, 1990) e estrutura da paisagem (Metzger, 1992 e Metzger, 1996).

As condições ambientais encontradas em Brotas, Torrinha e suas cercanias define uma paisagem onde se destaca a heterogeneidade de coberturas e usos das terras (vide Figura 5-7 e Figura 7-4) e a ausência de uma malha industrial desenvolvida e de grandes conurbações. Há elementos significativos da vegetação original preservados, inseridos em uma matriz de usos agrossilvipastoris muito diversificada (vide Figura 5-9, Figura 5-10, Figura 5-11, Figura 5-12 e Figura 5-13), e espalhados por anfiteatros de grande beleza cênica definidos pelas escarpas de cuevas arenítico-basálticas e de reverso (vide Figura 5-7a), morros testemunhos e várzeas nas planícies aluvionares (vide Figura 5-8c). Essa quadro ambiental favorável em termos de Estado de São Paulo certamente está relacionado ao fato da região não estar em nenhuma das rotas de desenvolvimento agrícola no século passado e industrial nesse século.

Esse caráter impar da área é tão relevante a ponto de incluí-la na Área de Proteção Ambiental (APA) Corumbataí-Botucatu-Tejupá, instituída em 8 de junho de 1983 através do Decreto Estadual 20.960, abrangendo uma área total de aproximadamente 650.000 ha distribuída em 23 municípios. Os remanescentes das formações fitogeográficas originais, onde se destacam os cerrados (vide Figura 5-8d), foram declarados, por esse mesmo instrumento jurídico, Zonas de Vida Silvestre. Esse tipo de unidade de conservação é ideal para promover um ordenamento territorial de áreas mistas, como as focalizadas em Brotas e Torrinha. A coexistência de atividades primárias e sistemas naturais permite articular a proteção e conservação da qualidade ambiental e dos ecossistemas com a melhoria da qualidade de vida da população local (Silva & Fornasari Filho, 1992). Entretanto, apesar desse panorama favorável, essa APA, como muitas outras, permanece no papel, sem ser implantada ou ao menos contar com sua espinha dorsal, o zoneamento ecológico-econômico, que é o instrumento técnico para o estabelecimento de normas de uso.

A seleção desta região como o objeto desta dissertação está relacionada à grande diversidade paisagística, com muitas coberturas e distintos usos das terras. Esta situação de grande variabilidade espacial em um território reduzido é muito adequada para testar os classificadores digitais de imagens de SR, pois minimiza diversos custos logísticos e operacionais tanto na aquisição, processamento e expressão de dados orbitais e de campo.

5.2 Coleta de dados

Os dados primários para a execução do projeto foram obtidos da base topográfica disponível e de imagens orbitais da área piloto nos município de Brotas e Torrinha, Estado de São Paulo, descritas a seguir.

Também foram reunidos dados complementares das fontes de informações disponíveis, tanto bibliográficas como cartográficas, sobre a região em questão, formando um conjunto de referência para auxiliar a interpretação das imagens orbitais.

5.2.1 Aquisição de mapas

A base cartográfica necessária para a execução do trabalho foi composta pelos seguintes elementos:

Mapa de vegetação natural

escala de 1:50.000 (São Paulo, 1989)

- Brotas SF-22-Z-B-III-4
- Itirapina SF-23-Y-A-I-3
- Ribeirão Bonito SF-22-Z-B-III-2
- São Carlos SF-23-Y-A-I-1

Mapas topográficos

escala de 1:10.000 (São Paulo, 1980)

- Brotas I SF-22-Z-B-III-4-NO-B
- Brotas II SF-22-Z-B-III-4-NE-A
- Brotas III SF-22-Z-B-III-4-NO-D
- Espraiado SF-22-Z-B-III-4-NE-C
- Fazenda Chão Flora SF-22-Z-B-III-4-NE-B
- Fazenda Chapadão SF-22-Z-B-III-4-NO-F
- Fazenda Magali SF-22-Z-B-III-4-SE-A
- Fazenda Marimbondo SF-22-Z-B-III-4-NE-F
- Fazenda Olivetti SF-22-Z-B-III-4-SO-B
- Fazenda Santa Cecília SF-22-Z-B-III-4-SE-C
- Fazenda Santa Isabel SF-22-Z-B-III-2-SE-F
- Fazenda São João SF-22-Z-B-III-2-SO-F
- Fazenda Varjão SF-22-Z-B-III-4-NE-E
- Gouveia SF-22-Z-B-III-2-SE-E
- Ribeirão Tamanduá SF-22-Z-B-III-4-NE-D
- São Sebastião da Serra SF-22-Z-B-III-4-SE-D
- Sítio São Sebastião SF-22-Z-B-III-4-SE-B

escala 1:50.000

- Brotas SF-22-Z-B-III-4 (Brasil, 1974)
- Itirapina SF-23-M-I-3 (Brasil, 1969)
- Ribeirão Bonito SF-22-Z-B-III-2 (Brasil, 1971)
- São Carlos SF-23-Y-A-I-1 (Brasil, 1971)

As informações sobre o solo foram obtidas do mapa de solos do levantamento semi-detalhado do Estado de São Paulo, quadricula de Brotas, na escala 1:50.000 (Almeida et al., 1981).

5.2.2 Aquisição de fotografias aéreas

Foram obtidas as fotografias aéreas pancromáticas mais recentes da área piloto, datadas do inverno de 1979, escala 1:35.000, produzidas pela Terrafoto, para auxiliarem a

interpretação dos padrões de cobertura terrestre nas imagens e a localização de áreas de interesse durante o trabalho de campo.

5.2.3 Aquisição de imagens orbitais

Foi adquirida, junto ao INPE, uma imagem do satélite Landsat-TM, órbita-ponto 220/75, quadrante C (96 x 96 km), de 29/8/92, com 7 bandas no formato digital, com 0% de cobertura de nuvens (vide Figura 5-2).

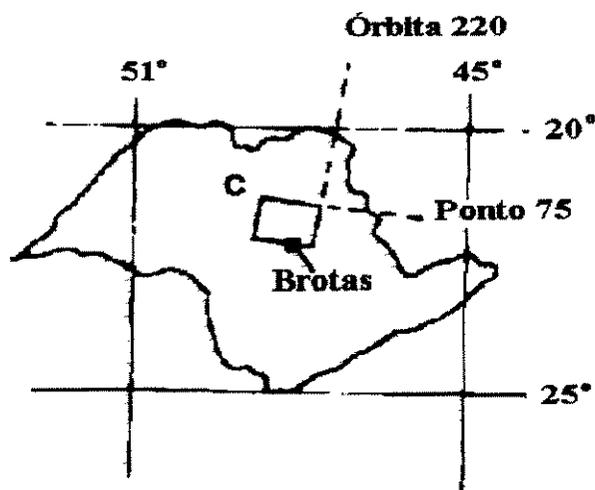


Figura 5-2 Localização geográfica da imagem da bacia hidrográfica do rio Jacaré-Pepira, São Paulo.

5.3 Caracterização da área de estudo

Através da bibliografia levantada foi possível elaborar uma descrição e análise sucinta dos principais condicionantes físicos, bióticos e antrópicos das paisagens existentes na área de estudo.

5.3.1 Hidrografia

O rio Jacaré-Pepira (Figura 5-3) tem cerca de 174 km de extensão e está localizado no centro geográfico do Estado de São Paulo, nascendo na Serra de São Pedro, localizada no planalto de Itaqueri ($47^{\circ} 55' W$ e $22^{\circ} 30' S$ a 980 m de altitude, Município de São Pedro) e desembocando no rio Tietê, na altura da represa de Ibitinga da CESP ($48^{\circ} 55' W$ e $21^{\circ} 51' S$ a 400m de altitude, Município de Ibitinga). Esta bacia hidrográfica drena uma área de 2483 km² abrangendo parte de 13 municípios (Metzger, 1992).

O rio percorre 16 km até a cota de 800 m, no lugarejo de Patrimônio de São Sebastião. Neste local existe uma barragem com 10 m de altura que forma uma represa (Maier, 1983). A barragem pertence a Centrais Elétricas de São Paulo (CESP) e acionava uma pequena usina hidrelétrica que atualmente está desativada. Após a barragem, o rio desce o reverso das cuestas basálticas apresentando-se muito encachoeirado, uma vez que

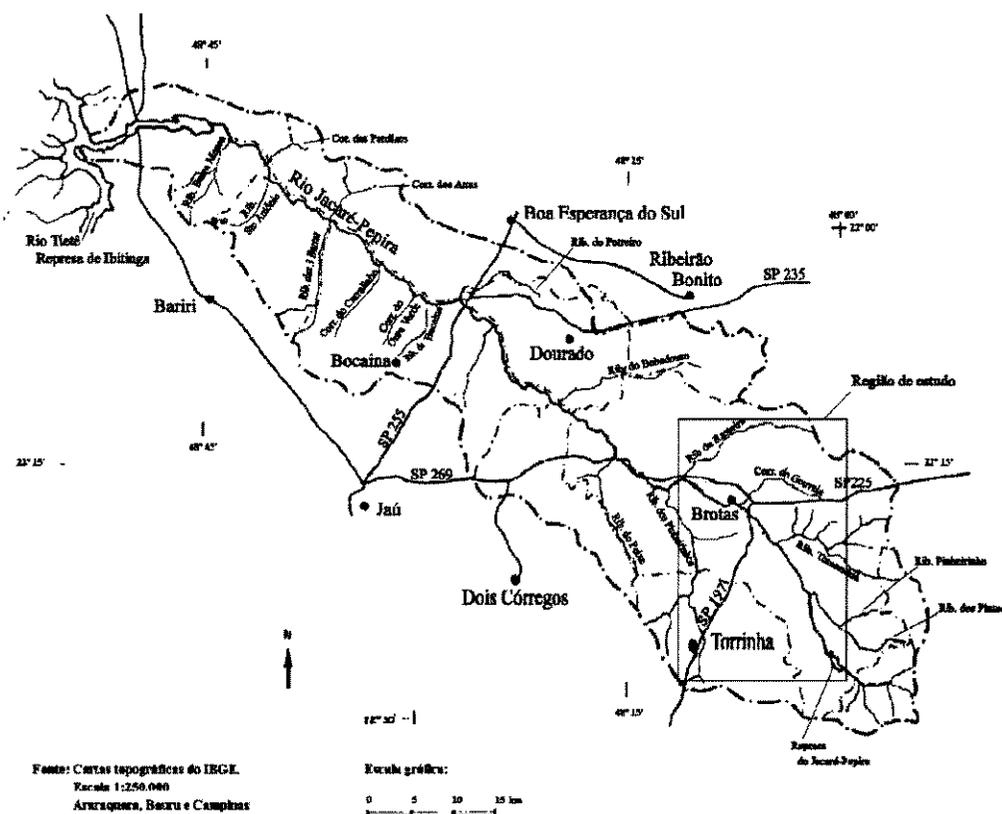


Figura 5-3. Bacia hidrográfica do rio Jacaré-Pepira, São Paulo, com alguns de seus principais afluentes (Adaptado de Aidar, 1992).

um desnível de 240m de altitude é superado em uma curta distância (25 km). Em seguida atravessa uma região plana, onde forma extensa várzea. Retornando ao encaixe em sua calha, o rio passa pelo perímetro urbano de Brotas, onde apresenta saltos e corredeiras que são atrações turísticas do município. Posteriormente o rio percorre cerca de 158 km, apresentando muitos meandros e áreas de várzea, antes de desaguar no rio Tietê, no município de Ibitinga ($48^{\circ} 55' W$ e $21^{\circ} 55' S$), a aproximadamente 400 m de altitude.

De acordo com as informações de Maier (1983) e a comunicação pessoal de C.A.Joly em 1991, o rio Jacaré-Pepira é um rio com águas não poluídas, apesar de receber esgotos domésticos *in natura* de quatro municípios. Este esgoto é assimilado e depurado pelo rio.

5.3.2 Geologia

A bacia do Jacaré-Pepira, tal qual grande parte do Estado de São Paulo, está incluída na Bacia do Paraná, uma unidade geotectônica cujo surgimento sobre a Plataforma Sul-Americana ocorreu a partir do Paleozóico (Devoniano Inferior ou Siluriano). Ao longo deste extenso período, uma camada espessa de sedimentos foi sendo acumulada. No

Jurássico Superior, graças à constante subsidência dessa Bacia, iniciaram-se novos processos tectônicos, completando com lavas basálticas e intrusões de diabásio o pacote que cobre a plataforma (Almeida & Mello, 1981).

Na bacia do rio Jacaré-Pepira, as rochas pertencem predominantemente a unidades litoestratigráficas em sua maioria datadas do Mesozóico. De acordo com o mapa geológico do Estado de São Paulo, escala 1:500.000, (Bistrichi et al., 1981), a seqüência cronológica da base para o topo presente na Bacia do Jacaré-Pepira está descrita sumariamente na Tabela 5-1.

Tabela 5-1. Coluna litoestratigráfica simplificada da bacia do rio Jacaré-Pepira, São Paulo.

Período geológico			Grupo	Formação	Rochas
Mesozóico	Triássico		São Bento ¹	Pirambóia (Trjp)	sedimentares - arenitos fluviais
	Jurássico a Cretáceo			Botucatu (Jkb)	sedimentares - arenitos eólicos
				Serra Geral (JKsg)	vulcânicas- derrames basálticos
				Intrusivas básicas tabulares (JKβ) ¹	vulcânicas - diques e lençóis diabásicos
	Bauru ¹		Adamantina (Ka)	sedimentares - arenitos fluviais	
			Itaqueri ³ (Kti)	sedimentares - arenitos argilosos	
Cenozóico	Quaternário	Holoceno		Aluviões (Qa) ²	sedimentos aluvionares

¹ rochas com pouca expressão na região.

² presente ao longo da calha do Jacaré-Pepira

³ posição cronológica controversa pois também pode ser considerada paleogênica

Na área de estudo (Figura 5-4), estão presentes praticamente todas as litologias constantes da Bacia do Jacaré-Pepira, com exceção da Formação Adamantina, restrita ao entorno da cidade de Dourado.

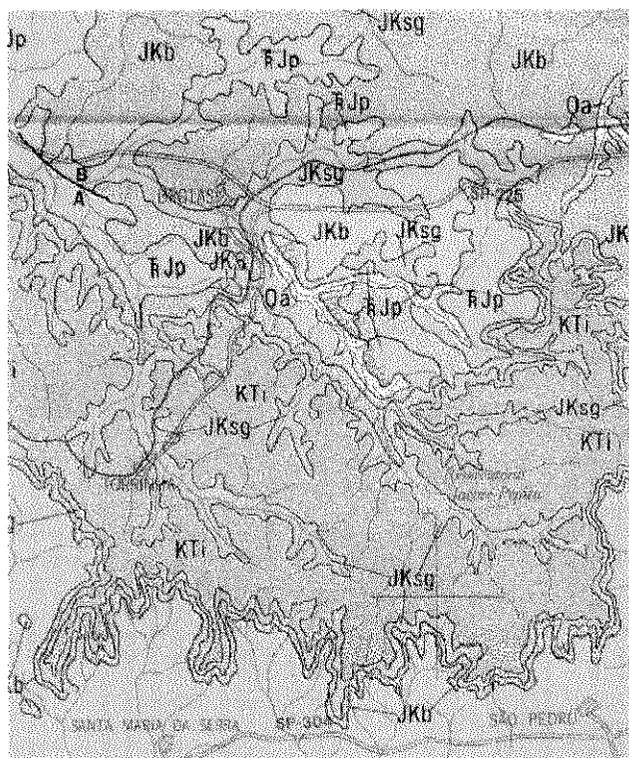
A montante da cidade de Brotas, na área de inundação de um trecho do rio Jacaré-Pepira (vide Figura 5-7c e Figura 5-8c), ocorrem sedimentos aluvionares, constituídos principalmente por argilas com alto teor de matéria orgânica.

Nessa mesma região, distanciando-se da calha do rio, localizam-se primeiramente rochas mais antigas, pertencentes a Formação Pirambóia, cujos arenitos são de deposição fluvial e têm granulação fina e média.

Descendo a coluna estratigráfica da Tabela 5-1, ocorrem os arenitos da Formação Botucatu, cuja granulação é de fina a média, de origem eólica e a estratificação cruzada. Na porção sul da área de estudo, estas litologias do Botucatu estão dispostas abaixo das rochas vulcânicas e básicas da Formação Serra Geral.

Esses derrames basálticos da Formação Serra Geral afloram encaixadas ao longo de dois trechos do rio Jacaré Pepira, a jusante de Brotas e na sua cabeceira, próximo ao Patrimônio, acompanhando a maioria de seus afluentes. Esta litologia basáltica define cuestras, acima das quais encontram-se os arenitos da Formação Itaqueri.

Na porção norte da área de estudo, a Formação Itaqueri está ausente, sendo que os arenitos da Formação Botucatu são a cobertura litológica predominante, entremeada por manchas de rochas das Formações Serra Geral e Pirambóia (Bistrichi et al., 1981).



Legenda:

- JKb - Formação Botucatu
- JKsg - Formação Serra Geral
- KTl - Formação Itaqueri
- TrJp - Formação Pirambóia
- Qa - Aluviões

Figura 5-4. Mapa geológico da área de estudo na bacia do rio Jacaré-Pepira, São Paulo (Adaptado de Bistrichi et al., 1981).

Esta distinção geológica entre os setores norte, o de colinas amplas, e sul, colinas médias, da área de estudo é marcante, como ilustra a imagem de satélite representada na Figura 7-4.

A seguir são descritas sucintamente as litologias presentes na área deste projeto.

5.3.2.1 Mezosóico

5.3.2.1.1 Grupo São Bento

A denominação de Grupo São Bento recai sobre um conjunto de arenitos e as rochas eruptivas da formação Serra Geral. Os arenitos recobrem as camadas paleozóicas e podem ser distinguidos em duas Formações: Pirambóia, componente inferior e de natureza predominantemente aquosa, e Botucatu, com origem eólica.

Essas duas formações seriam resultado de um ciclo único de deposição sedimentar, no qual as condições climáticas passaram de quentes e úmidas para desérticas. Ao final desse ciclo, quando ainda perduravam condições desérticas, iniciou-se a atividade vulcânica

basáltica da Formação Serra Geral, cobrindo o pacote sedimentar (Almeida & Melo, 1981).

5.3.2.1.1.1 Formação Pirambóia - TRJp

É uma seqüência de camadas arenosas composta por arenitos finos a médios, avermelhados, siltico-argilosos, de estratificação cruzada ou plano-paralela; níveis de folhelhos e arenitos argilosos de cores variadas e raras intercalações de natureza areno-conglomerática. Os registros de ondas ou correntes são freqüentes, ao contrário da ocorrência de fósseis conchastróides e ostracóides.

Datações radiométricas dos derrames basálticos, o registro fossilífero, a litologia e a estrutura dessa Formação indicam uma provável origem mesozóica pré-Jurássico Superior, associada a depósitos principalmente fluviais e de transbordamento, cuja formação ocorreu em um ambiente continental úmido, oxidante, com canais meandrantés, planícies de inundação e lagoas efêmeras e esporádicas (Almeida & Melo, 1981 e Petri & Fulfaro, 1988).

No Estado de São Paulo, esses arenitos estão expostos acompanhando toda a faixa de ocorrência dos depósitos sedimentares mesozóicos na Depressão Periférica, com exceção do vale do rio Grande. Na área de estudo deste projeto, estão presentes ao longo do Jacaré-Pepira e na parte norte (vide Figura 5-4).

5.3.2.1.1.2 Formação Botucatu - JKb

A Formação Botucatu é formada quase integralmente por arenitos avermelhados de boa seleção e granulação fina a média e uniforme. Essas características e sua estrutura sedimentar estratificada cruzada tangencial de médio a grande porte eólicos, apontam para origem eólica em um ambiente de sedimentação desértico, associado a dunas móveis (Almeida & Melo, 1981 e Petri & Fulfaro, 1988).

Entretanto, principalmente na porção basal, a homogeneidade dos sedimentos é interrompida por ocorrências seqüenciais e localizadas de depósitos fluviais arenos-conglomeráticos, de siltitos e argilitos lacustres, caracterizando deposição subaquosa em fácies torrenciais e estagnadas (Petri & Fulfaro, 1988).

O período cronológico provável da Botucatu é o Jurássico Médio-Superior e o Cretáceo Inferior pré-Aptiano. Entretanto, os limites possíveis vão do Triássico Superior, das camadas superiores da Formação Rosário do Sul, contemporânea da Pirambóia e com distribuição geográfica restrita ao Rio Grande do Sul, ao Neocomiano, correspondente aos primeiros derrames da Formação Serra Geral (Almeida & Melo, 1981).

Deste modo, esse grande deserto de areia, ou *erg*, advindo da acumulação de grandes dunas, com variados subambientes e acentuada aridez, perdurou até o vulcanismo basáltico. Contudo, nestas áreas de acumulação eólica também estavam presentes: 1. desertos pedregosos com fragmentos finos, ou *regs* aluviais, devido à deposição de material proveniente das regiões cristalinas periféricas e dos afloramentos de sedimentos existentes, em condições de alta energia; 2. lagoas pequenas e esporádicas nas planícies, possibilitando a formação de sedimentação hidroclásticos finos (Almeida & Melo, 1981).

Esses arenitos, em São Paulo, afloram em uma faixa ininterrupta, abaixo das serras basálticas e em suas escarpas, atingindo sua maior expressão em termos de área entre os rios Piracicaba e Mogi-Guaçu. Por sua vez, na área de estudo selecionada na Bacia do Jacaré-Pepira, o Botucatu ocupa a maior parte da região norte (vide Figura 5-4).

5.3.2.1.1.3 Formação Serra Geral - JKsg e intrusivas básicas

As rochas dessa Formação são vulcânicas toleíticas. Sua origem está relacionada a um conjunto de derrames de basaltos com coloração cinza e negra e textura afanítica. Há intercalações de arenitos intertrapeanos, com granulação fina a média e estratificação cruzada tangencial, análogos aos da Formação Botucatu, onde estão associados corpos intrusivos vulcânicos, principalmente diques e *sills* de diabásio.

Associados aos derrames de lavas, atuando como conduto em muitas situações, estão presentes um grande número de diques. São aproximadamente verticais, geralmente simples e com paredes paralelas, ocupando fendas de tração. As vezes cortam derrames e

estão conjugados a *sills*. Os *sills* são muito freqüentes nas rochas paleozóicas da Depressão Periférica e nos arenitos mesozóicos, dispendo-se geralmente de modo concordante com as rochas encaixantes.

As determinações radiométricas das rochas da Serra Geral, indicam um vulcanismo situado cronologicamente entre 147 e 119 milhões de anos passados, *id est* entre o Jurássico Superior e o Cretáceo Inferior pré-Aptiano, atingindo freqüência máxima entre 130 e 120 milhões de anos (Almeida & Melo, 1981).

Os basaltos da Formação Serra Geral provavelmente surgiram a partir de lava muito líquida cujo espalhamento ocorreu de forma rápida através de geoclastes e falhas menores, em um processo contínuo, migrando de sistema de fraturas para outro, em um ambiente desértico, conforme sugerem a homogeneidade das efusões, sua extensa distribuição, sua contemporaneidade com diques, a morfologia das dunas conservada isoladamente em intercalações eólicas, a ausência de material erodido dos basaltos no seio da Formação e a raridade de produtos piroclásticos (Almeida & Melo, 1981).

No Estado de São Paulo, os derrames estão expostos na face superior de morros testemunhos e das escarpas das cuestas basáltica (vide Figura 5-7a), penetrando pelos vales dos rios do Planalto Ocidental e podendo estender-se por grandes áreas nos planaltos de rebordo dessas cuestas, como no nordeste do Estado. Na Depressão Periférica, especificamente na região nordeste do Estado, é muito comum a presença de corpos intrusivos tabulares concordantes. Os diques de diabásio estão disseminados de modo geral, tanto em rochas sedimentares da bacia do Paraná como nas cristalinas pré-cambrianas (Almeida & Melo, 1981). Na área de estudo focalizada por este projeto, na bacia do Jacaré-Pepira, as rochas da Formação Serra Geral ocorrem principalmente ao longo das calhas de drenagem (vide Figura 5-4 e Figura 5-8b).

5.3.2.1.2 Grupo Bauru

O fim das efusões de lavas da Formação Serra Geral foi seguido por um soerguimento epirogênico generalizado da Plataforma Sul-Americana em todo o Brasil. Paralelamente, o segmento norte da Bacia do Paraná atuou como uma área negativa em relação à parte central da bacia e aos soerguimentos marginais, possibilitando a formação de áreas deprimidas localizadas, onde o Grupo Bauru depositou-se no Cretáceo Superior. Muito comum em boa parte de São Paulo, estes arenitos recobrem as lavas basálticas do Planalto Ocidental.

Apesar deste grupo estar presente na Bacia do Jacaré-Pepira através de sua Formação Adamantina, localizada próxima da cidade de Dourado, a existência deste grupo na área piloto deste projeto é duvidosa na medida que a posição litoestratigráfica da Formação Itaqueri é controversa, sendo considerada por alguns autores como pertencente ao Grupo Bauru e portanto de natureza cretácea (Bistrichi et al., 1981) ou por outros associada às coberturas cenozóicas (Almeida & Melo, 1981).

5.3.2.1.2.1 Formação Itaqueri - Kti

A litologia desta Formação caracteriza-se principalmente pela alternância dos seguintes constituintes: arenitos com cimento argiloso predominante e geralmente lentes alongadas de folhelhos e conglomerados polimíticos de extensão reduzida, muitas vezes ferruginizados (Ponçano, 1981). São arenitos de granulação variável, desde muito fina até grosseira e heterogênea, apresentando heterogeneidade mineralógica e porcentagem variável de pigmento ferruginoso. As estruturas deposicionais também são variáveis, desde estratificação plano-paralela ou cruzada nos arenitos até arranjo caótico nos conglomerados. Estes descritores indicam um ambiente de sedimentação fluvial em condições de alta energia e de alterações repentinas na velocidade de transporte dos materiais (Almeida & Barbosa, 1953).

Ponçano (1981) discute detalhadamente todas as dificuldades para se localizar adequadamente as rochas do Itaqueri na coluna litoestratigráfica da Bacia do Paraná. Resumidamente, sem desconsiderar a possibilidade de uma posição intermediária, a localização desta cobertura sedimentar pós-trapiana em uma faixa descontínua no reverso da cuesta basáltica, como pode ser constatado na bacia do Jacaré-Pepira, sugere abordá-la tanto como basal ao Grupo Bauru ou como subsequente, articulada ao soerguimento e arqueamento pós-cretáceo do leste paulista.

Na área piloto, ocorre apenas na região sul, cobrindo a maior parte de sua extensão, sendo interrompida pelas formações do Grupo São Bento, localizadas ao longo dos rios (vide Figura 5-4).

5.3.2.2 Cenozóicas

5.3.2.2.1 Aluviões (Qa)

Estes sedimentos aluvionais são principalmente compostos por areias inconsolidadas de granulação variável, argilas e cascalheiras fluviais subordinadamente, em depósitos de calha e ou terraços.

Ao longo do Jacaré-Pepira, essas planícies aluvionares são bem desenvolvidas, localizando-se na área piloto entre a cidade de Brotas e o reservatório do Patrimônio, a jusante da cuesta basáltica formada pelas rochas da Formação Serra Geral (vide Figura 5-4).

5.3.3 Geomorfologia

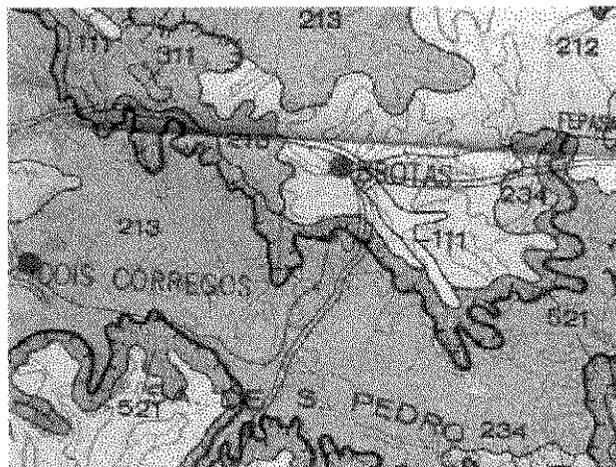
De acordo com o mapa geomorfológico do Estado de São Paulo (Ponçano et al., 1981), a bacia hidrográfica do Jacaré-Pepira possui grande parte de seu curso sobre a província das Cuestas Basálticas, tendo apenas o seu terço final sobre a província do Planalto Ocidental. As nascentes localizam-se no chamado reverso das cuestas basálticas que são uma sucessão de grandes plataformas estruturais (em basalto e arenito), inclinadas para o interior, em direção à calha do rio Paraná. Os vários riachos da nascente (vide Figura 5-8a e b) drenam por escarpas festonadas (vide Figura 5-7a e b), com desnível de 140m. Em seguida o rio passa por uma planície aluvial (vide Figura 5-7c e Figura 5-8c), continuando o seu trajeto por uma seqüência de colinas amplas, colinas médias e mesas basálticas, relevos

típicos da região de Cuestas Basálticas, para novamente atingir uma planície aluvial, já no Planalto Ocidental.

Os relevos de cuestas sustentados por rochas basálticas são uma das paisagens de maior beleza cênica das terras paulistas. Essa província caracteriza-se por uma morfologia marcada por duas feições dominantes (vide Figura 5-7b e Figura 5-9a): a escarpa e o reverso das cuestas. A primeira é um relevo alcantilado situado no divisor oriental com província da Depressão Periférica (vide Figura 5-7a). A segunda é formada uma seqüência de grandes plataformas suavizadas, com caimento noroeste, em basalto e arenito Botucatu ou Bauru (vide Figura 5-11b e c).

A atuação da dissecação sobre esses planaltos estruturais originou anfiteatros erosivos e festonamentos bem desenvolvidos, embora o entalhamento seja variável em função das particularidades de cada bacia hidrográfica (Maier, 1983).

A área de estudo está localizada em um desses planaltos estruturais, denominado Itaqueri. Tal como os outros isolados nas cuestas basálticas, seus limites são definidos por escarpas locais geralmente conhecidas por serras, as quais nesse caso são as do Itaqueri, de São Pedro, do Tabuleiro (vide Figura 5-5), de Brotas e da Cachoeira. Essas feições são encimadas por coberturas arenosas, restos dissecados da província do Planalto Ocidental (Ponçano et al. 1981).



Legenda:

111 - Planícies aluviais	234 - Morretes alongados e espigões
212 - Colinas amplas	521 - Escarpas festonadas
213 - Colinas médias	311 - Mesas basálticas

Figura 5-5. Mapa geomorfológico da área de estudo na bacia do Rio Jacaré-Pepira, São Paulo. (Adaptado de Carneiro et al. 1981)

Segundo Carneiro et al. (1981), os sistemas de relevo de ocorrência na área selecionada como objeto dessa dissertação são de três tipos.

O primeiro são os **relevos de agradação** composto por planícies aluviais (código 111) situadas junto as margens do rio Jacaré-Pepira e de alguns de seus afluentes e submetidas ao seu regime de inundações (vide Figura 5-5, Figura 5-7c, Figura 5-8c e Figura 5-9a).

O segundo grupo são os **relevos de degradação em planaltos dissecados** (vide Figura 5-7d, Figura 5-11b e Figura 5-12d). Estão presentes as seguintes feições: **colinas amplas** (código 212) com interflúvios de área geralmente superior a 4 km², topos extensos e aplainados, vertentes de perfis retilíneos a convexos, vales abertos, drenagem de baixa densidade e padrão subdendrítico; **colinas médias** (código 213) com interflúvios de áreas em torno

de 1 a 4 km², topos aplainados, vertentes de perfis convexos a retilíneos, vales abertos a

fechados, drenagem de média a baixa densidade e padrão sub-retangular e **morretes alongados e espigões** (código 234) com interflúvios sem orientação preferencial, topos angulosos a achatados, vertentes ravinadas com perfis retilíneos, vales fechados, drenagem de média a alta densidade e padrão dendrítico. Como pode ser constatada na Figura 5-5, a contribuição desse tipo de relevo para definição da paisagem da região é a mais pronunciada.

O terceiro tipo é formado pelos **relevos de transição** (vide Figura 5-7a, b e Figura 5-9a), representado por escarpas festonadas (código 521), com feições desfeitas em anfiteatros separados por espigões, topos angulosos, vertentes com perfis retilíneos, vales fechados, drenagem de alta densidade e padrão subparalelo a dendrítico. São essas feições particulares que conferem a grandiosidade e a beleza do cenário natural privilegiado de Brotas, Torrinha e vizinhança.

5.3.4 Pedologia

Segundo o levantamento pedológico de Almeida et al. (1981), na área focalizada neste trabalho, constata-se a presença dos seguintes solos, isolados ou reunidos em diversas associações: latossolos, podzólicos, terra roxas, areias quartzosas, solos litólicos e hidromórficos.

O grupo dos **latossolos** recobre a maior parte da área de estudo e pode ser diferenciado nos seguintes tipos: **roxo**, em duas subcategorias, o distrófico, com horizonte A moderado ou proeminente e da unidade Barão Geraldo (LRd), presente apenas no sopé da escarpa nas proximidades da Usina Alcooleira Paraíso, e o eutrófico, com horizonte A moderado ou chernozênico e pertencente a unidade Ribeirão Preto (Lrd), presente apenas em uma faixa simétrica a escarpa, novamente próxima à Usina Paraíso, em associação com a terra roxa estruturada (vide Figura 5-11c); **vermelho escuro**, álicos, com horizonte A moderado, os de textura média associados às unidades Dois Córregos (LE-1) e Hortolândia (LE-2) e os de textura argilosa à unidade Limeira (LE-3) e **vermelho amarelo**, os mais comuns na área considerada (vide Figura 5-11b), álicos, com horizonte A moderado e textura média, unidades Coqueiro (LV-2), Laranja Azeda (LV-3) e Três Barras (LV-4) onde o horizonte A é proeminente. O latossolo vermelho amarelo é a classe com maior frequência de voçorocas, algumas com dimensões extraordinárias como a fotografada na Figura 5-13a.

Os solos **podzólico** verificados na área de estudo têm a maior parte de suas manchas localizadas segundo o contorno das escarpas. São do tipo **vermelho amarelo** e caracterizam-se por argilas de atividade baixa. Podem ser distinguidos em eutrófico ou álico, abruptico, horizonte A moderado espesso, textura arenosa-média ou média-argilosa, unidade Serrinha (PV-1); eutrófico ou álico, abruptico, horizonte A moderado espesso, textura arenosa-média ou média-argilosa, unidade Monte Cristo (PV-2) e distrófico, horizonte A moderado, textura média ou arenosa-média, unidade Canela (PV-3).

As manchas de **terra roxa estruturada** estão situadas ao longo de alguns segmentos das escarpas na área de estudo. São solos eutróficos ou distróficos, com horizonte A moderado e podem ser separados em duas unidades: Estruturada (TE-1) e Itaguaçu (TE-2), um intermediário para latossolo roxo,

A segunda expressão em termos espaciais na área selecionada são as **areias quartzosas** profundas, álicas, com horizonte A moderado (AQ). Acompanhando toda a linha das escarpas, estão distribuídos duas categorias de **solos litólicos**, eutróficos ou distróficos. O Li-1 tem o horizonte A moderado e textura média e está posicionada sobre um substrato arenito Botucatu e Pirambóia. Com um horizonte A moderado ou chernozênico e textura argilosa, o substrato do Li-2 é basalto ou diabásico. Em pontos isolados nas margens do ribeirão do Gouveia e da Cachoeira e ao longo do ribeirão Tamanduá e de um trecho da calha do rio Jacaré-Pepira, a partir da montante da cidade de Brotas até a foz do ribeirão do Pinheirinho, ocorrem **solos hidromórficos**, do tipo gley pouco húmico e gley húmico com ou sem solo orgânico (Hi).

5.3.5 Climatologia

O clima da região de Brotas está classificado como subtropical úmido com um inverno seco e um verão chuvoso Cwa, segundo o Sistema de Köppen (Setzer, 1966). De acordo com Nicolini (1990), em um trabalho para o município de Jaú situado próximo de Brotas, a precipitação média anual da região varia entre 1100 a 1400 mm. Analisando o balanço hídrico dos dados da estação meteorológica do IAC em Jaú em uma série temporal de 26 anos (vide Figura 5-6), apresentado por Aidar (1992), o pico máximo de chuva é no mês de dezembro, atingindo aproximadamente 222 mm e o mínimo é em agosto, com cerca de 25 mm de precipitação. Entre os meses de novembro e começo de abril há um excedente hídrico médio no solo de 353 mm. A partir daí prevalece o consumo da água armazenada no solo até o final de setembro, com exceção de um breve período em junho onde há uma pequena reposição do estoque hídrico. O déficit conseqüente da utilização da água já surge

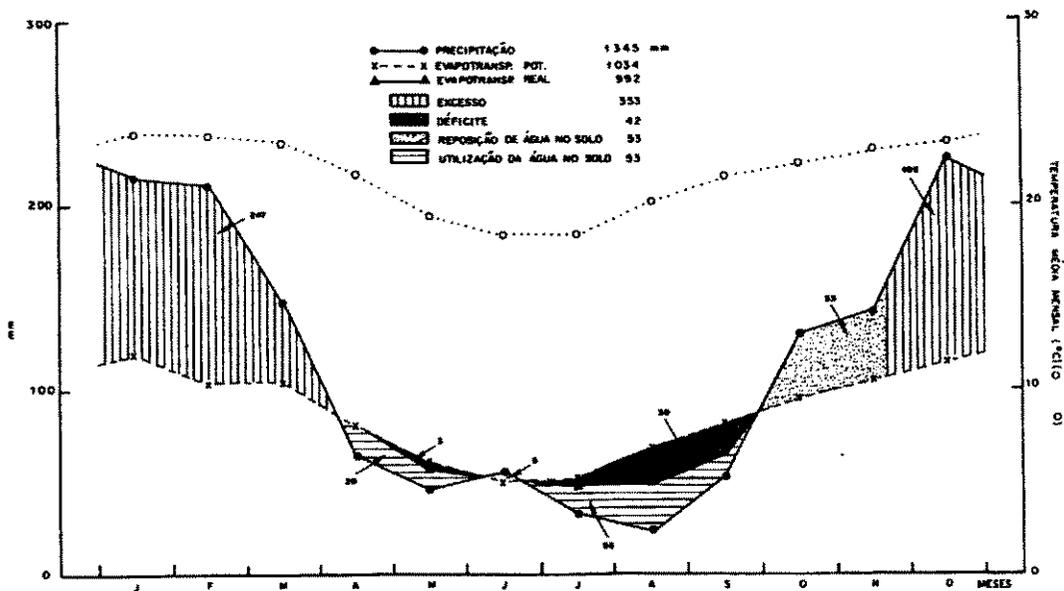


Figura 5-6. Balanço hídrico do período de 1956 - 1982 para a região de Jaú. Dados da estação experimental do IAC, 22°17'S, 48°34'W, altitude de 580m, capacidade de retenção hídrica do solo de 125 mm. (Adaptado de Aidar, 1992).

em abril e maio, porém com expressão reduzida. Seu valor médio máximo de 42 mm é definido de julho a setembro, após uma interrupção associada a recuperação verificada em junho. A reposição integral desta água ocorre de outubro a meados de novembro, com o início da estação das chuvas.

A temperatura média anual está entre 21,8 °C e 23,0 °C, sendo fevereiro o mês com temperatura média mais alta (25,1 °C) e junho com temperatura média mais baixa (18,7 °C) (Nicolini, 1990). A temperatura mínima absoluta registrada nesta região, durante o período de 1949 a 1978, foi de -0,5 °C, sendo que neste período, foram registradas a ocorrência de oito geadas.

Segundo o sistema alternativo de Nimer (1989), o clima da bacia é tropical quente e úmido com estação seca de 3 meses (junho - agosto). O verão é quente com temperatura média no mês mais quente superior a 22 °C. A média anual varia principalmente entre 18 e 20 °C. No inverno, a temperatura média é inferior a 18 °C em pelo menos um mês, a variação térmica fica entre 15 e 18 °C no mês mais frio (junho-julho) e frequência anual média de ocorrência de geadas na região é de 0 a 3 dias.

