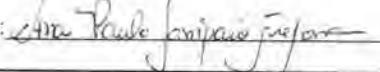


SERVIÇO DE PÓS-GRADUAÇÃO DO ICMC-USP

Data de Depósito: 07.12.2001

Assinatura: 

Determinação da cobertura de solo em fotografias aéreas do Projeto ARARA

Lúcio André de Castro Jorge

Orientador: *Prof. Dr. Onofre Trindade Jr.*

Dissertação apresentada ao Instituto de Ciências Matemáticas e de Computação - ICMC-USP, como parte dos requisitos para obtenção do título de Mestre em Ciências de Computação e Matemática Computacional.

USP – São Carlos
Dezembro/2001

A Comissão Julgadora:

Prof. Dr. Onofre Trindade Jr.

Onofre Trindade Jr
62777
RBL

Prof. Dr. João do Espírito Santo Batista Neto

Prof. Dr. Roberto de Alencar Lotufo

A Deus, pela vida,
à Nossa Senhora, pela luz no caminho,
aos meus Pais, que sempre me incentivaram,
à minha esposa Paula, que com muito amor me apoiou,
ao meu Filho Eduardo, que me fez não desistir,
e em memória aos meus queridos Avós, que certamente de longe
compartilham comigo desta conquista.

AGRADECIMENTOS

Um trabalho como este não é fruto de um esforço isolado, mas sim, o resultado do apoio e incentivo que recebemos de todos os que nos cercam no dia a dia.

Agradeço a todos que de uma forma ou de outra colaboraram com: a amizade, o apoio, o incentivo, as idéias, a orientação, as críticas, a compreensão....

Agradeço em especial ao meu orientador Dr. Onofre Trindade Júnior, pela orientação, incentivo e amizade.

À EMBRAPA Instrumentação Agropecuária pelo apoio financeiro, operacional e institucional, principalmente ao Dr. Luis Alberto Colnago pelo constante incentivo. Também, ao Dr. Paulo Estevão Cruvinel pela oportunidade e apoio.

Sumário

LISTA DE FIGURAS.....	VI
LISTA DE SIGLAS.....	VIII
RESUMO.....	IX
ABSTRACT.....	X
CAPÍTULO 1.....	1
INTRODUÇÃO.....	1
<i>1.1 Motivação.....</i>	1
<i>1.2 Estrutura do Trabalho.....</i>	6
CAPÍTULO 2.....	2
SIARCS E ARARA NA AVALIAÇÃO DA COBERTURA DO SOLO.....	2
<i>2.1 Considerações Iniciais.....</i>	2
<i>2.2 SIARCS.....</i>	2
<i>2.3 Projeto ARARA.....</i>	5
<i>2.4 Considerações Finais.....</i>	9
CAPÍTULO 3.....	11
SEGMENTAÇÃO DE IMAGENS	11
<i>3.1 Considerações Iniciais.....</i>	11
<i>3.2 Segmentação e Classificação de Imagens.....</i>	11
<i>3.3 Técnicas de segmentação.....</i>	12
<i>3.3.1 Limiarização</i>	13
<i>3.3.2 Detecção de Bordas.....</i>	16
<i>3.3.3 Segmentação por Regiões.....</i>	18
<i>3.3.4 Segmentação por Textura</i>	19
<i>3.3.5 Segmentação por Cor</i>	21
<i>3.3.6 Segmentação por Redes Neurais</i>	25
<i>3.3.7 Algoritmos Genéticos (AG)</i>	29
<i>3.3.8 Modelos de Contornos Ativos – Snakes.....</i>	30
<i>3.3.9 Métodos Baseados na Teoria Fuzzy.....</i>	31
<i>3.3.10 Segmentação por Análise Multivariada</i>	32
<i>3.4 Considerações Finais.....</i>	37
CAPÍTULO 4.....	39
TESTE E IMPLEMENTAÇÃO.....	39

<i>4.1 Considerações Iniciais</i>	39
<i>4.2 Especificação do Sistema</i>	39
4.2.1 Diagrama de Fluxo de Dados.....	40
4.2.2 Pré-processar a imagem.....	44
4.2.3 Selecionar Classes.....	45
4.2.4 Calcular Transformada de Hotelling.....	46
4.2.5 Classificar com Mahalanobis.....	47
4.2.6 Treinar Rede MLP	47
4.2.7 Classificar com MLP.....	52
4.2.8 Treinar Rede de Kohonen e Gerar Mapa Auto-Organizado	53
4.2.9 Classificar com ART2	62
<i>4.3 Considerações Finais</i>	67
CAPÍTULO 5.....	68
RESULTADOS OBTIDOS.....	68
<i>5.1 Considerações Iniciais</i>	68
<i>5.2 Imagens Utilizadas</i>	68
<i>5.3 Avaliação qualitativa da cobertura do solo pelos diferentes métodos</i>	71
<i>5.4 Avaliação quantitativa da cobertura do solo pelos diferentes métodos</i>	77
<i>5.5 Considerações Finais</i>	78
CAPÍTULO 6.....	80
CONCLUSÕES.....	80
<i>6.1 Considerações Iniciais</i>	80
<i>6.2 Conclusões</i>	80
<i>6.1 Proposta de trabalhos futuros</i>	81
REFERÊNCIAS BIBLIOGRÁFICAS	82
APÊNDICE A	92
MODELOS DE REPRESENTAÇÃO DAS CORES	92