5.3.6 Paisagens

Contrapondo-se ao dramático quadro de erradicação da cobertura vegetal original do Estado de São Paulo caracterizado por Victor (1975) e atualizada por iniciativa da Fundação SOS-Mata Atlântica (SOS Mata Atlântica, 1993), a bacia do rio Jacaré-Pepira manteve-se à margem do processo de fragmentação intenso verificado em todo o sudeste da Federação, conservando uma qualidade ambiental melhor do que as regiões adjacentes. Segundo a análise do uso das terras elaborada através de uma classificação de imagem Landsat-TM5 por Metzger (1992), a cobertura florestal (vide Figura 5-7a, b, Figura 5-8a, b e c) respondia por 8,2 % da área da bacia em 1990, valor elevado se comparado com a proporção de 1,18 % de florestas no interior do Estado de São Paulo (SOS Mata Atlântica, 1993). Além disso, há 9,64 % de campos úmidos e várzeas (vide Figura 5-8c e Figura 5-12b) situadas principalmente ao longo da calha principal da bacia. Entretanto, apesar dessa situação privilegiada no âmbito do panorama estadual, a vegetação natural da bacia e a paisagem como um todo foi bastante afetada pelo ciclo econômico do café no século passado, cujos remanescentes ainda podem ser encontrados em muitos pontos da região (vide Figura 5-11a), e da cana de açúcar nas últimas décadas (vide Figura 5-11b, c e d), entremeada pela expansão das pastagens sobre as culturas decadentes em solos empobrecidos (vide Figura 5-9 e Figura 5-13c). Enquanto os remanescentes de cerrados mais abertos são quase inexistentes, os fragmentos florestais (incluindo as matas e os cerradões, como o representado na Figura 5-8d) encontram-se com níveis de degradação variáveis em muitos pontos. Apenas a montante do rio Jacaré-Pepira ainda podem ser encontrados corredores de mata ciliar com uma continuidade maior (vide Figura 5-8c).

Desse modo, a paisagem da bacia do Rio Jacaré-Pepira é eminentemente rural e caracteriza-se pela ocorrência de mosaicos de manchas de formações florestais e savânicas (cerrado *sensu lato*, cerradões, mata mesófila semi-decídua e mata ciliar) sobrepostas sob uma matriz agrossilvipastoril. Essa configuração ficou muito bem caracterizada pelo mesmo trabalho de Metzger (1992), uma vez que 44,46 % da área é ocupada por pastagens; 29,13

% por agricultura; 4,82 % por reflorestamentos industriais (vide Figura 5-9d) ou particulares, muitas vezes abandonados como o ilustrado pela Figura 5-13b; 0,42 % por conurbações (vide Figura 5-12d), contraposto aos 8,2 % de florestas e 9,64 % de campos úmidos. Há ainda 1,26 % relativos a corpos d'água (vide Figura 5-12c) e 2,14 % de áreas analisadas, completando os 2314,81 Km² avaliados, correspondendo a 93,20 % da superfície total da bacia. A classe das pastagens, por ser a de maior expressão, abriga uma série de situações distintas desde pastagens abandonadas e sujas até as manejadas por sistemas de produção desenvolvidos, como fica patente na Figura 5-9. Outro uso das terras cuja participação na configuração da paisagem tem sido ampliada muito nos últimos anos é a citricultura, onde ainda podem ser detectados muitos estágios distintos em razão de sua expansão ser recente (vide Figura 5-10). Nesse trabalho também fica muito bem evidenciada a distribuição heterogênea dos tipos de cobertura das terras segundo unidades geomorfológicas em função das diferenças edáficas entre o reverso da cuesta, planície e vertente (vide Figura 6-2).

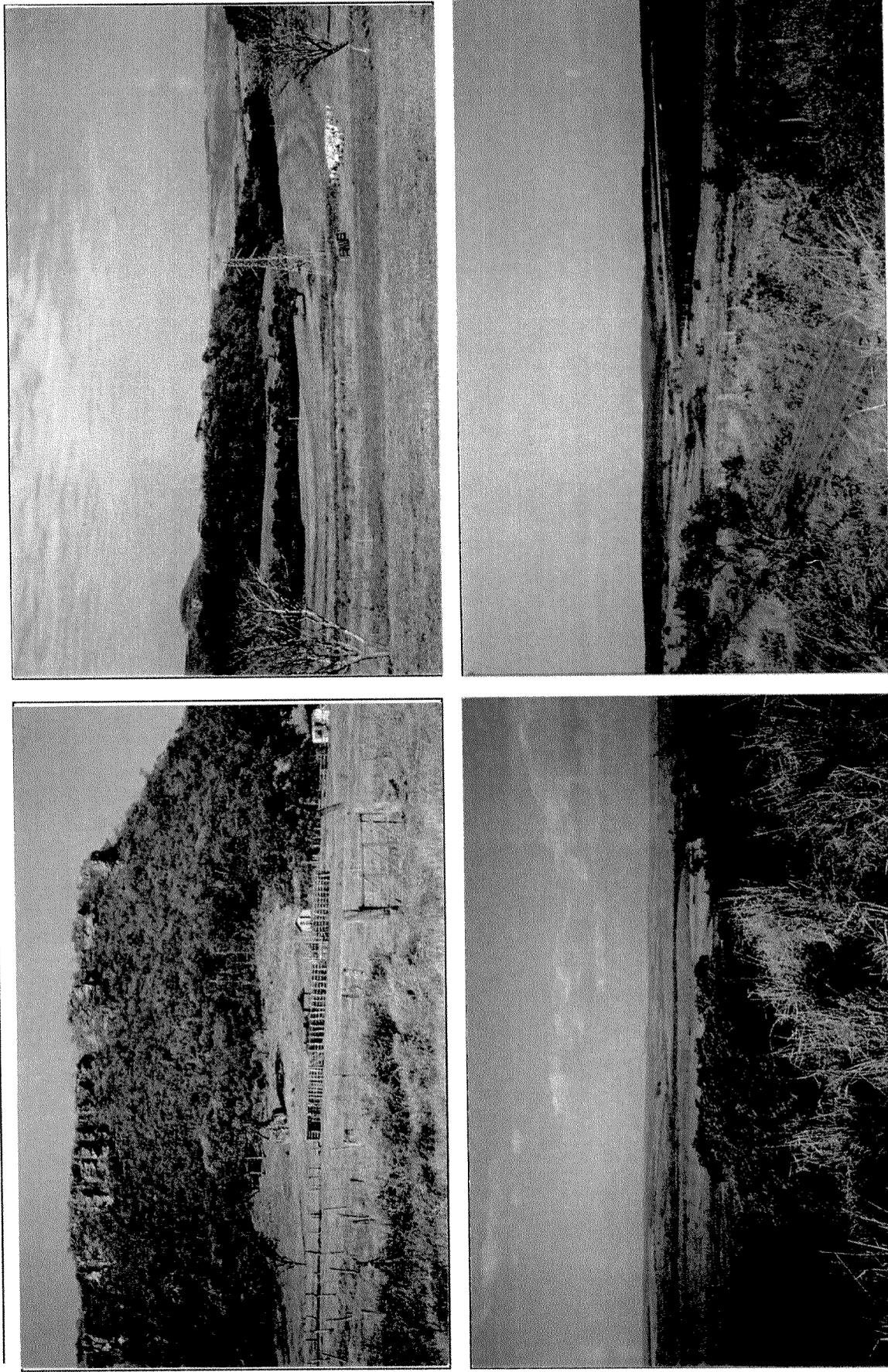


Figura 5-7. Paisagens da área de estudo na bacia do rio Jacaré-Pepira, Brotas e Torrinha, SP, com destaque para as florestas semi-decíduas localizadas junto as escarpas (*a* e *b*), a planície ao longo da calha do Jacaré-Pepira (*c*) e as colinas amplas, típicas de parte da região (*d*).

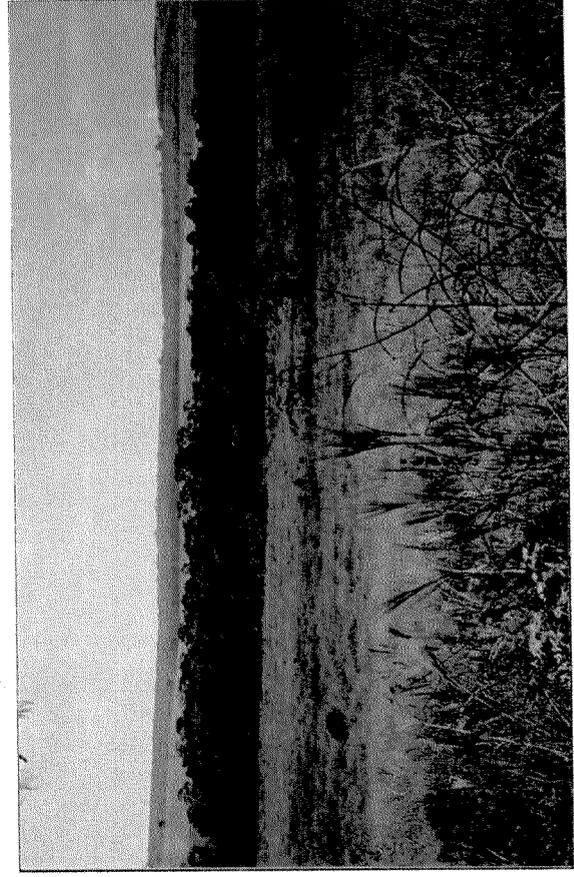
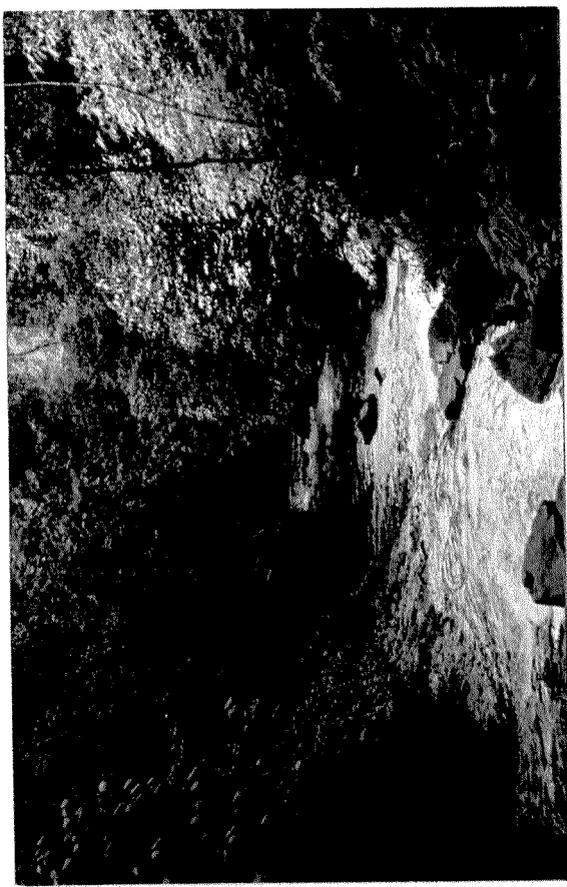
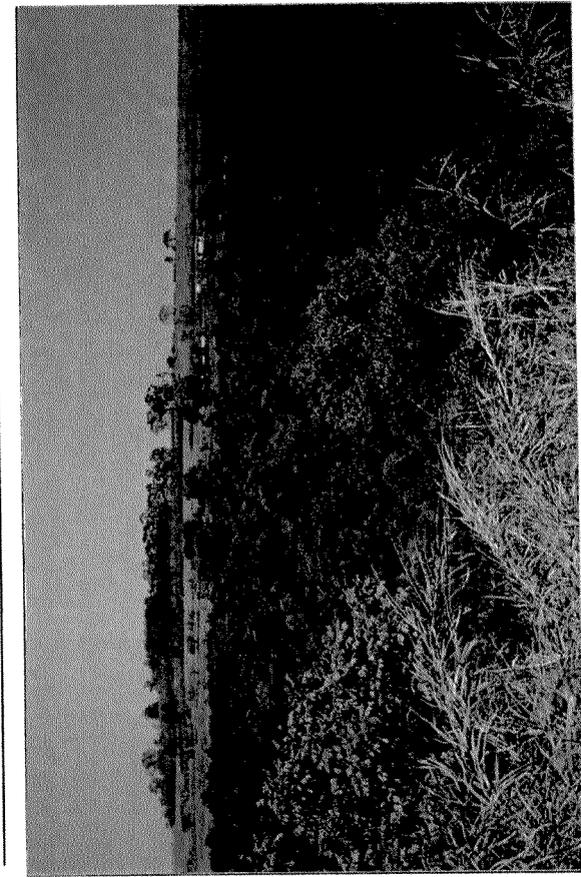


Figura 5-8. Formações vegetais comuns na área de estudo na bacia do rio Jacaré-Pepira, município de Brotas e Torrinha, SP, incluindo florestas semi-decíduas (*a*), ripárias (*b* e *c*) e cerradões (*d*).

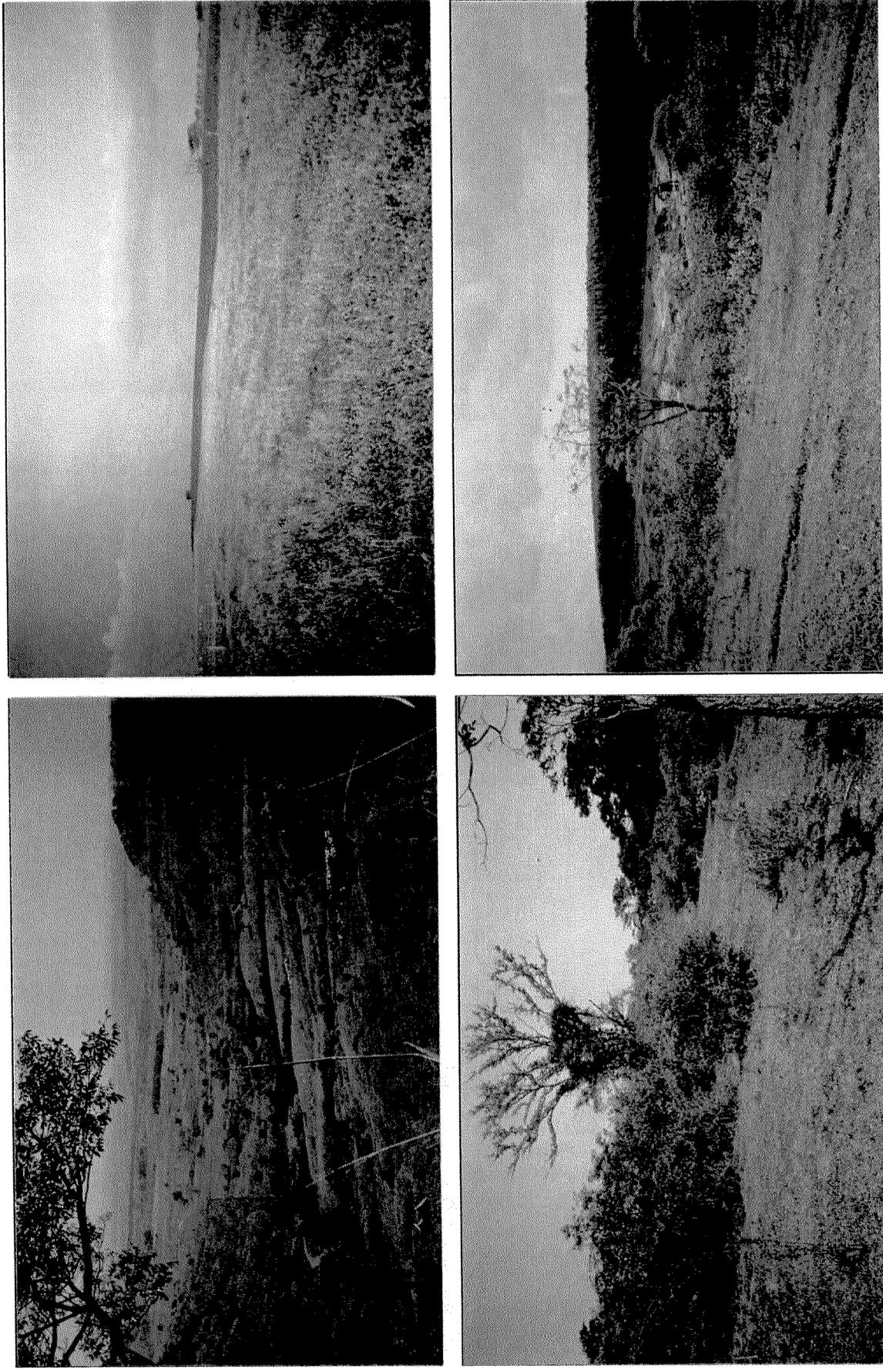


Figura 5-9. Usos das terras na área de estudo na bacia do rio Jacaré-Pepira, município de Brotas e Torrinhã, SP, incluindo pastagens com diferentes sistemas de produção (*a, b, c e d*), além de reflorestamentos de eucalipto (*d*).

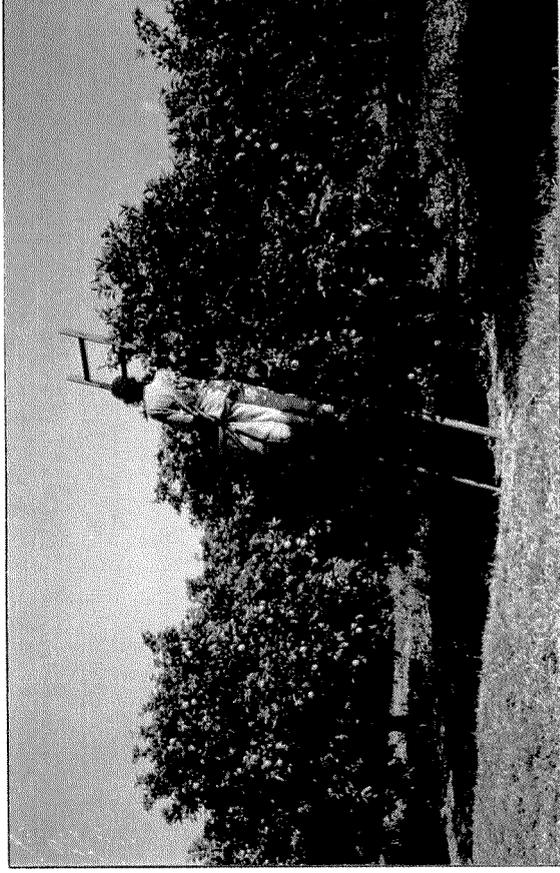
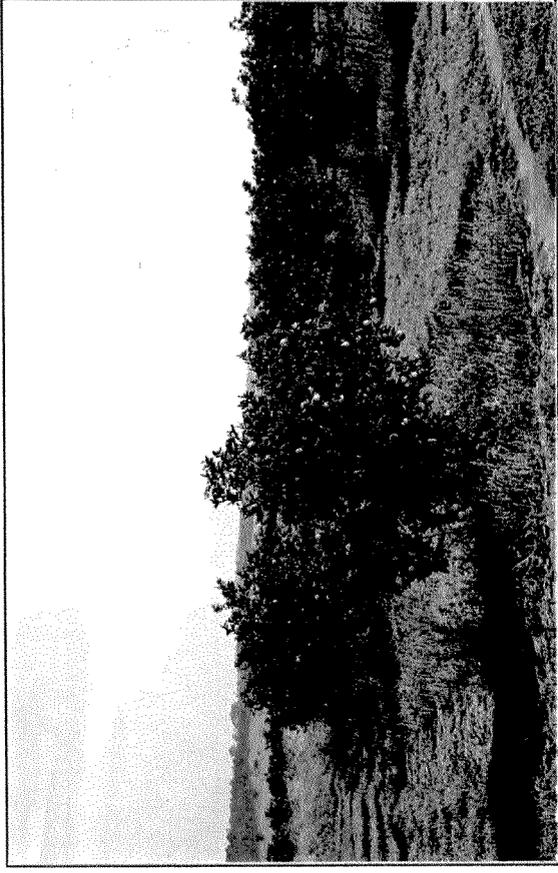


Figura 5-10. Citricultura, um dos usos das terras proeminentes na área de estudo na bacia do rio Jacaré-Pepira, município de Brotas e Torrinhã, SP, incluindo diferentes estágios de desenvolvimento (*a, b e c*).

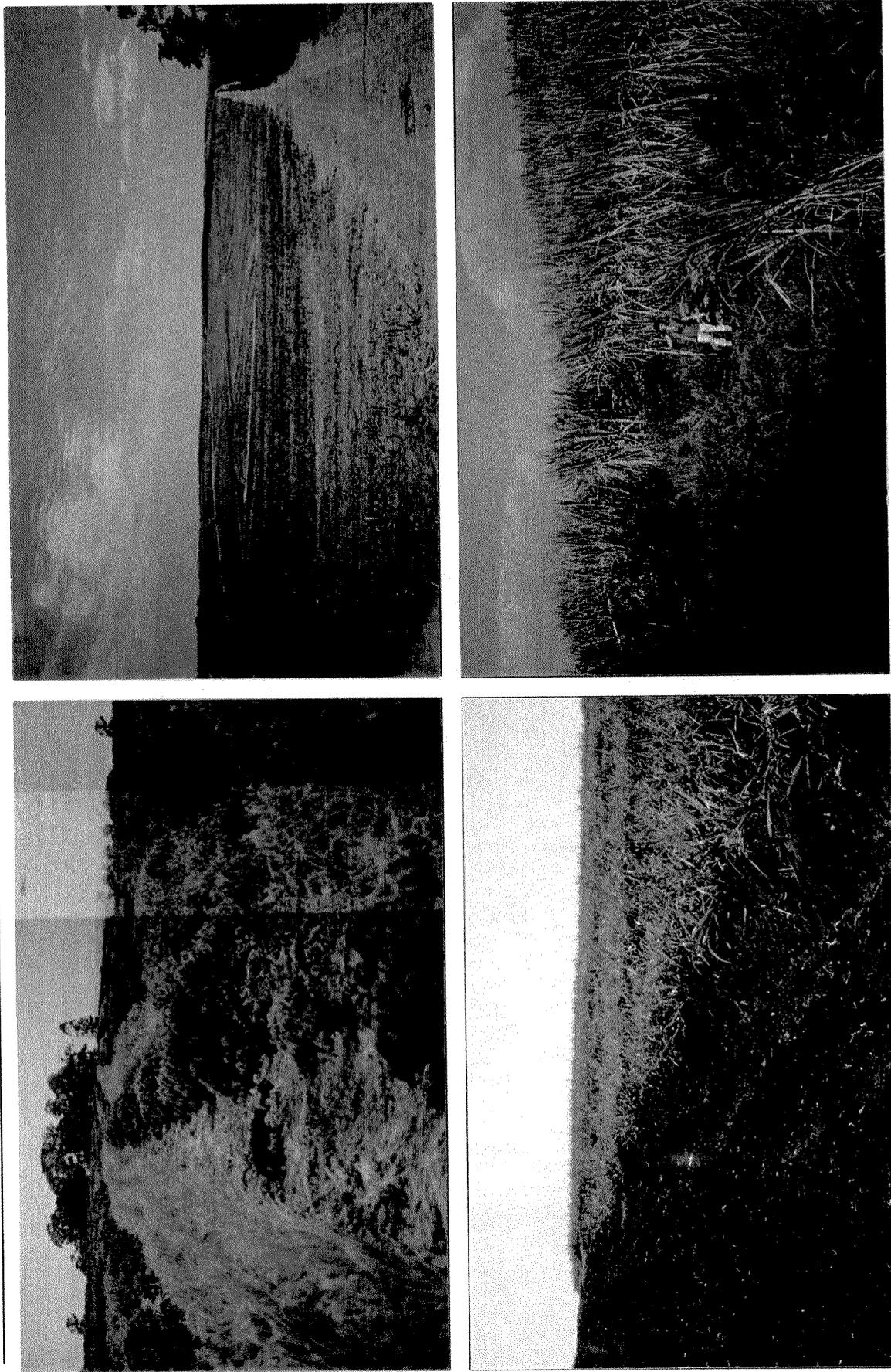


Figura 5-11. Usos das terras na área de estudo na bacia do rio Jacaré-Pepira, município de Brotas e Torrinhã, SP, incluindo consórcios como arroz e café (a) e canavieicultura em diferentes estágios de desenvolvimento (b, c e d).

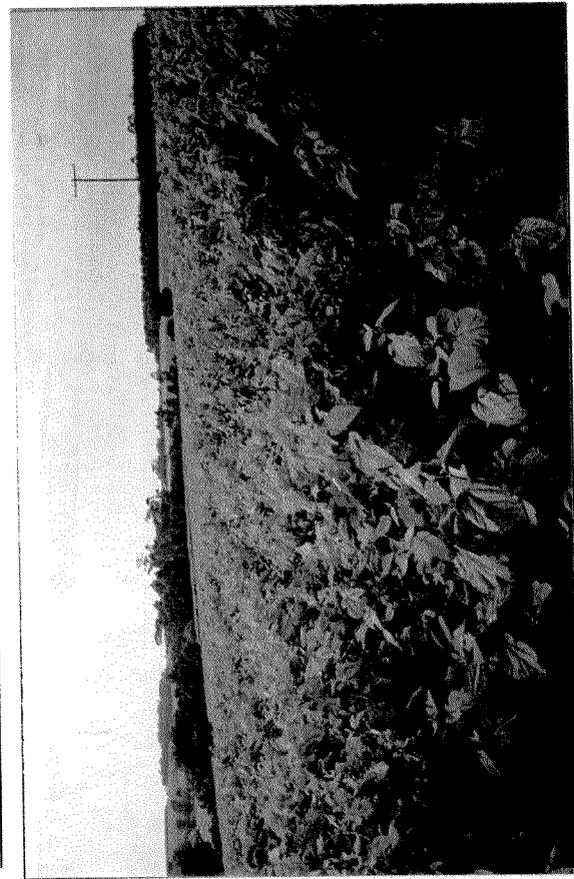


Figura 5-12. Cobertura e usos das terras na área de estudo na bacia do rio Jacaré-Pepira, município de Brotas e Torrinhã, SP, incluindo cultivo da amora para sericultura (a), várzeas (b), corpos d'água (c) e conurbações (d).

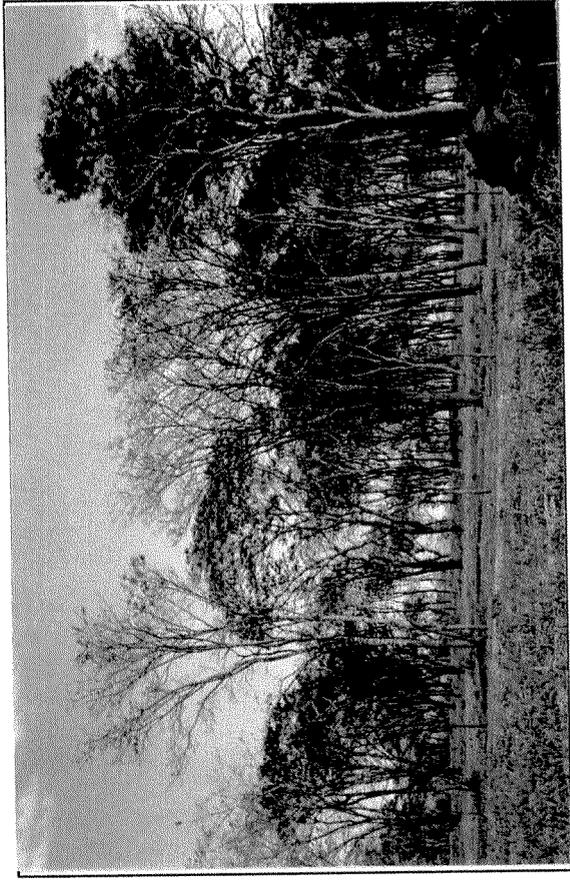
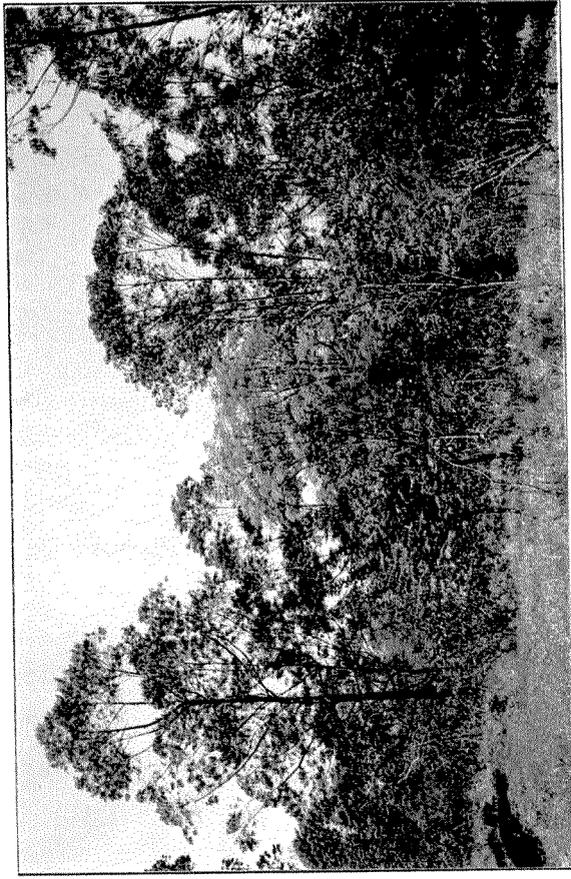
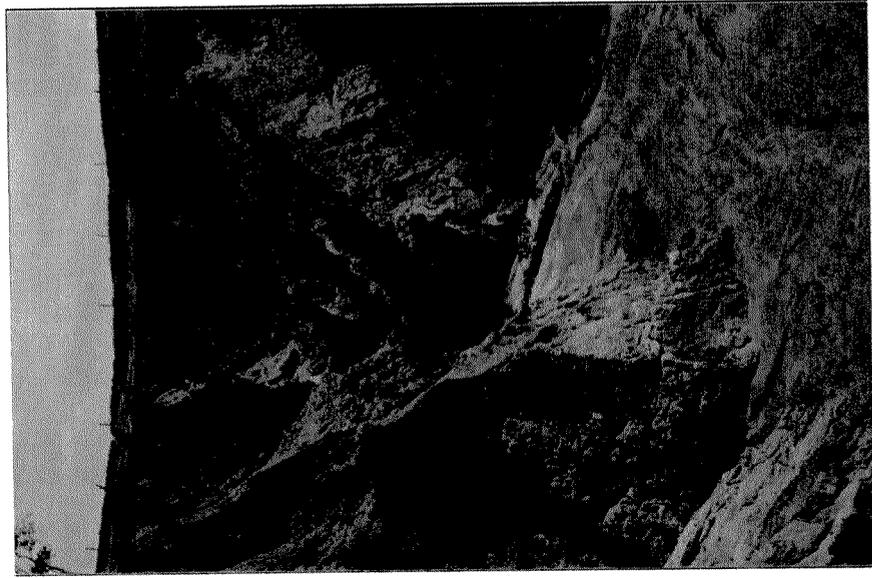


Figura 5-13. Cobertura e usos das terras na área de estudo na bacia do rio Jacaré-Pepira, município de Brotas e Torrinha, SP, incluindo grandes voçorocas (*a*), reflorestamentos de eucaliptos abandonados (*b*) e pastagens com grande densidade de árvores (*c*).

5.4 Infra-estrutura física

Nas atividades de laboratório foram utilizados os seguintes equipamentos e aplicativos do Instituto de Biologia e, principalmente, do Laboratório de Processamento de Informações Geo-referenciadas (LAPIG) do Instituto de Geociências, ambos da UNICAMP:

Equipamentos:

- estações gráficas Sun, modelos SparcServer 4-470, Sparcstation 1+ e IPX, equipadas com unidade leitora de fitas *streamer*, de fitas DAT e de discos óticos regraváveis;
- rede local, conectando PC e estações gráficas do LAPIG à rede UNICAMP e a Internet;
- mesa digitalizadora formato A₁;
- sistema para gravação de filmes 35 mm (Hardcopy), marca Matrix;
- ploter eletrostática, marca Versatec, modelo 8936;
- impressora térmica, marca Tektroniks, modelo Phase IIstdx;
- microcomputadores PC-AT 286, 486 e 586;
- leitor ótico, impressoras jato de tinta e laser;

Aplicativos:

- sistemas de processamento digital de imagens I²S (I²S, 1992) e ER Mapper, versões 4.0 e 4.1 (ER Mapper, 1993);
- sistemas integrados de análise de dados espaciais e multi-espectrais Idrisi, versão 4.01 (Eastman, 1992a e 1992b) e GRASS, versão 4.0 (USA, 1991) e 4.1(USA, 1993);
- sistema de projeto auxiliado por computador AutoCAD, versão 11 (AutoCAD, s.d.)
- sistema de processamento de texto Microsoft Word, versão 6.0 e
- sistema de editoração de imagens Coral Draw, versão 4.0 e 5.0.

Nas atividades de campo, foram utilizados os seguintes equipamentos:

- sistemas de posicionamento global (GPS), marca Magellan e marca Trimble
- conjunto completo de cartas topográficas, escala 1:50.000 e 1:10.000, referentes à área de estudo, descritos no item 5.2.1;
- impressões eletrostáticas das composição coloridas das imagem orbitais;
- fotografias 35 mm coloridas dos pontos amostrais, obtidas através de uma câmera fotográfica acoplada a um sistema para visualização de monitores de vídeo.

6. Métodos

6.1 Estrutura e sistemas para o processamento de imagens

Este projeto foi inicialmente desenvolvido junto ao Laboratório de Processamento de Informações Georreferenciadas (LAPIG) do Instituto de Geociências (IG) da UNICAMP, sob co-orientação do Dr. Álvaro P. Crósta, uma vez que o Departamento de Botânica do Instituto de Biologia da UNICAMP não dispunha de infra-estrutura para o processamento de imagens orbitais, nem tampouco de especialistas nesta área.

Posteriormente, com a disponibilidade de duas estações de trabalho no IB-UNICAMP, constatou-se a possibilidade de transferir parte das atividades do projeto para o Museu de História Natural. Para viabilizar essa alternativa, devido à inexistência de administradores de sistemas e redes no IB, foi necessário a realização de um treinamento especializado através de cursos de gerenciamento de sistemas oferecido pelo Centro de Computação/UNICAMP e um estudo dos manuais técnicos. Com a aquisição desses conhecimentos, foi possível instalar os equipamentos, implantar os sistemas e conectar o IB à rede Uninet e a Internet, beneficiando todo o conjunto de docentes e pós-graduandos da instituição. Finda a viabilização do uso das estações de trabalho do IB, a manutenção e gerenciamento do sistema e da rede foi executada por um período de 14 meses, concomitantemente com o desenvolvimento dessa dissertação. Com a sobrecarga das tarefas de gerenciamento da rede devido ao crescimento da demanda de utilização interna no IB e a desatualização dos computadores em termos de desempenho, tornou-se impossível prosseguir a consecução das atividades do projeto, as quais voltaram a ser efetuadas no LAPIG.

Os primeiros tratamentos e análises das imagens foram realizados com o sistema de processamento de imagens I²S System 600 (I²S, 1992), o único disponível no LAPIG na época. Trata-se de um sistema bastante poderoso, pois além de conter muitos algoritmos e funções, permite um maior controle do emprego das diferentes técnicas e métodos pelo analista e a implementação de novos algoritmos. Contudo, a sua interface com o usuário não é muito amigável e sua utilização requer uma base de conhecimento muito fundamentada.

Posteriormente os dados e análises em desenvolvimento foram migrados para o sistema ER Mapper 4.0 (ER Mapper, 1993), por ocasião de sua aquisição pelo LAPIG. Trata-se de um aplicativo australiano, cujo manipulação é bastante fácil e o armazenamento de dados no disco é otimizado pela sua estrutura de algoritmos.

Entretanto essa opção pelo ER Mapper não implicou em um abandono completo do I²S. No decorrer do trabalho, pelo fato de ambos os aplicativos apresentarem prós e contras, diversas vezes foram utilizados recursos de ambos segundo a conveniência de cada atividade, demonstrando o potencial de soluções multiplataformas na manipulação de imagens.

Além desses, outros aplicativos foram empregados para o processamento e geração dos dados de referência terrestre e cálculo da exatidão dos classificadores selecionados.

O GRASS (“Geographic Resources Analysis Support System” - USA, 1991 e USA, 1993) permitiu o tratamento das informações sobre o uso das terras levantadas pelo plano de amostragem, preparando-as para a avaliação do desempenho dos classificadores. É um sistema integrado de análise de dados espaciais e multiespectrais de domínio público desenvolvido pelo “U.S. Army Corps of Engineers, Construction Engineering Research Laboratory”. Nesse estágio, também foi utilizado um sistema de projeto auxiliado por computador AutoCAD (AutoCAD, s.d.) para digitalização dos croquis de campo.

A avaliação da exatidão das classificações efetuadas foi obtida pelo Idrisi (Eastman, 1992a e 1992b), um outro sistema integrado de processamento de imagens e de informações geográficas de base matricial, desenvolvido pela “Graduate School of Geography” da Clark University.

6.2 Procedimentos metodológicos

O tipo e a seqüência dos procedimentos utilizados nessa dissertação estão sintetizados na Figura 6-1. Foram quatro principais estágios: *Preparação*, *Classificação*, *Coleta de dados de referência* e *Análise da exatidão*. Em cada um deles, conforme ilustrado pelo fluxograma apresentado, diversas técnicas e métodos foram empregados em atividades interativas, cujos resultados foram analisados de forma a corrigir o curso do projeto durante a sua consecução. Na *Preparação*, as imagens foram tratadas de modo a facilitar a distinção de padrões de cobertura e uso das terras. Este tratamento, juntamente com os trabalhos de campo preliminares, permitiram a formulação da primeira versão da legenda temática. Na segunda etapa, *classificação*, seis diferentes algoritmos foram utilizados para categorização dos pixels da imagem. Concomitantemente, foi executada a fase de *coleta de dados de referência*, que respondeu essencialmente pela elaboração e aplicação de um plano de amostragem visando a coleta dos dados de referência. Por fim, os subsídios desses dois estágios anteriores foram conjugados para a implementação da etapa *análise da exatidão*, possibilitando avaliar o desempenho dos algoritmos e da classificação como um todo. Conforme indicado no fluxograma, os diversos procedimentos não estavam conectados de forma apenas unidirecional, articulando-se com etapas precedentes através de circuitos de retroalimentação. Cada uma desses quatro componentes é descrito com detalhes nos próximos itens.

6.3 Preparação das imagens

Antes de executar qualquer processamento com as imagens orbitais, foi necessário acessar as informações contidas nas fitas. Os arquivos fornecidos pelo INPE em fitas “streamer” (padrão TM-BSQ), devido ao seu formato de gravação diferente dos geralmente utilizados pelos sistemas de processamento digital, exigiram a utilização de um programa auxiliar para a extração dos meios de gravação do INPE em uma configuração compreensível aos outros sistemas.

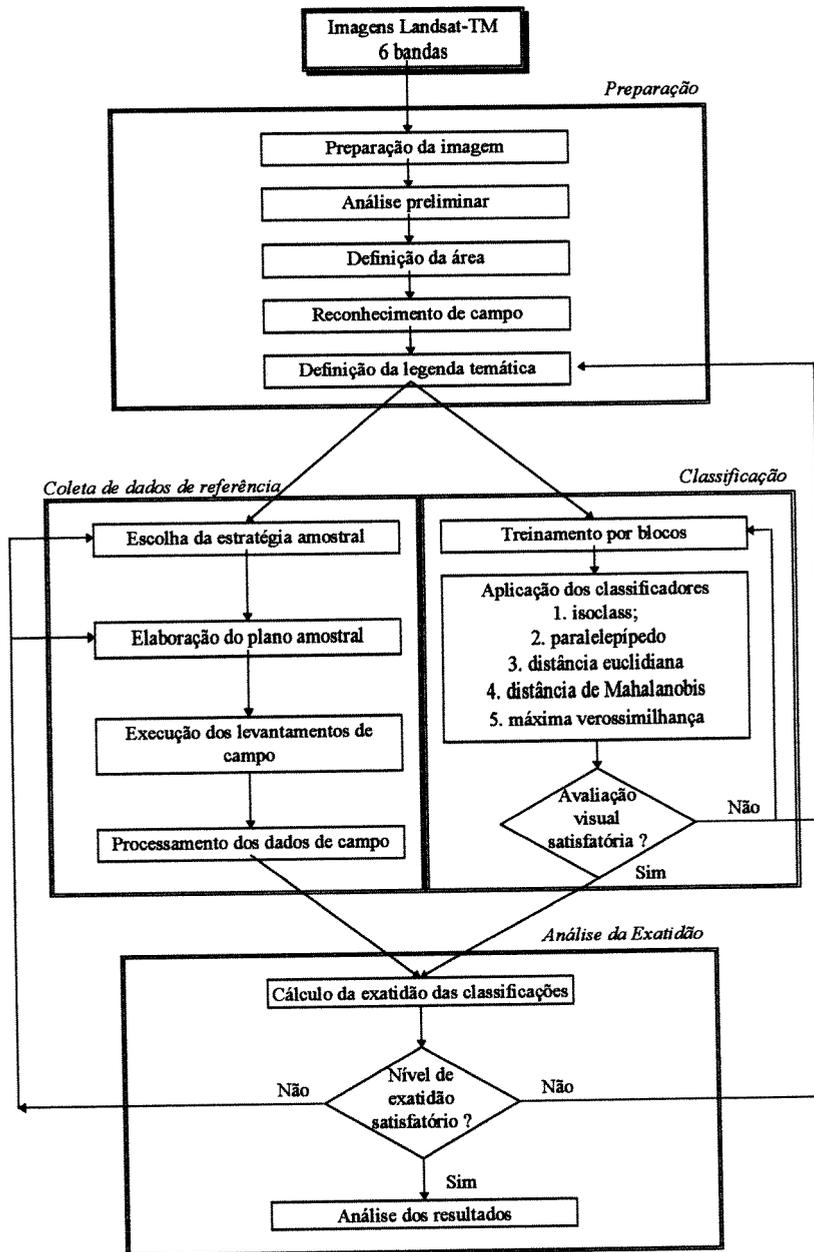


Figura 6-1. Fluxograma dos procedimentos de análise adotados para avaliação dos classificadores espectrais de imagens Landsat-TM na região de Brotas e Torrinha, SP.

Como a implementação desta função de leitura no ER Mapper é recente, o procedimento de extração adotado foi o programa “Lestr”, desenvolvido no INPE. Sua aplicação requer o fornecimento de algumas especificações sobre o arquivo pelo usuário (vide item 11.1). O tempo de leitura de cada fita, contendo uma banda de uma cena, foi de aproximadamente 35 minutos em estações de trabalho.

Após a leitura, para facilitar a manipulação da imagem, foi criado um arquivo único através da fusão das 6 bandas: 1, 2, 3, 4, 5 e 7, com o auxílio do I²S. Em seguida, o arquivo gerado foi convertido para o formato ER Mapper.

6.3.1 Georreferenciamento das imagens

O primeiro passo para executar o registro e a correção geométrica é o estabelecimento de pontos de controle. Neste trabalho foram selecionados 12 pontos distribuídos por toda a imagem, geralmente correspondendo a cruzamentos entre rodovias, ferrovias e cursos d'água, observáveis nos mapas topográficos 1:10.000 (São Paulo, 1980).

Por meio desses pontos estabeleceu-se uma relação entre os dados linha-coluna da imagem e o sistema de projeção UTM dos mapas. Nesta etapa, o polinômio de transformação afim de primeiro grau ("first-degree affine") foi aplicado. Concomitantemente, por meio da raiz quadrada do desvio quadrático médio (RMS) determinou-se a acurácia do processo. A etapa de reamostragem foi executada somente após a obtenção de um valor de RMS inferior a um pixel.

O algoritmo de reamostragem aplicado à imagem foi o "vizinho mais próximo". As informações necessárias para a implementação do georreferenciamento foram o *datum* geodético, a projeção do mapa e o tipo de coordenada.

6.3.2 Seleção de bandas

A estruturação do sistema visual humano em torno de três cores primárias requer a seleção de conjuntos de três bandas dentre as geradas pelo sensor, capazes de exprimir o maior conteúdo de informação espectral. Para a escolha desse tripleto foi aplicado inicialmente o método do fator de índice ótimo ("optimum index factor - OIF" proposto por Chavez et al., 1982). A técnica consiste em avaliar a quantidade de informações contidas em cada banda, analisando suas variâncias através dos respectivos desvios padrões e seus coeficientes de correlação. O OIF é calculado para as várias possibilidades de combinações de três bandas, dividindo a soma dos desvios padrões das bandas integrantes do tripleto pela soma dos valores absolutos dos coeficientes de correlação para as mesmas bandas. Sua computação é representada a seguir:

$$\text{OIF}_{(a,b,c)} = \frac{\sum (D_a + D_b + D_c)}{\sum (C_{ab} + C_{ac} + C_{bc})}$$

onde *a*, *b* e *c* são três bandas quaisquer; *D* é o desvio padrão e *C* é o coeficiente de correlação.

O tripleto com o maior valor de OIF teoricamente contém mais informações e uma menor redundância espectral. Contudo, esta técnica deve ser utilizada apenas como orientação na escolha das bandas, pois como baseia-se somente nos parâmetros quantitativos da cena, não leva em consideração parâmetros qualitativos (Hernandes, 1994). Desse modo, muitas vezes o critério para analisar o melhor tripleto não pode ser somente a quantidade de informações espectrais, mas sim também sua qualidade, ou seja, garantir a inclusão das informações de interesse nas bandas selecionadas. Por conseguinte, uma análise visual das composições coloridas com as quatro primeiras combinações apontadas pelo OIF foi executada para verificar se esses tripletos apresentavam informações adequadas ao reconhecimento dos elementos terrestres de interesse, principalmente no que se refere as categorias de uso das terras e de vegetação natural.

O método OIF foi aplicado considerando apenas as bandas 1, 2, 3, 4 e 5, uma vez que a banda 6 apresenta incompatibilidade de escala espacial e a 7 não apresentou bons resultados na visualização do uso das terras na região.

6.3.3 Realce radiométrico.

A fim de melhorar o contraste dos dados das imagens, um método de aumento não-linear, conhecido por equalização de histograma ou achatamento (Crósta, 1992), foi aplicado sobre a imagem. Esta técnica adota o princípio de que o contraste de uma imagem seria otimizada se todos os 256 possíveis níveis de intensidade fossem igualmente utilizados. A utilidade de um aumento de contraste do tipo equalização é realçar as imagens com extensas áreas de intensidade uniforme, associadas a histogramas com um largo pico central. Desse modo, é possível revelar detalhes de contraste previamente não perceptíveis.

6.3.4 Definição da área piloto

Com a finalidade de maximizar o aproveitamento da heterogeneidade da paisagem existente na região de Brotas, foi delimitada uma área de estudo para avaliar o desempenho dos classificadores para determinação do uso das terras em dois tipos de feições topográficas e geomorfológicas presentes nas cercanias de Brotas e Torrinha (vide figura 6-2): os relevos de degradação em planaltos dissecados com colinas amplas, presentes na Planície do Jacaré-Pepira, e os relevos com colinas médias, presentes nos Planalto de Itaqueri, conforme ilustrado simplificada pela figura 5-5 e descritas no item 5.3.3 (Carneiro et al., 1981 e Ponçano et al., 1981).

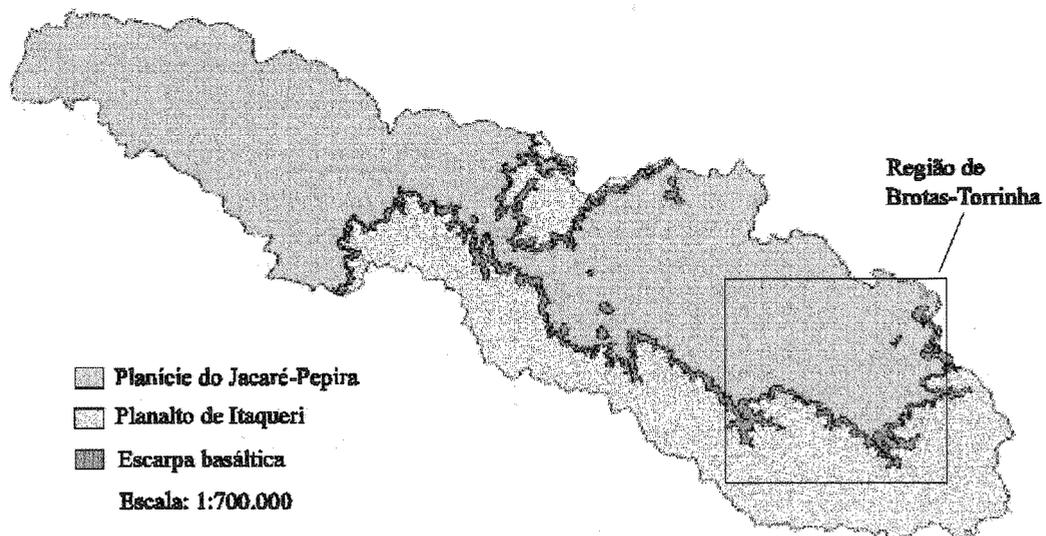


Figura 6-2. Grandes unidades topográficas da bacia do rio Jacaré-Pepira. (Adaptado de Metzger, 1992).

Deste modo, a área piloto abrange uma parte com relevo de colinas mais suaves e com grandes empreendimentos do setor florestal, citrícola, álcool-açucareiro e de pecuária bovina (vide figura 5-7 d e b; figura 5-8 a, b e d; figura 5-10; figura 5-11 b, c e d), e outra

com relevo de colinas mais acentuadas, dominadas por pequenas propriedades, parcelar agrícola reduzido e variadas coberturas e usos das terras (vide figura 5-9 c e figura 5-12 a). Sua forma irregular (vide figura 6-4) foi definida de modo a abranger estas duas paisagens, cobrindo uma extensão de 573 km² e localizada entre as coordenadas 22° 10' 55'' S - 47° 57' 4'' W e 22° 28' 43'' S - 48° 12' 44'' W.

A delimitação dessa área piloto sobre a imagem foi estabelecida através da criação de uma máscara sobreposta à imagem, definida pelas bacias de drenagem de dois afluentes do rio Jacaré-Pepira, os ribeirões Pinheirinho e da Rasteira (vide figura 6-4).

6.3.5 Reconhecimento da área de estudo

Em razão da visita ao campo ser indicado como o método mais adequado à coleta de dados de referência (Congalton & Biging, 1992), optou-se pela sua utilização em detrimento do simples uso de mapas e/ou fotos aéreas para o reconhecimento de padrões de cobertura e uso das terras.

O trabalho de campo foi dividido essencialmente em 2 partes: elaboração de uma legenda temática do uso das terras adequada às peculiaridades da região de Brotas e Torrinha e execução do plano de amostragem dos dados de referência da cobertura terrestre para a análise de exatidão.

A primeira etapa de campo, a qual é descrito a seguir, teve como objetivo o reconhecimento e familiarização com a área de estudo e definição das categorias a serem abordadas na legenda de identificação e classificação do uso das terras. A escolha do período de estiagem para a realização dos levantamentos de campo foi feita de modo a coincidir com a época da aquisição da imagem de satélite e para aproveitar as boas condições de tráfego da maioria das estradas nessa época. Foram utilizadas fotografias aéreas de parte da região de estudo, imagens orbitais Landsat-TM em dois formatos (fotografias geradas pelo sistema para gravação de filmes 35 mm e impressões na escala de 1:50.000 produzidas por traçador gráfico eletrostático), mapas topográficos na escala de 1:10.000 e 1:50.000 e um GPS.

Durante os percursos, os tipos de uso das terras observados eram anotados, suas coordenadas registradas e os polígonos delineados sobre a imagem impressa ou fotografada (vide figura 6-3). Como esses dados também serviram para a identificação de áreas de treinamento capazes de alimentar os classificadores digitais das imagens de satélite, foram selecionadas áreas representativas em termos de tamanho e variabilidade estrutural para cada uma das coberturas e usos da terra levantados por uma legenda temática preliminar.

Devido à defasagem temporal de um ano, entre a tomada da imagem orbital e a coleta dos dados de campo, foram realizadas entrevistas com proprietários das áreas rurais, a fim de levantar o histórico do uso das terras e obter os tipos correspondentes ao período da aquisição da imagem.

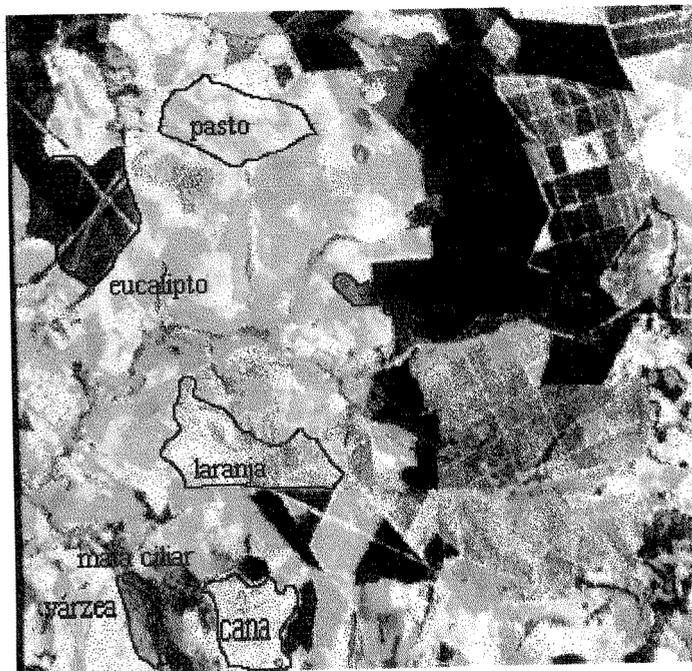


Figura 6-3. Imagem utilizada para reconhecimento de campo dos tipos de uso das terras da região de Brota e Torrinha, SP.

O grande inconveniente enfrentado nessa etapa foi a desatualização das informações contidas nos mapas topográficos do IBGE (Brasil, 1969; Brasil, 1971 e Brasil, 1974) e IGC (São Paulo, 1980) tornando o esforço de reconhecimento maior. Essas dificuldades foram contornadas com o auxílio do GPS e as composições coloridas. Além disso, a existência de culturas perenes e de fragmentos de vegetação natural ajudou muito a localização dos padrões de cobertura terrestre.

6.3.6 Legenda temática do uso das terras

A localização, a expressão espacial na região dos diferentes tipos de usos e as características dos sistemas de produção adotados foram devidamente anotadas. Essa base de dados possibilitou a formulação da primeira legenda temática do uso das terras. Sua composição foi sendo sucessivamente aperfeiçoada pela ampliação do conhecimento acumulado sobre a ocupação da região. Esse procedimento conduziu à definição de uma tipologia do uso das terras de um contingente inicial de 36 classes identificadas para uma primeira versão da legenda com 13 categorias e uma segunda com 9 (vide Tabela 6-1).

Tabela 6-1. Desenvolvimento da legenda temática para a classificação do uso das terras para a área de estudo na bacia do rio Jacaré Pepira, Brotas e Torrinha, SP.

Coberturas e usos das terras	Legenda n° 1	Legenda n° 2
1. campo alagado		
2. várzea	1. várzea	1. várzea
3. cerrado		
4. cerrado com afloramento	2. cerrado	
5. cerradão	3. cerradão	
6. mata ciliar	4. mata ciliar	
7. capão de mata		
8. mata de encosta		
9. mata	5. mata	2. vegetação natural
10. pasto limpo		
11. pasto suja	6. pasto	
12. pasto com poucas árvores		
13. pasto com árvores	7. pasto com arvores	3. pasto
14. laranjal jovem		
15. laranja em formação	8. laranjal jovem	4. laranjal jovem
16. laranjal formado	9. laranjal formado	5. laranjal formado
17. canavial com até 2 meses		
18. canavial intermediário		
19. canavial com mais de 9 meses	10. canavial	6. canavial
20. eucaliptal cortado		
21. eucaliptal novo		
22. eucaliptal adulto	11. eucaliptal	7. eucaliptal
23. solo exposto		
24. solo em preparo	12. solo	8. solo
25. corpo d'água	13. corpo d'água	9. corpo d'água
26. horticultura		
27. cultura anual	cultura anual ¹	
28. bambuzal		
29. cafezal		
30. cultura de abacate		
31. amoral		
32. pinheiral (pinus)		
33. outras culturas perenes	cultura perene ¹	
34. edificação rural e pomar		
35. área urbana		
36. estrada		

¹ Na constituição dos dados de referência, a legenda n° 1 teve seu número de categorias expandido para 15 pois as culturas anuais e perenes foram incluídas devido a sua presença em algumas amostras.

Para efeito do processo de classificação, foram testadas duas legendas temáticas do uso das terras: a primeira constituída de 13 classes e a segunda com 9 classes (vide Tabela 6-1). A diferença básica está na forma de tratamento das formações vegetais naturais e das pastagens. Ao contrário da legenda n° 1, onde se procurou incluir de forma abrangente a diversidade de coberturas existentes na região, a legenda n° 2 abordou a paisagem de forma mais homogênea, explorando a eventual possibilidade de um incremento no desempenho dos algoritmos de classificação. No caso da classificação MaxVer, foi utilizado também uma

legenda temática contendo 7 classes, obtida com a fusão entre as classes de citricultura, entre a de várzea e a de áreas naturais, com a finalidade de explorar a variação existente entre o número de classes e o grau de exatidão.

Nesse contexto de elaboração da legenda temática, a classe canavicultura foi objeto de um tratamento mais pormenorizado devido à sua expressão espacial na região. Após a implantação do Programa Pró-Álcool, esse cultivo atingiu uma grande importância na economia brasileira, tornando-se responsável por intensas alterações no meio rural brasileiro, com conseqüências variadas. Entre estas pode-se citar as ecológicas, onde se destaca a homogeneização da estrutura das paisagens através do deslocamento ou extinção de grande parte das culturas agrícolas e a eliminação de remanescentes florestais. Além disso, a prática do fogo ainda muito comum no sistema de cultivo da cana ou o aproveitamento dos resíduos do processo de destilação do álcool são atividades geradoras de impactos ambientais significativos. Paralelamente, entre as conseqüências sociais, ressalta-se as modificações da estrutura fundiária relacionadas com a incorporação de pequenas propriedades em grandes latifúndios e seus intensos corolários na organização sócio-econômica de vastas regiões. A expressão desse cultivo e de seus problemas associados motivou uma análise da viabilidade da utilização de imagens Landsat-TM para identificação espectral das variedades e da idade dos cultivos. Esses dados serviriam como subsídios para a constituição de um sistema de acompanhamento e gerenciamento da produção, dos recursos e de seus impactos ambientais, calcado em uma base de dados geocodificados e estruturado no ambiente de um sistema de informações geográficas.

A modelagem do problema foi realizada na área de ação da Usina Paraíso, Brotas, a qual franqueou o acesso ao seu banco de dados corporativo para auxiliar a fase de treinamento dos classificadores.

Os materiais utilizados para esta análise foram: mapa de localização das fazendas cultivadas pela Usina, mapas detalhados dos talhões da cada fazenda nas escalas 1:10.000 e 1:20.000, tabela de controle agrônômico (incluindo o número e área do talhão, tipo de solo, variedade, espaçamento, data de plantio e de corte, número do corte, produção e rendimento) e composições coloridas composta pelas bandas 4, 5 e 3 (RGB) impressas em traçador eletrostático.

6.4 Classificação

Com o objetivo de classificar o uso das terras na área de estudo segundo as legendas temáticas elaboradas, foram aplicados à imagem os seguintes algoritmos: paralelepípedo, distância mínima e variantes, máxima verossimilhança e isodata.

6.4.1 Seleção dos sítios de treinamento

Para executar com sucesso uma classificação digital é necessário que as categorias de interesse estejam cuidadosamente selecionadas. Neste contexto é fundamental que o analista considere a diferença básica entre classes de informação e classes espectrais (Jensen, 1996). As classes de informação são aquelas definidas pelo analista e as classes espectrais são aquelas inerentes aos dados da imagem. Um analista deve tentar traduzir as classes espectrais em classes de informações para satisfazer os requisitos da pesquisa, aplicando

seus conhecimentos sobre as características espaciais e espectrais do sistema sensor e relacionando esses parâmetros para cada tipo e proporção do alvo encontrado na cena.

Analisando-se os dados reunidos durante a campanha de reconhecimento de campo, foram selecionadas áreas na imagem representativas das classes de uso das terras presentes na área piloto. A captura dessas áreas de treinamento na imagem foi feita através da visualização da imagem na tela e da delimitação das áreas através do cursor. Os polígonos referentes aos tipos de cobertura foram cuidadosamente delineados evitando a inclusão de pixels próximos às margens e procurando conservar um padrão homogêneo para cada classe. O tamanho médio dessas áreas ficou em torno de 600 pixels.

Outro ponto considerado nesta fase foi a dispersão das amostras de treinamento, devido ao fato da maioria das classes de uso das terras não serem representadas por uma única classe espectral. Diante dessa situação, optou-se por tratar estas variações como classes distintas. Assim foram criadas subclasses em algumas categorias de uso das terras, as quais foram fundidas após a execução dos algoritmos de classificação.

As classes mais problemáticas, com ocorrência rara na imagem e geradoras de confusão com as classes mais representativas, foram eliminadas dos dados de treinamento. Portanto optou-se por aceitar erros de classificação de uma classe rara na área de estudo a fim de preservar a exatidão da classificação das classes espectrais similares presentes em áreas extensas.

Após a seleção das áreas de treinamento, foram calculadas a média e a variância dos números digitais para verificar se os grupos de dados eram espectralmente distintos. As amostras de classes distintas com sobreposições acentuadas foram recompiladas, eliminando-se uma das duas, de modo a formar os grupos finais de estatísticas usadas no processo de classificação.

6.4.2 Algoritmos supervisionados

Para aferir a eficiência dos classificadores supervisionados selecionados nas circunstâncias locais da área de estudo, foram utilizados as mesmas amostras de treinamento e portanto a mesma estatística das classes de informação presentes na legenda temática formulada.

Os classificadores utilizados foram: paralelepípedo, 3 tipos de distância mínima (Euclidiana padrão e baseada nos desvios padrões, e de Mahalanobis) e máxima verossimilhança (MaxVer). Um comentário mais aprofundado sob essas técnicas é apresentado no item 4.5.3.1.3.

O método do paralelepípedo determina, a partir dos dados de treinamento, os valores máximo e mínimo de NDs para cada uma das classes temáticas em cada um dos atributos. Esses valores extremos possibilitam a delimitação de paralelepípedos no espaço de atributos n-dimensional. Os pixels desconhecidos são designados a uma determinada classe, de acordo com sua posição em relação às regiões de decisão definida por cada paralelepípedo. Há casos nos quais alguns pixels são considerados como desconhecidos, pela impossibilidade de localizá-los em um dos paralelepípedos definidos.

Para executar a classificação, o algoritmo de distância mínima utiliza as médias dos valores espectrais (vetores médios), em cada banda para cada categoria, obtidas a partir dos dados de treinamento, para determinar a identidade de um pixel. A distância mínima entre os valores de um pixel desconhecido e a média de cada uma das categorias define a classe à qual pertence o pixel.

É possível calcular esta distância usando a forma euclidiana (Swain & Davis, 1978). Esse tipo permite o analista especificar uma distância mínima do pixel em relação à média da classe aceitável, podendo fazer com que um pixel não seja designado mesmo que sua posição seja próxima a da média de uma determinada categoria. Além da distância mínima euclidiana padrão, foi aplicada uma variante baseada nos desvios padrões.

Este classificador também pode se valer de equações de Mahalanobis para determinar a distância mínima. A diferença básica entre a distância euclidiana e de Mahalanobis reside na utilização das informações contidas na matriz de variância-covariância (“pooled covariance information”).

O classificador MaxVer avalia quantitativamente a variância e a covariância das categorias dos padrões de resposta espectral para classificar um pixel desconhecido. Para tanto, assume que a nuvem de pontos formada pelos dados de treinamento tem uma distribuição Gaussiana ou normal. Sob esta suposição, a distribuição do padrão de resposta de uma categoria pode ser completamente descrita pelo vetor médio e a matriz de covariância. Dados estes parâmetros, é possível calcular a probabilidade estatística de um dado pixel pertencer à uma determinada classe. A função de densidade de probabilidade associa a probabilidade de um pixel à uma classe. Após avaliar a probabilidade de cada categoria, o pixel pode ser designado para a classe mais semelhante (valor de probabilidade maior), ou rotulado como desconhecido, se os valores de probabilidade estão abaixo dos limites definidos pelo analista. É possível definir a probabilidade de ocorrência *a priori* das classes. Nesta pesquisa foram adotados pesos iguais a 0.04545 por limitações do pacote ER- Mapper.

6.4.3 Algoritmos não supervisionados

Essa categoria de técnicas de classificação esquadrinha o espaço de atributos definido pelas distribuições dos dados espectrais a procura de grupos de números digitais similares. A expectativa inerente a esses métodos é que alguns desses agrupamentos estejam associados às propriedades de uma determinada classe de cobertura terrestre, permitindo sua identificação *a posteriori*. Desse modo, a aplicação de classificadores não supervisionados é recomendada nos casos de indisponibilidade de uma caracterização do tema ou como uma primeira avaliação da imagem sob análise. Para uma descrição mais detalhada dos procedimentos e características dessa modalidade de programas, vide o item 4.5.3.1.2.

Nessa dissertação foi utilizado o classificador “Isodata”, para o qual alguns parâmetros auxiliares precisam ser fornecidos pelo analista para estabelecer as séries, suas divisões e fusões. A aplicação do algoritmo se baseou nos seguintes parâmetros:

- número inicial de 30 classes;
- número mínimo de pixels por classe de 0,01 (%);

As classes com um número de elementos menor que a porcentagem definida são eliminadas e seus pixels destinados a outras classes.

- limite de 4,5 para o desvio padrão;

Quando o desvio padrão excedia o estipulado e o número de pixels na classe era superior ao dobro do número mínimo de elementos (%), a classe era dividida, adotando como centros das duas classes o anterior mais e menos o desvio padrão. Os valores entre 4.5 e 7 são os mais usados para limitar o desvio padrão.

- distância mínima entre as médias das classes de 3,2;

As classes com distâncias menores são unidas, sendo os valores mais comumente usados 3,2 a 3,9.

A implementação desse algoritmo no sistema ER-Mapper não requer o número máximo de agrupamentos passíveis de fusão em qualquer iteração (Niblack, 1985). Entretanto outros parâmetros precisam ser estipulados:

- número total de iterações igual a 1.000;

Quando as médias das classes se estabilizam antes do número especificado ser alcançado, o processo é finalizado. O valor empregado foi o padrão oferecido pelo sistema.

- número de células estáveis durante uma iteração de 98 (%);

Esse parâmetro limita a mobilidade dos pixels entre as classes em cada fase da classificação.

- intervalo de amostragem de linhas e colunas de 1;

Esse tipo de amostragem permite a redução do tempo de processamento, ao escolher somente uma amostra da imagem para efeito dos cálculos. O valor selecionado determina o processamento de todas as linhas da imagem.

- número máximo de classes de 25;

Quando o número especificado é muito grande, o processo somente cessa quando o número de iterações ou a porcentagem imutável desejada for alcançada.

- valor de separação de 0;

Este valor é somado à média da classe anterior antes da sua divisão, definindo os centros das duas novas classes.

6.5 Coleta dos dados de referência

Essa etapa está marcada pela formulação e consecução do plano de amostragem para viabilizar o levantamento dos dados de referência sobre o uso das terras na região de Brotas e Torrinha necessários a aferição do desempenho dos classificadores espectrais.

6.5.1 Plano de amostragem

A partir da análise dos dados da região foi observada a existência de duas grandes feições distintas (vide figura 6-2 e item 6.3.4). Assim foram determinados dois estratos na região de estudo. O primeiro deles apresenta um relevo suave e uma ocupação dominada por grandes empreendimentos do setor florestal, citrícola e açúcar e álcool localizada ao norte. O segundo estrato possui relevo bastante ondulado, dominado por pequenas propriedades ao sul.

O delineamento amostral para os dois estratos foi gerado por uma estratégia de amostragem sistemática aleatória não-alinhada.

Foi definido um tamanho da amostra de 600 x 600 m no sentido de garantir, na campanha de campo, a acuidade da observação visual do analista durante a coleta dos dados de referência sobre a cobertura terrestre. O número de amostras, quarenta e seis, foi determinado no sentido de representar um total de 2.7% da área piloto (Rocha, 1992), adotando-se a abordagem baseada na porcentagem da área total de estudo (Taylor, 1989) ou da imagem digital (Congalton, 1988b). Dentre a variação no valor de porcentagem verificada na literatura, optou-se por essa porcentagem, no sentido de manter o equilíbrio entre a exequibilidade da campanha de campo e a conveniência estatística.

A localização geográfica dos levantamentos foi obtida através de uma grade com células de 3.600 x 3.600 m, sobreposta à imagem da área de estudo. Dentro de cada uma das 46 células foi sobreposta uma nova grade com subcélulas de 600 x 600 m. Para cada uma das 46 células, foi sorteada uma das 36 subcélulas (vide figura 6-4).

Cada área da imagem correspondente a uma amostra de 600 x 600 m foi fotografada de modo a ocupar uma posição central em uma grade de 25 subcélulas e facilitar sua localização no campo (vide figura 6-5). Além disso, a posição geográfica de cada amostra foi estabelecida nos mapas de 1:10.000. Sobre uma fotocópia do mapa da área referente à amostra foi fixado um quadrado de papel vegetal com tamanho equivalente às dimensões da amostra, formando-se assim o material de campo necessário para coleta de área de cobertura das terras (vide figura 6-5). Este protocolo de coleta de dados foi levado ao campo para obter os dados reais de cobertura terrestre, possibilitando a comparação entre o resultado da classificação e a realidade da cobertura terrestre da região.

Foram realizadas doze viagens para coleta de dados nos sítios de teste, para serem utilizados como referência na montagem da matriz de erro, correspondendo a segunda etapa do trabalho de campo mencionado no item 6.3.5. Os tipos de uso das terras das 46 amostras de 600 x 600 m foram coletadas segundo a legenda temática previamente definida.

O procedimento de campo adotado nessa etapa foi o de localizar os pontos amostrais com o auxílio de fotografias e impressões em traçador eletrostático da composição colorida da imagem Landsat-TM na escala de 1:50.000, de mapas topográficos 1:10.000 e 1:50.000 e de um aparelho de GPS.

Quando uma amostra era localizada, as informações do tipo de cobertura eram delineadas sobre um "overlay" transparente na escala de 1:10.000 (vide figura 6-5). Cabe reênfatizar que a coleta dos dados de referência foi feita no período de estiagem

de modo a coincidir com a época da aquisição da imagem de satélite.

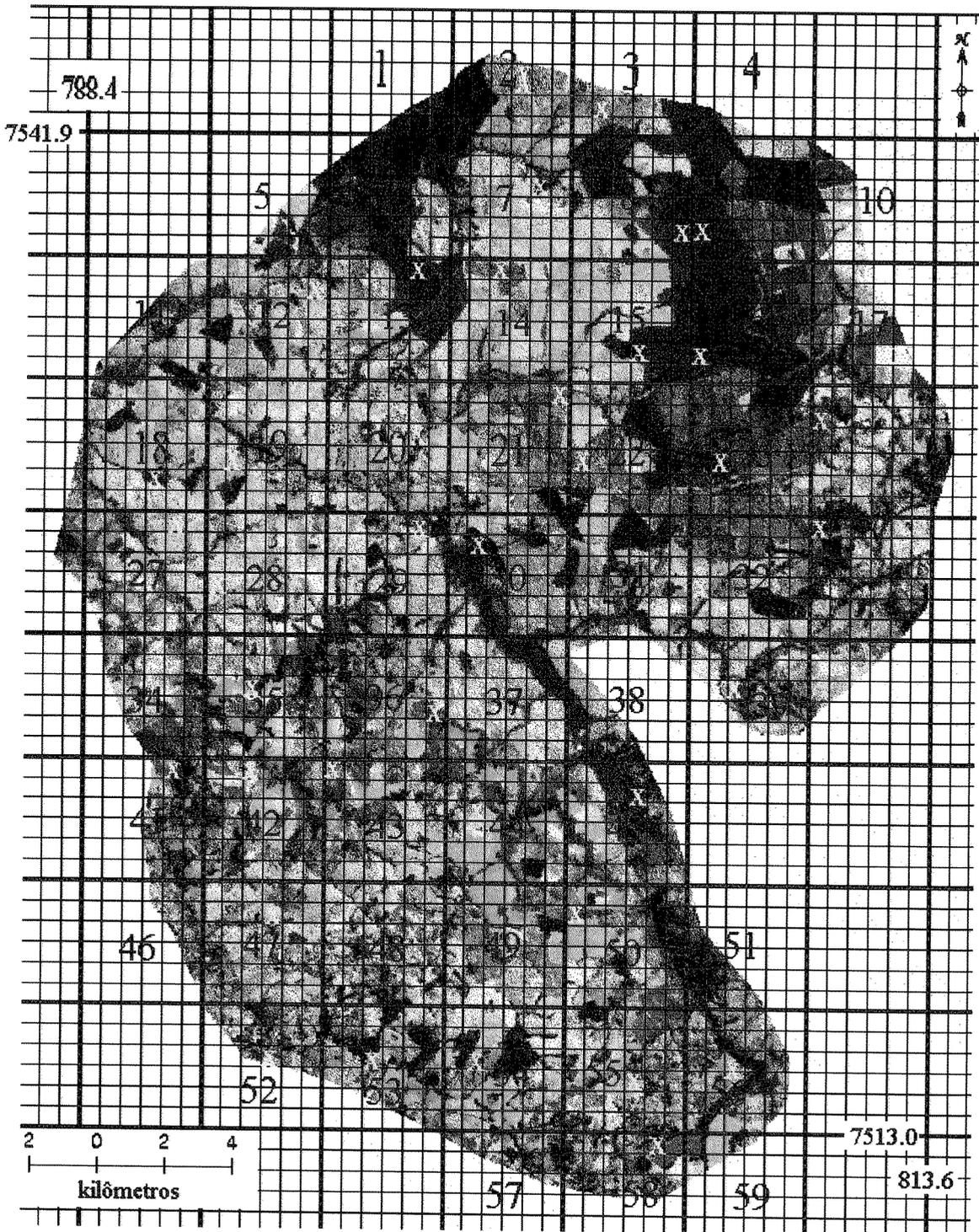


Figura 6-4. Imagem preparada para o emprego da estratégia de amostragem dos dados de referência terrestre sobre o uso das terras na região de Brotas e Torrinha, SP.

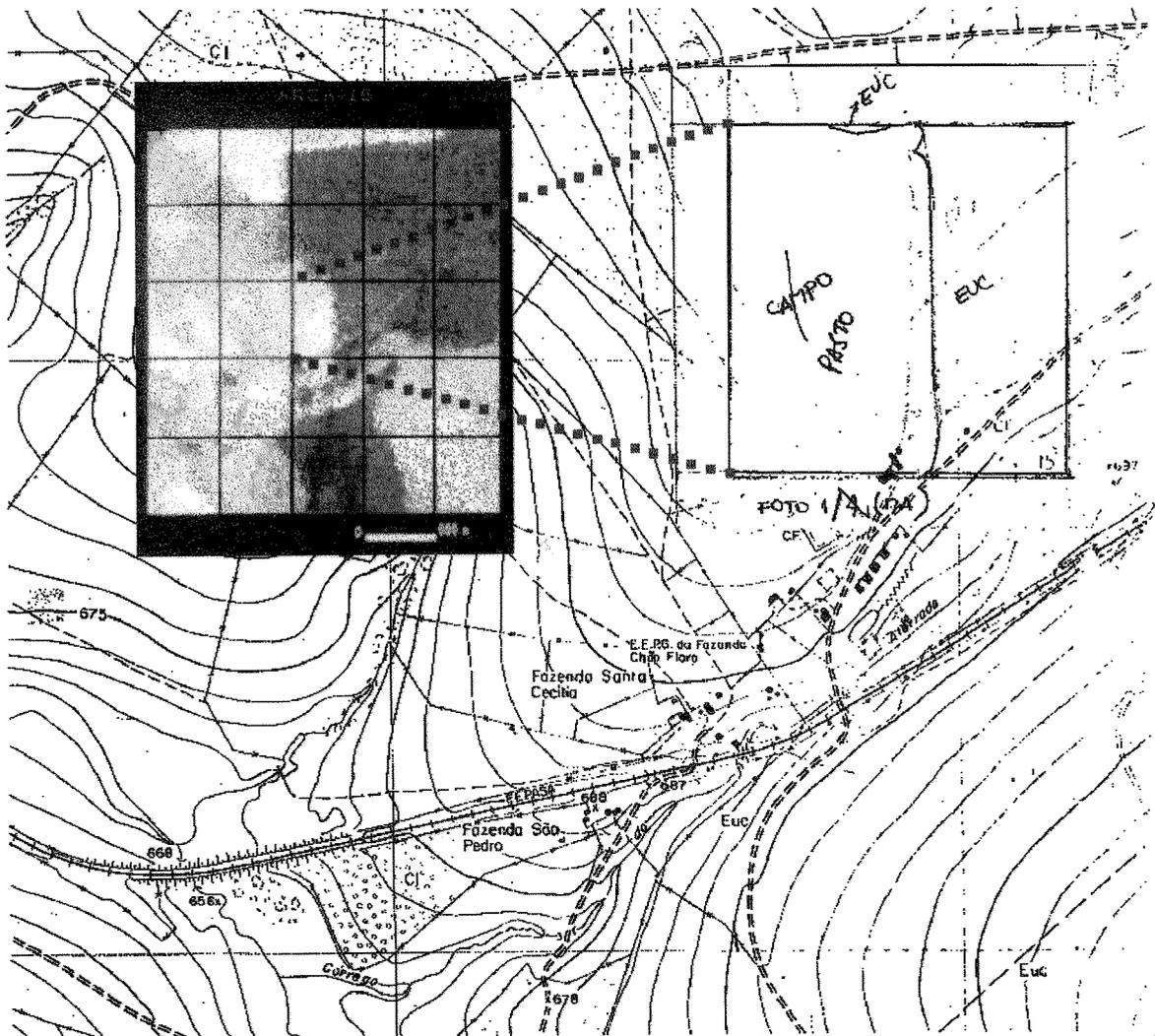


Figura 6-5. Exemplo de protocolo de coleta de dados de referência terrestre em um ponto amostral, constituído do “overlay” com o uso das terras observado e fotografia do trecho da imagem correspondente com a grade amostral, ambos sobre a base cartográfica na escala 1:10.000.

6.5.2 Processamento dos dados de campo

Os dados de uso das terras mapeados nos “overlays” (vide figura 6-5) foram inicialmente revisados e homogeneizados no que se refere às linhas dos polígonos e aos rótulos das categorias identificadas, segundo as quinze classes da legenda temática nº 1 (vide tabela 6-1). Em seguida cada um dos 46 levantamentos coletados durante as incursões da segunda etapa de campo foi digitalizado através de um sistema do aplicativo AutoCAD. Estes dados, em formato vetorial, foram inseridos em um sistema de informações geográficas, o GRASS, onde os polígonos foram rotulados, designando-se as categorias correspondentes. A seguir os arquivos foram transformados para formato matricial e tiveram sua resolução convertida de 1 para 30 metros, de forma a coincidir com a resolução espacial das imagens Landsat-TM. Posteriormente, os valores do sistema de coordenadas de cada

amostra foram alterados de modo a posicioná-los um ao lado do outro. Desse modo, cada arquivo pode ser unido formando um mosaico, o qual representa os dados de referência terrestre ("ground truth") utilizados na geração das matrizes de erro (vide Figura 6-5).

6.6 Análise da exatidão

Nesta etapa, as mesmas grades utilizadas no procedimento de amostragem (vide item 6.5.1) foram sobrepostas às imagens classificadas.

As parcelas das cinco imagens classificadas, correspondentes às amostras de campo, foram recortadas uma a uma, em um total de 46 por tratamento, através de recursos do ER- Mapper. Em razão das dificuldades verificadas nesse sistema de processamento de imagens para a criação de mosaicos a partir desses recortes, todos os arquivos das 46 parcelas retiradas de cada imagem classificada foram migrados para o aplicativo GRASS, onde foram unidos em mosaicos, posteriormente utilizados para a montagem das matrizes de erro (vide figura 7-11, figura 7-12, figura 7-13 e figura 7-14).

Após a geração dos mosaicos com os dados de referência sobre o uso das terras ("verdades terrestres") e as informações geradas pelos classificadores paralelepípedo, distância mínima (3 formas), máxima verossimilhança e isodata foram construídas matrizes de erros com o auxílio do aplicativo IDRISI, tanto para as classificações com a legenda nº 1, como para as com a legenda nº 2 (vide tabela 6-1). Para maiores detalhes desse procedimento, vide o item 4.5.3.2.2.

A partir dessas matrizes, os coeficientes de exatidão foram obtidos, incluindo a exatidão global e individual e o coeficiente de concordância Kappa global e individual. (vide item 4.5.3.2).

Uma vez que a exatidão global não considera como o erro está distribuído através das categorias individuais, foram calculados a exatidão individual das classes de uso das terras. Para se evitar confusão com os termos "exatidão do usuário" e "exatidão do produtor", apontadas por Janssen & van der Well (1994), calculou-se a exatidão individual das categorias a partir dos erros de comissão e omissão. As relações entre as duas abordagens para esses cálculos são demonstradas a seguir:

exatidão do usuário (%) = 100% - erro de comissão (%) e

exatidão do produtor (%) = 100% - erro de omissão (%).

7. Resultados

Os resultados desta pesquisa são a seguir apresentados preferencialmente sob a forma de figuras e tabelas, dispostas segundo uma estrutura de tópicos, coerente com a descrição dos procedimentos metodológicos (vide item 6.2.).

7.1 Preparação das imagens

7.1.1 Seleção de bandas

A Tabela 7-1 e a Tabela 7-2 apresentam os dados estatísticos das seis bandas da imagem Landsat-TM (1, 2, 3, 4, 5 e 7), órbita-ponto 220/75, quadrante C, 92 por 92 km, de 29/8/92, com 0% de cobertura de nuvens, utilizada para esta pesquisa .

Tabela 7-1. Dados estatísticos de bandas da imagem Landsat-TM, 29/8/92, órbita-ponto 220/75 e quadrante C.