Lista de Figuras

<i>Figura 2.1: Diagrama de contexto do SIARCS.....</i>	<i>3</i>
<i>Figura 2.2: Diagrama de Fluxo de Dados do SIARCS - Nível 1</i>	<i>4</i>
<i>Figura 2.3: Imagem de cobertura do solo para uma cultura de soja, a segmentação através de filtro de limiarização interativo e o resultado binário.....</i>	<i>4</i>
<i>Figura 2.5: Aeronave desenvolvida no projeto ARARA – Fase II.....</i>	<i>7</i>
<i>Figura 2.6: Sistema típico Fase II</i>	<i>8</i>
<i>Figura 2.7: Sistema Fase I utilizando helicóptero.....</i>	<i>9</i>
<i>Figura 3.1: (a) Imagem original de manchas em um tecido e seu (b) histograma.....</i>	<i>14</i>
<i>Figura 3.2: (a) Imagem segmentada com limiarização na faixa de 5-129 e (b) histograma com a faixa demarcada</i>	<i>15</i>
<i>Figura 3.3: Imagem de duas laranjas: (a) original e (b) segmentada do fundo.</i>	<i>15</i>
<i>Figura 3.4: Histogramas (a) "R", (b) "G" e (c) "B" da imagem da figura 3.3.</i>	<i>16</i>
<i>Figura 3.5: (a) Imagem resultante da detecção de bordas na figura 1a por limiarização e (b) a segmentação das regiões a partir das bordas.....</i>	<i>18</i>
<i>Figura 3.6: (a) Imagem resultante de um operador gradiente para detecção de bordas na figura 1a e (b) a segmentação das regiões a partir das bordas.....</i>	<i>18</i>
<i>Figura 3.7: Construção da barragem no Watershed</i>	<i>19</i>
<i>Figura 3.8: Modelo típico de um neurônio.....</i>	<i>26</i>
<i>Figura 4.1: Diagrama de Fluxo de Dados</i>	<i>43</i>
<i>Figura 4.2: Interface para operação de filtragem da imagem.....</i>	<i>44</i>
<i>Figura 4.3: (a) Imagem original e (b) imagem suavizada.....</i>	<i>45</i>
<i>Figura 4.4: Interface para Seleção de Classes.....</i>	<i>45</i>
<i>Figura 4.5: Topologia da rede MLP utilizada.....</i>	<i>50</i>
<i>Figura 4.6: Variação de erro com diferentes valores de η e α, com duas camadas intermediárias.....</i>	<i>51</i>
<i>Figura 4.7: Variação de erro com diferentes valores de η e α, com uma camada intermediária</i>	<i>52</i>
<i>Figura 4.8: Variação do erro quadrático para a fase de teste para 4 diferentes arquivos, na mesma topologia e valores de η e α.....</i>	<i>53</i>
<i>Figura 4.9: Topologia da Rede de Kohonen.....</i>	<i>54</i>
<i>Figura 4.10: Critério de vizinhança para atualização de pesos</i>	<i>55</i>
<i>Figura 4.11: Entrada de parâmetros na Rede de Kohonen.....</i>	<i>57</i>
<i>Figura 4.12: Imagem teste.....</i>	<i>57</i>
<i>Figura 4.13: Matrizes RGB da figura 4.12.....</i>	<i>58</i>
<i>Figura 4.14: Mapa de saída com entradas RGB da figura 4.13, com $\alpha = 0,7$, erro=0,001 e ciclos=1000 ...</i>	<i>59</i>
<i>Figura 4.15 : Histograma de cores no sistema HOT</i>	<i>59</i>
<i>Figura 4.16: Matriz H da imagem da figura 4.14.....</i>	<i>59</i>

<i>Figura 4.17: Mapa de Kohonen para $\alpha = 0.9$, ciclos=100 e erro= 0.01.....</i>	60
<i>Figura 4.18: Mapa de Kohonen $\alpha = 0.7$, ciclos= 1000, erro= 0.001.....</i>	60
<i>Figura 4.19: Imagem de cobertura depois do filtro de suavização.....</i>	61
<i>Figura 4.20: Mapa de Kohonen para a imagem da figura 4.17, com $\alpha = 0.7$. erro= 0,001 e ciclos= 1000..</i>	61
<i>Figura 4.21: Detalhe de parte do mapa 100x100 gerado para a imagem da figura 4.17.....</i>	62
<i>Figura 4.23: Detalhe da arquitetura ART2</i>	65
<i>Figura 4.24: Saída da rede ART2 para a imagem da figura 4.19.....</i>	66
<i>Figura 5.1: Imagens obtidas com fotos químicas no helicóptero.....</i>	69
<i>Figura 5.2: Imagens obtidas acima de 10m do chão, em vídeo no avião.....</i>	69
<i>Figura 5.3: Imagens obtidas com câmera digital no avião.</i>	70
<i>Figura 5.4: Imagens com filmagem em vídeo pela torre de filmagem</i>	70
<i>Figura 5.5: (a) Imagem obtida com helicóptero (b) com a suavização, (c) a segmentação manual com o SIARCS, a segmentação automática (d) com o discriminador linear, (e) com rede MLP, treinada com três amostras de cada padrão, (f) com rede de Kohonen e rotulada pela mesma MLP, (g) com ART2, $\rho = 0,9$ e (h) a mesma ART2, $\rho = 0,7$.....</i>	72
<i>Figura 5.6: (a) Imagem obtida com helicóptero (b) com a suavização, (c) a segmentação manual com o SIARCS, a segmentação automática (d) com o discriminador linear, (e) com rede MLP, treinada com três amostras de cada padrão, (f) com rede de Kohonen e rotulada pela mesma MLP, (g) com ART2, $\rho = 0,9$ e (h) a mesma ART2, $\rho = 0,7$.....</i>	73
<i>Figura 5.7: (a) Imagem obtida com avião (b) com a suavização, (c) a segmentação manual com o SIARCS, a segmentação automática (d) com o discriminador linear, (e) com rede MLP, treinada com três amostras de cada padrão, (f) com rede de Kohonen e rotulada pela mesma MLP, (g) com ART2, $\rho = 0,9$ e (h) a mesma ART2, $\rho = 0,7$</i>	74
<i>Figura 5.8: (a) Imagem obtida com torre de filmagem (b) com a suavização, (c) a segmentação manual com o SIARCS, a segmentação automática (d) com o discriminador linear, (e) com rede MLP, treinada com três amostras de cada padrão, (f) com rede de Kohonen e rotulada pela mesma MLP, (g) com ART2, $\rho = 0,9$ e (h) a mesma ART2, $\rho = 0,7$.....</i>	75
<i>Figura 5.9: (a) Imagem obtida com avião (b) com a suavização, (c) a segmentação manual com o SIARCS, a segmentação automática (d) com o discriminador linear, (e) com rede MLP, treinada com três amostras de cada padrão, (f) com rede de Kohonen e rotulada pela mesma MLP, (g) com ART2, $\rho = 0,9$ e (h) a mesma ART2, $\rho = 0,7$.....</i>	76
<i>Figura A-1: Sistema CIE-XYZ</i>	94
<i>Figura A-2: Cubo RGB</i>	95
<i>Figura A-3: Modelo típico perceptivo.....</i>	95
<i>Figura A-4: Sistema HSV.....</i>	96
<i>Figura A-5: Duplo cone para sistema HSB</i>	97