Bandas	1	2	3	4	5	7
Mínimo	52.000	19.000	8.000	11.000	3.000	1.000
Máximo	141.000	121.000	255.000	255.000	219.000	221.000
Média	66.048	30.263	36.769	49.519	84.333	30.941
Mediana	66.000	30.000	37.000	48.000	87.000	32.000
Desvio padrão	3.881	3.507	7.295	7.193	24.181	13.352
Desvio padrão (n-1)	3.881	3.507	7.295	7.193	24.181	13.352
Coefficiente de correlação	4.433	1.015	0.266	0.155	0.082	0.049
Coefficiente de covariância	798.183	56.061	27.063	9.817	3.178	1.049

Tabela 7-2. Matriz de correlação de bandas da imagem Landsat-TM, 29/8/92, órbita-ponto 220/75 e quadrante C.

Bandas	1	2	3	4	5	7
1	1.000	0.996	0.982	0.984	-0.215	0.900
2	0.996	1.000	0.992	0.978	-0.210	0.923
3	0.982	0.992	1.000	0.955	-0.201	0.953
4	0.984	0.978	0.955	1.000	-0.223	0.857
5	-0.215	-0.210	-0.201	-0.223	1.000	-0.170
7	0.900	0.923	0.953	0.857	-0.170	1.000

As duplas de bandas menos correlacionadas foram as 7-5, 3-5, 2-5, 1-5 e 5-4, conforme indica a Tabela 7-2.

Os resultados obtidos a partir da aplicação do OIF são apresentados na Tabela 7-3.

Considerando os quatro primeiros tripletos hierarquizados pelo OIF como os com a maior quantidade de informações, foram testadas as possíveis combinações RGB de cada grupo. As melhores composições coloridas foram: 451, 513, 521 e 453 (vide Figura 7-1, Figura 7-2, Figura 7-3 e Figura 7-4).

Tabela 7-3. Fatores de índice ótimo (OIF) e a soma dos desvios padrões e dos valores absolutos dos coeficientes de correlação para as possíveis combinações entre as bandas 1, 2, 3, 4 e 5 da imagem Landsat-TM, 29/8/92, órbita-ponto 220/75 e quadrante C.

Combinação de bandas	Soma dos desvios padrões	Soma dos coef. de correlação	OIF	Ordem
145	100.291	1.422	70.527	1
135	94.415	1.399	67.506	2
125	90.798	1.421	63.882	3
345	86.515	1.379	62.735	4
245	82.898	1.411	58.764	5
235	77.022	1.403	54.892	6
134	75.341	2.922	25.787	7
124	71.724	2.958	24.247	8
123	65.848	2.971	22.165	9
234	57.949	2.926	19.808	10

Na composição 451, as diferentes categorias do uso das terras são bem identificadas. O eucalipto assume uma cor vermelha bastante intensa, tornando-se esverdeado em função da heterogeneidade do plantio. Os laranjais são identificados por uma mescla de tonalidades verde, laranja e rosa. O cerrado, o cerradão e a mata aparecem em vermelho e em associações com o verde escuro, quando atingidos por processos de degradação. As matas de encosta, ciliar e as várzeas apresentam-se em vermelho mais claro. Os canaviais podem ser bem identificados nas áreas homogêneas de cor rósea e vermelha. As pastagens aparecem em matizes variando do azul, verde a ciano, confundindo-se com as solos expostos também caracterizadas por ciano e branco. Nesta combinação, os tons mais escuros de azul estão associados à umidade do solo. Os corpos d'água assumem a cor verde escuro.

Na composição 513, a cor verde representa as várias formações florestais (eucalipto, mata, cerrados etc) com suas variações tonais associadas à biomassa da seguinte maneira: quanto mais escuro maior quantidade de matéria verde. Os laranjais aparecem em verde-avermelhado e os pastos em tons pasteis, indo do amarelo ao roxo. Nesse caso, a cor azul está associada a solos expostos.

Na composição 521, as coberturas vegetais naturais aparecem nas cores verdes ao preto, cuja variação tonal está associada à biomassa das categorias. O pasto é representado por tons de roxo e rosa. As cores esverdeadas variando até o amarelo e branco estão associadas a áreas com solo exposto, tais como locais desprovidos de cobertura vegetal ou em preparo para o cultivo. Os corpos d'água aparecem em verde.

A composição 453 apresenta características bastante semelhantes a 451, porém a definição das categorias é mais nítida. Nesta composição o eucalipto é definido apenas pela cor vermelha intensa. As matas de encosta, ciliar e as várzeas tingem-se de vermelho, como o eucalipto, diferenciando-se deste pela textura rugosa típica dos padrões de vegetação natural. Os canaviais aparecem em magenta ou vermelho e os solos expostos em azul escuro e branco. Nesta combinação, os pastos ornem-se de vários matizes, tais como amarelo, rosa e ciã. Os corpos d'água são identificados por preto.

7543.8 + 787.6



Figura 7-1. Composição colorida das bandas 4/5/1-RGB da imagem Landsat-TM, de 29/8/92, cobrindo a área de estudo na bacia do rio Jacaré Pepira, Brotas e Torrinha, SP.

7543.8+^{787.6}



2 0 2 4
quilômetros

814.1
7511.3+

Figura 7-2. Composição colorida das bandas 5/1/3-RGB da imagem Landsat-TM, de 29/8/92, cobrindo a área de estudo na bacia do rio Jacaré Pepira, Brotas e Torrinha, SP.

7543.8+ 787.6



2 0 2 4
kilômetros

814.1
7511.3+

Figura 7-3. Composição colorida das bandas 5/2/1-RGB da imagem Landsat-TM, de 29/8/92, cobrindo a área de estudo na bacia do rio Jacaré Pepira, Brotas e Torrinha, SP.

7543.8+^{787.6}



Figura 7-4. Composição colorida das bandas 4/5/3-RGB da imagem Landsat-TM, de 29/8/92, cobrindo a área de estudo na bacia do rio Jacaré Pepira, Brotas e Torrinha, SP.

7.1.2 Legenda temática do uso das terras

Uma breve caracterização das principais categorias de cobertura e uso das terras constituintes da legenda temática adotada para a região de Brotas e Torrinha, incluindo alguns aspectos dos sistemas de produção empregados, é exposta a seguir:

- pastagens

Os pastos são basicamente formados por espécies exóticas como: *Brachiaria brizantha* (Hochst.) Stapf (braquiárião), *Brachiaria decumbens* Stapf (braquiárinha), *Panicum maximum* Jacq. (capim-colonião), *Pennisetum purpureum* Schum. (capim-napier), *Melinis minutiflora* Beauv. (capim-gordura) e *Paspalum notatum* Flügge (gramabatatais ou São Carlos). Há pastagens bem conservadas, com sistemas de cultivo intensivo, geralmente ocupadas pelas duas espécies de braquiárias citadas, com extensas áreas, principalmente na parte sul da região de estudo, e também pastos sujos com uma densidade elevada de plantas invasoras, incluindo arbustos desenvolvidos como a *Peschiera fuchsiaefolia* Miers. (leiteiro) ou *Sida* sp. (guanxuma) etc (vide Figura 5-9). Uma particularidade da área é a presença de pastos com grande densidade de árvores, com destaque para o faveiro (*Pterodon emarginatus* Vog.) conforme representado pela Figura 5-13c.

- canaviais

Esta cultura é bastante representativa na região, abastecendo usinas açucareiras como a Usina Paraíso, Brotas, da Barra, Piracicaba, etc. principalmente para a produção de álcool (vide Figura 5-11b, c e d).

A Usina Paraíso, responsável pela maior parte da área plantada nos municípios de Brotas e Torrinha, adota 2 variedades de cana de açúcar. Uma delas caracteriza-se por um ciclo de vida maior (18 meses), sendo normalmente plantada em janeiro e cortada nos meses de junho e julho. A outra é a cana de 12 meses, cujo plantio ocorre normalmente nos meses de julho e agosto e o corte em junho e julho.

A análise da viabilidade da aplicação de imagens Landsat-TM para identificação espectral das variedades e da idade dos cultivos, proposta no item 6.3.6, foi processada através da localização de todos os talhões de cana, em uma impressão da composição colorida. Com esses dados espaciais e os agrônômicos fornecidos pela Usina Paraíso gerou-se uma matriz descritiva das classes de informação relativas à canavicultura existentes na área focalizada, durante o período da aquisição da imagem (vide Tabela 7-4). Essa matriz foi elaborada a partir das variedades de cana, dos tipos de solos e da idade do cultivo. Foi analisada a associação entre as 14 classes de informação definidas pelas três variáveis arroladas e as classes espectrais geradas a partir da aplicação do algoritmo de classificação não supervisionada Isodata. A ausência de padrões bem definidos, o elevado número de classes de informação e a presença de muitas sobreposições e omissões entre as classes impediu o estabelecimento de uma relação entre as variedades de cana cultivada e os dados espectrais. Desse modo, através da análise visual constatou-se a impossibilidade de detectar com exatidão as diferentes variedades de cana em suas fases de crescimento através do simples uso da classificação espectral das imagens Landsat -TM.

Tabela 7-4. Idade dos plantios de cana de açúcar segundo suas variedades e tipos de solos, nas fazendas cultivadas pela Usina Paraíso na área piloto (região de Brotas e Torrinha, SP) durante o período de aquisição da imagem orbital Landsat-TM em agosto de 1992.

Idade das variedades cultivadas (meses)	Tipos de solos (sistema da Usina Paraíso)				
	1	2	3	4	5
RB 72454			8 e 10		7, 8 e 13
RB 765418		7			
RB 785148					13
SP 701143					11, 12, 13 e 14
SP 711406		10, 10.5 e 16	9	9.5 e 10.5	9.5, 10 e 11
SP 716163	7	4	8		11.5
Diversas					18

- laranjais

A principal diferença do sistema de produção dessa importante cultura em Brotas e Torrinha em relação a outras áreas do Estado é o fato de na sua maioria serem plantações formadas por produtores bem sucedidos de pólos tradicionais do setor como Limeira e Matão. São grandes propriedades com sistemas de cultivo muito intensivos e organizados (vide Figura 5-10). Entretanto, o processo de expansão desse cultivo na região foi paralisado nos últimos anos devido à crise do setor de produção de suco.

- eucaliptais

As plantações de eucaliptos são bastante extensas na região, principalmente nos solos mais pobres da parte norte da área de estudo. Grande parte dos talhões são pertencentes às indústrias de papel Ripasa de Limeira e Champion de Mogi-Guaçu (vide Figura 5-9d). Mas além desses plantios mais bem cultivados, há uma extensa área heterogênea e mal conservada, situada em uma porção edaficamente muito pobre, conforme documentado pela Figura 5-13b.

- solos expostos

No período de aquisição da imagem, esta categoria é muito freqüente e está principalmente relacionada à cultura de cana de açúcar. Fazem parte desse conjunto solos em pousio com ou sem resíduos de cultura, preparados para o plantio na estação das chuvas seguinte ou com culturas em estágio iniciais de desenvolvimento (vide Figura 5-11b) e solos erodidos sem cobertura vegetal, com a presença localizada de voçorocas (vide Figura 5-13a).

- várzeas

Essas áreas planas e próximas aos corpos d'água são inundadas durante a estação das chuvas. Estão conspicuamente presentes ao longo da calha do rio Jacaré-Pepira, formando extensos campos limpos muito característicos (vide Figura 5-7c e Figura 5-12b).

- matas e cerrados

A cobertura florestal, praticamente erradicada em todo o Estado de São Paulo, foi poupada em alguns locais na bacia do rio Jacaré-Pepira. As condições desses remanescentes em termos florísticos, estruturais e dimensionais é variada, mas são um dos patrimônios mais importantes dessa Área de Proteção Ambiental (vide Figura 5-7a, b e Figura 5-8a).

As matas ciliares ocupam solos úmidos próximos aos rios e contém espécies tolerantes à anoxia radicular durante os períodos de inundação (vide Figura 5-8b e c). Na região, essas formações aparecem ao longo da maioria dos rios e córregos, porém em geral a dimensão da faixa florestada não corresponde à largura definida pela legislação e sua composição florística e estrutura estão bastante degradadas.

Nas partes mais elevadas das encostas e no topo das colinas, alguns remanescentes da antiga cobertura de savanas existente na região podem ser localizados. Em quase sua totalidade são cerradões (vide Figura 5-8d) com estrutura definida por pelo menos dois estratos, fisionomia fechada, altura em torno de 10 m e composição alterada pela retirada seletiva de madeira de melhor qualidade, ocorrência de queimadas e perturbações mais drásticas ocorridas por volta do começo desse século.

- corpos d'água

Na região investigada por essa dissertação, essa série engloba rios, lagoas naturais e artificiais, incluindo açudes nas propriedades (vide Figura 5-12c) e a represa do Patrimônio.

- outras culturas

Uma série de tipos de uso das terras detectados em fragmentos isolados e muito pequenos não foi considerada independentemente (vide Tabela 6-1). Por exemplo, devido à similaridade de suas características espectrais com as pastagens e a pouca expressão espacial, as culturas anuais, como sorgo e milho, não receberam tratamento isolado. Sua distinção necessitaria uma abordagem multitemporal capaz de acompanhar as mudanças fenológicas ao longo das diferentes estações do ano.

Outro caso é o de culturas perenes, com cobertura restrita à determinados locais, como abacate, amora (vide Figura 5-12a) e café (vide Figura 5-11a). São parcelares agrícolas reduzidos e distribuídos principalmente no sul e leste da área de estudo. Apesar de serem inicialmente manipuladas como classes específicas, esses cultivos foram retirados da análise pois acarretavam erros muito grandes no resultado dos classificadores.

7.2 Classificação

As imagens classificadas através dos algoritmos de distância mínima euclidiana típico, distância mínima euclidiana baseada em desvios padrões, distância mínima de Mahalanobis, máxima verossimilhança e isodata são apresentadas respectivamente na Figura 7-5, Figura 7-6, Figura 7-7, Figura 7-8 e Figura 7-9.

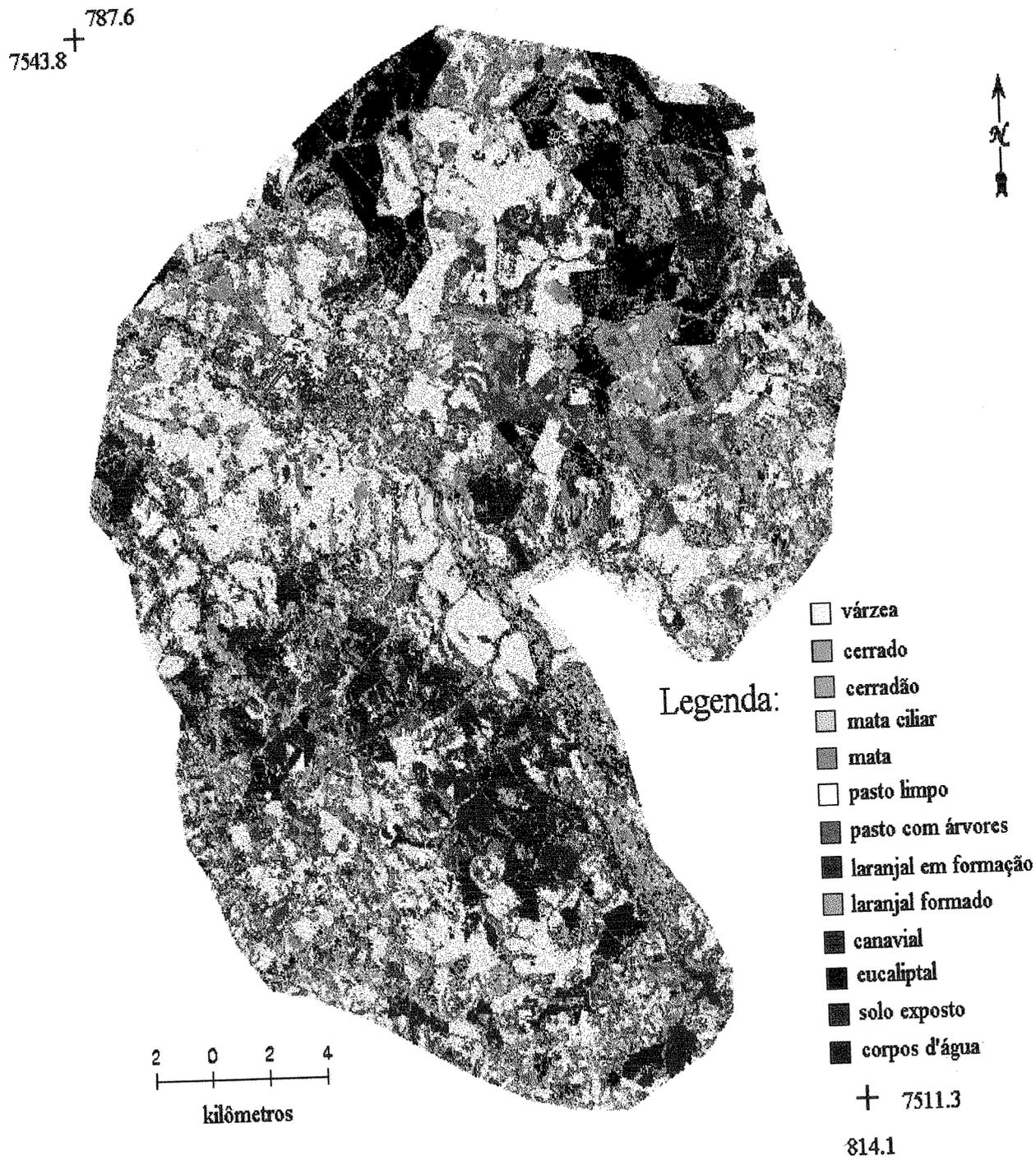


Figura 7-5. Imagem Landsat-TM da área de estudo na bacia do rio Jacaré Pepira, Brotas e Torrinha, SP, classificada pelo algoritmo de distância mínima euclidiano.

7543.8 + 787.6

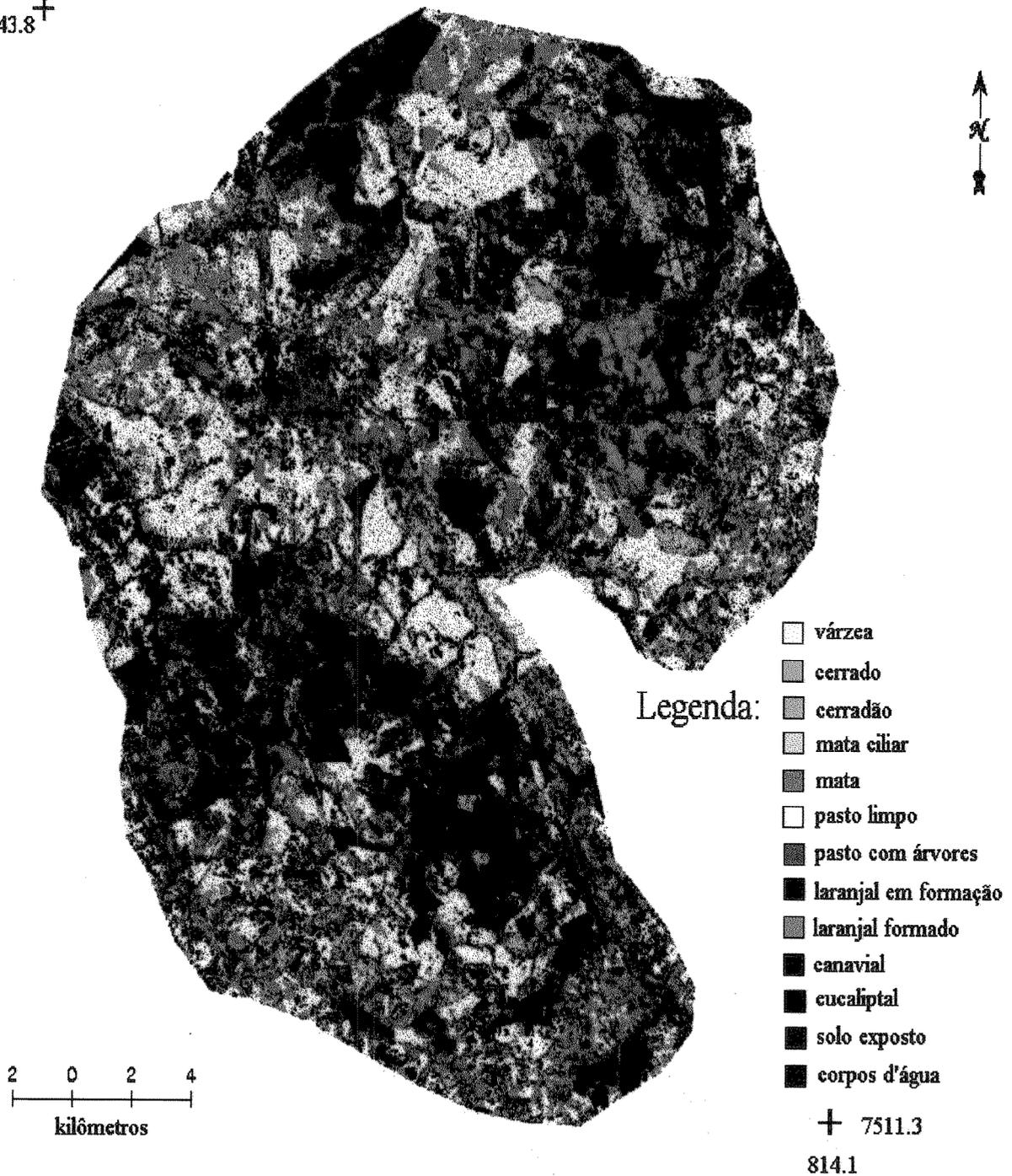
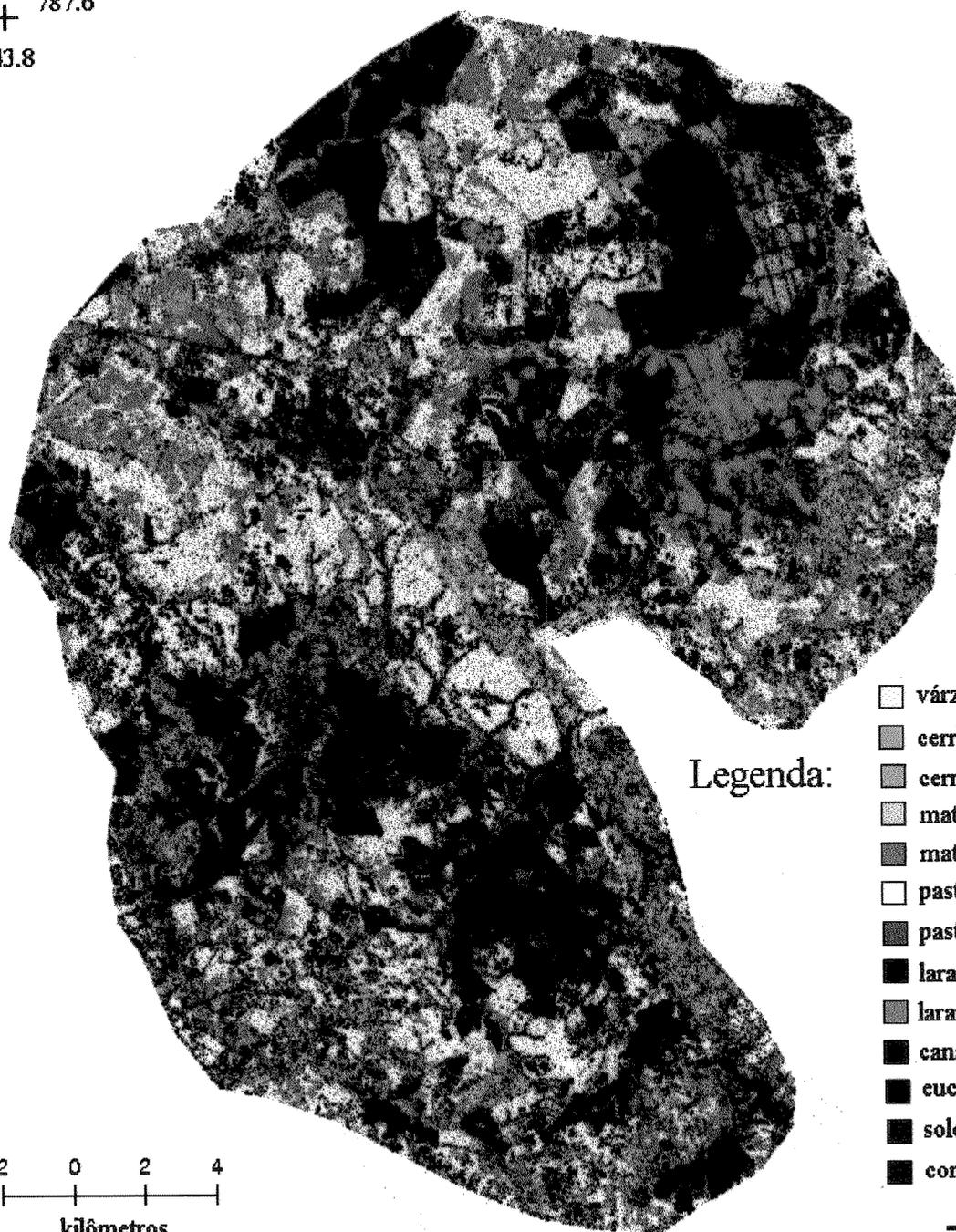


Figura 7-6. Imagem Landsat-TM da área de estudo na bacia do rio Jacaré Pepira, Brotas e Torrinha, SP, classificada pelo algoritmo de distância mínima euclidiano baseado nos desvios padrões.

+ 787.6
7543.8



Legenda:

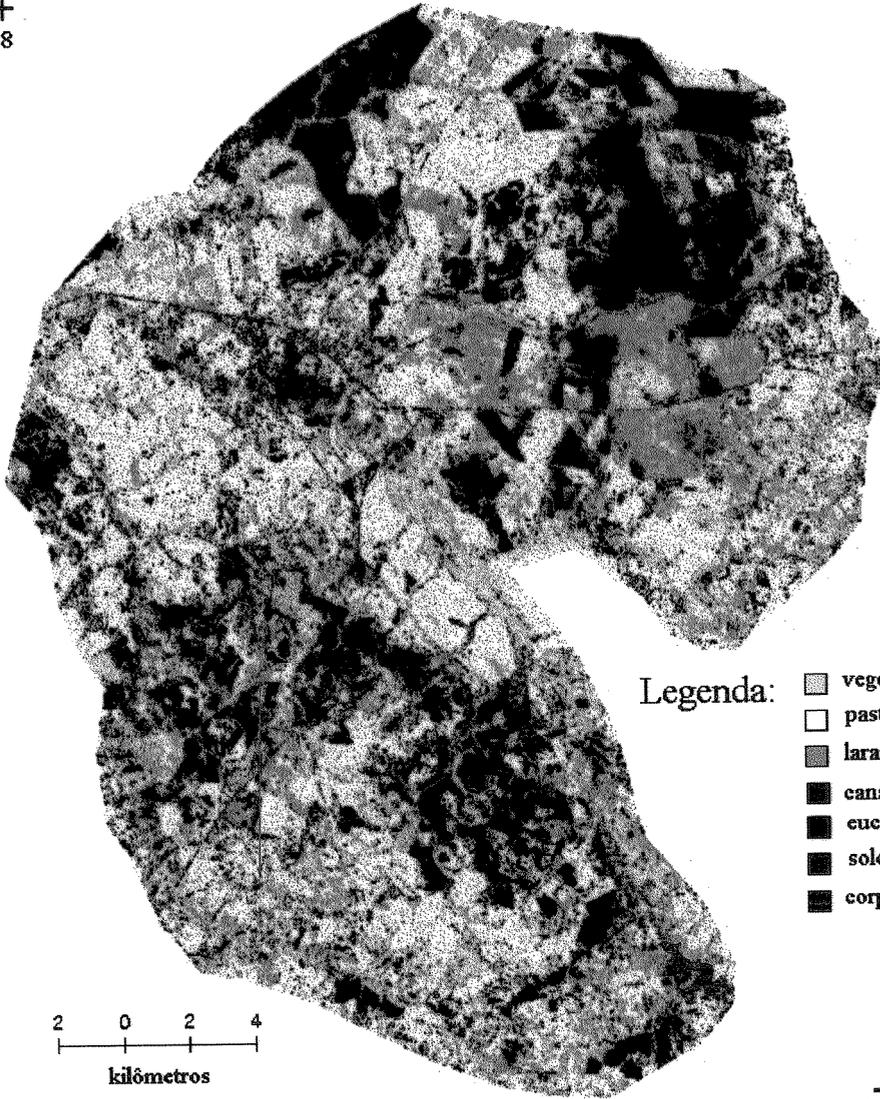
- várzea
- cerrado
- cerradão
- mata ciliar
- mata
- pasto limpo
- pasto com árvores
- laranjal em formação
- laranjal formado
- canavial
- eucaliptal
- solo exposto
- corpos d'água

2 0 2 4
kilômetros

+ 7511.3
814.1

Figura 7-8. Imagem Landsat-TM da área de estudo na bacia do rio Jacaré Pepira, Brotas e Torrinha, SP, classificada pelo algoritmo de máxima verossimilhança.

787.6
+
7543.8



- Legenda:
-  vegetação natural
 -  pasto
 -  laranjal
 -  canavial
 -  eucaliptal
 -  solo exposto
 -  corpos d'água

2 0 2 4
kilômetros

+ 7511.3
814.1

Figura 7-9. Imagem Landsat-TM da área de estudo na bacia do rio Jacaré Pepira, Brotas e Torrinha, SP, classificada pelo algoritmo isodata.

7.3 Coleta dos dados de referência

Os dados sobre o uso das terras mapeados durante o trabalho de campo segundo a legenda nº 1 acrescida das classes cultura permanente e temporária (vide Tabela 6-1), após serem revistos, uniformizados, digitalizados e manipulados em um sistema de informação geográfica, são apresentados sob a forma de um mosaico na Figura 7-10. Da esquerda para a direita, de cima para baixo e em ordem crescente, estão dispostos os 46 pontos de coleta da verdade terrestre definidos pela estratégia amostral adotada. Os quadriláteros incompletos correspondem às amostras sorteadas no limite da área de estudo selecionada. Cada célula desse mosaico corresponde a uma área de 600 por 600 m, expressa em pixels de 30 m.

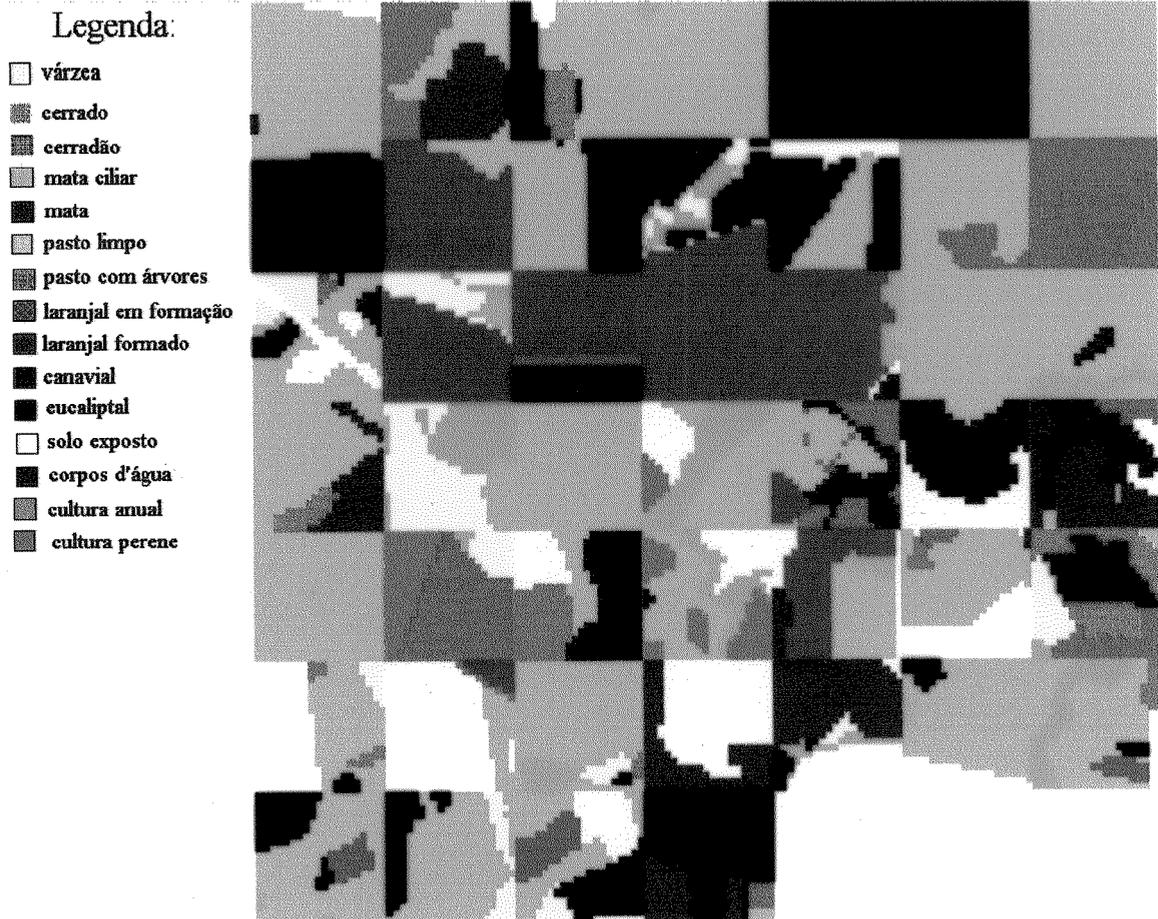


Figura 7-10. Mosaico das amostras do uso das terras coletadas na área de estudo na bacia do rio Jacaré Pepira, Brotas e Torrinha, SP.

7.4 Análise da exatidão

As 46 células de cada imagem classificada, correspondentes às amostras da verdade terrestre, sofreram o processo de recorte e fusão para geração dos mosaicos de classificação, os quais são apresentados na Figura 7-11, Figura 7-12, Figura 7-13 e Figura 7-14.

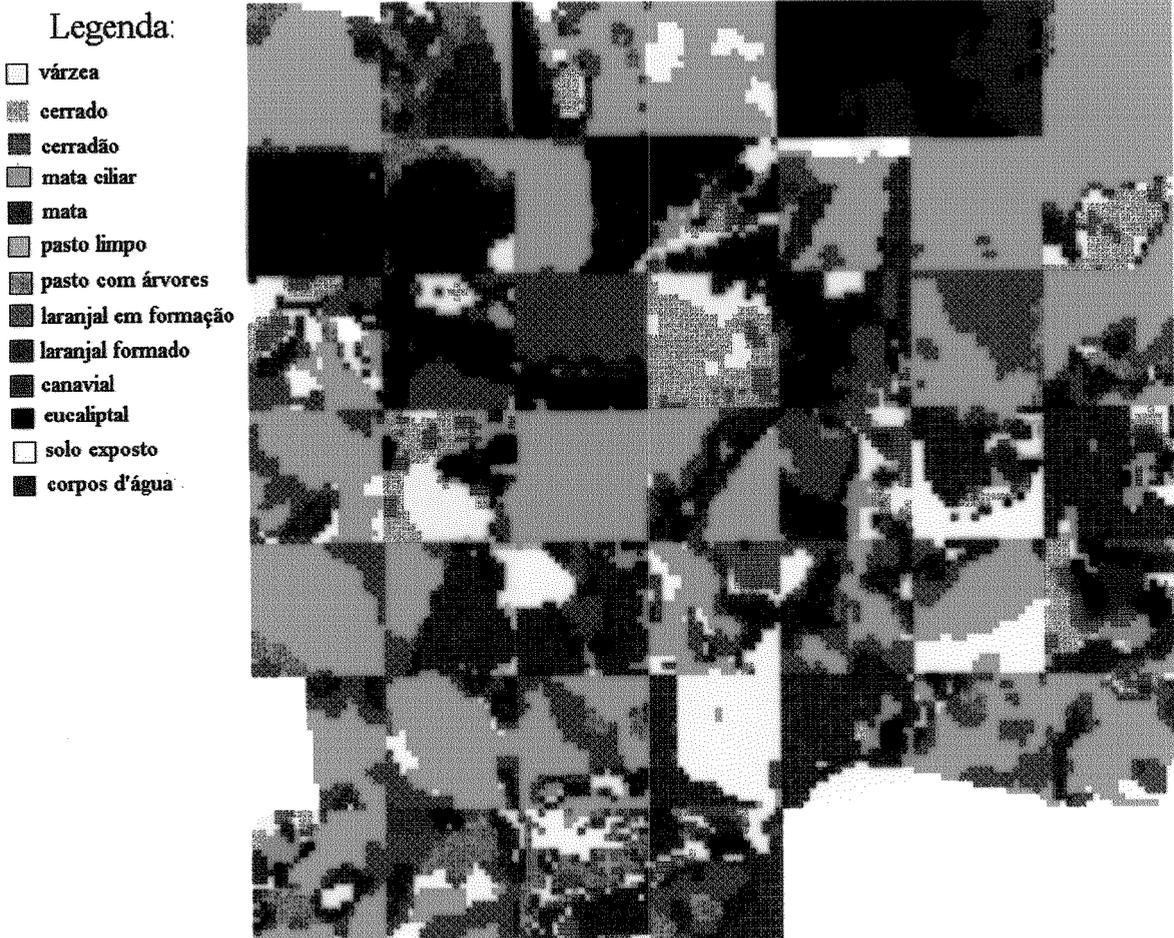


Figura 7-11. Mosaico de células da imagem Landsat-TM classificada pelo algoritmo de distância mínima euclidiano, correspondentes aos pontos amostrais na área de estudo na bacia do rio Jacaré Pepira, Brotas e Torrinha, SP.

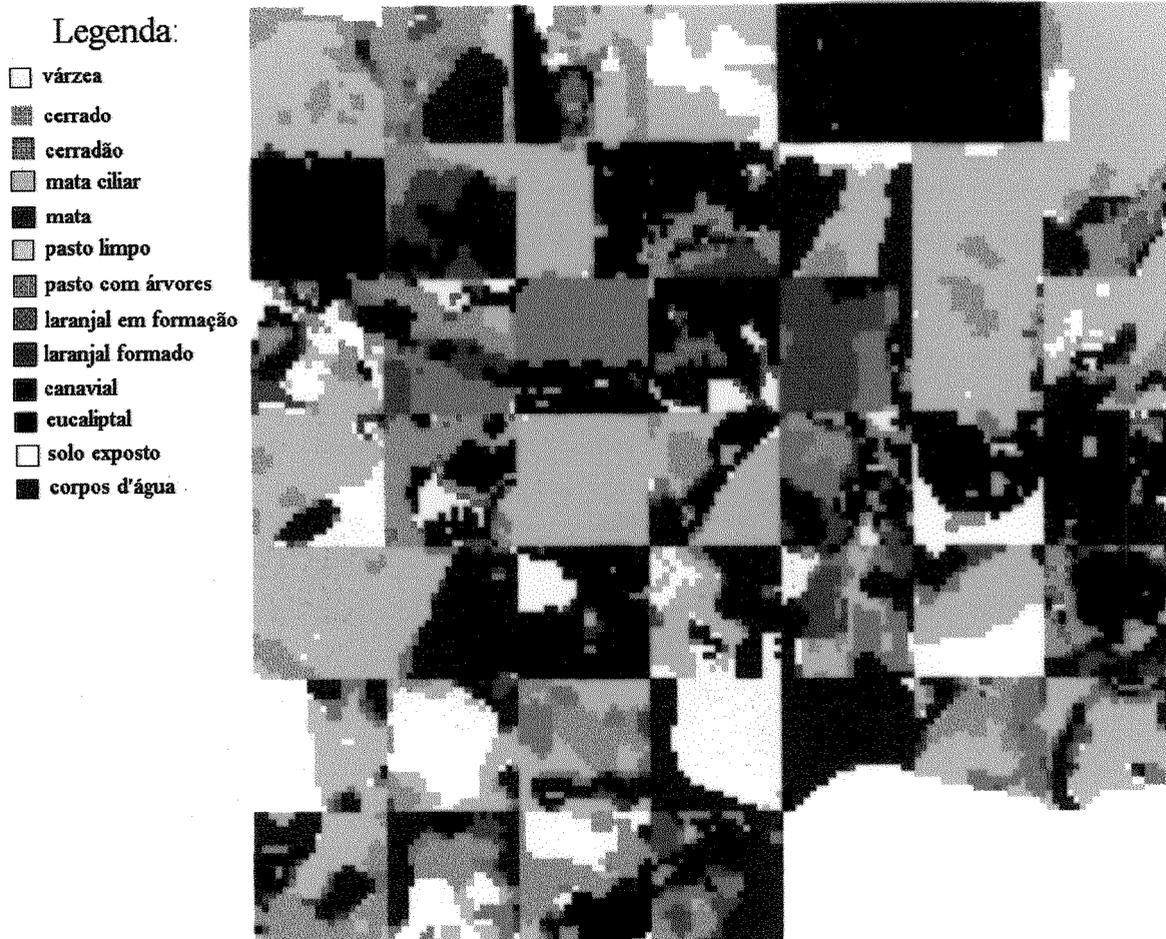


Figura 7-12. Mosaico de células da imagem Landsat-TM classificada pelo algoritmo de distância mínima euclidiano baseado nos desvios padrões, correspondentes aos pontos amostrais na área de estudo na bacia do rio Jacaré Pepira, Brotas e Torrinha, SP.

Legenda:

- várzea
- ▒ cerrado
- ▒ cerradão
- ▒ mata ciliar
- mata
- ▒ pasto limpo
- ▒ pasto com árvores
- ▒ laranjal em formação
- ▒ laranjal formado
- canavial
- eucaliptal
- solo exposto
- corpos d'água

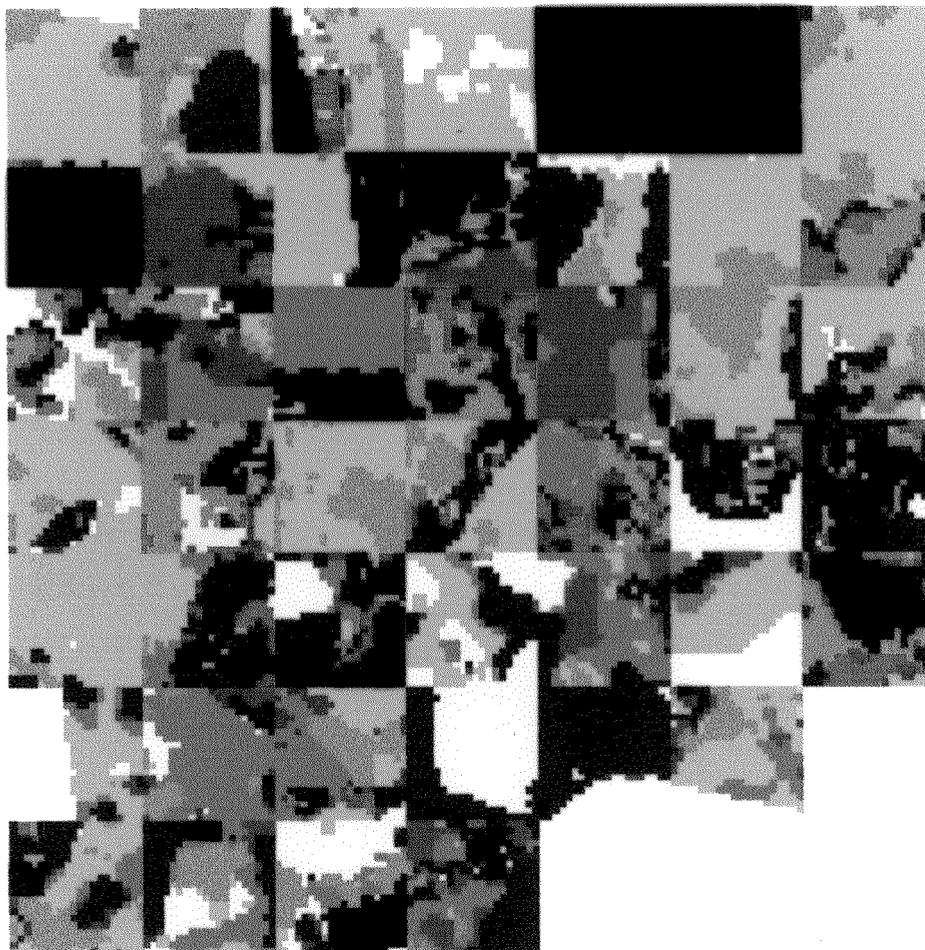


Figura 7-13 Mosaico de células da imagem Landsat-TM classificada pelo algoritmo de distância mínima de Mahalanobis, correspondentes aos pontos amostrais na área de estudo na bacia do rio Jacaré Pepira, Brotas e Torrinha, SP.

- Legenda:
- várzea
 - ▨ cerrado
 - ▩ cerradão
 - ▧ mata ciliar
 - ▦ mata
 - ▥ pasto limpo
 - ▤ pasto com árvores
 - ▣ laranjal em formação
 - ▢ laranjal formado
 - canavial
 - eucaliptal
 - solo exposto
 - ▀ corpos d'água

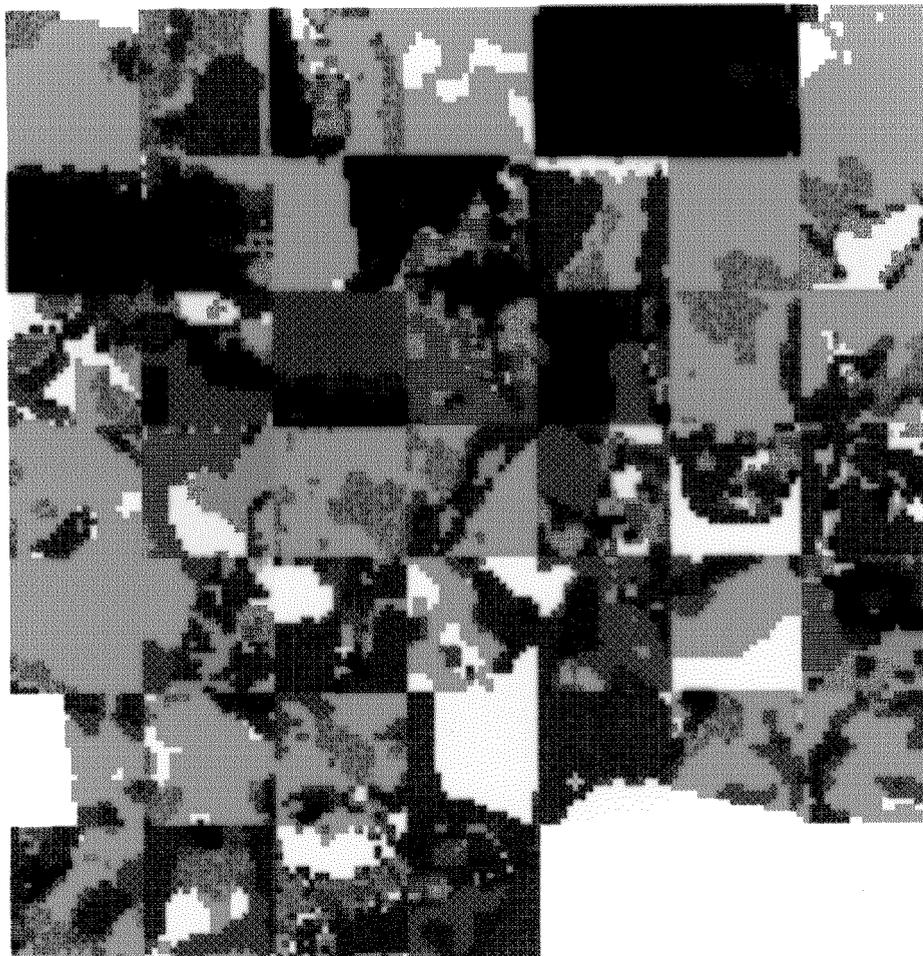


Figura 7-14. Mosaico de células da imagem Landsat-TM classificada pelo algoritmo de máxima verossimilhança, correspondentes aos pontos amostrais na área de estudo na bacia do rio Jacaré Pepira, Brotas e Torrinha, SP.

As matrizes de erros, elaboradas a partir do cruzamento do mosaico das amostras de referência com os de células das imagens classificadas, a exatidão global, os índices de concordância Kappa global e individual para cada categoria de uso das terras são descritos por uma série de trinta tabelas. As nove associadas a aplicação do algoritmo de máxima verossimilhança são apresentadas na seqüência da Tabela 7-6 a Tabela 7-14. As vinte e uma outras tabelas correspondentes a análise de exatidão do desempenho dos outros classificadores estão inclusas no Anexo 11.2.

Como esses coeficientes para análise de exatidão foram aplicados para os cinco algoritmos de classificação utilizados, adotando diferentes legendas temáticas, a discriminação dos resultados gerados está sumarizada na Tabela 7-5, para facilitar a localização dos resultados de cada tratamento.

Tabela 7-5. Sumário das análises de exatidão executadas para cada algoritmo de classificação.

Algoritmos de classificação	Coeficientes de exatidão	Chaves de classificação		
		7 classes	9 classes	13 classes
dist. mínima euclidiana	matriz de erro		Tabela 11-4	Tabela 11-1
	Exatidão global		Tabela 11-5	Tabela 11-2
	Kappa		Tabela 11-6	Tabela 11-3
dist. mínima euclidiana -desvios padrões	matriz de erro		Tabela 11-10	Tabela 11-7
	Exatidão global		Tabela 11-11	Tabela 11-8
	Kappa		Tabela 11-12	Tabela 11-9
dist. mínima de Mahalanobis	matriz de erro		Tabela 11-16	Tabela 11-13
	Exatidão global		Tabela 11-17	Tabela 11-14
	Kappa		Tabela 11-18	Tabela 11-15
máxima verossimilhança	matriz de erro	Tabela 7-12	Tabela 7-9	Tabela 7-6
	Exatidão global	Tabela 7-13	Tabela 7-10	Tabela 7-7
	Kappa	Tabela 7-14	Tabela 7-11	Tabela 7-8
Isoclass	matriz de erro		Tabela 11-19	
	Exatidão global		Tabela 11-20	
	Kappa		Tabela 11-21	

Tabela 7-6. Matriz de erro entre os dados de referência terrestre e da classificação da imagem pelo algoritmo de máxima verossimilhança segundo a legenda temática nº 1.

Classes mapeadas	Classes de referência terrestre																Total	Erro C
	1	2	3	4	5	6	7	8	9	10	11	12	14	15	16			
0	0	0	0	0	0	41	0	0	4	0	0	3	0	0	0	48		
1	135	3	136	10	27	28	2	0	4	2	7	4	2	0	0	360		
2	54	39	138	31	100	13	38	0	155	16	35	4	5	0	0	628		
3	206	31	34	89	78	13	0	0	198	2	66	6	1	0	0	724		
4	7	27	58	106	31	8	5	0	1	7	37	3	0	0	0	290		
5	1	15	9	12	187	7	0	0	0	0	78	4	1	0	0	314		
6	8	3	16	35	8	4778	154	92	22	94	30	290	0	5	14	5549		
7	32	1	8	3	7	821	195	0	15	6	2	129	11	0	32	1262		
8	18	15	32	53	15	693	23	43	541	53	75	43	1	11	11	1627		
9	33	17	21	36	19	145	36	4	846	30	93	37	8	4	3	1332		
10	116	58	218	233	215	321	16	23	153	669	466	119	39	124	34	2804		
11	15	12	3	20	33	14	0	0	2	4	1557	5	0	0	0	1665		
12	23	2	2	3	12	308	0	75	14	49	3	908	7	10	28	1444		
14	0	0	0	0	0	0	0	0	0	0	7	0	85	0	0	92		
Total	648	223	675	631	732	7190	469	237	1955	932	2456	1555	160	154	122	18139		
Erro O	0,79	0,83	0,95	0,83	0,74	0,34	0,58	0,82	0,57	0,28	0,37	0,42	0,47	1	1	0,47		

Legenda:

Linhas: dados da imagem classificada;

Colunas: dados de referência terrestre;

Erro O: erro de omissão (expresso em proporção); Erro C: erro de comissão (expresso em proporção);

Classes de usos das terras: 0. pixels não classificados; 1. várzea; 2. cerrado; 3. cerrado; 4. mata ciliar; 5. mata; 6. pastagem limpa; 7. pastagem com árvores; 8.

laranja jovem; 9. laranja formado; 10. canavial; 11. eucaliptal; 12. solo exposto; 14. corpos d'água; 15. cultura anual e 16. cultura perene.

Tabela 7-7. Índice de exatidão global da imagem classificada pelo algoritmo de máxima verossimilhança segundo a legenda temática nº 1.

Exatidão global: 0,44
Intervalo de confiança - 90%: +/- 0,0061 (0,4656 - 0,4778)
Intervalo de confiança - 95%: +/- 0,0073 (0,4645 - 0,4790)
Intervalo de confiança - 99%: +/- 0,0096 (0,4622 - 0,4813)

Tabela 7-8. Índice Kappa global e individual entre os dados de referência terrestre e da classificação da imagem pelo algoritmo de máxima verossimilhança segundo a legenda temática n° 1.

Imagem classificada		Imagem de referência	
Categorias	KIA	Categorias	KIA
não rotulados	0,00	várzea	0,19
várzea	0,35	cerrado	0,15
cerrado	0,05	cerradão	0,01
cerradão	0,01	mata ciliar	0,15
mata ciliar	0,34	mata	0,24
mata	0,58	pastagem	0,52
pastagem	0,77	pastagem com árvores	0,37
pastagem com árvores	0,13	laranja jovem	0,10
laranja jovem	0,01	laranja formado	0,39
laranja formado	0,59	canavial	0,67
canavial	0,20	eucaliptal	0,60
eucaliptal	0,93	solo exposto;	0,55
solo exposto;	0,59	corpos d'água	0,53
corpos d'água	0,92	cultura anual	0,00
		cultura perene	0,00
Índice de concordância Kappa global: 0,44			

Tabela 7-9. Matriz de erro entre os dados de referência terrestre e da classificação da imagem pelo algoritmo de máxima verossimilhança segundo a legenda temática n° 2.

Classes mapeadas	Classes de referência terrestre											Total	Erro C
	1	2	6	8	9	10	11	12	14	15	16		
0	0	0	41	0	4	0	0	3	0	0	0	48	1
1	135	176	30	0	4	2	7	4	2	0	0	360	0,63
2	268	985	84	0	354	25	216	17	7	0	0	1956	0,50
6	40	81	5948	92	37	100	32	419	11	5	46	6811	0,13
8	18	115	716	43	541	53	75	43	1	11	11	1627	0,97
9	33	93	181	4	846	30	93	37	8	4	3	1332	0,36
10	116	724	337	23	153	669	466	119	39	124	34	2804	0,76
11	15	68	14	0	2	4	1557	5	0	0	0	1665	0,06
12	23	19	308	75	14	49	3	908	7	10	28	1444	0,37
14	0	0	0	0	0	0	7	0	85	0	0	92	0,08
Total	648	2261	7659	237	1955	932	2456	1555	160	154	122	18139	
Erro O	0,79	0,56	0,22	0,82	0,57	0,28	0,37	0,42	0,47	1	1		0,38

Legenda:

Linhas: dados da imagem classificada;

Colunas: dados de referência terrestre;

Erro O: erro de omissão (expresso em proporção);

Erro C: erro de comissão (expresso em proporção);

Classes de usos das terras: 0. pixels não classificados; 1. várzea; 2. mata; 6. pastagem; 8. laranja jovem; 9. laranja formado; 10. canavial; 11. eucaliptal; 12. solo exposto; 14. corpos d'água; 15. cultura anual; 16. cultura perene.

Tabela 7-10. Índice de exatidão global da imagem classificada pelo algoritmo de máxima verossimilhança segundo a legenda temática n° 2.

Exatidão global: 0,38
Intervalo de confiança - 90%: +/- 0,0059 (0,3779 - 0,3898)
Intervalo de confiança - 95%: +/- 0,0071 (0,3768 - 0,3909)
Intervalo de confiança - 99%: +/- 0,0093 (0,3746 - 0,3932)

Tabela 7-11. Índice Kappa global e individual entre os dados de referência terrestre e da classificação da imagem pelo algoritmo de máxima verossimilhança segundo a legenda temática nº 2.

Imagem classificada		Imagem de referência	
Categorias	KIA	Categorias	KIA
não rotulados	0,00	várzea	0,19
várzea	0,35	mata	0,37
mata	0,43	pastagem	0,64
pastagem	0,78	laranja jovem	0,10
laranja jovem	0,01	laranja formado	0,39
laranja formado	0,59	canavial	0,67
canavial	0,20	eucaliptal	0,60
eucaliptal	0,93	solo exposto	0,55
solo exposto	0,59	corpos d'água	0,53
corpos d'água	0,92	cultura anual	0,00
		cultura perene	0,00
Índice de concordância Kappa global=0,51			

Tabela 7-12. Matriz de erro entre os dados de referência terrestre e da classificação da imagem pelo algoritmo de máxima verossimilhança segundo uma redução da legenda temática nº 2.

Classes mapeadas	Classes de referência terrestre										Total	Erro C
	2	6	9	10	11	12	14	15	16			
0	0	41	4	0	0	3	0	0	0	48	1	
2	1564	114	358	27	223	21	9	0	0	2316	0,32	
6	254	6799	578	153	107	462	12	16	57	8438	0,19	
9	126	185	846	30	93	37	8	4	3	1332	0,36	
10	840	360	153	669	466	119	39	124	34	2804	0,76	
11	83	14	2	4	1557	5	0	0	0	1665	0,06	
12	42	383	14	49	3	908	7	10	28	1444	0,37	
14	0	0	0	0	7	0	85	0	0	92	0,08	
Total	2909	7896	1955	932	2456	1555	160	154	122	18139		
Erro O	0,46	0,14	0,57	0,28	0,37	0,42	0,47	1	1		0,31	

Legenda:

Linhas: dados da imagem classificada;

Colunas: dados de referência terrestre;

Erro O: erro de omissão (expresso em proporção);

Erro C: erro de comissão (expresso em proporção);

Classes de usos das terras: 0. pixels não classificados; 2. áreas naturais; 6. pastagem; 9. laranja; 10. canavial; 11. eucaliptal; 12. solo exposto; 14. corpos d'água; 15. cultura anual; 16. cultura perene.

Tabela 7-13. Índice de exatidão global da imagem classificada pelo algoritmo de máxima verossimilhança segundo uma redução da legenda temática nº 2.

Exatidão global: 0,31
Intervalo de confiança - 90%: +/- 0,0057 (0,3092 - 0,3205)
Intervalo de confiança - 95%: +/- 0,0068 (0,3081 - 0,3216)
Intervalo de confiança - 99%: +/- 0,0089 (0,3059 - 0,3237)

Tabela 7-14. Índice Kappa global e individual entre os dados de referência terrestre e da classificação da imagem pelo algoritmo de máxima verossimilhança segundo uma redução da legenda temática nº 2.

Imagem classificada		Imagem de referência	
Categorias	KIA	Categorias	KIA
não rotulados	0,00	áreas naturais	0,47
áreas naturais	0,61	pastagem	0,74
pastagem	0,66	laranjal	0,39
laranjal	0,59	canavial	0,67
canavial	0,20	eucaliptal	0,60
eucaliptal	0,93	solo exposto	0,55
solo exposto	0,59	corpos d'água	0,53
corpos d'água	0,92	cultura anual	0,00
		cultura perene	0,00
Índice de concordância Kappa global=0,58			

8. Discussão

A exatidão é o resultado de todos os erros presentes nas diversas fases de sua determinação.

Os principais aspectos dessa dissertação foram discutidos conforme a estrutura apresentada no item 6.2.

Esta dissertação foi desenvolvida privilegiando aspectos metodológicos relacionados com avaliação do desempenho dos classificadores espectrais através da determinação de sua exatidão. Entretanto, há uma demanda significativa por pesquisas sobre exatidão da classificação de imagens em problemas imediatos à gestão dos recursos naturais, a semelhança do trabalho desenvolvido por Bailey et al. (1994) para o Serviço de Parques Nacionais do EUA. Seria interessante explorar a potencialidade da região, tanto em termos da sua representatividade e heterogeneidade no âmbito do Estado de São Paulo como nas facilidades oferecidas pela base de dados gerada por outros projetos científicos realizados na região (item 5.1). Um ponto a ser sempre observado na execução de projetos, é a diferença nas abordagens e procedimentos entre os estudos orientados para aperfeiçoamentos metodológicos e os dirigidos para equacionar problemas ambientais. Os primeiros, muitas vezes, lidam com situações ideais e um número reduzido de variáveis controladas. Geralmente não são abordados adequadamente os problemas relacionados com a aplicabilidade dos métodos, tais como os gradientes espaciais e temporais dos elementos terrestres, a disponibilidade logística, as limitações dos cronogramas de execução reduzidos, as restrições condicionadas pelos interesses políticos, econômicos e administrativos. Assim a adequabilidade dos procedimentos em situações gerenciais deve ser avaliada em termos da base científica, justificativa e implementação prática (restrições de custo e de tempo).

8.1 Preparação das imagens

8.1.1 Seleção de bandas

O método OIF aplicado neste trabalho indicou os tripletos 145, 135 e 125 como as três combinações com maior quantidade de informações espectrais (vide Tabela 7-3). Teoricamente, os OIFs com valores mais altos representam os tripletos com maior quantidade de informação. Contudo após análise visual das composições coloridas executadas para as quatro primeiras combinações apontadas pelo OIF (vide Figura 7-1, Figura 7-2, Figura 7-3 e Figura 7-4), optou-se pela utilização das bandas 4,5,3 (RGB). Observou-se que apresentavam um maior contraste visual entre os elementos terrestres, principalmente no que se refere à vegetação (vide Figura 7-4). Esse resultado está coerente com vários trabalhos onde este triplete foi considerado nos estudo da cobertura terrestre (Townshend et al., 1983; Townshend, 1984; Atkinson et al., 1985; Fuller et al., 1989; Fuller & Parsell, 1990; Karteris, 1990; Janssen & Middelkoop, 1992; Tardin et al., 1992;

Tennakoon et al., 1992; Congalton et al., 1993; Fuller et al. 1994; Nelson et al., 1994; Tuomisto et al., 1994; Jusoff et al., 1995; Sader et al., 1995 e Sharma et al., 1995).

Muitas outras pesquisas determinaram outras combinações como as melhores. Por exemplo, Bouchardy & Rodriguez (1992) apontam as bandas 4, 3 e 1 como as mais reveladoras da extrema heterogeneidade de uma área agrícola. Hernandez (1994) obteve os maiores OIFs nos tripletos 345 e 457. Essa variação se deve ao fato do índice OIF ser apenas um indicador da seleção, sendo a escolha final das bandas determinada pelo tipo de informação desejada pelo analista, como alertado por Sheffield (1985).

Para este trabalho, evitou-se utilizar procedimentos de redução de dimensionalidade dos dados espectrais, por serem pouco usuais em avaliações dos usos das terras e por transformarem as variáveis originais em um sistema de referência diferente. As modificações variam de acordo com as características da imagem, tendendo a complicar a interpretação dos resultados finais.

Muitas vezes é difícil para um analista sem prática em lidar com dados orbitais, aplicar sua experiência prévia sobre a relação entre os elementos da superfície terrestre e seus respectivos padrões de cores em uma composição colorida gerada a partir das imagens não originais (Sheffield, 1985; Fung & Ledrew, 1987 e Chavez & Kwarteng, 1989).

As combinações coloridas utilizando bandas originais produzem imagens com cores a primeira vista incomuns, mas a designação das cores não é dependente da cena. Assim, em oposição às cores dependente da cena de alguns métodos de redução de dimensionalidade, como principais componentes ou razão de bandas, o intérprete rapidamente aprende a associar as cores com as condições do terreno, fundamental para a atuação em órgãos de planejamento ambiental, onde são necessárias, ferramentas de fácil aprendizado e rápido manuseio. Portanto, há vantagens na utilização das bandas originais em relação às razões de banda ou outros tipos de combinações, facilitando a padronização e difusão do uso de SR em estudos ambientais, ecológicos e biológicos.

O método OIF (Chavez et al., 1982) vem ao encontro desta linha de análise. Uma das suas vantagens está no fato de levar em consideração tanto o desvio padrão como a correlação entre as bandas. Se somente o desvio padrão ou as variâncias das bandas fossem usadas para a seleção do triplete, poderia se incorrer na seleção de bandas com o maior desvio padrão, porém com a maior correlação, desconsiderando um segmento das informações espectrais contidas na imagem.

Conese & Maselli (1993) sugerem o uso da análise de informação mútua para a seleção de bandas Landsat TM para a análise de vegetação, pois essa técnica pode ser aplicada na determinação de quais bandas são mais informativas em relação a uma variável ecológica escolhida. Deste modo, obtém-se a porcentagem de informação expressa em cada banda sobre um certo tema, evitando a necessidade de verificar se as combinações com maior quantidade de informações espectrais correspondem as mais adequadas à análise de um determinado tema. Esse método se baseia nos conceitos do sistema de entropia (Kullback, 1959). Em geral, esse conceito é utilizado para quantificar a informação expressa por uma certa variável sobre um sistema. Este enfoque já foi aplicado em estudos de sensoriamento remoto para avaliar a informação contida em imagens Landsat -TM para fins de compressão de dados (Chen et al., 1987).

8.1.2 Definição da área piloto

A escolha da região de Brotas-Torrinha como objeto de uma dissertação na área de sensoriamento remoto mostrou-se adequada por diferentes aspectos.

Primeiro, a grande diversidade de atividades primárias, sistemas de produção, feições geomorfológicas, paisagens e fragmentos de áreas naturais, permite amostrar uma série de situações em um território delimitado, reduzindo a alocação de tempo e recursos logísticos nas campanhas de levantamento de dados e conseqüentemente diminuindo os custos dos projetos de pesquisa.

Além dessa pluralidade no âmbito dos agroecossistemas, ao contrário das regiões temperadas, no Brasil não há restrições sazonais drásticas às atividades agrossilvipastoris, permitindo a ocorrência concomitante de vários estágios de desenvolvimento de uma mesma cultura. Este fator, cuja importância não pode ser desconsiderada pelo intérprete de produtos suborbitais e orbitais de sensoriamento remoto no cenário nacional, pode ser encontrado com facilidade na região de estudo, principalmente na canavicultura.

Além das vantagens acima abordadas, sua localização próxima a universidades e instituições de pesquisa com programas na área de sensoriamento remoto e ecologia, como UNICAMP e EMBRAPA em Campinas, UNESP em Rio Claro, USP e UFSCar em São Carlos e INPE em São José dos Campos, é um outro ponto favorável à seleção desta região como área teste.

Por último, um fator favorável existente na região de Brotas e Torrinhã é a existência de uma base topográfica em uma escala cartográfica relativamente grande (1:10.000), o que facilita muito a localização no terreno dos pontos amostrais e das manchas dos diferentes tipos de uso das terras presentes na imagem.

8.1.3 Padronização entre legenda temática do uso das terras e dados espectrais.

Esses aspectos fisiográficos, humanos e logísticos conferem à região uma grande complexidade ambiental. Obter bons resultados na classificação de imagens orbitais nestas condições é justamente um dos desafios do SR.

A definição de legendas temáticas muito detalhadas, visando classificar imagens Landsat-TM não é recomendada devido as restrições espaciais e espectrais intrínsecas do sensor. Desse modo a discriminação inadequada de sistemas de classificação da cobertura e uso das terras com muitos níveis hierárquicos não está associada a problemas estruturais dos classificadores. Entretanto com o advento e disseminação iminentes de produtos com alta resolução espacial e espectral (Crósta, 1996), esse cenário será muito transformado, possibilitando, por exemplo, o desenvolvimento de aplicações para os níveis III e IV do sistema de classificação de cobertura e uso das terras do U.S.G.S., cuja relação com a escala espacial é respectivamente entre 1:80.000 e 1:20.000 e maior que 1:20.000 (Anderson et al., 1976)

A limitação das imagens Landsat-TM para a classificação espectral está associada ao fato dos elementos terrestres serem um conjunto de fenômenos com variados graus de diferenciação, identificados como uma classe simples e homogênea. Esse fato pode ser bem constatado nas classes de vegetação natural analisadas no item 8.2. Para evitar esse tipo de

barreira, a maioria dos autores estuda a exatidão da classificação baseando-se em um número reduzido de macroclasses. A maioria dos trabalhos com exatidão de classificação relativamente alta lida com poucas classes de uso/cobertura das terras (em média 6), reduzindo a sobreposição espectral (vide Tabela 8-1). São exemplos dessa abordagem os seguintes trabalhos: 6 classes (Fisher, 1994); 5 classes (Shimoda et al., 1988), 4 classes (Hoffer, 1975, Stenback & Congalton, 1990), 3 classes (Story & Congalton, 1986). Além disso, níveis elevados de exatidão também podem ser verificados em procedimentos de classificação alternativos (Conese & Maselli, 1992; Green et al., 1993; Bauer et al., 1994 e Jusoff & Manaf, 1995).

Tabela 8-1. Síntese de aspectos selecionados de alguns trabalhos sobre o cálculo da classificação do uso das terras a partir de imagens orbitais.

Autores	Classes de cobertura/uso	Classificador	Exatidão global	Kappa=	Observações
Hoffer, 1975	4 classes: floresta de coníferas, floresta decídua, agricultura e água	algoritmos não supervisionado, supervisionado modificado e análise de agrupamentos modificada.	G=0.7663 G=0.7845 G=0.7136 G=0.8592	K=0.605 (0.00073) K=0.586 (0.00087) K=0.476 (0.001099) K=0.718 (0.00076)	não supervisionado com 10 grupos não supervisionado com 20 grupos supervisionado modificado an. de agrupamentos modificada
Story & Congalton, 1986	3 classes: floresta, água e área urbana	classificador misto com não supervisionado e supervisionado	G=63%	K=49.0	
Chuvieco & Congalton, 1988	8 classes: área urbana, residencial, campos d'água, pinheirais, vegetação arbustiva, vegetação arbustiva + solo exposto, culturas irrigadas e culturas não irrigadas.	máxima verossimilhança	G=48.1 G=57.5	K=49.8	
Shimoda et al., 1988	5 classes: vegetação, área urbana, arroz irrigado, água e outros	máxima verossimilhança	G=60% a 70%		
Stenback & Congalton, 1990	4 classes de floresta de coníferas: dossel espesso com e sem sub-bosque, dossel moderado com e sem sub-bosque	algoritmo não-supervisionado	G=86.0	K=72.8	
Rocha, 1992	11 classes: cultura anual e pasto, canavieira, citricultura 1, citricultura 2, pinheiral, eucalipto, solo exposto 1, solo exposto 2, mata ciliar, várzea e campos d'água	máxima verossimilhança análise de agrupamentos	G=64.48 G=63.91	K=58.79	
Comese & Maselli, 1992	5 classes: pradaria, área reflorestada, área agrícola, vegetação arbustiva densa e vegetação arbustiva esparsa	máxima verossimilhança tradicional e com matriz de transição não definido		K=0.5082, 0.5702 e 0.3735 K=0.5107, 0.6541 e 0.6032	
Congalton & Green, 1993	3 classes: flor. de coníferas com mais de 70% de oclusão no dossel e mais de 10% de oclusão no estrato de 32 in. de dap) ou mais; flor. de coníferas com 70% de oclusão no dossel e mais de 10% de oclusão no estrato de 21 a 31 in. de dap.		G=82%	Não especificado	
Green et al., 1993	6 classes: floresta, vegetação não florestal, área construída desenvolvida, solo exposto ou áreas abandonadas, água e nuvens	classificador misto com análise de agrupamentos e máxima verossimilhança	G=86%		
Gong & Howarth, 1990b	12 classes: áreas residenciais, áreas industrial/comercial, pavimentações, solo exposto, gramados e árvores, gramíneas cultivadas, árvores decíduas, coníferas, área cultivada, pasto e culturas novas, solo exposto e água	máxima verossimilhança	G=86.1 G=76.7		inclusão de informação estrutural

Bauer et al., 1994	11 classes: florestas de terras baixas; álamo/bétula; florestas boreais; coníferas das terras altas; abeto balsâmico/ abeto branco; coníferas das terras baixas; arbustos + desmatamentos + gramados; agricultura; construções, água e pântanos	classificador misto: análise de agrupamentos orientada	G=63 a 75%	Kappa = 0.56 a 0.76.	
Maselli et al., 1994	5 classes: floresta de coníferas, floresta decídua, cereais, oliveiras e área urbana	máxima verossimilhança	G=0.577	k=0.374	amostragem estratificada aleatória
Fisher, 1994	6 classes: água, construção, asfalto, áreas abertas, solo exposto, florestas/várzeas		G=77%		
Jusoff & Manaf, 1995	6 classes: floresta, área urbana, seringais, óleo de coco, solo exposto e campo	classificador misto com máxima verossimilhança e análise de agrupamentos	G=90%		

A Tabela 8-2 demonstrou o aumento dos valores de coeficiente de concordância Kappa global e conseqüentemente da exatidão de classificação, quando o número de classes foi reduzido de 13 para 9 e no caso do Maxver para 7.

Tabela 8-2. Coeficientes de concordância Kappa global obtidos pelos classificadores Maxver; distância mínima de Mahalanobis, euclidiana simples e baseada em desvios padrões e isodata.

Classificador	Kappa - 13 classes	Kappa - 9 classes	Kappa - 7 classes
Máxima verossimilhança	0.4362	0.5147	0.5800
Distância Mahalanobis	0.4197	0.5037	
Distância euclidiana	0.3972	0.4458	
Distancia euclidiana-DP	0.3681	0.4361	
Isodata	-	-	0.4841

Esta análise foi de grande utilidade na definição de uma legenda temática apropriada para as especificidades da paisagem rural do entorno de Brotas e Torrinha. Portanto para a utilização de imagens Landsat-TM em áreas complexas, como a região estudada, o uso de macroclasses é o mais indicado.

A definição da legenda é um processo dinâmico, exigindo adequações constantes durante os procedimentos de classificação e análise de sua exatidão. Essa fase de definição da legenda temática para a classificação é fundamental para orientar o analista na opção pelo tipo e formato de imagem a ser empregado (sensor, analógico ou digital), pelo classificador adequado e em outras decisões subseqüentes no processo de avaliação dos usos das terras, para atender às demandas dos administradores municipais e empresários da região.

8.2 Classificação

8.2.1 Seleção das amostras de treinamento

Uma conseqüência da dificuldade de estabelecer sistemas de classificação adequados *a priori* é a introdução de incertezas no processo de amostragem das áreas de treinamento. O caráter dinâmico da legenda temática para a categorização da cobertura e uso das terras implica em constantes alterações no processo de localização, identificação e delimitação das amostras de treinamento. Classes em levantamento podem ser fundidas ou divididas conforme a consolidação da legenda.

Além disso, a ausência de um diagrama de dispersão ("scattergram") na versão do sistema de processamento digital utilizado na execução dessa dissertação prejudicou a verificação e possível redução de pixels indesejáveis, o que poderia evitar a escolha de classes com grande sobreposição espectral durante o treinamento.

Outro fator prejudicial para o processo de coleta de áreas de treinamento e conseqüentemente para a classificação correta das séries, é o reduzido tamanho de algumas categorias como as formas de vegetação natural. Por esse motivo as classes melhor

classificadas são aquelas com maior extensão contínua, como os reflorestamentos de eucaliptos, as pastagens e as plantações de cana de açúcar.

8.2.2 Identificação dos padrões de uso das terras

Analisando as Tabela 7-6, Tabela 7-9 e Tabela 7-12, pôde-se observar que o algoritmo de máxima verossimilhança identificou corretamente grande parte dos bosques de eucalipto e das pastagens.

No caso do eucalipto alguns pixels foram designados para a classe cana, enquanto outros foram incluídos nas classes de vegetação natural da região, onde quase todos os fragmentos caracterizam-se por algum grau de perturbação, como por exemplo os cerrado e as matas ciliares. O maior problema na identificação do eucalipto está associado às áreas com indivíduos jovens ou próximos às áreas de transição com outras classes de usos, tais como, mata ciliar e pastagens, onde são mais comuns as interferências na resposta espectral da classe.

As alterações na estrutura geradas pela grande quantidade de gramíneas invasoras nas áreas de vegetação natural e pelo corte seletivo de madeira devem ser responsáveis pela confusão verificada entre esta e a categoria da cana de açúcar. Além disso, as vegetações naturais foram muito confundidas entre si, apesar de apresentarem pouca inclusão de outras categorias de cobertura terrestre. As respostas espectrais das faixas estreitas de mata ciliar inseridas nos reflorestamentos de eucalipto apresentam uma influência da reflectância do eucalipto, como abordado anteriormente, confundindo o classificador. Além disso as matas ciliares da região estão degradadas e descaracterizadas.

As outras categorias de uso das terras foram abordadas por esse classificador da seguinte maneira: as áreas de pastagens, pastagens com árvores e corpos d'água foram bem definidas; os laranjais apresentaram uma grande confusão entre seus estágios de desenvolvimento; o solo exposto foi bem identificado havendo a inclusão de área urbana, que não foi contemplada na fase de coleta de áreas de treinamento.

O algoritmo de classificação distância mínima de Mahalanobis (vide Tabela 11-13 e Tabela 11-16), apresentou um desempenho muito semelhante ao de máxima verossimilhança. Entretanto foi verificada uma diferença significativa em relação às classes de vegetação natural, com uma menor sensibilidade as variações de cobertura vegetal, destinando a maior parte das áreas de cerrado, mata ciliar e várzea para a classe mata.

Em comparação ao Maxver, os classificadores baseados na distância euclidiana (vide Tabela 11-1, Tabela 11-4, Tabela 11-7 e Tabela 11-10) tiveram uma eficiência muito inferior. As áreas de pasto foram designadas erroneamente para a classe solo em preparo. Apesar de continuar identificando corretamente as áreas de cana, foram incluídas nesta classe outras categorias como laranja, eucalipto e vegetação natural. Extensas áreas de eucalipto foram classificadas como mata e cana. Os pixels correspondentes aos cerrados e cerradões foram designados para a classe várzea.

O algoritmo paralelepípedo apresentou um desempenho no processo de rotulação dos pixels, muito inferior ao demais, sendo a imagem resultante da classificação eliminada das análises posteriores.

As classes espectrais identificadas pelo classificador não-supervisionado isoclass cujos padrões puderam ser associados às classes de cobertura são: eucaliptais, vegetação natural, canaviais, solo exposto, corpos d'água e laranjais.

Seguindo os resultados dos classificadores supervisionados, a sobreposição entre as respostas espectrais das classes vegetação natural e eucaliptais ocorreu principalmente nas áreas de contato entre essas classes. Outros erros de designação foram observados entre os laranjais e as áreas de vegetação natural de extensão espacial reduzida e entre os canaviais e as categorias de vegetação natural. Em relação as áreas de solo exposto, estes foram adequadamente detectadas.

Quanto ao tempo de processamento requisitado pelos algoritmos, o paralelepípedo, apesar de seu desempenho fraco na rotulação dos pixels, foi o mais rápido (7 min.), seguido pelos algoritmos de distância mínima euclidiana simples e baseado nos desvios padrões (respectivamente 8 e 11 min.). O classificador distância mínima de Mahalanobis e o de máxima verossimilhança apresentaram velocidade de processamento semelhantes, 20 e 19 minutos, respectivamente.

De acordo com as considerações acima, os coeficientes de concordância Kappa mostraram que o Maxver gerou os melhores resultados de classificação seguido da distância mínima de Mahalanobis, distância mínima euclidiana simples e baseada no desvio padrão (vide 6). O isodata apresentou limitações na geração de classes espectrais correspondentes às classes de informações utilizadas. Conseqüentemente foi possível calcular o coeficiente de concordância apenas para 7 classes. Neste caso, o Maxver continua apresentando os melhores resultados.

Os processamentos envolvendo a legenda temática nº 1 (13 classes) apresentaram os eucaliptais, corpos d'água e pastos como as classes com os menores erros de comissão. Quanto aos erros de omissão, as classes cana, pasto e eucalipto, apresentaram os menores valores para as classificações abrangendo o Maxver, distância mínima euclidiana simples e de Mahalanobis. Contudo, no caso da distância mínima euclidiana baseada no desvio padrão, o eucalipto apresentou um valor maior de erro em comparação a classe solo.

Para os processamentos com a legenda temática nº 2 (9 classes), as categorias de uso das terras com melhores resultados em termos dos erros de comissão e omissão permaneceram as mesmas.

Analisando as tabelas de contingência (vide itens 7.4 e 11.2), pode-se notar a existência de tipos de erros de designação mais freqüentes entre algumas classes.

No caso das classes pastagem, solo exposto, laranjal em formação e canal, as confusões ocorrem principalmente devido às variações dos sistemas de produção adotados na região e seus reflexos na biomassa e ecofisiologia dos estágios fenológicos das culturas. Por exemplo, as áreas de citros observadas em Brotas apresentavam entre as linhas de

plântio práticas de proteção do solo distintas (vide Figura 5-10). Em alguns laranjais constatou-se cobertura de espécies de gramíneas utilizadas na formação dos pastos, com diferentes quantidades de fitomassa. Em outras propriedades, o solo era mantido exposto. Outro exemplo marcante é a cana de açúcar cuja resposta espectral varia segundo as fenofases da cultura. Quando pequena, o sinal é semelhante ao do pasto, com forte influência da reflectância do solo.