Lista de Siglas

AG	Algoritmo Genético
AP	Agricultura de Precisão
ARARA	Aeronaves de Reconhecimento Assistidas por Rádio e Autônomas
ART	Teoria da Ressonância Adaptativa
BP	<i>Backpropagation</i>
CS	Cobertura do Solo
GIS	Sistemas Geográficos de Informações
GPS	Sistema de Posicionamento Global
HSI	Componentes H, S e I do sistema HSI
HOT	Componentes H, O e T do sistema Hotelling
MLP	Perceptron Multi-camada
PNN	Rede Neural Probabilística
RGB	Componentes R, G e B do sistema RGB
RN	Rede Neural
RNA	Rede Neural Artificial
SIARCS	Sistema Integrado para Análise de Raízes e Cobertura do Solo
SOM	Mapas Auto-Organizados
TH	Transformada de Hotelling
UAV	<i>Unmanned Aerial Vehicle</i>

Resumo

Dados importantes para o acompanhamento de uma área agrícola podem ser avaliados através de imagens aéreas. Dentre estes, destaca-se como um dos mais significativos, a identificação e a classificação da cobertura do solo.

A grande dificuldade reside na não disponibilidade de metodologias apropriadas para a análise e a classificação dos padrões de cobertura, principalmente para monitoramento de pequenas propriedades. Imagens de cobertura são imagens complexas, com padrões difíceis de serem definidos. Os padrões variam para cada tipo de solo, dependem das condições de iluminação ambiente, da resolução da imagem, do tipo de planta e resíduos orgânicos sobre o solo, dentre outros fatores.

A extração de atributos de cada pixel é de extrema importância na diferenciação das regiões. Neste trabalho, apresenta-se uma revisão das principais técnicas de segmentação de imagens digitais que serviram de base para a escolha dos métodos utilizados.

A cor foi a característica discriminante utilizada com o objetivo de segmentar de forma automática diferentes padrões de cobertura do solo. Foram testados métodos clássicos de análise como a transformada de Hotelling e o discriminante linear de Mahalanobis. Também foram estudadas técnicas não convencionais, como as Redes Neurais, principalmente pela possibilidade de implementação em hardware específico de alto desempenho. Foram selecionados modelos de redes supervisionadas e não supervisionadas.

Os resultados obtidos indicam a viabilidade de utilização das técnicas avaliadas neste trabalho na segmentação de imagens aéreas e mostram suas limitações e vantagens principais.

Abstract

Important data for the monitoring of agricultural areas can be obtained from aerial images. One of the most valuable data is the classification and measurement of soil covering.

The main problem is the lack of proper methodology for the analysis and classification of soil covering patterns, mainly for the monitoring of small farms. Images used for soil covering analysis are complex, with patterns of difficult identification. Patterns depend on, among other variables, soil type, lighting conditions, image resolution, crop type and type of organic material over the soil.

The recognition of single pixel properties is of maximum importance to identify regions on an image. In this work, the main techniques used to segment digital images are revisited, some of them implemented and tested.

Color is the discriminating characteristic of the image used in this work to automatically segment the different patterns of soil covering. Classical methods, such as the Hotelling transform and the linear discriminator of Mahalanobis, were evaluated. Also have been evaluated non-conventional techniques such as Neural Networks, both supervised and non-supervised models. Neural Networks have special importance because they can be implemented using custom-designed high performance hardware.

Results show that the techniques used in this work are adequate to segment aerial images and make clear its main advantages and limitations.

Capítulo 1

Introdução

1.1 Motivação

Na agricultura, observar a cultura é a melhor forma de avaliar o crescimento e a produtividade da área de cultivo. O melhor sensor do ambiente é a própria plantação uma vez que o estado de nutrição da planta, a disponibilidade de água, presença de pragas e doenças determinam o seu desenvolvimento.

Na prática agrícola convencional têm sido feitas por amostragens a análise de solo, a avaliação da cultura e a verificação de doenças ou de pragas. Estas amostras representam talhões assumidos como uniformes. Com base nessa interpretação, são aplicadas quantidades uniformes de insumos como fertilizantes e defensivos. No entanto, as propriedades do solo e as infestações variam de local para local dentro da mesma lavoura. Em virtude do tamanho das propriedades e por consequência das lavouras, a amostragem feita tradicionalmente não leva em conta a variabilidade espacial [MOL97].

Recentes avanços tecnológicos mostram que é possível alocar insumos com base nas necessidades de cada unidade do campo. Este é um método de administração cuidadosa e detalhada do solo e da cultura para adequar as diferentes condições encontradas em cada pedaço de lavoura, tendo em vista a desuniformidade intrínseca dos solos. Neste caso, o manejo é feito por unidades muito pequenas, observando a variabilidade espacial. Essa técnica tem sido denominada de Agricultura de Precisão. [SMI91], [BEV96].

O conceito de agricultura de precisão (AP), embora simples de ser entendido, requer para sua adoção, um conjunto bastante complexo de tecnologias, como: sensoriamento, manipulação de dados, controle, localização e informação geográfica [MOL97].

O sensoriamento consiste em medir com a devida precisão as quantidades que definem a variabilidade e que normalmente dizem respeito à cultura e ao solo. Pode ser dividido em

três grupos: sensoriamento por meio do deslocamento do sensor no campo, com a finalidade de obter medidas continuamente; sensoriamento por coleta de amostras com base em um quadriculado, que define o tamanho da amostra; e sensoriamento remoto, que consiste na obtenção de dados a distância.

O sensoriamento remoto tem sido largamente utilizado para documentar e identificar muitos problemas e soluções na agricultura. Diversas aplicações têm sido realizadas, principalmente para o estudo dos recursos terrestres e do meio ambiente. A fotografia aérea, radar aéreo e imagem de satélite são os meios mais comuns de realizar este tipo de sensoriamento.

Geralmente a análise das imagens aéreas é capaz de detectar propriedades importantes para o acompanhamento da área agrícola, tais como [MOR97]:

- ✓ **Propriedades dos solos:** as características físicas, como teores de matéria orgânica, textura, permeabilidade, dos diversos tipos de solos podem ser correlacionados com a resposta espectral registrada em imagens provenientes dos campos. Ainda que influenciadas por outras variáveis como tratos culturais, umidade, uma adequada seleção de bandas espectrais e resolução espacial, as imagens podem proporcionar uma boa caracterização dos solos dentro de um campo de cultivo, gerando a identificação de zonas homogêneas de propriedades físicas do solo. Uma identificação deste tipo pode ser uma ferramenta poderosa para a interpolação de parâmetros de amostras de solo coletadas em campo, ajustando os modelos de interpretação de maneira a se produzir um excelente mapa de sítios homogêneos, os quais servem de base para a definição de tratos culturais e aplicações diferenciadas dentro do campo.
- ✓ **Detecção de ataques de pragas e infestações:** em certos intervalos do espectro eletromagnético (bandas espectrais) é possível detectar mudanças fisiológicas e estruturais em plantas, mesmo antes destas alterações estarem visíveis a olho nu (ex. redução de espaços interfoliares ou diminuição de teores de clorofila). Tais alterações estão muitas vezes associadas a ataques de organismos ou infestações de insetos. Essas informações quando associadas a outros dados, por exemplo, em um sistema geográfico de informações (GIS), podem orientar programas de aplicação de defensivos agrícolas e práticas de tratos culturais. Desta forma tem-se o uso otimizado e controlado dos insumos e das interferências nos cultivos.

- ✓ **Stress hídrico:** estimativas de evapotranspiração e índice de deficiência hídrica podem ser acompanhados pela análise de imagens multiespectrais combinando duas imagens de bandas espectrais diferentes, normalmente uma na porção do infravermelho e outra na porção visível do espectro eletromagnético. Através desta combinação se obtém o índice de vegetação que, quando combinado com dados provenientes de sensores instalados no solo, contribuem para um excelente ajuste de mapas de diagnóstico de déficit hídrico. A partir de tais mapas podem ser tomadas decisões relativas à quantidade, freqüência e qualidade de água a ser aplicada em sistemas de irrigação.
- ✓ **Monitoramento de culturas:** acompanhar o desenvolvimento de um campo é uma das ações imprescindíveis para a tomada de decisões que envolvem a AP. Ajustes em tratos culturais, irrigação e mesmo replantios podem ser executados pela detecção de falhas nos talhões da cultura ou diferenças no desenvolvimento. Imagens de alta resolução podem prover tais dados e auxiliar o acompanhamento do desenvolvimento de um campo de cultivo, uma vez que as falhas e as diferenças de padrões são claramente percebidas por um analista experiente. Por vezes, tais feições não são detectadas em campo a não ser em estágios muito avançados, onde a possibilidade de intervenção diminui. A aquisição das imagens deve ser feita ao longo do tempo de tal forma que o desenvolvimento da cultura possa ser acompanhado, o que exige certa flexibilidade das plataformas de coleta. Os cultivos de ciclo longo são os principais beneficiários deste tipo de acompanhamento, uma vez que há mais tempo para a execução de ações corretivas, além de demandarem um intervalo de observações mais longo. No entanto, mesmo cultivos de ciclo médio podem ser acompanhados por sensores em plataformas flexíveis quanto à operação, como aeromodelos e balões.
- ✓ **Zoneamento de sítios:** um dos aspectos de grande importância em AP é a determinação de zonas homogêneas de manejo. Só a partir deste zoneamento pode-se ter uma idéia clara das nuances do terreno e consequentemente das necessidades de cada local do campo de cultivo. Os métodos mais comuns têm sido baseados em mapeamentos de produtividade de cada colheita e uma intensa amostragem de solos. As imagens podem auxiliar neste processo uma vez que as classificações de imagens multiespectrais guardam uma excelente correlação com alguns índices usados para determinar zonas específicas de manejo.
- ✓ **Conservação do solo:** muitos cultivos principais e secundários são avaliados em relação a sua taxa de cobertura do solo visando, principalmente, a proteção do mesmo

contra o impacto das gotas de chuva. Estas causam, dentre outros efeitos, um selamento superficial do solo e processos de erosão. Imagens de alta resolução permitem calcular a cobertura do solo com precisão e alimentam modelos de previsibilidade de erosão permitindo assim uma ação preventiva e corretiva. Estas técnicas também são úteis para manejos como o plantio direto onde se torna importante a determinação de matéria seca sobre o solo.

Pode-se dizer que os sensores que detectam a energia eletromagnética e a registra em filmes ou na forma digital são fundamentais em AP. Eles são os instrumentos básicos em sensoriamento remoto, que a bordo de plataformas aéreas ou orbitais proporcionam dados e possibilitam a detecção de fenômenos e o acompanhamento de alvos determinados. As fotos aéreas obtidas por câmeras instaladas em aviões são uma excelente fonte de dados pela precisão, controle e qualidade inerentes ao processo.

Imagens, com filmes de 35mm, têm sido utilizadas para medir a cobertura de resíduos [SLO77]. Fotos obtidas por aeromodelos e câmeras digitais, para avaliação de plantas de cobertura na entrelinha de uma cultura de citrus e numa área de pastagem, mostram ser de grande interesse, principalmente quando a área de amostragem é grande [JOR99], [MER00].

Pode-se observar que o método fotográfico apresenta precisão, desde que a imagem possua boa resolução e sejam corrigidas suas distorções [SAL88]. É também utilizado para avaliar o desenvolvimento de árvores, demonstrando que o mapeamento de áreas de florestas pode ser feito de forma precisa com fotografias aéreas [MER00]. Também é utilizado no monitoramento de reservas biológicas objetivando a conservação do solo [MAR00].

Rezende et al. [REZ00] utilizaram fotos aéreas de baixo custo, mesmo com todo o progresso do sensoriamento remoto a partir de satélites artificiais, pois estas são de extrema utilidade não só na fotogrametria como também em aplicações onde se exige foto-interpretação.

A utilização de imagens aéreas fornece observações permanentes, reduz o tempo de observação no campo e não tem sua utilização impedida nos dias nublados e quando há movimentação provocada pelo vento. No entanto, existe distorção da imagem, que aumenta do centro para a periferia.