A classe vegetação natural envolve todas as categorias encontradas na região exceto as várzeas. As características dessas formações estão bastante alteradas apesar de serem áreas com estatuto de conservação definido. A utilização sem planejamento ou sustentabilidade, o desmatamento, o corte seletivo, a incidência de queimadas e o processo de sucessão secundária em parcelas agrícolas abandonadas ou em pousio conduziram os fragmentos remanescentes da cobertura vegetal original a uma elevada heterogeneidade em sua composição florística, fisionomia, estrutura e dinâmica ecológica, caracterizando um quadro típico de todo o Estado de São Paulo. As manchas inseridas na matriz agrícola são compostas por um misto de fases sucessionais, apresentando muitas vezes espécies invasoras entremeadas aos espécimes da vegetação original. Estas características impõem às formações originalmente muito diferentes (campo cerrado, cerrado, cerradão, mata semi-decídua, mata decídua, mata perenifólia e mata ciliar) respostas espectrais análogas, gerando confusões até mesmo com as culturas instaladas na região, como é o caso do eucalipto, cana de açúcar e citros. Outro fator complicador para a caracterização dessas formações é a ocorrência de áreas de cultivo abandonadas, onde algumas espécies de cerrado se reinstalam (vide Figura 5-13 b).

Os reservatórios e os cursos d'água de grande porte são bastante escassos na região de estudo, além de apresentarem um processo de assoreamento bastante avançado. O rio Jacaré-Pepira, o maior corpo d'água da região, está coberto na maioria de sua extensão por algum tipo de vegetação, induzindo alguns erros de designação do pixels para essa classe.

8.2.3 Desempenho dos classificadores

O classificador espectral Maxver, considerado um dos mais eficientes e mais utilizados, está implementado nos principais programas de processamento digital de imagens digitais. São freqüentemente empregados como rotinas padrão para a classificação dos dados orbitais, servindo como termo de comparação para avaliar a eficiência de outros algoritmos (Ince, 1987 e Venkateswarlu & Sing, 1995a). Entretanto seus resultados nessa dissertação não atenderam totalmente às expectativas, como demonstra os resultados apresentados pela Tabela 8-3.

Esse comportamento não é raro (Shimoda et al., 1988; Conese & Maselli, 1992; Rocha, 1992 e Maselli et al., 1994). Tais algoritmos são extremamente sensíveis às anomalias de distribuição e ao tamanho das amostras de treinamento.

O algoritmo Maxver, ou qualquer outro de natureza paramétrica, quando aplicado em paisagens de elevada heterogeneidade espacial, onde a forma da distribuição espectral das categorias de cobertura terrestre não obedece ao critério da normalidade, não apresenta

bons resultados na classificação e na estimativa de áreas, especialmente quando há intensas discrepâncias na extensão das classes (Maselli et al., 1990 e 1992). Isto pode explicar sua baixa eficiência na identificação das classes várzea, mata, cerradão, mata ciliar e laranja em formação.

Tabela 8-3. Tabela dos coeficientes Kappa global e exatidão global com seus intervalos de confiança para os 5 classificadores, considerando as chaves de classificação de 7, 9 e 13 categorias.

Classificadores	Kappa global	Exatidão global	Exatidão global - intervalo de confiança		
			90%	95%	99%
euclidiana simples 9	0,45	0,44	0,4376 - 0,4498	0,4365 - 0,4509	0,4342 - 0,4532
euclidiana simples 13	0,40	0,50	0,4940 - 0,5062	0,4929 - 0,5074	0,4906 - 0,5097
euclidiana - d. padrão 9	0,44	0,45	0,4410 - 0,4531	0,4398 - 0,4543	0,4375 - 0,4566
euclidiana - d. padrão 13	0,37	0,53	0,5271 - 0,5393	0,5260 - 0,5405	0,5237 - 0,5428
Mahalanobis 9	0,50	0,39	0,3870 - 0,3989	0,3859 - 0,4001	0,3836 - 0,4023
Mahalanobis 13	0,42	0,49	0,4844 - 0,4966	0,4832 - 0,4978	0,4809 - 0,5001
Máx. verossimilhança 7	0,58	0,31	0,3092 - 0,3203	0,3081 - 0,3216	0,
Máx. verossimilhança 9	0,51	0,38	0,3779 - 0,3898	0,3768 - 0,3909	0,3746 - 0,3932
Máx. verossimilhança 13	0,44	0,47	0,4656 - 0,4778	0,4645 - 0,4790	0,4622 - 0,4813

Ince (1987) em uma revisão sobre o uso de Maxver para a classificação de dados de Landsat-TM em paisagens agrícolas, mostrou que a exatidão e os erros obtidos pelo Maxver podem variar consideravelmente dependendo da formação estatística dos dados de treinamento.

Baseado nestas observações, classificadores não-paramétricos como distância mínima podem ser implementados, em substituição aos paramétricos. Analisando a Tabela 11-15, Tabela 11-18, Tabela 7-8, Tabela 7-11 e Tabela 7-14), verifica-se que o classificador distância mínima de Mahalanobis apresentou valores maiores em relação ao Maxver para as classes solo exposto, laranja formada, mata, várzea, pasto com árvores, cerrado, laranja em formação e cerradão. O classificador de distância mínima euclidiana não mostrou a mesma superioridade, provavelmente pela sua limitação em lidar com as classes de distribuição alongada. O algoritmo Mahalanobis pode ser utilizado quando o usuário dispõe de um número reduzido de amostras, evitando a dependência entre a eficiência do classificador e a exatidão da estimativa do vetor médio e da matriz de covariância. Além disso, seu processamento tem certa sensibilidade às classes com distribuição alongada (Venkateswarlu & Singh, 1995b).

Uma demonstração dessas vantagens relativas foi feita por Booth & Oldfield (1989) ao testarem 4 esquemas de classificação supervisionada em imagens Landsat-TM: distância mínima, árvore de decisão, Maxver e distância mínima modificada ("deviant distance algorithm"). Nesse estudo empírico, embora o Maxver tenha gerado a classificação com a melhor exatidão, o uso de um algoritmo mais rápido como o de distância mínima seguido de uma filtragem modal forneceu resultados com exatidão similar e com um tempo menor de processamento.

Contudo, tanto os procedimentos paramétricos como os não paramétricos têm mostrado muitas limitações intrínsecas em lidar com os dados de baixa resolução dos sensores disponíveis, impedindo uma exploração satisfatória dos dados de sensoriamento remoto para algumas propostas práticas de classificação da cobertura da terra.

Além disso, a necessidade de preceder a utilização dos dados de sensoriamento remoto com uma etapa de definição de procedimentos metodológicos adequados cria empecilhos para sua aplicabilidade em questões imediatas. Normalmente os órgãos ambientais, governamentais e não governamentais, atuantes na temática do meio ambiente, precisam obter informações para programas de monitoramento, manejo e/ou conservação ambiental em tempo relativamente reduzido e com custos controlados. Esse condicionante gera uma demanda por soluções operacionais que não impliquem em uma fase de desenvolvimento e testes longos.

A pesquisa tem tentado aperfeiçoar as técnicas de classificação, criando novos algoritmos e procedimentos para lidar com as dificuldades de classificar os usos das terras em áreas complexas. Muitas dessas iniciativas tem explorado outras dimensões das imagens orbitais, além do tradicional componente espectral (Bauer et al., 1977; Dawbin & Evans, 1988; Hodgson, 1988; Booth & Oldfield, 1989; Bryant, 1989; Bolstad & Lillesand, 1991; Guo & Haigh, 1994; Huang & Mausal, 1994; Venkateswarlu & Singh, 1995a e Venkateswarlu & Singh, 1995b).

Métodos incorporando aspectos como textura, tonalidade e contexto na rotulação dos pixels já são uma realidade irreversível (vide item 4.5.3). Uma tendência muito comum verificada nesses esforços para a geração de novos procedimentos de identificação de padrões é a formulação de estratégias híbridas, incorporando e articulando diferentes técnicas, de modo que uma supra as deficiências da outra.

Um exemplo são algumas soluções associando as vantagens de estratégias paramétricas e não paramétricas. Uma delas é através da inserção de uma probabilidade anterior derivada de um processo não paramétrico em um classificador paramétrico convencional.

Entretanto, nenhum desses caminhos já está suficientemente consolidado a ponto ser incorporado como uma técnica padrão nos pacotes de processamento digital de imagem consagrados.

8.3 Coleta dos dados de referência

A definição de esquemas de amostragem apropriados para aferir a exatidão dos mapas temáticos é um tema controvertido, onde existem diferentes opiniões. Nesse contexto, alguns pontos do planejamento da amostragem adotado devem ser destacados.

A área de 600 m por 600 m de cada sítio de amostragem mostrou-se adequada por permitir e facilitar a visualização simultânea de todos os elementos da cobertura terrestre presentes na parcela de campo e a coleta de uma amostra com tamanho representativo.

O tamanho da amostragem (2,7% da imagem), *a priori*, pareceu suficiente, se considerarmos o pequeno número de zeros encontrados na matriz de erro e a proporção entre o número de células da matriz e o tamanho das amostras coletadas. Além disso, o número de unidades obtido, para cada classe, respeita o valor mínimo de 50 sugerido por Congalton (1991). Rocha (1992) analisando a influência do tamanho da amostra (0.50 % - 2.50% da área total) sobre os coeficientes de concordância Kappa e exatidão global, concluiu que não houve uma variação significativa dos valores médios desses índices de exatidão, existindo, entretanto, uma variação não-linear em suas variâncias. Seria interessante, nos esquemas de amostragem, incluir pesos para as classes, proporcionais a importância e a representatividade das mesmas no contexto espacial da região.

Os dois grandes estratos definidos inicialmente em função das feições geomorfológicas não apresentaram resultados muito diferentes no tocante aos índices de exatidão das classificações quando comparados com o desempenho dos algoritmos na área de estudo inteira. Portanto, optou-se por um tratamento unificado da cena.

Esse comportamento dos classificadores não é suficiente para descartar o método de amostragem baseado em estratos, principalmente sua qualidade de possibilitar uma abordagem balanceada da diversidade existente no universo amostral, designando um número de amostras independente da proporção de ocorrência do fenômeno, como por exemplo uma classe de uso das terras com grande extensão na área de estudo (Frontier, 1983).

Poderiam ser testadas outras formas de estratificação, contribuindo para verificar a influência de outras variáveis na exatidão da classificação. Por exemplo, a opção por um critério geomorfológico poderia ser mantida, avaliando-se a alternativa de um maior número de tipos de paisagem, ao invés dos dois utilizados.

A amostragem aleatória sistemática não-alinhada utilizada foi adequada pois assegurou a representatividade e a confiabilidade nos resultados. A definição da localização dos pontos de amostragem baseada em uma grade sobreposta sobre a imagem, conduziu à coleta dos dados de referência por toda a imagem, atingindo quase todos os tipos de cobertura e as respectivas transições entre os vários padrões de ocupação.

Outro ponto fundamental alcançado foi a independência entre os dados destinados a análise da exatidão e as informações para o treinamento dos classificadores, evitando a inclusão de tendências nos resultados. Entretanto, essa utilização de dados provenientes de amostras diferentes para as análises de exatidão e preparação das classificações poderia implicar em um aumento de tempo e de custo do projeto. Entretanto, durante as incursões à região, demonstrou-se a exequibilidade de utilizar um único esforço de campo para tais atividades, otimizando a utilização dos recursos logísticos, financeiros e humanos. Assim, as coletas de dois grupos de dados distintos para os diferentes procedimentos pode ser executada durante as mesmas incursões a campo.

Para a coleta dos dados de referência terrestre, o acerto na logística de campo foi fundamental. O protocolo de coleta de dados do ponto amostral se mostrou bastante funcional, sendo de fácil manuseio e apresentando informações preciosas para a localização

dos elementos terrestres. Além disso, permitiu a padronização e a organização dos dados. Entretanto, o material utilizado para a confecção desses protocolos, papel vegetal e fotocópia do mapa topográfico, sofreu deformações, constituindo uma fonte de erro adicional. Assim sendo, deveriam ser substituídos por materiais de melhor qualidade, principalmente no aspecto de estabilidade, como o “tekron” ou acetato, e evitado o uso de fotocópias e mapas desgastados pela utilização no deslocamento da equipe no campo.

O uso de impressões da imagens e de um aparelho de GPS para a determinação dos sítios de amostragem possibilitou a superação das dificuldades impostas pela desatualização da base cartográfica nos deslocamentos na área de estudo.

O mapeamento dos dados de referência terrestre no campo é um processo sujeito a alguns tipos de erros.

A localização do vértice inicial da amostra, mesmo nos locais com bons pontos de referência, sempre implica em algum erro imponderável e pequeno.

Definida a posição do ponto amostral, há uma deformação inerente na reprodução da forma e dimensão dos parcelares de uso das terras e fragmentos de cobertura vegetal natural no protocolo de coleta de dados. Estimar no campo um quadrilátero de 600 por 600 m a partir de alguns pontos de observação é sempre uma avaliação qualitativa influenciada pela topografia, estrutura vertical da vegetação e pelas facilidades de acesso aos vértices e limites da amostra. Esse erro na avaliação dos limites da amostra a partir do vértice de observação ocorre em grande parte dos pontos, com intensidades variáveis em função da possibilidade de percorrer o entorno da amostra. A única forma de agregar uma mensuração aproximada são as raras ocasiões nas quais a localização do ponto coincide com a orientação de alguma estrada, tornando possível utilizar o odômetro do veículo.

Além desse erro, nas situações em que a localização precisou ser feita através do GPS, deve-se considerar a imprecisão do equipamento. Nesses casos, a posição do levantamento pode estar até aproximadamente 100-150 m deslocado da sua coordenada real. Entretanto, o uso do sistema de posicionamento global foi restrito a situações específicas onde não havia nenhuma referência segura, como no caso de um ponto em uma área de pastagem justaposta a um laranjal novo.

Na grande maioria das vezes, a definição da amostra era feita através de referências inconfundíveis como cruzamentos de estradas; parcelas adjacentes com padrões de reflexão espectral muito contrastantes, como pasto e eucalipto; construções agrícolas; acidentes geográficos como lagos, açudes, barragens, curvas, confluências de rios etc.

8.4 Análise da exatidão

A exatidão média dos classificadores (vide 5) foi muito menor quando comparada aos resultados obtidos em alguns trabalhos como Green et al. (1993), Congalton & Green (1993), Bauer et al. (1994) e Jusoff & Manaf (1995) (vide 7).

A suposição adotada é que todas as diferenças entre o mapa gerado a partir da classificação espectral da imagens Landsat-TM e os dados de referência sejam devidas a

classificação. Essa opção reflete algumas das situações dos projetos envolvendo sensoriamento remoto e sistemas de informações geográficos vivenciadas em empresas privadas, instituições governamentais e ONGs.

Contudo, é interessante analisar a complexidade dos procedimentos envolvidos na obtenção da exatidão da classificação, auxiliando no entendimento e no aperfeiçoamento dos resultados desse trabalho. A exatidão é a somatória de todos os erros presentes nas diversas fases de sua determinação. A confiança nos dados derivados de técnicas de sensoriamento remoto, tema que vem recebendo a cada ano maior atenção por parte dos pesquisadores e usuários de produtos suborbitais e orbitais, é freqüentemente representado por uma matriz de erro e as análises derivadas. Em muitas situações atribui-se todas as diferenças entre o mapa gerado a partir da classificação de sensoriamento remoto e os dados de referência à erros decorrentes do processo de rotulação dos padrões espectrais. Porém, outros fatores podem ser significantes e diminuir substancialmente a exatidão representada pela matriz. Muitos deles já foram discutidos nos itens anteriores como por exemplo, a influência da legenda temática ou as distorções do protocolo de campo. A seguir serão discutidos alguns outros pontos considerados importantes:

- Georreferenciamento.

Os procedimentos adotados para executar a georreferência das imagens (seleção dos 12 de pontos de controle, aplicação da função de interpolação linear e reamostragem baseada no método do “vizinho mais próximo”), foram adequados, como indica o valor do RMS, correspondente a 0,3 pixel (9 m). Este resultado está de acordo com o estudo de Smith & Kovalick (1985). Os efeitos de diferentes técnicas de reamostragem sobre os resultados de classificações foram comparados, concluindo-se que não existem grandes diferenças se a técnica do “vizinho próximo” é executado antes ou depois da classificação. Taylor (1989) havia recomendado essa técnica para tratar imagens multiespectrais por causar um impacto menor sobre a imagem, preservando a relação entre as bandas.

O método de reamostragem aplicado nos procedimentos de georreferência deve ser escolhido cuidadosamente, sem negligenciar o caráter dos dados e a relação do tamanho do pixel original e do referenciado, uma vez que esse processo envolve uma transformação do espaço de atributos e do tipo de distribuição das classes espectrais.

- Variação temporal do uso das terras.

As mudanças na cobertura terrestre entre as datas da aquisição dos dados orbitais e da coleta dos dados de referência no campo foram significativas na obtenção da exatidão de classificação, uma vez que dificultou a determinação segura da correspondência entre os elementos terrestres e suas respostas espectrais. A incerteza das informações pode induzir o analista a erros na interpretação e no delineamento dos dados. Assim a coleta de dados de campo deveria ser realizada o mais próximo possível da data da imagem. A utilização de entrevistas com os proprietários das terras, apesar de minimizar os erros, aumentou o tempo de permanência da equipe no campo e conseqüentemente o custo do projeto.

- Delineamento e digitalização dos dados de referência terrestre.

A inclusão de erros na fase de delineamento dos dados de referência deve ser sempre considerada em qualquer estudo cartográfico. Neste trabalho, como em alguns outros, procedimentos manuais foram adotados, tanto para a delimitação como para a digitalização dos dados de referência. Uma das desvantagens desse procedimento é a sujeição dos dados à tendenciosidade e arbitrariedade do analista. Apesar da inconsistência na interpretação humana ser um fato notório, especialmente para áreas heterogêneas onde os polígonos apresentam dimensões pequenas e formas irregulares, ainda é um fator de difícil controle. Medidas de variação na interpretação deveriam ser implementadas nas primeiras etapas de coleta de dados, a fim de padronizar os procedimentos e minimizar as tendenciosidades pessoais. Isto poderia ser feito utilizando pequenos subgrupos de pessoas da equipe, os quais coletariam os dados em alguns sítios comuns. A concordância entre essas informações seriam analisadas e utilizada na padronização.

Apesar de serem fatores importantes, a formação de subgrupos e a escolha de materiais mais estáveis para o mapeamento da verdade terrestre, nem sempre podem ser executados devido a um terceiro fator, igualmente importante para a consecução de qualquer pesquisa - a disponibilidade de recursos humanos e financeiros.

- Técnica de amostragem.

Segundo Congalton (1991), as técnicas de amostragem usualmente utilizadas para a determinação de exatidão não são adequadas, tomando como exemplo o trabalho de Fitzpatrick-Lins (1978b). Contudo, as técnicas de amostragem variam de acordo com as características dos elementos terrestres de interesse. Isto significa que não existe uma única estratégia amostral adequada, sendo inadequada aquela que não representa a população sob análise. Deve-se considerar para a escolha de qualquer técnica, as variáveis espaciais, espectrais e contextuais das classes de interesse, além do custo-benefício de tais opções. A implementação da amostragem aleatória sistemática não-alinhada foi possível graças à facilidade de acesso à região e a existência de bases cartográficas. Assim pode-se fazer usos das vantagens dessa estratégia de amostragem - a consistência e a não tendenciosidade das estimativas da população - sem empecilhos em termos de verificação de campo dos pixels selecionados, comum nas regiões de difícil acesso ou sem informações contextuais. Porém, foram constatadas deficiências em amostrar adequadamente categorias de áreas reduzidas, introduzindo deste modo erros na designação dos pixels. Para suprir tais deficiências, poder-se-ia utilizar estratos definidos por tipos de cobertura das terras, como defendido por Rudd (1971), Zonneveld (1974), van Genderen et al. (1978), Ginevan (1979) e Bailey et al. (1994).

- Conversões entre formatos.

As conversões de formatos (vetor/matriz) também devem ser consideradas nas determinações de exatidão de classificação, uma vez que introduzem erros de limites nos polígonos. Houve alterações nas informações dos polígonos, correspondentes ao uso das terras digitalizados, decorrente da incompatibilidade entre o caráter analógico dos arquivos vetoriais e o digital das imagens. Assim, a transformação de escala dos mosaicos, de 1m

para 30m, contribuiu no erro final dos dados. Os erros aumentam a medida que os poligonos referentes as classes diminuem de tamanho ou apresentam formatos estreitos.

9. Conclusões e Recomendações

A composição colorida com melhor contraste visual entre os elementos terrestres, principalmente em relação as classes de vegetação natural, foi a 453 (RGB). Esse resultado contrasta com o obtido pelo método OIF, onde esse tripleto foi a quarta combinação com maior quantidade de informações espectrais.

Os classificadores com os maiores valores do coeficiente Kappa foram máxima verossimilhança e distância mínima de Mahalanobis. Contudo a dimensão desses valores indica a inadequabilidade desses algoritmos espectrais para o mapeamento de uso das terras em áreas densamente ocupadas com uma grande variedade de categorias de ocupação dos solos, com parcelares agrícolas reduzidos, e com culturas e sistemas de produção distintos.

A redução do número de classes de usos das terras resultou no aumento dos índices de exatidão de classificação. Entretanto essa diminuição de classes provavelmente deprimirá a adequabilidade dos produtos gerados em relação às demandas dos administradores públicos e dos empresários, atuantes na região, por informações minuciosas capazes de subsidiar o processo de tomada de decisão.

A estratégia de amostragem selecionada e o plano de amostragem executados atenderam os requisitos necessários à aferição da exatidão de classificação:

- a área de 600 m por 600 m de cada sítio de amostragem permitiu a visualização simultânea de todos os elementos da cobertura terrestre presentes na parcela de campo e a coleta de uma amostra total com tamanho representativo, correspondente a 2,7% da imagem;
- a amostragem aleatória sistemática não-alinhada assegurou a representatividade e a confiabilidade dos resultados, uma vez que a coleta dos dados de referência foi realizada por toda a imagem, atingindo quase todos os tipos de cobertura e as respectivas transições entre os vários padrões de ocupação;
- o coeficiente de concordância Kappa mostrou-se mais eficiente do que a exatidão global para aferir a acurácia dos classificadores espectrais utilizados;
- o coeficiente de Kappa condicional também teve um comportamento melhor em relação aos erros de omissão e comissão na avaliação dos algoritmos analisados.

Diante da pluralidade de situações agroecológicas existentes ainda não há uma série consagrada de procedimentos eficientes e simples para a classificação digital de imagens orbitais e cálculo de sua exatidão. Obter um desempenho ótimo para todas as classes ou para todas as áreas de uma região é um desafio a ser superado com muitos esforços para aperfeiçoar e integrar diferentes técnicas de classificação e as respectivas análises de sua exatidão. Além disso, a disponibilização dos produtos de alta resolução

espacial e espectral ou de radar, em franca expansão, com o lançamento de novos sensores, auxiliará em muito na busca da solução dessas questões.

Os testes comparativos entre essas novas técnicas e tecnologias, em um contexto de variáveis controladas e confiabilidade das informações, são um caminho produtivo para determinar precisamente os pontos críticos nas condutas de classificação dos usos das terras, particularmente em áreas complexas; na padronização de métodos e na incorporação de dados sólidos sobre o uso das terras no cotidiano de projetos sobre avaliação e monitoramento de impactos ambientais, planejamento agrícola e mudanças globais.

Considerando a importância da temática do cálculo da exatidão no desenvolvimento da classificação digital de imagens orbitais e a experiência reunida na consecução desta pesquisa, pode-se sugerir algumas recomendações:

- determinar, caso a caso, a viabilidade da aplicação do sensoriamento remoto aos diferentes tipos de problemas do gerenciamento dos recursos naturais, considerando-se as especificidades de cada legenda temática e das condições naturais de cada área focalizada;
- garantir a confiabilidade das informações produzidas por mapeamento com emprego de sensoriamento remoto e sistemas de informações geográficas através da análise da exatidão;
- implementar procedimentos para determinar a exatidão durante as fases preliminares de projetos de classificação da cobertura e uso das terras;
- organizar os projetos de classificação do uso das terras com cronogramas de execução limitados, evitando uma defasagem muito grande entre a aquisição dos dados remotos e dos dados de referência terrestre e minimizando os custos com a realização de entrevistas com proprietários rurais;
- evitar a inclusão de categorias com pouca representatividade espacial e de frequência na composição da legenda temática;
- formular estratégias híbridas na classificação da cobertura e uso das terras, explorando abordagens texturais, apoio de dados auxiliares, segmentadores, etc, dando prioridade para os algoritmos de máxima verossimilhança e distância mínima de Mahalanobis nos casos da aplicação de classificadores espectrais;
- selecionar dados de treinamento espectralmente homogêneos, utilizando diagramas de dispersão sempre que possível;
- adotar a estratégia de amostragem aleatória sistemática não-alinhada, especialmente em áreas complexas onde seja possível distinguir padrões *a priori*;
- planejar a coleta dos dados de campo de modo a contemplar os três tipos de dados necessários para: o reconhecimento de padrões de cobertura e uso das terras para o delineamento da legenda temática, a determinação de áreas de treinamento para

alimentar a execução dos classificadores e o mapeamento *in loco* do uso das terras para servir como verdade terrestre no cálculo da exatidão;

- estruturar campanhas integradas de levantamento de dados de referência terrestre, sem uma dedicação exclusiva à obtenção de uma das três categorias de informações necessárias;
- empregar, como protocolo para mapeamento dos dados de referência terrestre, a sobreposição de um material estável como acetato sobre uma reprodução de boa qualidade da base cartográfica, acompanhada de uma impressão ou fotografia da imagem;
- padronizar o cálculo da exatidão temática e sua expressão através de matrizes de erros, índice de concordância Kappa, erros de omissão e comissão;
- utilizar matrizes de erro quadráticas;
- considerar todos os componentes geradores de incertezas associados a classificação dos usos das terras na análise da exatidão.

10. Bibliografia

- AHEARN, S.C.; WEE, C. Data space volumes and classification optimization of SPOT and Landsat data. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.57, n.1, p.61-65, 1991.
- AHERN, F.J.; SIROIS, J. Reflectance enhancements for the thematic mapper: an efficient way to produce images of consistently high quality. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.55, n.1, p.61-67, 1989.
- AICKIN, M. Maximum likelihood estimation of agreement in the constant predictive probability model, and its relation to Cohen's kappa. **Biometrics**, Alexandria, v.46, n.2, p. 293-302, 1990.
- AIDAR, M.P.M. **Ecologia do araribá (*Centropogon tomentosus* Guill. ex Benth., Fabaceae) e o ecótono mata ciliar da bacia do rio Jacaré-Pepira, São Paulo**. Campinas: IB-UNICAMP, 1992. 99p. Dissertação de Mestrado em Ecologia.
- ALMEIDA, C.L.F.; OLIVEIRA, J.B.; PRADO, H. **Levantamento pedológico semidetalhado do Estado de São Paulo: quadrícula de Brotas: 1. Mapa de solos**. Campinas: Instituto Agrônomo de Campinas, 1981. Esc. 1:100.000. mapa, impr., color.
- ALMEIDA, F.F.M. de; BARBOSA, O. Geologia das quadrículas de Piracicaba e Rio Claro. **Boletim Div. Geol. Min.**, Rio de Janeiro, v.1943. p.1-96, 1953. Apud ALMEIDA; MELO (1981).
- ALMEIDA, F.F.M. de; MELO, M.S.A. Bacia do Paraná e o vulcanismo Mesozóico. In: ALMEIDA, F.F.M.; HASUI, Y.; PONÇANO, W.L.; DANTAS, A.S.L.; CARNEIRO, C.D.R.; MELO, M.S.; BISTRICHI, C.A. **Mapa Geológico do Estado de São Paulo. Nota Explicativa**. São Paulo: Instituto de Pesquisas Tecnológicas do Estado de São Paulo, 1981. p.46-81. v.1. Esc. 1:500.000.
- ANDERSON, J.R.; HARDY, E.E.; ROACH, J.T.; WITMER, R.E. **A land use and land cover classification system for use with remote sensor data**. Washington: U.S. Government Printing Office, 1976. 28p. U.S. Geological Survey Professional Paper 964. Apud LILLESAND & KIEFER (1994).
- ARONOFF, S. Classification accuracy: a user approach. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.48, n.8, p.1299-1307, 1982a.
- ARONOFF, S. The map accuracy report: a user's view. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.48, n.8, p.1309-1312, 1982b.
- ARONOFF, S. The minimum accuracy value as an index of classification accuracy. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.51, n.1, p.99-111, 1985.
- ARONOFF, S. **Geographic information systems: a management perspective**. Ottawa: WDL, 294p. 1989.

- ATKINSON, P.; CUSHNIE, J.L.; TOWNSHEND, J.R.G.; WILSON, A. improving Thematic Mapper land cover classification using filtered data. **International Journal of Remote Sensing**, London, v.6, n.6, p.955-961, 1985.
- AUTOCAD. **AutoCAD release 11 reference manual**. s.l.: Autodesk, 640p. 1990.
- BAILEY, M.; CONVIS, C.; DAVIS, F.; LI, X.; GOODIN, K.; GROSSMAN, D. **Accuracy assessment procedures**. Redlands: USDI-NBS/NPS/ESRI/University of California-NCGIA/The Nature Conservancy, 108p. 1994. Documento preliminar.
- BARBOSA, M.N. & CUNHA, R.C. Sensoriamento remoto, a terra vista do espaço. **Revista Brasileira de Tecnologia**, Brasília, v.18, n.4., p.39-46,1987.
- BARRET, J.P.; NUTT, M.E. **Survey sampling in the environmental sciences: a computer approach**. Wenthworth: Compress, 1979. 319p. Apud CONGALTON (1988).
- BAUER, M.E.; BURK, T.E.; EK, A.R.; COPPIN, P.R.; LIME, S.D.; WALSH, T.A.; WALTERS, D.K.; BEFORT, W.; HEINZEN, D.F. Satellite inventory of Minnesota forest resources. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.60, n.3, p.287-298, 1994.
- BAUER, M.E.; SILVA, L.F., HOFFER, R.M.; BAUMGARDNER, M.F. Agricultural scene understanding. West Layfayette: LARS-Purdue University,1977. LARS Contract Report 112667. Apud BRYANT (1989).
- BENSON, A.S.; DRAEGER, W.C.; PETTINGER, L.R. Ground data collection and use. **Photogrammetric Engineering**, Bethesda, v.37, p.1159-1167, 1971.
- BERRY, B.J.L. **Sampling, coding and storing flood plain data**. Washington: U.S. Department of Agriculture, 1962. Agriculture Handbook n. 237. Apud ROSENFELD et. al. (1982).
- BERRY, B.J.L.; BAKER, A.M. Geographic sampling. Spatial analysis: a reader. In: BERRY, B.J.L.; MARBLE, D.F. (Ed.) **Statistical Geography**. Englewood Cliffs: Prentice Hall, 1968. p.91-100. Apud CONGALTON (1988).
- BISHOP, Y.; FIENBERG, S.; HOLLAND, P. **Discrete multivariate analysis - theory and practice**. Cambridge: MIT Press, 1975. 575p. Apud CONGALTON et al. (1983).
- BISTRICH, C.A.; CARNEIRO, C.D.R.; DATAS, A.S.L.; PONÇANO, W. L.; CAMPANHA, G.A. de C.; NAGATA, N.; ALMEIDA, M.A. de; STEIN, D.P.; MELLO, M.S. de; CREMONINI, O.A. **Mapa Geológico do Estado de São Paulo**. São Paulo: Instituto de Pesquisas Tecnológicas do Estado de São Paulo, 1981. v.2. Esc. 1:500.000.
- BOLSTAD, P.V.; LILLESAND, T.M. Rapid maximum likelihood classification. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.57, n.1, p.67-74, 1991.

- BOLSTAD, P.V.; LILLESAND, T.M. Rule-based classification models: flexible integration of satellite imagery and thematic spatial data. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.58, n.7, p.965-971, 1992.
- BOOTH, D.J.; OLDFIELD, R.B. A comparison of classification algorithms in terms of speed and accuracy after the application of a post-classification modal filter. **International Journal of Remote Sensing**, London, v.10, n.7, p.1271-1276, 1989.
- BOUCHARDY, J.Y.; RODRIGUEZ, B.Z. Improved classification of satellite images pixels: presentation of sample methodology using masks. **Photo-Interprétation**, Paris, v.5/6, p. 205-208, 1992.
- BRASIL. Ministério do Planejamento e Coordenação Geral. Fundação Instituto Brasileiro de Geografia. Departamento de Cartografia. **Carta do Brasil**. Rio de Janeiro: IBGE, 1969. Esc. 1:50.000. Color.
- BRASIL. Ministério do Planejamento e Coordenação Geral. Fundação Instituto Brasileiro de Geografia. Departamento de Cartografia. **Carta do Brasil**. Rio de Janeiro: IBGE, 1971. Esc. 1:50.000. Color.
- BRASIL. Ministério do Planejamento e Coordenação Geral. Fundação Instituto Brasileiro de Geografia. Departamento de Cartografia. **Carta do Brasil**. Rio de Janeiro: IBGE, 1974. Esc. 1:50.000. Color.
- BRENNAN, R.L.; PREDIGER, D.J. Coefficient kappa: some uses, misuses, and alternatives. **Educational and Psychological Measurement**, Durham, v.41, n.3, p.687-699, 1981.
- BRYANT, J.A. Fast classifier for image data. **Pattern Recognition**, Elmsford, v.22, n.1, p.45-48, 1989.
- BÜTTNER, G.; HAJÓS, T.; KORÁNDI, M. Improvements to the effectiveness of supervised training procedures. **International Journal of Remote Sensing**, London, v.10, n.6, p.1005-1013, 1989.
- CAMPBELL, J. Spatial correlation effects upon the accuracy of supervised classification of land cover. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.47, n.3, p.355-363, 1981.
- CAMPBELL, J.B. **Introduction to remote sensing**. New York: Guilford Press, 1987. 551p.
- CAMPBELL, J.B. **Introduction to remote sensing**. New York: Guilford Press, 1996. 622p. 2 ed.
- CAMPBELL, N.A. Robust procedure in multivariate analysis. I: Robust covariance estimation. **Applied Statistics**, London, v.29, n.3, p.231-237, 1980.
- CARD, D.H. Using known map category marginal frequencies to improve estimates of thematic map accuracy. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.48, n.3, p.431-439, 1982.

- CARNEIRO, C.D.R.; BISTRICH, C.A; PONÇANO, W. L.; ALMEIDA, M.A. de. **Mapa Geomorfológico do Estado de São Paulo**. São Paulo: Instituto de Pesquisas Tecnológicas do Estado de São Paulo, 1981. v.2. Esc. 1:1.000.000. mapa, impr., color.
- CASTLEMAN, K.R. **Digital image processing**. Englewood Cliffs: Prentice Hall, 1979. 429p.
- CEPAM. **Consórcio do rio Jacaré-Pepira, relato de um experiência piloto**. São Paulo: CEPAM, 1990. 131p.
- CETIN, H.; LEVANDOWSKI, D.W. Interactive classification and mapping of multi-dimensional remotely sensed data using n-dimensional probability density functions (nPDF). **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.57, n.12, p.1579-1587, 1991.
- CETIN, H.; WARNER, T.A.; LEVANDOWSKI, D.W. Data classification, visualization, and enhancement using n-dimensional probability density functions (nPDF): AVIRIS, TIMS, TM, and geophysical applications. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.59, n.12, p.1755-1764, 1993.
- CHANDRASEKHAR, M.G. Indian remote sensing programme - future scenario. In: LAND SATELLITE INFORMATION IN THE NEXT DECADE CONFERENCE, 1995, Vienna. **Proceedings...** Bethesda: ASPRS, 1995. p.I-1-I-10.
- CHAVEZ Jr., P.S.; BERLIN, G.L.; SOWERS, L.B. Statistical method for selecting Landsat MSS ratios. **Journal of Applied Photographic Engineering**, Springfield, n.8, p.:23-30, 1982.
- CHAVEZ Jr., P.S.; KWARTENG, A.Y. Extracting spectral contrast in Landsat Thematic Mapper data using selective principal component analysis. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.55, n.3, p.339-348, 1989.
- CHEN, T.M.; STAELIN, D.H.; ARPIS, R.B. Information content analysis of Landsat image data for compression. **IEEE Transactions on Geoscience and Remote Sensing**, New York, v.25, p.499-501, 1987. Apud CONESE & MASELLI (1993).
- CHRISMAN, N.R. Modeling error in overlaid categorical maps. In: GOODCHILD, M.; GOPAL, S. (Ed.) **The accuracy of spatial databases**. London: Taylor and Francis, 1989. 290p. c. 2.
- CHRISMAN, N.R. The error component in spatial data. In: MAGUIRE, D.J.; GOODCHILD, M.F.; RHIND D.W. (Ed.) **Geographical information systems: principles and applications**. London: Longman, 1991. p. 165-174.
- CHUVIECO, E.; CONGALTON. R.G. Using cluster analysis to improve the selection of training statistics in classifying remotely sensed data. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.54, n.9, p.1275-1281, 1988.
- COCHRAN, W.G. **Sampling techniques**. New York: John Wiley & Sons, 1977. 428p.

- COHEN, J. A coefficient of agreement for nominal scales. **Educational and Psychological Measurement**, Durham, v.20, n.1, p.37-46, 1960.
- COHEN, J. Weighted Kappa nominal scale agreement with provision for scaled disagreement or parcial credit. **Psychological Bulletin**, Washington, v.70, n.4, p.213-220, 1968.
- COLWELL, R.N. **Manual of Remote Sensing**. Falls Church: American Society of Photogrammetry, 1983. 2400p.
- CONESE, C.; MARACCHI, G.; MASELLI, F.; MIGLIETTA, F. Statistical variance decomposition for the classification of TM scenes. In: CAPPELLINI, V.; COSTANDINIDES, A.G. (Ed.) **Digital signal processing**. Amsterdam: Elsevier, 1987. 922-931p. Apud CONESE & MASELLI (1993).
- CONESE, C.; MASELLI, F. Use of multitemporal information to improve classification performance of TM scenes in complex terrain. **ISPRS Journal of Photogrammetry & Remote Sensing**, Amsterdan, v.46, n.4, p.187-197, 1991.
- CONESE, C.; MASELLI, F. Use of error matrices to improve area estimates with maximum likelihood classification procedures. **Remote Sensing of Environment**, New York, v.40, n.2, p.113-124, 1992.
- CONESE, C.; MASELLI, F. Selection of optimum bands from TM scenes through mutual information analysis. **ISPRS Journal of Photogrammetry & Remote Sensing**, Amsterdan, v.48, n.3, p.2-111, 1993.
- CONGALTON, R.G. Using spatial autocorrelation analysis to explore errors in maps generated from remotely sensed data. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.54, n.5, p.587-592, 1988a.
- CONGALTON, R.G. A comparison of sampling schemes used in generating error matrices for assessing the accuracy of maps generated from remotely sensed data. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.54, n.5, p.593-600, 1988b.
- CONGALTON, R.G. A review of assessing the accuracy of classifications of remotely sensed data. **Remote Sensing of Environment**, New York, v.37, n. 1, p.35-46, 1991.
- CONGALTON, R.G.; BIGING, G.S. A pilot study evaluating ground reference data collection efforts for use in forest inventory. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.58, n.12, p.1669-1671, 1992.
- CONGALTON, R.G.; GREEN, K. A pratical look at the sources of confusion in error matrix generation. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.59, n.5, p.641-644, 1993.
- CONGALTON, R.G.; GREEN, K.; TEPLY, J. Mapping old growth forests on National Forest and Park lands in the Pacific Northwest from remotely sensed data.

- Photogrammetric Engineering & Remote Sensing**, Bethesda, v.59, n.4, p.529-535, 1993.
- CONGALTON, R.G.; MEAD, R.A. A quantitative method to test for consistency and correctness in photointerpretation. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.49, n.1, p.69-74, 1983.
- CONGALTON, R.G.; ODERWALD, R.G.; MEAD, R.A. Assessing landsat classification accuracy using discrete multivariate analysis statistical techniques. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.49, n.12, p.1671-1678, 1983.
- CONGALTON, R.G.; ODERWALD, R.G.; MEAD, R.A. Assessing landsat classification accuracy using discrete multivariate analysis statistical techniques. *Erratum*. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.50, n.10, p.1477, 1984.
- CONGALTON, R.G.; REKAS, A. COMPAR: a computerized technique for in-depth comparison of remotely sensed data. In: ANNUAL MEETING OF AMERICAN SOCIETY OF PHOTOGRAMMETRY AND REMOTE SENSING, 51, 1985. **Proceedings...** Washington: American Society of Photogrammetry and Remote Sensing, 1985. p.98-106. Apud STORY & CONGALTON (1986).
- CONNERS; R.W.; HARLOW, C.A. A theoretical comparison of texture algorithms. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, New York, v.2, n.3, p.204-222, 1980. Apud FUNG & CHAN (1994).
- CORVES, C; PLACE, C.J. Mapping the reliability of satellite-derived landcover maps - an example from the Central Brazilian Amazon Basin. **International Journal of Remote Sensing**, London, v.15, n.6, p.1283-1294, 1994.
- CRIST, E.P.; CICONE, R.C. Application of the tasseled cap concept to simulated thematic mapper data. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.50, n.3, p. 343-352, 1984.
- CROSS, A.M.; MASON, D.C.; DURY, S.J. Segmentation of remotely sensed images by a split and merge process. **International Journal of Remote Sensing**, London, v.9, n.8, p.1329-1345, 1988.
- CRÓSTA, A.P. **Processamento digital de imagens de sensoriamento remoto**. Campinas: IG-UNICAMP, 1992. 170p.
- CRÓSTA, A.P. Século XXI em alta resolução. **Fator Gis**, Curitiba, v.4, n.15, p.14-17, 1996.
- CURRAN, P.J. **Principles of remote sensing**. London: Longman, 1985. 256p.
- CURRAN, P.J. The semivariogram in remote sensing: an introduction. **Remote Sensing of Environment**, New York, v.24, n.3, p.493-507, 1988.
- CURRAN, P.J.; WILLIAMSON, H.D. Sample size for ground and remotely sensed data. **Remote Sensing of Environment**, New York, v.20, n.1, p.31-41, 1986.

- DALLEMAND, J.F.; TARDIN, A.T.; BATISTA, G.T.; CHEN, S.C. Sensoriamento remoto e agricultura. **Ciência Hoje**, Rio de Janeiro, v.8, n.43, p.45-51, 1988.
- DEAN, M.E.; HOFFER, R.M. An evaluation of Thematic Mapper simulator data for mapping forest cover. In: INTERNATIONAL SYMPOSIUM ON MACHINE PROCESSING OF REMOTELY SENSED DATA, 8, 1982, West Lafayette. **Proceedings...** West Lafayette: LARS-Purdue University, 1982. p.300-307. Apud TOWNSHEND (1984).
- DICKS, S.E.; LO, T.H.C. Evaluation of thematic map accuracy in a land-use and land-cover mapping program. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.56, n.9, p.1247-1252, 1990.
- DOTTAVIO, C.L.; DOTTAVIO, F.D. Potential benefits of new satellite sensors to wetland mapping. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.50, n.5, p.599-606, 1984.
- DREYER, P. Classification of land cover using optimized neural nets on SPOT data. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.59, n.5, p.617-621, 1993.
- DUDA, R.D.; HART, P.E. **Pattern classification and scene analysis**. New York: John Wiley & Sons, 1973. 482p.
- DUTRA, L.V.; MASCARENHAS, N.D.A. Some experiments with spatial features extraction methods in multispectral classification. **International Journal of Remote Sensing**, London, v.5, n.2, p.302-313, 1984.
- DWIVIDI, R.S.; RAO, B.R.M. The selection of the best possible Landsat-TM band combination for delineating salt-affected soils. **International Journal of Remote Sensing**, London, v.13, n.11, p.2051-2058, 1992.
- DYE, R.; WOOD, L. Resolution improvement by multi-temporal data merging. **ISPRS Journal of Photogrammetry & Remote Sensing**, Amsterdam, v.44, n.1, p.14-20, 1989.
- DYMOND, J.R. An improved Skidmore/Turner classifier. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.59, n.5, p.623-626, 1993.
- EASTMAN, J.R. **Idrisi. Version 4.01. Technical Reference**. Worcester: Clark University, 1992a. 178p. v.1. Review 1.
- EASTMAN, J.R. **Idrisi. Version 4.01. Technical Reference**. Worcester: Clark University, 1992b. 213p. v.2. Review 1.
- ER MAPPER. **ER Mapper 4.0 Reference manual**. West Perth: ER-Mapper, 1993. não paginado.
- ESTES, J.E.; HAJIC, E.J.; TINNEY, L.R. Fundamentals of image analysis: analysis of visible and thermal infrared data. In: COLWELL, R.N. **Manual of Remote Sensing**. Falls Church: American Society of Photogrammetry, 1983. p.987-1124. v.1.

- EVERITT, B.S. Moments of the statistics Kappa and weighted Kappa. **British Journal of Mathematical and Statistical Psychology**, Leicester, v.21,n.1, p.97-103, 1968.
- FENSTERMAKER, L.K. A proposed approach for national to global scale error assessment. In: GIS/LIS, 1991, Atlanta. **Proceedings...** Washington: American Society of Photogrammetry and Remote Sensing, 1991. p.293-300. Apud CONGALTON & GREEN (1993).
- FIDALGO, E.C.C. **Exatidão no processo de reconhecimento de padrões de vegetação natural a partir de imagens TM-Landsat em áreas de mata atlântica, no Estado de São Paulo**. São José dos Campos: INPE, 1994. 83p. Dissertação de Mestrado em Sensoriamento Remoto. Documento preliminar.
- FIENBERG, S. **The analysis of cross-classified categorical data**. Cambridge: MIT Press, 1980. 198 p. Apud CONGALTON (1983).
- FIORELLA, M.; RIPPLE, W.J. Determining successional stage of temperate coniferous forests with Landsat satellite data. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.59, n.2, p.239-246, 1993.
- FITZPATRICK-LINS, K. Accuracy and consistency comparisons of land use and land cover maps made from high-altitude photographs and landsat MSS imagery. **Journal of Research**, Reston, v.6, n.1, p.23-40, 1978a. Apud CONGALTON (1988). United States Geological Survey.
- FITZPATRICK-LINS, K. Accuracy of selected land use and land cover maps in Greater Atlanta Region, Georgia. **Journal of Research**, Reston, v.6, n.2, p.169-173, 1978b. Apud CONGALTON (1988). United States Geological Survey.
- FITZPATRICK-LINS, K. Comparison of sampling procedures and data analysis for land-use and land-cover map. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.47, n.3, p.343-351, 1981.
- FLEISS, J.L.; COHEN, J.; EVERITT, B.S. Large sample standard errors of Kappa and weighted Kappa. **Psychological Bulletin**, Washington, v.72, n.5, p.323-327, 1969.
- FOODY, G.M. On the compensation for chance agreement in image classification accuracy assessment. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.58, n.10, p.1459-1460, 1992.
- FOODY, G.M. Land cover classification by an artificial neural network with ancillary information. **International Journal of Geographical Information Systems**, London, v.9, n.5, p.527-542, 1995a.
- FOODY, G.M. Using prior knowledge in artificial neural network classifications with a minimal training set. **International Journal of Remote Sensing**, London, v.16, n.2, p.301-312, 1995b.
- FORMAN, R.T.T.; GODRON, M. **Landscape ecology**. New York: John Wiley & Sons, 1986. 619p.

- FRANKLIN, S. E.; WILSON, B.A. Spatial and spectral classification of remote sensing imagery. **Computer & Geosciences**, v.17, n.8, p.1151-1172, 1991
- FRANKLIN, S.E.; PEDDLE, D.R. Classification of SPOT HRV imagery and texture features. **International Journal of Remote Sensing**, London, v.11, n.3, p.551-556, 1990.
- FREUND, J.E.; WILLIAMS, F.J. **Elementary business statistics**. Ingelwood Cliffs: Prentice-Hall, 1972. Apud ROSENFELD et. al. (1982).
- FRIEDMAN, S.Z. **Mapping urbanized area expansion through digital image processing of Landsat and conventional data**. Pasadena: Jet Propulsion Laboratory- CALTECH-NASA, 1980. Publicação 79-113. Apud JENSEN (1986).
- FRONTIER, S. **Stratégies d'échantillonnage em écologie**. Paris: Masson, 1983. 494p. (Collection d'Ecologie, 17).
- FULLER, R.M.; GROOM, G.B.; JONES, A.R. The land cover map of Great Britain: an automated classification of Landsat Thematic Mapper data. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.60, n.5, p.553-562, 1994.
- FULLER, R.M.; PARSELL, R.J. Classification of TM imagery in the study of land use in lowland Britain: practical considerations for operational use. **International Journal of Remote Sensing**, London, v.11, n.10, p.1901-1917, 1990.
- FULLER, R.M.; PARSELL, R.J.; OLIVER, M.; WYATT, G. Visual and computer classifications of remotely-sensed images: a case study of grasslands in Cambridgeshire. **International Journal of Remote Sensing**, London, v.10, n.1, p.193-210, 1989.
- FUNG, T.; CHAN, K.C. Spatial composition of spectral classes: structural approach for image analysis of heterogeneous land-use and land-cover types. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.60, n.2, p.173-180, 1994.
- FUNG, T.; LEDREW, E. Application of principal components analysis to change detection. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.53, n.12, p.1649-1658, 1987.
- GALLEGO, J.; DELINCÉ, J. Crop area estimation through area frame sampling and remote sensing. In: **MARS PROJECT**. Ispra: Institute for Remote Sensing Applications-JRC, 1991.
- GARCIA, M.C.; ALVAREZ, R. TM digital processing of a tropical forest region in southeastern Mexico. **International Journal of Remote Sensing**, London, v.15, n.8, p.1611-1632, 1994.
- GINEVAN, M.E. Testing land-use map accuracy: another look. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.45, n.10, p.1371-1377, 1979.
- GOLDBERG, M.; SHLIEN, S. A four-dimensional histogram approach to the clustering of Landsat data. In: **SYMPOSIUM ON MACHINE PROCESSING OF REMOTELY**

- SENSED DATA, 1977. **Proceedings...** West Lafayette: Purdue University, 1977. p.250-259. Apud MATHER (1987).
- GOLDBERG, M.; SHLIEN, S. A clustering scheme for multispectral images. **IEEE Transactions on Systems, Man and Cybernetics**, New York, v.8, n.2, p.86-92, 1978.
- GONG, P.; HOWARTH, P.J. Performance analyses of probabilistic relaxation methods for land-cover classification. **Remote Sensing of Environment**, New York, v.30, n., p.33-42, 1989.
- GONG, P.; HOWARTH, P.J. Assessment of some factors influencing multispectral land-cover classification. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.56, n.5, p.597-603, 1990a.
- GONG, P.; HOWARTH, P.J. The use of structural information for improving land cover classification accuracies at the rural-urban fringe. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.56, n.1, p.67-73, 1990b.
- GONG, P.; HOWARTH, P.J. Frequency-based contextual classification and gray-level vector reduction for land use identification. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.58, n.4, p.432-437, 1992.
- GONG, P.; MARCEAU, D.J.; HOWARTH, P.J. Comparison of spatial feature extraction algorithms for land use classification with SPOT-HRV data. **Remote Sensing of Environment**, New York, v.40, n.2, p.137-151, 1992.
- GOODCHILD, M.F.; BIGING, G.S.; CONGALTON, R.G.; LANGLEY, P.G.; CHRISMAN, N.R.; DAVIS, F.W. **Final report of the accuracy assessment task force**. Santa Barbara: University of California-NCGIA, 1994. Apud BAILEY et al. (1994).
- GORDON, D.K.; PHILLIPSON, W.R. A texture enhancement procedure for separating orchard from forest in Thematic Mapper data. **International Journal of Remote Sensing**, London, v.7, n.2, p.301-304, 1986.
- GREEN, E.J.; STRAWDERMAN, W.E.; AIROLA, T.M. Assessing classification probabilities for thematic maps. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.59, n.5, p.635-639, 1993.
- GUO, L.J. Balance contrast enhancement technique and its application in image colour composition. **International Journal of Remote Sensing**, London, v.12, n.10, p.2133-2151, 1991.
- GUO, L.J.; HAIGH, J.D. A three-dimensional feature space iterative clustering method for multi-spectral image classification. **International Journal of Remote Sensing**, London, v.15, n.15, p.633-644, 1994.
- GUTTMAN, I.; WILKS, S.S.; HUNTER, J.S. **Introductory engineering statistics**. New York: John Wiley & Sons, 1982. 580p.

- HAACK, B.; BRYANT, N.; ADAMS, S. Landsat Thematic Mapper digital information content for agricultural environments. In: ANNUAL MEETING OF AMERICAN SOCIETY OF PHOTOGRAMMETRY AND REMOTE SENSING, 1986, Falls Church. **Proceedings...** Washington: American Society of Photogrammetry and Remote Sensing, 1986. p.260-268. Apud HAACK & JAMPOLER (1995).
- HAACK, B.; JAMPOLER, S. Colour composite comparisons for agricultural assessments. **International Journal of Remote Sensing**, London, v.16, n.9, p.1589-1598, 1995.
- HAAS, R.H. Landsat Thematic Mapper products for rangeland assessment. **Geocarto International**, Hong Kong, v.7, p.27-33, 1992.
- HAAS, R.H.; WALTZ, F.A. Evaluation of Thematic Mapper data for natural resources assessment. In: PECORA REMOTE SENSING SYMPOSIUM, 8, 1983, Sioux Falls. **Proceedings...** Washington: U.S. Geological Survey, 1983. p.122-133. Apud HAACK & JAMPOLER (1995).
- HARALICK, R.M. Automatic remote sensors image processing: In: ROSENFELD A. (Ed.) **Digital Picture Analysis**. Berlin: Springer-Verlag, 1976. p.5-63.
- HARALICK, R.M. Statistical and structural approaches to texture. **Proceedings of the Institute of Electrical and Electronics Engineers**, Tokyo, v.67, n.5, p. 786-804, 1979.
- HARALICK; R.M.; FU, K. Pattern recognition and classification. In: COLWELL, R.N. **Manual of Remote Sensing**. Falls Church: American Society of Photogrammetry, 1983. p.793-805. v.1.
- HARDIN, P.J. Parametric and nearest-neighbor methods for hybrid classification: a comparison of pixel assignment accuracy. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.60, n.12, p.1439-1448, 1994.
- HAY, A.M. Sampling designs to test land-use map accuracy. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.45, n.4, p.529-533, 1979.
- HEALY, M.J.R. Multivariate normal plotting. **Applied Statistics**, London, v.17, p.157-161, 1968.
- HELLDEN, U. 1980. **A test of Landsat 2 imagery and digital data for thematic mapping illustrated by an environmental study in northern Kenya**. Sweden: Natural Geography Institute-Land University, 1980. Report n. 47. Apud ROSENFELD & FITZPATRICK-LINS (1986).
- HEPNER, G.F.; LOGAN, T.; RITTER, N.; BRYANT, N. Artificial neural network classification using a minimal training set: comparison to conventional supervised classification. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.56, n.4, p.469-473, 1990.
- HERNANDES, G.L.S. **Processamento digital de imagens Landsat/TM e Geoscan/AMSS na caracterização lito-estrutural e delimitação de alteração**

- hidrotermal na área do depósito aurífero de riacho dos Machados (MG).** Campinas: IG-UNICAMP, 1994. 120p. Dissertação de Mestrado em Geociências.
- HIGASHI, M. ADEOS. In: LAND SATELLITE INFORMATION IN THE NEXT DECADE CONFERENCE, 1995, Vienna. **Proceedings...** Bethesda: ASPRS, 1995. p.I-20-I-24.
- HILL, J.; MÉGIER, J. Regional land cover and agricultural statistics and mapping in the Département Ardèche, France, by use of Thematic Mapper data. **International Journal of Remote Sensing**, London, v.9, n.10-11, p.1573-1595, 1988.
- HILL, R.A.; FOODY, G.M. Separability of tropical rain-forests types in the Tambopata-Candamo Reserved Zone, Peru. **International Journal of Remote Sensing**, London, v.15, n.13, p.2687-2693, 1994.
- HIXSON, M.M.; DAVIS, B.J.; BAUER, M.E. Sampling Landsat classifications for crop area estimation. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.47, n.9, p.1343-1348, 1981.
- HIXSON, M.M.; SCHOLZ, D.; FUHS, N.; AKIYAMA, T. Evaluation of several schemes for classification of remotely sensed data. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v. 46, n.12, p.1547-1553, 1980.
- HODGSON, M.E. Reducing the computational requirements of the minimum-distance classifier. **Remote Sensing of Environment**, New York, v.25, n1., p.117-128, 1988.
- HOFFER, R. **Natural resource mapping in mountainous terrain by computer analysis of ERTS-1 satellite data.** West Lafayette: Purdue University, 1975. 124p. LARS Research Bulletin, n. 919. APUD CONGALTON et al (1983).
- HOFFER, R.; FLEMING, M. **Mapping vegetative cover by computer aided analysis of satellite data.** Purdue University. West Lafayette: Purdue University, 1978. 10p. LARS Technical Report, n.011178. APUD CONGALTON et al (1983).
- HOFFER, R.M. Biological and physical considerations in applying computer-aided analysis techniques to remote sensor data. In: SWAIN,P.H., DAVIS, S.M. (Ed.) **Remote sensing: the quantitative approach.** New York: McGraw-Hill, 1978. 227-289p.
- HOFFER, R.M.; JOHANNSEN, C.J. Ecological potentials in spectral signature analysis. In: JOHNSON, P.L. (Ed.) **Remote sensing in ecology.** Athens: University of Georgia Press, 1969. Apud SWAIN & DAVIS (1978).
- HORD, R.M. **Digital image processing of remotely sensed data.** New York: Academic Press, 1982. 256p. Apud JENSEN (1986).
- HORD, R.M.; BROONER, W. Land-use map accuracy criteria. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.42, n.5, p.671-677, 1976.
- HUANG, K; MAUSAL, P.M. Comparing a Piecewise linear classifier with Gaussian Maximum-Likelihood and Parallelepiped classifiers in terms of accuracy and speed.

- Photogrammetric Engineering & Remote Sensing**, Bethesda, v.60, n.11, p.1333-1338, 1994.
- HUDSON, W.D.; RAMM, W. Correct formulation of the Kappa coefficient of agreement. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.53, n.4, p.421-422, 1987.
- HUTCHINSON, C.F. Techniques for combining Landsat and ancillary data for digital classification improvement. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.48, n.1, p.123-130, 1982.
- INCE, F. Maximum likelihood classification, optimal or problematic? A comparison with the nearest neighbour classification. **International Journal of Remote Sensing**, London, v.8, n.12, p.1829-1838, 1987.
- INTERNATIONAL SYMPOSIUM ON PRIMARY DATA ACQUISITION, 1990, Manaus. **Proceedings...** São José dos Campos: INPE/ISPRS/SELP/RSBC, 1990. 2v. v.28. parte 1.
- I²S. **System 600 image processing software. User's guide.** Milpitas: International Imaging Systems, 1992. não paginado.
- JANSSEN, L.L.F.; JAARSMA, M.N.; VAN DER LINDEN, E.T.M. Integrating topographical data with remote sensing for land cover classification. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.56, n., p.1503-1506, 1990.
- JANSSEN, L.L.F.; MIDDELKOOP, H. Knowledge-based crop classification of a Landsat Thematic Mapper image. **International Journal of Remote Sensing**, London, v.13, n.15, p.2827-2837, 1992.
- JANSSEN, L.L.F.; VAN DER WEL, F.J.M. Accuracy assessment of satellite derived land-cover data: a review. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.60, n.4, p.419-426, 1994.
- JENSEN, J.R. **Introductory digital image processing - a remote sensing perspective.** Englewood Cliffs: Prentice-Hall, 1986. 379p.
- JENSEN, J.R. **Introductory digital image processing - a remote sensing perspective.** Englewood Cliffs: Prentice-Hall, 1996. 316p.2 ed.
- JOLY, C.A. Biodiversity of the gallery forest and its role in soil stability in the Jacaré-Pepira water, state of São Paulo Brasil. In: JENSEN, A.E. (Ed.) Ecotones at the river basin scale-global land/water interactions. ECOTONES REGIONAL WORKSHOP, 1992, Barmera, South Australia. **Proceedings...** Adelaide: UNESCO-South Australian Department of Environment and Natural Resources, 1994. p.40-66.
- JORIA, P.; JORGENSON, J.C. Comparison of three methods of mapping tundra with Landsat digital data. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.59, n.2, p.239-246, 1993.

- JUSOFF, K.; MANAF, M.R.A. Satellite remote sensing of deforestation in the Sungai Buloh forest reserve, Peninsular Malaysia. **International Journal of Remote Sensing**, London, v.16, n.11, p.1981-1997, 1995.
- JUSTICE, C.O.; TOWNSHEND, J.R.G. Integrating ground data with remote sensing. In: TOWNSHEND, J.R.G. (Ed.). **Terrain analysis and remote sensing**. Boston: Allen and Unwin, 1981. p.38-58.
- JUSTICE, C.O.; TOWNSHEND, J.R.G. A comparison of unsupervised classification procedures using LANDSAT MSS data for an area of complex surface conditions in Basilicata, southern Italy. **Remote Sensing of Environment**, New York, v.12, n.4, p.407-420, 1982.
- KARTERIS, M.A. The utility of digital Thematic Mapper data for natural resources classification. **International Journal of Remote Sensing**, London, v.11, n.9, p.1589-1598, 1990.
- KARTIKEYAN, B.; GOPALAKRISHNA, B.; KALUBARME, M.H.; MAJUMDER, K.L. Contextual techniques for classification of high and low resolution remote sensing data. **International Journal of Remote Sensing**, London, v.15, n.5, p. 1037-105, 1994.
- KEEFER, B.J.; SMITH, J.L.; GREGOIRE, T.G. Modeling and evaluating the effects of stream mode digitizing errors on map variables. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.57, n.7, p.957-963, 1991.
- KENK, E.; SONDEHEIM, M.; YEE, B. Methods for improving accuracy of Thematic Mapper ground cover classifications. **Canadian Journal of Remote Sensing**, Ottawa, v.14, n.1, p.17-31, 1988.
- KISH, L. **Survey sampling**. New York: John Wiley & Sons, 1965. 643p.
- KULLBACK, S. **Information theory and statistics**. New York: John Wiley & Sons. 1959. Apud CONESE & MASELLI (1993).
- KUSHWAHA, S.P.S.; KUNTZ, S.; OESTEN, G. Applications of image texture in forest classification. **International Journal of Remote Sensing**, London, v. 15, n. 11, p. 2273-2284, 1994
- LABOVITZ, M.L.; MASUOKA, E.J. The influence of autocorrelation on signature extraction - an example from a geobotanical investigation of Cotter Basin, Montana. **International Journal of Remote Sensing**, London, v.5, n.5, p.315-332, 1984.
- LANDGREBE, D.A. The quantitative approach: concept and rationale. In: SWAIN, P.H., DAVIS, S.M. (Ed.) **Remote sensing: the quantitative approach**. New York: McGraw-Hill, 1978. 1-20p.
- LANGEREUX, P. Orbimage vise um marche mondial de 500 MF. **Air & Comos**, Paris, v.33, n.1541, p. 47-48, 1995.
- LARDIER, C. Les satellites-espions americains. **Air & Comos**, Paris, v.33, n.1527, p. 50-52, 1995a.

- LARDIER, C. L'imagerie 'haue résolution' déclassifiée. **Air & Comos**, Paris, v.33, n.1510, p. 36-37, 1995b.
- LARK, R.M. Sample size and class variability in the choice of a method of discriminant analysis. **International Journal of Remote Sensing**, London, v. 15, n. 7, p. 1551-1555, 1994.
- LAUER, D.T.; WALTZ, F.A. Landsat-4 investigation of Thematic Mapper and Multispectral Scanner applications. In: SYMPOSIUM ON THE LANDSAT-4 SCIENTIFIC CHARACTERIZATION-EARLY RESULTS, 1983, Greenbelt. **Proceedings...** Washington: Goddard Space Flight Center-NASA, 1983. Apud HAACK & JAMPOLER (1995).
- LAUVER, C.L.; WHISTLER, J.L. Hierarchical classification of Landat TM imagery to identify natural grassland areas and rare species habitat. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.59, n.5, p.627-634, 1993.
- LIEBERG, S.A. **Tolerância a inundaç o e aspectos demogr ficos de *Inga affinis* DC.** Campinas: IB-UNICAMP, 1990. 104p. Dissertaç o de Mestrado em Ecologia.
- LILLESAND, T.M.; KIEFER, R.W. **Remote sensing and image interpretation.** New York: John Wiley & Sons, 1994. 750p.
- LING, H.S., ROSENFELD, G.H, FITZPATRICK-LINS, K. **Sample selection for thematic map accuracy testing.** Reston: United States Geological Survey, 1979. Circular. Apud CONGALTON (1988).
- LOBO, P.C. **Toler ncia   inundaç o de plantas de *Talauma ovata* St.Hil. e aspectos de seu comportamento em uma mata ciliar da bacia do rio Jacar -Pepira, Brotas, S o Paulo.** Campinas: IB-UNICAMP, 1993. 98p. Dissertaç o de Mestrado em Biologia Vegetal.
- LOBO, P.C.; JOLY, C.A. Ecofisiologia da germinaç o de sementes de *Talauma ovata* St. Hil. (Magnoliaceae), uma esp cie t pica de matas de brejo. **Rev. Bras. Bot.**, S o Paulo, v.18, 1996. (no prelo)
- LUBCHENCO, J.; OLSON, A.M.; BRUBAKER, L.B.; CARPENTER, S.R.; HOLLAND, M.M.; HUBBELL, S.P.; LEVIN, S.A; MACMAHON, J.A.; MATSON, P.A.; MELILLO, J.M.; MOONEY, H.A.; PETERSON. C.H.; PULLIAM, H.R.; REAL, L.A.; REGAL, P.J.; RISSER, P.G. The sustainable biosphere initiative: an ecological research agenda. **Ecology**, Washington, v.72, p.371-412, 1991.
- LUNETTA, R.S.; CONGALTON, R.G.; FENSTERMAKER, L.K.; JENSEN, J.R; MCGWIRE, K.C.; TINNEY, L.R. Remote sensing and geographic information system data integration: error sources and research issues. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.57, n.6, p.677-687, 1991.
- MA, Z.; OLSON Jr., C.E. A measurement of spectral overlap among cover types. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.55, n.10, p.1441-1444, 1989.

- MAHAR, K.; AFIFI, M.S. Linear and correlation analysis for computerized identification of categories in Landsat images. **International Journal of Remote Sensing**, London, v.16, n.12, p.2277-2284, 1995.
- MAIER, M.H. **Geoecologia, hidrografia, hidroquímica, clima e processos antrópicos na bacia do rio Jacaré-Pepira (SP)**. São Carlos: Depto de Ciências Biológicas-UFSCar, 1983. 305p. Tese de Doutorado em Ciências.
- MANGOLINI, M.; ARINO, O. ERS-SAR and Landsat-TM multitemporal fusion for crop statistics. **Earth Observation Quarterly**, Noorwijk, n.51, p.11-15, 1996.
- MARBLE, D.F. Geographic information systems: an overview. In: PEQUET, D.J.; MARBLE, D.F. **Introductory readings in geographic information systems**. London: Taylor & Francis, 1990. p.8-17.
- MARBLE, D.F.; PEQUET, D.J. Geographic information systems and remote sensing. In: COLWELL, R.N. **Manual of Remote Sensing**. Falls Church: American Society of Photogrammetry, 1983. p.923-958. v.1.
- MARQUES, M.C.M. **Estudos auto-ecológicos do guanandí (*Calophyllum brasiliense* Camb., Clusiaceae) em uma mata ciliar do município de Brotas**. Campinas: IB-UNICAMP, 1994. 91p. Dissertação de Mestrado em Biologia Vegetal.
- MASELLI, F.; CONESE, C.; DE FILIPPIS, T.; ROMANI, M. Integration of ancillary data into a maximum-likelihood classifier with nonparametric priors. **ISPRS Journal of Photogrammetry & Remote Sensing**, Amsterdam, v.50, n.2, p.2-11, 1995.
- MASELLI, F.; CONESE, C.; PETKOV, L. Use of probability entropy for the estimation and graphical representation of the accuracy of maximum likelihood classifications. **ISPRS Journal of Photogrammetry & Remote Sensing**, Amsterdam, v.49, n.2, p.13-20, 1994.
- MASELLI, F.; CONESE, C.; PETKOV, L.; RESTI, R. Inclusion of prior probabilities derived from a nonparametric process into the maximum-likelihood classifier. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.58, n.2, p.201-207, 1992.
- MASELLI, F.; CONESE, C.; ZIPOLI, G.; PITTAU, M.A. Use of error probabilities to improve area estimates based on maximum likelihood classifications. **Remote Sensing of Environment**, New York, v.31, n.2, p.155-160, 1990.
- MASELLI, F.; RODORF, A.; CONESE, C. Fuzzy classification of spatially degraded Thematic Mapper data for the estimation of sub-pixel components. **International Journal of Remote Sensing**, London, v.17, n.3, p.537-551, 1996.
- MATHER, A.S. **Land use**. Harlow: Longman, 1986.286p.
- MATHER, P.M. **Computer processing of remotely sensed images**. Chichester: John Wiley & Sons, 1987. 352p.

- MATSON, P.A.; USTIN, S.L. The future of remote sensing in ecological studies. **Ecology**, Washington, v.72, n.6, p.1917, 1991.
- MAUSEL, P.W.; KRAMBER, W.J.; LEE, J.K. Optimum band selection for supervised classification of multispectral data. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.56, n.1, p.55-60, 1990.
- MAYER, K.E; FOX, L. Identification of conifer species grouping from LANDSAT digital classifications. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.47, n. 11, p.1607-1614, 1981.
- MEAD, R.A.; MEYER, M. Landsat digital data application to forest vegetation and land use classification in Minnesota. In: SYMPOSIUM ON MACHINE PROCESSING OF REMOTELY SENSED DATA, 1977. **Proceedings...** West Lafayette: Purdue University, 1977. p. 270-280. APUD CONGALTON et al. (1983).
- MEAD, R.A.; SZAJGIN, J. Landsat classification accuracy assessment procedures. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.48, n.1, p.139-141, 1982.
- MEER, F. van der. Spectral unmixing of Landsat Thematic Mapper data. **International Journal of Remote Sensing**, London, v.16, n.16, p.3189-3194, 1995.
- MERCHANT, J.W. Employing LANDSAT MSS data in land use mapping: observations and considerations. In: RICHARSON, D.F. (Ed). Remote sensing: as an input to Geographic Informations Systems in the 1980s. PECORA REMOTE SENSING SYMPOSIUM, 7, 1982, Sioux Falls. **Proceedings...** Washington: United States Geological Survey, 1982. p.71-91. Apud CURRAN (1985).
- MEREMBECK, B.F.; BORDEN, F.Y.; PODWYSOCKI, M.H.; APLEGATE, D.N. Application of canonical analysis to multispectral scanner data. In: ANNUAL SYMPOSIUM ON COMPUTER APPLICATIONS IN THE MINERAL INDUSTRIES, 14, 1977, s.l. **Proceedings...** New York: Society of Mining Engineers-American Institute in Mining, Metallurgical and Petroleum Engineers, 1977. Apud SHEFFIELD (1985).
- MEREMBECK, B.F.; TURNER, B.J. Directed canonical analysis and the performance of classifiers under its associated linear transformation. **IEEE Transactions on Geoscience and Remote Sensing**, New York, v.18, n.2, p.190-196, 1980.
- METZGER, J.P. **Evaluation de la diversité du paysage par traitements numériques de données Thematic Mapper**. Le cas du bassin Jacaré-Pepira (Sud-Est du Brésil). Toulouse: Centre d'Ecologie des Ressources Renouvelables-Université Paul Sabatier, 1992. 57p. DEA "Ecologie des systèmes aquatiques continentaux".
- METZGER, J.P. **Structure du paysage et diversité des peuplements ligneux fragmentés du rio Jacaré-Pepira (Sud-Est du Brésil)**. Toulouse: Centre d'Ecologie des Ressources Renouvelables-Université Paul Sabatier, 1996. 216p. Tese de Doutorado em Ecologia da Paisagem.

- MIKHAIL, E.M.; ACKERMAN, F. **Observations and least squares**. New York: University Press of America, 1976. Apud BAILEY (1994).
- MOIK, J.G. Digital processing of remotely sensed images. Washington: NASA, 1980. 330p. NASA SP-341. Apud CURRAN (1985).
- MONMONIER, M. **How to lie with maps**. Chicago: The University of Chicago Press, 1991. 176 p.
- NAIMAN, R.J.; DÉCAMPS, H. (Ed.) **The ecology and management of aquatic-terrestrial ecotones**. Paris: UNESCO-Parthenon, 1990. 316p.
- NAIMAN, R.J.; DÉCAMPS, H.; FOURNIER, F. Role of land/inland water ecotones in landscape management and restoration: a proposal for collaborative research. **MAB Digest**, Paris, v.4, 1989. Apud JOLY (1994).
- NAIMAN, R.J.; HOLLAND, M.M.; DÉCAMPS, H.; RISSER, P.G. A new UNESCO programme: research and management of land/inland water ecotones. **Biology International**, v.17, p.107-136, 1988. Apud NAIMAN & DÉCAMPS (1990).
- NARENDRA, P.M.; GOLDBERG, M. A non-parametric clustering scheme for Landsat. **Pattern Recognition**, Elmsford, v.9, n.4, p.207-215, 1977.
- NELSON, B.W.; KAPOV, V.; ADAMS, J.B.; OLIVEIRA, W.J.; BRAUN, O.P.G.; AMARAL, I.L. Forest disturbance by large blowdowns in the Brazilian Amazon. **Ecology**, Washington, v.75, n.3, p.853-858, 1994.
- NIBLACK, W. **An introduction to digital image processing**. Englewood Cliffs: Prentice Hall, 1985. 215p.
- NICOLINI, E. M. **Composição florística e estrutura fitossociológica do estrato arbóreo em mata mesófila semidecídua no município de Jahu, S.P.** Rio Claro: Instituto de Biociências-UNESP, 1990. Dissertação de Mestrado.
- NIMER, E. **Climatologia do Brasil**. Rio de Janeiro: IBGE, 1989. 421p.
- NOVO, E.M.L.M. **Sensoriamento remoto - princípios e aplicações**. São Paulo: Edgard Blucher, 1989. 308p.
- OSAWA, Y.; TODA, K.; WAKABAYASHI, H.; HAMAZAKI, T.; HIGARASHI, T. Japan's land observation satellite program and advanced land observing satellite (ALOS). In: **LAND SATELLITE INFORMATION IN THE NEXT DECADE CONFERENCE**, 1995, Vienna. **Proceedings...** Bethesda: ASPRS, 1995. p.I-25-I-30.
- PAOLA, J.D.; SCHOWENGERDT, R.A. A review and analysis of backpropagation neural networks for classification of remotely-sensed multi-spectral imagery. **International Journal of Remote Sensing**, London, v.16, n.16, p.3033-3058, 1995.
- PETRI, S.; FÚLFARO, V.J. **Geologia do Brasil**. São Paulo: EDUSP, 1988. 631p.
- PONÇANO, W.L. As coberturas cenozóicas. In: ALMEIDA, F.F.M.; HASUI, Y.; PONÇANO, W.L.; DANTAS, A.S.L.; CARNEIRO, C.D.R.; MELO, M.S.;

- BISTRICHI, C.A. **Mapa Geológico do Estado de São Paulo. Nota Explicativa.** São Paulo: Instituto de Pesquisas Tecnológicas do Estado de São Paulo, 1981. p.82-96. v.1. Esc. 1:500.000.
- PONÇANO, W.L.; CARNEIRO, C.D.R.; BISTRICHI, C.A.; ALMEIDA, F.F.M.; PRANDINI, F.L. **Mapa Geomorfológico do Estado de São Paulo. Nota Explicativa.** São Paulo: Instituto de Pesquisas Tecnológicas do Estado de São Paulo, 1981. 94p. v.1. Esc. 1:1.000.000.
- PRATT, W.K. **Digital image processing.** New York: John Wiley & Sons, 1978. 750p.
- PRISLEY, S.P.; SMITH, J.L. Using classification error matrices to improve the accuracy of weighted land-cover models. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.53, n.9, p.1259-1263, 1987.
- RHODE, W.G. Digital image analysis techniques for natural resource inventories. In: NACIONAL COMPUTER CONFERENCE, 1978. **Proceedings...** [S.l.: s.n.], 1978. p. 43-106. Apud CONGALTON (1988).
- RIBEIRO, E. dos S.C. **Análise variográfica e processamento digital de imagens aplicado à área da suite intrusiva de Itu (SP).** Campinas: IG-UNICAMP, 1994. 160p. Dissertação de Mestrado em Geologia.
- RICHARDS, J.A. Thematic mapping from multispectral image data using the principal components transformation. **Remote Sensing of Environment**, New York, v.16, n.1, p.35-46, 1984.
- RICHARDS, J.A. **Remote sensing digital image analysis.** Berlin: Springer-Verlag, 1986. 281p.
- ROCHA, J.V. **The influence of ground survey size on accuracy of area estimates from satellite images.** Silsoe: Cranfield Institute of Technology-Silsoe College, 1992. 244p. v.1. Tese de Doutorado em Filosofia.
- ROSE, P.W.; ROSENDAHL, P.C. Classification of Landsat data for hydrological application, Everglades National Park. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.49, n.4, p.505-511, 1983.
- ROSENFELD, G.H. Sampling design for estimating change in land use and land cover. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.48, n.5, p.793-801, 1982.
- ROSENFELD, G.H.; FITZPATRICK-LINS, K. A coefficient of agreement as measure of thematic classification accuracy. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.52, n.2, p.223-227, 1986.
- ROSENFELD, G.H.; FITZPATRICK-LINS, K.; LING, H.S. Sampling for thematic map accuracy testing. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.48, n.1, p.131-137, 1982.

- ROSENFELD, G.H.; MELLEY, M.L. Applications of statistics to thematic mapping. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.46, n.10, p.1287-1294, 1980.
- ROUGHGARDEN, J.; RUNNING, S.W.; MATSON, P.A. What does remote sensing do for ecology. **Ecology**, Washington, v.72, n.6, p.1918-1922, 1991.
- RUDD, R.D. Macro land use mapping with simulated space photos. **Photogrammetric Engineering, Bethesda**, v.37, p.365-372, 1971.
- SABINS, F., **Remote Sensing: Principles and Interpretation**. New York: W.H. Freeman, 1977. 449p. Apud MASELLI et al. (1992).
- SADER, S.A.; AHL, D.; LIOU, W.S. Accuracy of Landsat-TM and GIS rule-based methods for forest wetland classification in Maine. **Remote Sensing of Environment**, New York, v.53, n.3, p.133-144, 1995.
- SALINO, A. **Flora pteridofítica das matas ciliares da bacia do Jacaré-Pepira, Estado de São Paulo, Brasil**. Campinas: IB-UNICAMP, 1993. 277p. Dissertação de Mestrado em Biologia Vegetal.
- SALIS, S.M. **Composição florística e estrutura de um remanescente de mata ciliar do rio Jacaré-Pepira, Brotas, São Paulo**. Campinas: IB-UNICAMP, 1990. 111p. Dissertação de Mestrado em Biologia Vegetal.
- SALIS, S.M.; SHEPERD, G.J.; JOLY, C.A. Floristic comparison of mesophytic semideciduous forests of the interior of the state of São Paulo, Southeast Brazil. **Vegetatio**, v.119, p.115-164, 1995.
- SALOMONSON, V.V.; SMITH, P.L.; PINK Jr., A.B.; WEBB, W.C.; LYNCH, T.J. An overview of progress in the design and implementation of Landsat-D systems. **IEEE Transactions on Geoscience and Remote Sensing**, New York, v.18, n.2, p.137-146, 1980.
- SÃO PAULO. Secretaria de Estado dos Negócios do Interior. Coordenadoria de Ação Regional. Instituto Geográfico e Cartográfico. **Plano cartográfico do Estado de São Paulo**. São Paulo: IGC-SP, 1980. Esc. 1:10.000. mapa, impr.
- SÃO PAULO. Secretaria do Meio Ambiente. Departamento Estadual de Proteção de Recursos Naturais. **Mapeamento de Vegetação Natural do Estado de São Paulo**. São Paulo: DPRN/IF/Fundação Florestal/INPE, 1989. Esc. 1:50.000. mapa, impr.
- SAXON, E.C. Multitemporal texture transformed Landsat imagery for mapping ecological gradients. In: AUSTRALIAN REMOTE SENSING CONFERENCE, 3, 1984, Brisbane. **Proceedings...** [S.l.: s.n.], 1984. p255-259. Apud KUSHWAHA et al. (1994).
- SCHAALE, M.; FURRER, R. Land surface classification by neural networks. **International Journal of Remote Sensing**, London, v.16, n.16, p.3003-3031, 1995.