A avaliação da cobertura do solo (CS), por plantas ou por resíduos, bem como o monitoramento da cultura pode ser realizado pela análise das diferenças de cores e de texturas presentes nas imagens [ELW76], [SAL88], [LI98]. Tradicionalmente, as análises qualitativa e quantitativa dos resíduos e plantas têm sido feitas por métodos fotogravimétricos e técnicas estatísticas. [CAL92].

Um método que proporciona precisão na medida da área é o método gravimétrico (ou método fotogravimétrico, quando fotos são usadas). Este método obtém a informação de área a partir do peso. Normalmente utiliza-se um papel recortado do mesmo tamanho da planta ou da área a ser medida, determinando-se o seu peso. O valor da área a ser medida é obtido comparando-se o peso referente a esta área com o peso de um pedaço do mesmo papel de uma área padrão, através de uma simples regra de três. Proporcionam precisão na medida da área, porém são técnicas extremamente trabalhosas quando aplicadas a um grande número de imagens. Também são limitados quando se deseja avaliar áreas específicas nas imagens caracterizadas pela alteração no padrão de cor.

Experiências têm mostrado que os métodos estatísticos, apesar de muito utilizados, são tediosos, particularmente quando um grande número de imagens devem ser avaliado. Além disto, necessitam de ajustes para cada espécie analisada e para cada tratamento [CAL92].

Soluções que apresentam um maior grau de automação têm como elemento comum a utilização de técnicas de processamento e análise de imagens digitais provenientes da tecnologia de interpretação de imagens de satélite e sensoriamento remoto [LIN83], [HAR88], [HAG90], [JOR93], [KAM94], [FRA95], [YON96], [MAR00].

Dentre os sistemas utilizados para análise da cobertura do solo destaca-se o sistema SIARCS (Sistema Integrado para Análise de Raízes e Cobertura do Solo) [EMB96], [CRE94], [CRE96]. Este sistema, desenvolvido na Embrapa Instrumentação Agropecuária, trabalha de forma semi-automática e interativa, possibilitando ao usuário efetuar a segmentação e análise das regiões de interesse. Além da cobertura de solo, vários monitoramentos podem ser processados através da ferramenta interativa de limiarização implementada.

A necessidade de métodos automáticos e independentes dos usuários, principalmente para utilização em sistemas autônomos de aquisição e análise de imagens é fator determinante em aplicações de AP. O uso do SIARCS se torna inviável a não ser que o mesmo seja modificado.

A grande motivação neste trabalho foi estudar técnicas de segmentação para desenvolvimento de um novo módulo de processamento e análise de imagens a ser incorporado no sistema SIARCS da Embrapa. Este poderá ser embarcado no aeromodelo do projeto ARARA (Aeronaves de Reconhecimento Assistidas por Rádio e Autônomas) [SOU98], [SOU99], [TRI00], do Laboratório de Computação de Alto Desempenho do Instituto de Ciências Matemáticas e de Computação, ICMC-USP.

A possibilidade de desenvolvimento de hardware dedicado, possibilitando o aumento na velocidade de processamento, foi determinante na escolha das técnicas utilizadas.

1.2 Estrutura do Trabalho

Após a introdução feita neste capítulo, onde foi apresentada a motivação deste trabalho, no capítulo 2 é apresentada uma introdução sobre os dois projetos envolvidos: o SIARCS e o ARARA.

No capítulo 3 são apresentadas as técnicas de segmentação de imagens, destacando-se os trabalhos de segmentação de imagens por cor, o uso das componentes principais e redes neurais.

No capítulo 4 são apresentadas, com maior detalhamento, as técnicas avaliadas e os algoritmos utilizados na sua implementação. Também são apresentados os testes feitos para verificação do funcionamento dos algoritmos.

No capítulo 5 podem ser observados os resultados obtidos nas imagens do projeto ARARA e uma avaliação final das técnicas utilizadas. As conclusões e propostas de trabalhos futuros podem ser vistas no capítulo 6.

Capítulo 2

SIARCS e ARARA na Avaliação da Cobertura do Solo

2.1 Considerações Iniciais

No capítulo anterior mostrou-se a necessidade de um sistema automático de segmentação e classificação dos padrões de cobertura de solo visando, principalmente, atender os requisitos da agricultura de precisão. Também, foram citados o sistema SIARCS, para avaliação da cobertura do solo, e o projeto ARARA, para o desenvolvimento de aeronaves para obtenção de imagens aéreas.

Neste capítulo, apresenta-se na seção 2.2 o sistema SIARCS e suas vantagens e limitações na avaliação da cobertura do solo. Também é feita uma breve introdução sobre o projeto ARARA na seção 2.3, descrevendo as aeronaves desenvolvidas e testadas, especialmente para a avaliação da cobertura do solo.

2.2 SIARCS

O software SIARCS (Sistema Integrado para Análise de Raízes e Cobertura do Solo), desenvolvido pela Embrapa Instrumentação Agropecuária [EMB96], foi concebido com o objetivo de estudar o sistema radicular das plantas bem como determinar a porcentagem de cobertura do solo por plantas e resíduos.

Este software trabalha de forma semi-automática e interativa, possibilitando ao usuário efetuar a segmentação e análise das regiões de interesse. Utiliza uma ferramenta interativa de limiarização da imagem, que trabalha basicamente com as cores da imagem.

A determinação da cobertura do solo através do SIARCS segue as seguintes etapas:

- ✓ Abertura da Imagem;
- ✓ Pré-processamento da imagem;

- ✓ Segmentação da imagem por classe de cobertura do solo;
- ✓ Determinação da Cobertura do Solo.

As relações entre cada uma destas etapas podem ser observadas no diagrama de contexto do SIARCS, apresentado na figura 2.1.