- SCHOWENGERDT, R.A. **Techniques for image processing and classification in remote sensing**. Orlando: Academic Press, 1983. 249p.
- SETZER, J. **Atlas climático e ecológico do Estado de São Paulo**. São Paulo: Comissão Interstadual da Bacia Paraná-Uruguai, 1966. 61p.
- SHARMA, S.A.; BHATT, H.P.; AJAI. Oilseed crop discrimination: selection of optimum bands and role of middle infrared. **ISPRS Journal of Photogrammetry & Remote Sensing**, Amsterdam, v.50, n.5, p.25-30, 1995.
- SHEFFIELD, C. Selecting band combinations from multispectral data. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.58, n.6, p.681-687, 1985.
- SHIMODA, H; FUKUE, K; YAMAGUCHI, R. ZHANG Z. ; SAKATA, T. Accuracy of landcover classification of TM and Spot data. In: IGARSS '88 SYMPOSIUM, 1988, Edimburgh. **Proceedings...** Frascati: ESA, 1988. p529-533. Ref. ESA SP-284
- SHORT, N.M. **The Landsat tutorial workbook: basics of satellite remote sensing**. Greenbelt: Goddard Space Flight Center-NASA, 1982. Reference Publication. n. 1078. Apud ROSENFELD & FITZPATRICK-LINS (1986).
- SILVA, L.F. Radiation and instrumentation in remote sensing. In: SWAIN,P.H., DAVIS, S.M. (Ed.) **Remote sensing: the quantitative approach**. New York: McGraw-Hill, 1978. 21-135p.
- SILVA, W.S. DA; FORNASARI FILHO, N. **Unidades de conservação ambiental e áreas correlatas no Estado de São Paulo**. São Paulo: IPT, 1992. 85p. (Boletim nº 63)
- SIMPÓSIO BRASILEIRO DE GEOPROCESSAMENTO, 1, 1990, São Paulo. **Geoprocessamento**. São Paulo: USP-EP, 1990. 351p.
- SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO -SBSR, 6., 1990, Manaus. **Anais...** São José dos Campos: INPE/SELPER/SBC, 1990. 4v.
- SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO -SBSR, 7., 1993, Curitiba. **Anais...** São José dos Campos: INPE/PR-SEMA/SELPER/SBC, 1993. 3v.
- SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO -SBSR, 7., 1996, Salvador. **Anais...** São José dos Campos: INPE/SELPER/SBC, 1996. 1CD-ROM.
- SKIDMORE, A.K. An expert system classifies eucalypt forest types using Thematic Mapper data and a digital terrain model. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.55, n.10, p.1149-1464, 1989a.
- SKIDMORE, A.K. Unsupervised training area selection in forests using a nonparametric distance measure and spatial informantion. **International Journal of Remote Sensing**, London, v.10, n.1, p.133-146, 1989b.
- SKIDMORE, A.K.; TURNER, B.J. Forest mapping accuracies are improved using a supervised nonparametric classifier with SPOT data. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.54, n.10, p.1415-1421, 1988.

- SKIDMORE, A.K.; TURNER, B.J. Map accuracy assessment using line intersect sampling. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.58, n.10, p.1453-1457, 1992.
- SMITH, J.; KOVALICK, W. A comparison of the effects of resampling before and after classification on the accuracy of the Landsat derived cover type map. In: INTERNATIONAL CONFERENCE ON ADVANCED TECHNOLOGY FOR MONITORING AND PROCESSING GLOBAL ENVIRONMENTAL INFORMATION, 1985, London. **Proceedings...** London: University of London, 1985. 6p. Apud STORY & CONGALTON (1986).
- SMITH, J.; PRISLEY, S.; WEIH, R. 1991 Considering the effect of spatial variability on the outcomes of forest management decisions. In: GIS/LIS, 1991, Atlanta. **Proceedings...** Washington: American Society of Photogrammetry and Remote Sensing, 1991. p. 286-292. Apud CONGALTON & GREEN (1993).
- SNEDECOR, G.; COCHRAN, W. **Statistical methods**. Ames: Iowa State University Press, 1989. 503p.
- SOS MATA ATLÂNTICA. **Evolução dos remanescentes florestais e ecossistemas associados do domínio da mata atlântica no Estado de São Paulo no período 1985-1990**. São Paulo: Fundação SOS Mata Atlântica/INPE, 1993. 46p.
- SPITZER, R.L.; COHEN, J.; FLEISS, J.L.; ENDICOTT, J. Quantification of agreement of psychiatric diagnosis. **Archives of General Psychiatry**, Chicago, v.17, p.83-87, 1967.
- SRINIVASAN, A.; RICHARDS, J.A. Knowledge-based techniques for multi-source classification **International Journal of Remote Sensing**, London, v.11, n.3, p.505-525, 1990.
- ST-ONGE, B.A.; CAVAYAS, F. Estimating forest stand structure from high resolution imagery using the directional variogram. **International Journal of Remote Sensing**, London, v.16, n.11, p.1999-2021, 1995.
- STEHMAN, S.V. Thematic map accuracy assessment from the perspective of finite population sampling. **International Journal of Remote Sensing**, London, v.16, n.3, p.589-593, 1995.
- STEHMAN, S.V. Estimating the Kappa coefficient and its variance under stratified random sampling. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.62, n.4, p.401-402, 1996.
- STENBACK, J.M.; CONGALTON. R.G. Using Thematic Mapper imagery to examine forest understory. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.56, n.9, p.1285-1290, 1990.
- STORY, M.; CAMPBELL, J.; BEST, G. An evaluation of the accuracies of five algorithms for machine processing of remotely sensed data. In: PECORA REMOTE

- SENSING SYMPOSIUM, 9, 1984, Sioux Falls. **Proceedings...** Washington: U.S. Geological Survey, 1984. p.399-405. Apud STORY & CONGALTON (1986).
- STORY, M.; CONGALTON, R.G. Accuracy assessment: a user's perspective. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.52, n.3, p.397-399, 1986.
- STRAHLER, A.H. The use of prior probabilities in maximum likelihood classification of remotely sensed data. **Remote Sensing of Environment**, New York, v.10, n.2, p.135-163, 1980. Apud CURRAN (1985).
- SWAIN, P.H. Fundamentals of pattern recognition in remote sensing. In: SWAIN, P.H., DAVIS, S.M. (Ed.) **Remote sensing: the quantitative approach**. New York: McGraw-Hill, 1978. 136-187p.
- SWAIN, P.H., DAVIS, S.M. **Remote sensing: the quantitative approach**. New York: McGraw-Hill, 1978. 395p.
- SYMPOSIUM ON RESOURCE AND ENVIRONMENTAL MONITORING, 7., 1994, Rio de Janeiro. **Proceedings...** São José dos Campos: INPE/ISPRS, 1994. 3v. v.30. parte 7.
- TARDIN, A.T.; ASSUNÇÃO, G.V.; SOARES, J.V. Análise preliminar de imagens TM visando a discriminação de café, citrus e cana-de-açúcar na região de Furnas, MG. **Pesq. agropec. bras.**, Brasília, v.27, n.9, p.1355-1361, 1992.
- TAYLOR, J.C. Evaluation of SPOT satellite images for measuring landscape features. In: **MONITORING LANDSCAPE CHANGES IN THE NATIONAL PARKS**. Silsoe: Cranfield Institute of Technology-Silsoe College, 1989. Relatório não publicado. Apud ROCHA (1992).
- TAYLOR, M.M. Principal components colour display of ERTS imagery. In: **EARTH RESOURCE TECHNOLOGY SATELLITE-1 SYMPOSIUM**, 3, 1974. **Proceedings...** Washington: Goddard Space Flight Center-NASA, 1974. p.1877. vol. 1B. Apud SHEFFIELD (1985).
- TENNAKOON, S.B.; MURTY, V.V.N.; EIUMNOH, A. Estimation of cropped area and grain yield of rice using remote sensing data. **International Journal of Remote Sensing**, London, v.13, n.3, p.427-439, 1992.
- THIBAUT, D.A. Land satellite information in the future. In: **LAND SATELLITE INFORMATION IN THE NEXT DECADE CONFERENCE**, 1995, Vienna. **Proceedings...** Bethesda: ASPRS, 1995. p.III-14-III-21.
- THOMAS, I.L.; CHING, N.P.; BENNING, V.M.; D'AGUANNO, J.A. A review of multi-channel indices of class separability. **International Journal of Remote Sensing**, London, v.8, n.3, p.331-350, 1987.
- TODD, W. J., GEHRING, D.G.; HAMAN, J.F. Landsat wildland mapping accuracy. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.46, n.4, p.509-520, 1980.

- TOMLINS, G.F. Canadian experience in wetland monitoring by satellite. In: SMITH, H. (Ed.). **Plants and the daylight spectrum**. Academic Press: London, 1981. p102 - 113. Apud CURRAN (1985).
- TOWNSHEND, J.R.G. Agricultural land-cover discrimination using thematic mapper spectral bands. **International Journal of Remote Sensing**, London, v.5, n.4, p.681-698, 1984
- TOWNSHEND, J.R.G. Land cover. **International Journal of Remote Sensing**, London, v.13, n.6-7, p.1319-1328, 1992.
- TOWNSHEND, J.R.G.; GAYLER, J.R.; HARDY, J.R.; JACKSON, M.J.; BAKER, J.R. Preliminary analysis of Landsat-4 Thematic Mapper products. **International Journal of Remote Sensing**, London, v.4, n.4, p.817-828, 1983.
- TOWNSHEND, J.R.G.; JUSTICE, C.O.; LI, W.; GURNEY, C.; MCMANUS, J. Global land cover classification by remote sensing: present capabilities and future possibilities. **Remote Sensing of Environment**, New York, v.35, n.4, p.243-255, 1991.
- TUOMISTO, H.; LINNA, A.; KALLIOLA, R. Use of digitally processed satellite images in studies of tropical rain forest vegetation. **International Journal of Remote Sensing**, London, v.15, n.8, p.a595-1610, 1994.
- TURK, G. GI index: a measure of the success of prediction. **Remote Sensing of Environment**, New York, v.8, p.65-75, 1979.
- TURNER, B.L.; SKOLE, D.; SANDERSON, S.; FISCHER, G.; FRESCO, L.; LEEMANS, R. **Land-use and land-cover change**. Stockholm/Geneva: IGBP/HDP, 1995. 132p. (IGBP Report No. 35).
- USA. U.S. Army Corps of Engineers. **GRASS 4.0. Geographical Resources Analysis Support System. user's reference manual**. Fairfax: CERL-DBA Systems, 1991. 515p.
- USA. U.S. Army Corps of Engineers. **GRASS 4.1. Geographical Resources Analysis Support System. user's reference manual**. Fairfax: CERL, 1993. 537p.
- VAN GENDEREN, J.L.; LOCK, B.F. Testing land-use map accuracy. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.43, n.9, p.1135-1137, 1977.
- VAN GENDEREN, J.L.; LOCK, B.F.; VASS, P.A. Remote sensing: statistical testing of thematic map accuracy. **Remote Sensing of Environment**, New York, v.7, n.1, p.3-14, 1978.
- VAUGHN, R.C. **Acceptance sampling**. Ames: Iowa State University Press, 1974. 233p. Apud GINVEVAN (1979).
- VENKATESWARLU, N.B. **Evaluation of classifier for pattern recognition**. Kanpur: Indian Institute of Technology, 1988. Tese de mestrado. Apud VENKATESWARLU & SINGH (1995).

- VENKATESWARLU, N.B.; RAJU, P.S.V.S.K. A new fast classifier for remotely-sensed data. **International Journal of Remote Sensing**, London, v.14, n.2, p.383-393, 1993.
- VENKATESWARLU, N.B.; SINGH, R.P. A fast maximum likelihood classifier. **International Journal of Remote Sensing**, London, v.16, n.2, p.312-320, 1995a.
- VENKATESWARLU, N.B.; SINGH, R.P. Various approaches to speed up Mahalanobis distance classifier. **International Journal of Remote Sensing**, London, v.16, n.16, p.3157-3163, 1995b.
- VEREGIN, H. **A taxonomy of error in spatial database**. Santa Barbara: University of California-NCGIA, 1989. 113p. Technical paper, n.89-12. Apud JANSSEN & VAN DER WEL (1994).
- VICTOR, M.A.M. **A devastação florestal**. São Paulo: Sociedade Brasileira de Silvicultura, 1975. 48p.
- VIÑAS, O.; BAULIES, X. 1:250.000 land use map of Catalonia (32.000 km²) using multitemporal Landsat-TM data. **International Journal of Remote Sensing**, London, v.16, n.1, p.129-146, 1995.
- WATSON, A.I.; VAUGHAN, R.A.; POWELL, M. Classification using the watershed method. **International Journal of Remote Sensing**, London, v.13, n.10, p.1881-1890, 1992.
- WHARTON, S.W. A context-based land-use classification for high-resolution remotely sensed data. **Journal of Applied Photographic Engineering**, Springfield, v.8, n.1, p.46-50, 1980.
- WHARTON, S.W.; TURNER, B.J. ICAP: An interactive cluster analysis procedure for analyzing remotely sensed data. **Remote Sensing of Environment**, New York, v.11, n., p.279-293, 1981.
- WICKLAND, D.E. Future directions for remote sensing in terrestrial ecological research. In: ASRAR, G. (Ed.) **Theory and applications of optical remote sensing**. New York: John Wiley & Sons, 1989. 691-724p.
- WICKLAND, D.E. Mission to planet Earth: the ecological perspective. **Ecology**, Washington, v.72, n.6, p.1923-1933, 1991.
- WILKINSON, G.G. A review of current issues in the integration of GIS and remote sensing data. **International Journal of Geographical Information Systems**, London, v.10, n.1, p.85-101, 1996.
- WILKINSON, G.G.; FOLVING, S.; KANELLOPOULOS, I.; McCORMICK, N.; FULLERTON, K.; MÉGIER, J. Forest mapping from a multi-source satellite data using neural network classifiers - an experiment in Portugal. **Remote Sensing Reviews**, New York, v.12, p.83-106, 1995.
- WILLIAMS, D.; IRONS, J.; LATTY, R.; MARKHAM, B.; NELSON, R.; STAUFFER, M.; TOLL, D. Impact of TM sensor characteristics on classification accuracy. In:

- INTERNATIONAL GEOSCIENCE AND REMOTE SENSING SYMPOSIUM, 1983, New York. **Proceedings...** New York: IEEE, 1983. p.51-59. v.1. Apud STORY & CONGALTON (1986).
- WOODCOCK, C.E.; STRAHLER, A.H. The factor of scale in remote sensing. **Remote Sensing of Environment**, New York, v.21, n.4, p.311-332, 1987.
- WOODCOCK, C.E.; STRAHLER, A.H.; JUPP, D.L.B. The use of variogram in remote sensing I: scene models and simulated images. **Remote Sensing of Environment**, New York, v.25, n.2, p.323-348, 1988a.
- WOODCOCK, C.E.; STRAHLER, A.H.; JUPP, D.L.B. The use of variogram in remote sensing II: real digital images. **Remote Sensing of Environment**, New York, v.25, n.3, p.349-379, 1988b.
- WU, J.K.; CHENG, D.S.; WANG, W.T.; CAI, D.L. Model based remotely-sensed imagery interpretation. **International Journal of Remote Sensing**, London, v.9, n.8, p.1347-1356, 1988.
- YOOL, S.R.; STAR, Y.L.; ESTES, J.E.; BOTKIN, D.B.; ECKHARDT, D.W.; DAVIS, F.W. Performance analysis of image processing algorithms for classification of natural vegetation in the mountains of Southern California. **International Journal of Remote Sensing**, London, v.7, n.5, p.683-702, 1986.
- ZHIZHUO, W. On the renaming of the discipline "photogrammetry". **ISPRS Journal of Photogrammetry & Remote Sensing**, Amsterdam, v.51, n. 1, p.1-4, 1996.
- ZHUANG, X.; ENGEL, B.A.; BAUMGARDNER, M.F.; SWAIN, P.H. Improving classification of crop residues using digital land ownership data and Landsat TM imagery. **Photogrammetric Engineering & Remote Sensing**, Bethesda, v.57, n.11, p.1487-1492, 1991.
- ZHUANG, X.; ENGEL, B.A.; LOZANO-GARCIA, D.F.; FERNÁNDEZ, R.N.; JOHANNSEN, C.J. Optimization of training data required for neuro-classification. **International Journal of Remote Sensing**, London, v.15, n.16, p.3271-3277, 1994.
- ZICKEL, C.S. **Fitossociologia e dinâmica do estrato herbáceo de dois fragmentos florestais do Estado de São Paulo**. Campinas: IB-UNICAMP, 1995. 129p. Tese de Doutorado em Ciências.
- ZONNEVELD, I.S. Aerial photography, remote sensing and ecology. **ITC Journal**, Enschede, n.4., p.553-560, 1974. Apud CONGALTON (1988).

11. Anexos

11.1 Rotina de execução do programa Lestr

O formato de gravação dos arquivos fornecidos pelo INPE em fitas “streamer” (padrão TM-BSQ) pode ser extraído para uma configuração compreensível a outros sistemas de processamento de imagens através do programa Lestr. As instruções para sua utilização são fornecidas a seguir através da tela de apresentação do programa:

NOME DA IMAGEM: gouveia

-64 -64 18 18WRS : 220.75

Data : 29- 8-1992

Satélite : 5

Quadrante : C

Fita atual : 3

no. de pixels : 3244

no. de linhas : 3088

Banda	Volume	Carga
-------	--------	-------

1	1	s
---	---	---

2	1	s
---	---	---

3	1	s
---	---	---

4	2	s
---	---	---

5	2	s
---	---	---

6	2	n
---	---	---

7	3	s
---	---	---

pixel inicial -> X1 : 0

pixel final -> X2 : 3243

linha inicial -> Y1 : 0

linha final -> Y2 : 3087

SELECIONE BANDAS DESEJADAS s/n

banda 1 ->n

banda 2 ->n

banda 3 ->n

banda 4 ->n

banda 6 ->n

banda 7 ->s

Numero de pixels com correção de skew =3439

Numero de linhas com correção de skew =3088

Lendo banda : gouveia_7.i

Observação: Os valores de x,y são os fornecidos acima sendo que o x2 e y2 devem ser diminuída de 1 unidade, pois x0 e y0 são 0.

11.2 Tabelas da análise da exatidão

As matrizes de erros, a exatidão global e os índices de concordância Kappa global e individual para cada categoria de uso das terras são apresentados em uma série de vinte e uma tabelas, correspondendo aos algoritmos de distância mínima euclidiana, distância mínima euclidiana baseada nos desvios padrões, distância mínima de Mahalanobis e isodata.

Tabela 11-1. Matriz de erro entre os dados de referência terrestre e da classificação da imagem pelo algoritmo de distância mínima euclidiana segundo a legenda temática nº 1.

Classes mapeadas	Classes de referência terrestre																Total	Erro C
	1	2	3	4	5	6	7	8	9	10	11	12	14	15	16			
0	0	0	0	0	0	3	0	0	0	0	0	41	13	0	0	57	1	
1	203	17	95	22	43	73	9	6	138	1	86	42	16	2	14	767	0,74	
2	143	47	129	45	45	15	5	3	246	5	44	16	4	0	0	747	0,94	
3	71	14	52	33	50	13	0	1	33	4	62	63	8	0	0	404	0,87	
4	9	29	20	70	78	10	1	0	3	32	48	5	0	0	0	305	0,77	
5	1	3	2	18	54	5	0	0	0	0	351	3	2	0	0	439	0,88	
6	6	1	24	49	9	4929	259	99	44	89	66	378	3	5	9	5970	0,17	
7	4	0	0	2	1	299	35	0	46	18	9	32	1	0	29	476	0,93	
8	38	20	33	63	20	1250	51	46	516	71	134	85	0	21	16	2364	0,98	
9	78	29	53	89	56	237	91	20	883	49	115	60	8	24	16	1808	0,51	
10	55	51	261	185	351	102	18	7	34	597	236	41	30	92	12	2072	0,71	
11	18	11	4	54	19	18	0	0	1	10	1297	6	1	0	0	1439	0,10	
12	22	1	2	1	6	236	0	55	11	56	2	783	3	10	26	1214	0,36	
14	0	0	0	0	0	0	0	0	0	0	6	0	71	0	0	77	0,08	
Total	648	223	675	631	732	7190	469	237	1955	932	2456	1555	160	154	122	18139		
Erro O	0,69	0,79	0,92	0,89	0,93	0,31	0,93	0,81	0,55	0,36	0,47	0,50	0,56	1	1		0,50	

Legenda:

Linhas: dados da imagem classificada;

Colunas: dados de referência terrestre;

Erro O: erro de omissão (expresso em proporção); Erro C: erro de comissão (expresso em proporção);

Classes de usos das terras: 0. pixels não classificados; 1. várzea; 2. cerrado; 3. cerrado; 4. mata ciliar; 5. mata; 6. pastagem limpa; 7. pastagem com árvores; 8. laranjal jovem; 9. laranjal formado; 10. canavial; 11. eucaliptal; 12. solo exposto; 14. corpos d'água; 15. cultura anual e 16. cultura perene.

Tabela 11-2. Índice de exatidão global da imagem classificada pelo algoritmo de distância mínima euclidiana segundo a legenda temática nº 1.

Exatidão global: 0,50
Intervalo de confiança - 90%: +/- 0,0061 (0,4940 - 0,5062)
Intervalo de confiança - 95%: +/- 0,0073 (0,4929 - 0,5074)
Intervalo de confiança - 99%: +/- 0,0096 (0,4906 - 0,5097)

Tabela 11-3. Índice Kappa global e individual entre os dados de referência terrestre e da classificação da imagem pelo algoritmo da distância mínima euclidiana segundo a legenda temática nº 1.

Imagem classificada		Imagem de referência	
Categorias	KIA	Categorias	KIA
não rotulados	0,00	várzea	0,28
várzea	0,24	cerrado	0,18
cerrado	0,05	cerradão	0,06
cerradão	0,10	mata ciliar	0,10
mata ciliar	0,20	mata	0,05
mata	0,09	pastagem	0,53
pastagem	0,71	pastagem com árvores	0,05
pastagem com árvores	0,05	laranja jovem	0,07
laranja jovem	0,01	laranja formado	0,39
laranja formado	0,43	canavial	0,59
canavial	0,25	eucaliptal	0,49
eucaliptal	0,89	solo exposto;	0,47
solo exposto;	0,61	corpos d'água	0,44
corpos d'água	0,92	cultura anual	0,00
		cultura perene	0,00
Índice de concordância Kappa global=0,40			

Tabela 11-4. Matriz de erro entre os dados de referência terrestre e da classificação da imagem pelo algoritmo de distância mínima euclidiana segundo a legenda temática nº 2.

Classes mapeadas	Classes de referência terrestre												Total	Erro C
	1	2	6	8	9	10	11	12	14	15	16			
0	0	0	3	0	0	0	0	41	13	0	0	57	1	
1	203	177	82	6	138	1	86	42	16	2	14	767	0,74	
2	224	689	49	4	282	41	505	87	14	0	0	1895	0,64	
6	10	86	5522	99	90	107	75	410	4	5	38	6446	0,14	
8	38	136	1301	46	516	71	134	85	0	21	16	2364	0,98	
9	78	227	328	20	883	49	115	60	8	24	16	1808	0,51	
10	55	848	120	7	34	597	236	41	30	92	12	2072	0,71	
11	18	88	18	0	1	10	1297	6	1	0	0	1439	0,10	
12	22	10	236	55	11	56	2	783	3	10	26	1214	0,36	
14	0	0	0	0	0	0	6	0	71	0	0	77	0,08	
Total	648	2261	7659	237	1955	932	2456	1555	160	154	122	18139		
Erro O	0,69	0,70	0,28	0,81	0,55	0,36	0,47	0,50	0,56	1	1		0,44	

Legenda:

Linhas: dados da imagem classificada;

Colunas: dados de referência terrestre;

Erro O: erro de omissão (expresso em proporção);

Erro C: erro de comissão (expresso em proporção);

Classes de usos das terras: 0. pixels não classificados; 1. várzea; 2. mata; 6. pastagem; 8. laranja jovem; 9. laranja formado; 10. canavial; 11. eucaliptal; 12. solo exposto; 14. corpos d'água; 15. cultura anual; 16. cultura perene.

Tabela 11-5. Índice de exatidão global da imagem classificada pelo algoritmo de distância mínima euclidiana a legenda temática nº 2.

Exatidão global: 0,44
Intervalo de confiança - 90%: +/- 0,0061 (0,4376 - 0,4498)
Intervalo de confiança - 95%: +/- 0,0072 (0,4365 - 0,4509)
Intervalo de confiança - 99%: +/- 0,0095 (0,4342 - 0,4532)

Tabela 11-6. Índice Kappa global e individual entre os dados de referência terrestre e da classificação da imagem pelo algoritmo de distância mínima euclidiana segundo a legenda temática nº 2.

Imagem classificada		Imagem de referência	
Categorias	KIA	Categorias	KIA
não rotulados	0,00	várzea	0,28
várzea	0,24	mata	0,22
mata	0,27	pastagem	0,57
pastagem	0,75	laranja jovem	0,07
laranja jovem	0,01	laranja formado	0,39
laranja formado	0,43	canavial	0,59
canavial	0,25	eucaliptal	0,49
eucaliptal	0,89	solo exposto	0,47
solo exposto	0,61	corpos d'água	0,44
corpos d'água	0,92	cultura anual	0,00
cultura perene	0,00	cultura perene	0,00
Índice de concordância Kappa global=0,45			

Tabela 11-7. Matriz de erro entre os dados de referência terrestre e da classificação da imagem pelo algoritmo da distância mínima euclidiana baseada nos desvios padrões segundo a legenda temática n° 1.

Classes mapeadas	Classes de referência terrestre																Total	Erro C
	1	2	3	4	5	6	7	8	9	10	11	12	14	15	16			
0	0	0	0	0	0	3	0	0	0	0	0	41	0	0	0	44	1	
1	70	0	23	4	5	6	3	0	61	0	7	1	0	0	0	180	0,61	
2	51	25	53	20	6	2	4	0	64	0	15	2	4	0	0	246	0,90	
3	172	29	93	83	47	23	0	0	61	6	61	8	3	0	0	525	0,82	
4	3	14	3	21	20	0	0	0	2	2	19	0	0	0	0	82	0,74	
5	7	17	17	82	138	11	0	0	2	2	625	0	1	0	0	900	0,85	
6	6	4	21	40	9	4581	207	121	46	27	37	132	0	6	12	5249	0,13	
7	50	8	13	12	11	965	140	0	149	31	11	159	2	0	45	1596	0,91	
8	18	6	15	27	8	519	31	13	464	29	38	10	0	3	3	1184	0,99	
9	32	16	24	29	25	120	35	5	563	21	66	14	4	1	4	959	0,41	
10	202	90	407	303	414	374	45	47	591	681	526	166	54	132	23	4055	0,83	
11	8	11	1	6	47	8	1	0	2	2	1043	12	1	0	0	1140	0,09	
12	29	3	5	4	2	574	3	51	17	131	4	1009	1	12	35	1880	0,46	
14	0	0	0	0	0	4	0	0	0	0	4	1	90	0	0	99	0,09	
Total	648	223	675	631	732	7190	469	237	1955	932	2456	1555	160	154	122	18139		
Erro O	0,89	0,89	0,86	0,97	0,81	0,36	0,70	0,95	0,71	0,27	0,58	0,35	0,44	1	1		0,53	

Legenda:

Linhas: dados da imagem classificada;

Colunas: dados de referência terrestre;

Erro O: erro de omissão (expresso em proporção); Erro C: erro de comissão (expresso em proporção);

Classes de usos das terras: 0. pixels não classificados; 1. várzea; 2. cerrado; 3. cerrado; 4. mata ciliar; 5. mata; 6. pastagem limpa; 7. pastagem com árvores; 8. laranjal jovem; 9. laranjal formado; 10. canavial; 11. eucaliptal; 12. solo exposto; 14. corpos d'água; 15. cultura anual e 16. cultura perene.

Tabela 11-8. Índice de exatidão global da imagem classificada pelo algoritmo da distância mínima euclidiana baseada nos desvios padrões segundo a legenda temática n° 1.

Exatidão global: 0,53
Intervalo de confiança - 90%: +/- 0,0061 (0,5271 - 0,5393)
Intervalo de confiança - 95%: +/- 0,0073 (0,5260 - 0,5405)
Intervalo de confiança - 99%: +/- 0,0096 (0,5237 - 0,5428)

Tabela 11-9. Índice Kappa global e individual entre os dados de referência terrestre e da classificação da imagem pelo algoritmo de distância mínima euclidiana baseada nos desvios padrões segundo a legenda temática nº 1.

Imagem classificada		Imagem de referência	
Categorias	KIA	Categorias	KIA
não rotulados	0,00	várzea	0,10
várzea	0,37	cerrado	0,10
cerrado	0,09	cerradão	0,11
cerradão	0,15	mata ciliar	0,03
mata ciliar	0,23	mata	0,15
mata	0,12	pastagem	0,49
pastagem	0,79	pastagem com árvores	0,23
pastagem com árvores	0,06	laranja jovem	-0,01
laranja jovem	0,00	laranja formado	0,25
laranja formado	0,54	canavial	0,65
canavial	0,12	eucaliptal	0,39
eucaliptal	0,90	solo exposto;	0,61
solo exposto;	0,49	corpos d'água	0,56
corpos d'água	0,91	cultura anual	0,00
		cultura perene	0,00
Índice de concordância Kappa global=0,37			

Tabela 11-10. Matriz de erro entre os dados de referência terrestre e da classificação da imagem pelo algoritmo de distância mínima euclidiana baseada nos desvios padrões segundo a legenda temática nº 2.

Classes mapeadas	Classes de referência terrestre											Total	Erro C
	1	2	6	8	9	10	11	12	14	15	16		
0	0	0	3	0	0	0	0	41	0	0	0	44	1
1	70	32	9	0	61	0	7	1	0	0	0	180	0,61
2	233	668	40	0	64	10	720	10	8	0	0	1753	0,62
6	56	118	5893	121	195	58	48	291	2	6	57	6845	0,14
8	18	56	550	13	464	29	38	10	0	3	3	1184	0,99
9	32	94	155	5	563	21	66	14	4	1	4	959	0,41
10	202	1214	419	47	591	681	526	166	54	132	23	4055	0,83
11	8	65	9	0	0	2	1043	12	1	0	0	1140	0,09
12	29	14	577	51	17	131	4	1009	1	12	35	1880	0,46
14	0	0	4	0	0	0	4	1	90	0	0	99	0,09
Total	648	2261	7659	237	1955	932	2456	1555	160	154	122	18139	
Erro O	0,89	0,70	0,23	0,95	0,71	0,27	0,58	0,35	0,44	1	1		0,45

Legenda:

Linhas: dados da imagem classificada;

Colunas: dados de referência terrestre;

Erro O: erro de omissão (expresso em proporção); Erro C: erro de comissão (expresso em proporção);

Classes de usos das terras: 0. pixels não classificados; 1. várzea; 2. mata; 6. pastagem; 8. laranja jovem; 9. laranja formado; 10. canavial; 11. eucaliptal; 12. solo exposto; 14. corpos d'água; 15. cultura anual; 16. cultura perene.

Tabela 11-11. Índice de exatidão global da imagem classificada pelo algoritmo de distância mínima euclidiana baseada nos desvios padrões segundo a legenda temática nº 2.

Exatidão global: 0,45
Intervalo de confiança - 90%: +/- 0,0061 (0,4410 - 0,4531)
Intervalo de confiança - 95%: +/- 0,0072 (0,4398 - 0,4543)
Intervalo de confiança - 99%: +/- 0,0095 (0,4375 - 0,4566)

Tabela 11-12. Índice Kappa global e individual entre os dados de referência terrestre e da classificação da imagem pelo algoritmo de distância mínima euclidiana baseada nos desvios padrões segundo a legenda temática nº 2.

Imagem classificada		Imagem de referência	
Categorias	KIA	Categorias	KIA
não rotulados	0,00	várzea	0,10
várzea	0,37	mata	0,22
mata	0,29	pastagem	0,63
pastagem	0,76	laranjal jovem	-0,01
laranjal jovem	0,00	laranjal formado	0,25
laranjal formado	0,54	canavial	0,65
canavial	0,12	eucaliptal	0,39
eucaliptal	0,90	solo exposto	0,61
solo exposto	0,49	corpos d'água	0,56
corpos d'água	0,91	cultura anual	0,00
		cultura perene	0,00
Índice de concordância Kappa global=0,44			

Tabela 11-13. Matriz de erro entre os dados de referência terrestre e da classificação da imagem pelo algoritmo de distância mínima de Mahalanobis segundo a legenda temática n° 1.

Classes mapeadas	Classes de referência terrestre																Total	Erro C
	1	2	3	4	5	6	7	8	9	10	11	12	14	15	16			
0	0	0	0	72	0	252	0	0	0	9	27	0	0	19	24	403	1	
1	72	0	7	5	11	2	0	0	0	2	0	2	0	0	0	101	0,29	
2	53	51	129	32	83	14	41	0	167	6	28	1	2	0	0	607	0,92	
3	178	14	127	71	68	14	1	0	139	8	44	5	0	0	0	669	0,81	
4	0	0	18	22	8	3	3	0	0	26	0	0	0	0	0	89	0,75	
5	1	22	23	64	208	6	0	0	0	60	3	0	0	0	0	387	0,46	
6	7	2	17	27	9	4425	151	96	25	93	30	272	0	5	2	5161	0,14	
7	76	1	34	5	9	1077	233	5	50	12	8	141	5	0	32	1688	0,86	
8	16	14	30	39	16	586	16	36	509	42	62	35	1	5	9	1416	0,97	
9	28	16	12	20	16	126	8	4	791	27	77	27	5	2	2	1161	0,32	
10	167	79	275	235	256	333	16	25	260	696	534	120	42	114	25	3177	0,78	
11	21	12	1	35	37	11	0	0	3	1567	7	1	1	0	0	1695	0,08	
12	26	3	2	4	11	338	0	71	14	45	2	917	5	9	28	1475	0,38	
14	3	0	0	0	0	3	0	0	0	7	0	0	97	0	0	110	0,12	
Total	648	223	675	631	732	7190	469	237	1955	932	2456	1555	160	154	122	18139		
Erro O	0,89	0,77	0,81	0,97	0,72	0,38	0,50	0,85	0,60	0,25	0,36	0,41	0,39	1	1		0,49	

Legenda:

Linhas: dados da imagem classificada;

Erro O: erro de omissão (expresso em proporção);

Erro C: erro de comissão (expresso em proporção);

Classes de usos das terras: 0. pixels não classificados; 1. várzea; 2. cerrado; 3. cerrado; 4. mata ciliar; 5. mata; 6. pastagem limpa; 7. pastagem com árvores; 8. laranja jovem; 9. laranja formado; 10. canavia; 11. eucaliptal; 12. solo exposto; 14. corpos d'água; 15. cultura anual e 16. cultura perene.

Colunas: dados de referência terrestre;

Tabela 11-14. Índice de exatidão global da imagem classificada pelo algoritmo de distância mínima de Mahalanobis segundo a legenda temática n°

Exatidão global: 0,49
Intervalo de confiança - 90%: +/- 0,0061 (0,4844 - 0,4966)
Intervalo de confiança - 95%: +/- 0,0073 (0,4832 - 0,4978)
Intervalo de confiança - 99%: +/- 0,0096 (0,4809 - 0,5001)

Tabela 11-15. Índice Kappa global e individual entre os dados de referência terrestre e da classificação da imagem pelo algoritmo de distância mínima de Mahalanobis segundo a legenda temática nº 1.

Imagem classificada		Imagem de referência	
Categorias	KIA	Categorias	KIA
não rotulados	0,00	várzea	0,11
várzea	0,70	cerrado	0,20
cerrado	0,07	cerradão	0,16
cerradão	0,16	mata ciliar	0,03
mata ciliar	0,22	mata	0,27
mata	0,52	pastagem	0,46
pastagem	0,76	pastagem com árvores	0,45
pastagem com árvores	0,12	laranja jovem	0,08
laranja jovem	0,01	laranja formado	0,36
laranja formado	0,64	canavial	0,69
canavial	0,18	eucaliptal	0,60
eucaliptal	0,91	solo exposto;	0,55
solo exposto;	0,59	corpos d'água	0,60
corpos d'água	0,88	cultura anual	0,00
		cultura perene	0,00
Índice de concordância Kappa global = 0,42			

Tabela 11-16. Matriz de erro entre os dados de referência terrestre e da classificação da imagem pelo algoritmo de distância mínima de Mahalanobis segundo a legenda temática nº 2.

Classes mapeadas	Classes de referência terrestre												Total	Erro C
	1	2	6	8	9	10	11	12	14	15	16			
0	0	72	252	0	0	0	9	27	0	19	24	403	1	
1	72	23	2	0	0	0	2	0	2	0	0	101	0,29	
2	232	949	82	0	306	14	158	9	2	0	0	1752	0,46	
6	83	104	5886	101	75	105	38	413	5	5	34	6849	0,14	
8	16	99	602	36	509	42	62	35	1	5	9	1416	0,97	
9	28	64	134	4	791	27	77	27	5	2	2	1161	0,32	
10	167	845	349	25	260	696	534	120	42	114	25	3177	0,78	
11	21	85	11	0	0	3	1567	7	1	0	0	1695	0,08	
12	26	20	338	71	14	45	2	917	5	9	28	1475	0,38	
14	3	0	3	0	0	0	7	0	97	0	0	110	0,12	
Total	648	2261	7659	237	1955	932	2456	1555	160	154	122	18139		
Erro O	0,89	0,58	0,23	0,85	0,60	0,25	0,36	0,41	0,39	1	1		0,39	

Legenda:

Linhas: dados da imagem classificada;

Colunas: dados de referência terrestre;

Erro O: erro de omissão (expresso em proporção);

Erro C: erro de comissão (expresso em proporção);

Classes de usos das terras: 0. pixels não classificados; 1. várzea; 2. mata; 6. pastagem; 8. laranja jovem; 9. laranja formado; 10. canavial; 11. eucaliptal; 12. solo exposto; 14. corpos d'água; 15. cultura anual; 16. cultura perene.

Tabela 11-17. Índice de exatidão global da imagem classificada pelo algoritmo de distância mínima de Mahalanobis segundo a legenda temática nº 2.

Exatidão global: 0,39
Intervalo de confiança - 90%: +/- 0,006 (0,3870 - 0,3989)
Intervalo de confiança - 95%: +/- 0,0071 (0,3859 - 0,4001)
Intervalo de confiança - 99%: +/- 0,0094 (0,3836 - 0,4023)

Tabela 11-18. Índice Kappa global e individual entre os dados de referência terrestre e da classificação da imagem pelo algoritmo de distância mínima de Mahalanobis segundo a legenda temática nº 2.

Imagem classificada		Imagem de referência	
Categorias	KIA	Categorias	KIA
não rotulados	0,00	várzea	0,11
várzea	0,70	mata	0,36
mata	0,48	pastagem	0,63
pastagem	0,76	laranjal jovem	0,08
laranjal jovem	0,01	laranjal formado	0,36
laranjal formado	0,64	canavial	0,69
canavial	0,18	eucaliptal	0,60
eucaliptal	0,91	solo exposto	0,55
solo exposto	0,59	corpos d'água	0,60
corpos d'água	0,88	cultura anual	0,00
		cultura perene	0,00
Índice de concordância Kappa global=0,50			

Tabela 11-19. Matriz de erro entre os dados de referência terrestre e da classificação da imagem pelo algoritmo isodata.

Classes mapeadas	Classes de referência terrestre												Total	Erro C
	1	2	6	8	9	10	11	12	14	15	16			
0	0	0	0	0	0	0	0	0	13	0	0	13	1	
2	436	1292	237	13	478	174	578	116	44	16	17	3401	0,62	
6	18	139	5825	105	173	95	135	659	3	9	20	7181	0,19	
9	120	266	1193	28	1215	85	139	76	8	12	39	3181	0,62	
10	24	467	98	7	60	478	197	29	12	103	3	1478	0,68	
11	7	68	8	0	1	18	1383	0	0	0	0	1485	0,07	
12	43	26	297	84	28	82	12	675	6	14	43	1310	0,48	
14	0	3	1	0	0	0	12	0	74	0	0	90	0,18	
Total	648	2261	7659	237	1955	932	2456	1555	160	154	122	18139		
Erro O	1	0,43	0,24	1	0,38	0,49	0,44	0,57	0,54	1	1		0,40	

Legenda:

Linhas: dados da imagem classificada;

Colunas: dados de referência terrestre;

Erro O: erro de omissão (expresso em proporção);

Erro C: erro de comissão (expresso em proporção);

Classes de usos das terras: 0. pixels não classificados; 2. vegetação natural; 6. pastagem; 9. laranjal formado; 10. canavial; 11. eucaliptal; 12. solo exposto; 14. corpos d'água; 15. cultura anual e 16. cultura perene.

Tabela 11-20. Índice de exatidão global da imagem classificada pelo algoritmo isodata.

Exatidão Global: 0,40
Intervalo de confiança - 90%: +/- 0,006 (0,3908 - 0,4027)
Intervalo de confiança - 95%: +/- 0,0071 (0,3896 - 0,4039)
Intervalo de confiança - 99%: +/- 0,0094 (0,3874 - 0,4061)

Tabela 11-21. Índice Kappa global e individual entre os dados de referência terrestre e da classificação da imagem pelo algoritmo isodata.

Imagem classificada		Imagem de referência	
Categorias	KIA	Categorias	KIA
não rotulados	0,00	várzea	0,00
vegetação natural	0,29	vegetação natural	0,47
pastagem	0,67	pastagem	0,60
laranja formado	0,31	laranja jovem	0,00
canavial	0,29	laranja formado	0,54
eucaliptal	0,92	canavial	0,47
solo exposto	0,47	eucaliptal	0,52
corpos d'água	0,82	solo exposto	0,39
		corpos d'água	0,46
		cultura anual	0,00
		cultura perene	0,00
Índice de concordância Kappa global: 0,48			