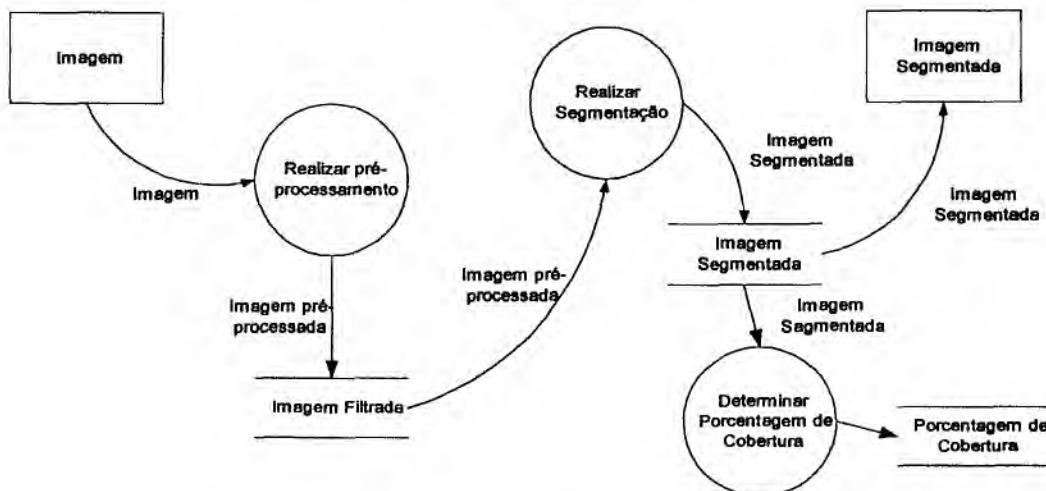


Figura 2.1: Diagrama de contexto do SIARCS

O diagrama de fluxo de dados apresentado na figura 2.2 mostra as funções para segmentação da imagem de cobertura do solo. Os resultados de cada uma das etapas mostradas neste diagrama podem ser vistos na figura 2.3. Nessa figura pode-se observar a imagem original, a imagem resultante da seleção de cores, a imagem binária e a interface de seleção interativa das cores.

A imagem pode estar nos formatos BMP e PCX, ambos com 24 bits, 8 bits ou menos. Se a imagem possui 24 bits é feita a redução das cores para uma escala de 8 bits (256 cores) através de um método que procura manter fieis as tonalidades das cores.

Em seguida é feito o processo de segmentação por limiarização, melhor descrito no próximo capítulo. É utilizado de forma interativa através da tabela de cores da imagem. Esta tabela possui 256 cores que podem ser visualizadas e selecionadas diretamente, conforme pode ser visto na figura 2.3. Cada quadrícula de seleção de cor representa uma cor da tabela. Interativamente são selecionadas as faixas de interesse, podendo ser observado o resultado sobre a imagem original. Obtém-se em seguida a imagem binária, que nada mais é que, a imagem com os pixels das cores selecionadas transformados em

preto e o restante em branco. Através da imagem binária é obtida a porcentagem de cobertura sobre o solo.

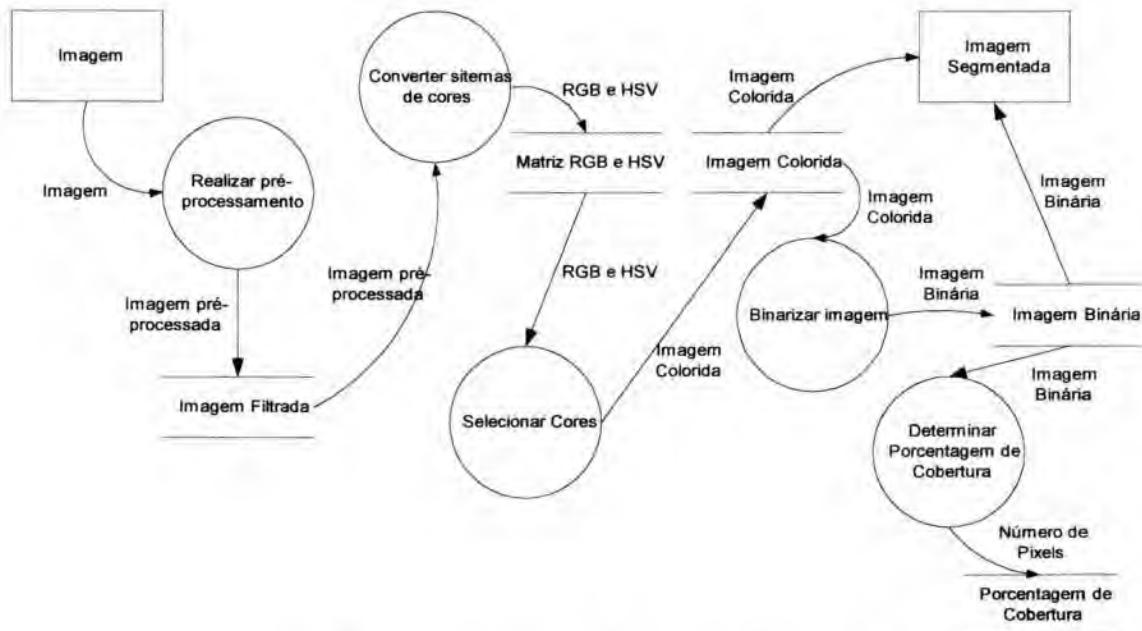


Figura 2.2: Diagrama de Fluxo de Dados do SIARCS - Nível 1

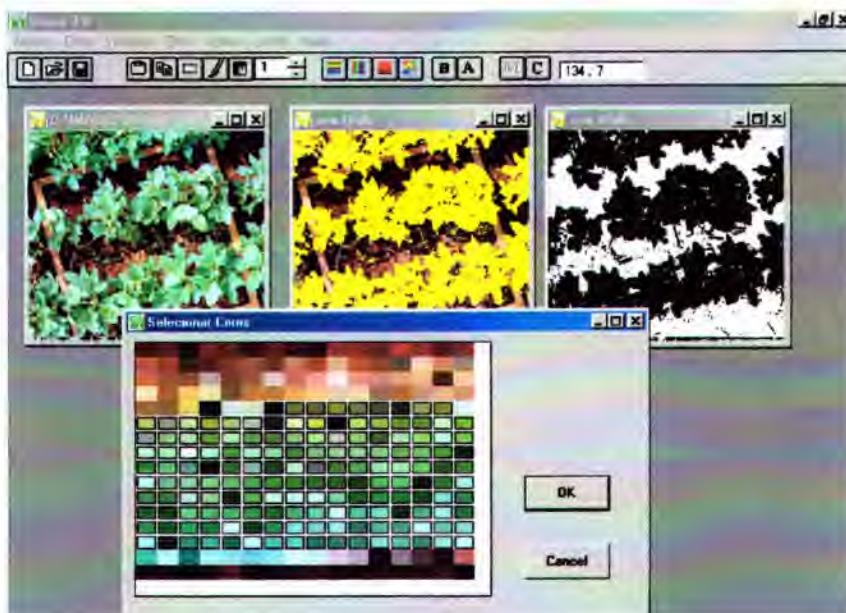


Figura 2.3: Imagem de cobertura do solo para uma cultura de soja, a segmentação através de filtro de limiarização interativo e o resultado binário

2.3 Projeto ARARA

Os meios tradicionais para a obtenção de imagens aéreas utilizam aeronaves tripuladas. O custo de operação dessas aeronaves torna o seu uso proibitivo em pequenas e médias propriedades agrícolas, principalmente quando são necessárias imagens obtidas com uma periodicidade muito elevada ou em grande quantidade. A utilização de meios não convencionais para a obtenção das imagens aéreas pode minimizar de forma significativa o seu custo. O uso de uma torre de filmagem possibilita a obtenção de imagens aéreas de alta resolução de pequenas áreas do solo a baixa altitude. Outra possibilidade consiste no uso de aeronaves em escala reduzida, como no projeto ARARA (Aeronaves de Reconhecimento Assistidas por Rádio e Autônomas).

Atualmente, os veículos aéreos não tripulados (UAV – *Unmanned Aerial Vehicle*) têm desempenhado um papel importante em diversas aplicações. Com a miniaturização dos componentes eletrônicos e o incremento do poder de processamento dos computadores, tornou possível a utilização de UAV's em diversas áreas. Projetos de baixo custo podem realizar tarefas até então somente possíveis com a utilização de aeronaves tripuladas. Os UAV's atuais podem chegar a lugares de difícil acesso, como encostas de montanhas e vales, podem desenvolver velocidades muito altas ou permanecer suspensos no ar.

No desenvolvimento da tecnologia de aviões, os UAV's têm sido utilizados como modelos de baixo custo para testes de novos conceitos aerodinâmicos. Na área militar, os UAV's são utilizados em missões de reconhecimento e prática de tiro ao alvo, sendo protótipos reutilizáveis, diferentemente dos mísseis que são utilizados uma única vez.

A fotografia aérea e a coleta de dados para monitoramento ambiental têm feito uso intensivo de UAV's e se beneficiado da miniaturização do GPS (*Global Positioning System*), das câmeras de vídeo, das câmeras fotográficas digitais, dos equipamentos de rádio e dos sensores atmosféricos.

A utilização de aeronaves em escala reduzida para monitoramento apresenta uma série de vantagens em relação à utilização de aeronaves convencionais. A principal delas está no custo de aquisição e operação do equipamento, sendo substancialmente menor que o de aeronaves convencionais. Outra grande vantagem é a possibilidade de operação em velocidade e altitudes menores, sem risco para o piloto e para o equipamento. Aeronaves em escala reduzida também apresentam maior flexibilidade de operação, pois dispensam pistas de pouso e são facilmente transportáveis, permitindo o deslocamento de toda a

equipe de monitoramento e equipamento por meios convencionais de transporte de passageiros.

O projeto ARARA está centrado na pesquisa de UAV's e no estudo da viabilidade da sua aplicação em diversos setores. Está dividido em quatro fases. Os resultados de cada fase representam sistemas com níveis de complexidade e custo crescentes, adequados para aplicações com nível crescente de abrangência geográfica e automação.

- ✓ **Fase I:** Uma câmera fotográfica digital foi adaptada em um aeromodelo convencional [SOU98], [SOU99]. A câmera é disparada por um sinal de rádio enviado através de um equipamento de Rádio-Controle, que também é utilizado para controlar o aeromodelo [PAL99]. Diversos parâmetros e condições de vôo foram determinados experimentalmente para a obtenção de imagens com qualidade aceitável para as aplicações previstas. Este tipo de sistema é adequado para aplicações onde o vôo da aeronave pode ser controlado de terra (pilotagem remota, com visão direta da aeronave). Um transmissor opcional de vídeo pode ser usado para fornecer imagens em tempo real em aplicações de monitoramento. O diagrama básico de um sistema tipo I pode ser visto na figura 2.4.



Figura 2.4: Sistema típico Fase I

- ✓ **Fase II:** Envolve o projeto e a construção de uma aeronave específica para o ARARA [RIB00], conforme apresentado na figura 2.5. A aeronave dispõe de uma grande variedade de sensores a bordo para controle do vôo, tais como GPS, altímetro, velocímetro, horizonte artificial, rotação do motor e nível de combustível. Duas câmeras foram incluídas; suas imagens são transmitidas a um computador na estação em solo. A câmera instalada na frente da aeronave permite que o piloto possa controlar a aeronave fora do seu alcance visual, enquanto que a câmera

voltada para baixo permite que o piloto possa selecionar a área de interesse. As imagens transmitidas pela aeronave são sobrepostas à imagem de um painel convencional de avião, no programa STT (Sistema de Telemetria e Telecomando do projeto ARARA). As condições de pilotagem da aeronave são semelhantes àquelas encontradas nos simuladores de vôo, tal como o *Flight Simulator* da Microsoft. Os instrumentos visualizados no computador são completamente funcionais, recebendo dados enviados pelos sensores a bordo da aeronave através de um modem acoplado no canal de áudio de um transmissor analógico de vídeo. Esta configuração permite vôos além do alcance visual do piloto e controle de rota e altitude do vôo. Uma vez que a área designada está sob visada da câmera apontada para baixo, o piloto pode disparar uma câmera fotográfica a bordo ou acionar um gravador de vídeo no solo. Atualmente estão em fase final de implementação os módulos eletrônicos da segunda versão da aeronave projetada [TRI00]. Neste novo projeto, especial atenção foi dada para a facilidade de construção e utilização do equipamento. Na figura 2.6 tem-se um sistema típico para a fase II.



Figura 2.5: Aeronave desenvolvida no projeto ARARA – Fase II



Figura 2.6: Sistema típico Fase II

- ✓ **Fase III:** Em fase de desenvolvimento, o hardware do sistema da fase III é similar ao da fase II. A diferença principal está no software necessário para realizar operações autônomas. O software foi dividido em dois módulos. O SINACOM (Sistema de Navegação e Controle de Manobras), que inclui o sub-módulo para projeto de missões e carga no submódulo embarcado de controle [BAR00]. O segundo módulo, o SECAV (Sistema de Estabilidade e Controle Automático de Vôo), recebe uma seqüência de manobras básicas de vôo do SINACOM e controla a aeronave para a realização desta seqüência, mantendo a sua estabilidade. Com uma aeronave desta fase, o usuário pode programar através do SINACOM a rota de vôo da aeronave e as tarefas necessárias (tais como fotografar e/ou filmar uma área pré-definida), para o cumprimento automático de uma missão. A aeronave retorna à base para um pouso por pára-quedas depois de realizar a missão programada.
- ✓ **Fase IV:** Esta fase inclui um sistema de processamento de imagens a bordo. A aeronave poderá identificar características tais como linhas de transmissão, estradas, cursos de rios, etc. Esta fase deverá permitir uma variedade de tarefas de monitoramento totalmente autônomas como a inspeção de linhas de transmissão de energia com a finalidade de localizar possíveis falhas, a contagem de animais em uma determinada área ou a localização de áreas sujeitas a desmatamento.

Além das atividades mencionadas acima, também está em desenvolvimento um sistema completamente digital para telemetria, telecomando e transmissão de vídeo para aeronaves da fase II, em substituição aos equipamentos analógicos até aqui utilizados [TRE00].

Paralelamente com o desenvolvimento das aeronaves e sistemas de controle, os algoritmos de processamento de imagens para aplicações da fase IV estão sendo estudados e implementados. Este procedimento possibilitará a utilização da aeronave em aplicações reais tão logo esta esteja completamente operacional.

Para os estudos de cobertura do solo em baixa altitude, cerca de 5 metros de altitude, foi implementado um aeromodelo do tipo helicóptero, conforme pode ser observado na figura 2.7.



Figura 2.7: Sistema Fase I utilizando helicóptero

2.4 Considerações Finais

No capítulo anterior foram apresentadas as aplicações na agricultura que motivaram este trabalho. Neste capítulo foram apresentados os projetos ARARA e SIARCS, ressaltando o interesse na segmentação e classificação de áreas diferenciadas pela cor.

A segmentação de imagens implementada no SIARCS é totalmente interativa, dependendo fortemente, do usuário do sistema. A sua aplicação em grandes áreas, principalmente de forma automática torna-se totalmente inviável.

A maior preocupação no desenvolvimento de uma aplicação que envolve a análise automática de imagens se encontra na escolha do método que resulta na melhor segmentação e classificação dos objetos. Muitas vezes, os resultados obtidos com um dado método são bons para uma aplicação e ruins para outras. Ou seja, a eficiência do método pode variar dependendo da aplicação.

A escolha adequada das ferramentas utilizadas para segmentar e classificar padrões em uma imagem é determinante na eficiência deste processo. Deve-se levar em consideração o potencial para discernimento dos objetos e as características da imagem a ser processada.

Técnicas de segmentação de imagens são apresentadas no próximo capítulo, destacando-se principalmente a importância da cor. Também é feita uma rápida abordagem de outras técnicas presentes na literatura, incluindo a aplicação de redes neurais na análise e segmentação de imagens.

Capítulo 3

Segmentação de Imagens

3.1 Considerações Iniciais

No capítulo anterior foram apresentados os projetos ARARA e SIARCS envolvidos neste trabalho. Percebe-se que um dos maiores desafios do processamento de imagens de cobertura é a segmentação da imagem para a obtenção de dados significativos. A grande dificuldade reside no fato de não se conhecer de antemão o número e o tipo de estruturas que se encontram na imagem. Estas estruturas são identificadas a partir de sua geometria, forma, topologia, textura, cor ou brilho, sendo escolhidas aquelas que possibilitam a melhor identificação.

Neste capítulo são apresentados na seção 3.2 os conceitos de segmentação e classificação de imagens. Na seção 3.3 são abordados os métodos clássicos de segmentação e é feita uma revisão rápida das técnicas de segmentação utilizadas mais recentemente. Destaca-se aqui a segmentação por cor e segmentação por Redes Neurais. Finalmente, na seção 3.4 são apresentadas algumas considerações sobre o material apresentado neste capítulo.

3.2 Segmentação e Classificação de Imagens

A segmentação de imagens é um processo que tipicamente partitiona o domínio espacial de uma imagem em subconjuntos mutuamente exclusivos, chamados regiões, onde cada região é uniforme e homogênea com respeito a algumas propriedades como tom ou textura e cujos valores diferem, em alguns aspectos e significados, das propriedades de cada região vizinha.

Um bom processo de segmentação deve resultar em regiões que sejam uniformes e homogêneas com respeito a alguma característica. Por exemplo, nível de cinza ou textura.

Regiões adjacentes devem ter valores significativamente diferentes com respeito à característica segmentada.

Segundo Nevatia [NEV86] a segmentação é a operação de separar os componentes de uma imagem em subconjuntos que correspondem a objetos ou regiões, e a classificação tem por objetivo classificar cada objeto em diferentes categorias. A segmentação pode ser implementada por diferentes técnicas, como por exemplo, a limiarização (*thresholding*), detecção de bordas e extração de regiões. Dentre estas abordagens, a extração de regiões pode ser resultante de uma operação de classificação dos pixels, que utiliza desde técnicas de agrupamento (*clustering*) até a simples classificação individual de cada pixel.

Segundo Jain [JAI89], os processos de classificação e segmentação são fortemente relacionados. A classificação pode proporcionar a segmentação e vice-versa. A classificação dos pixels de uma imagem é uma forma de rotular seus componentes e pode resultar na segmentação dos vários objetos da mesma. Um exemplo, em sensoriamento remoto, é a classificação de dados multiespectrais para cada pixel, resultando na segmentação de várias regiões com diferentes culturas ou com mata nativa, dentre outras. Similarmente, em reconhecimento de caracteres, a segmentação de imagens resulta na classificação ou identificação de cada objeto.

Os métodos clássicos de segmentação estão divididos em três grandes grupos: limiarização, segmentação por bordas e segmentação por regiões. Outras abordagens tem sido utilizadas, como o uso de redes neurais, algoritmos genéticos, técnicas de classificação e restauração por métodos estocásticos [SON99], etc.

3.3 Técnicas de segmentação

Em geral, uma imagem pode ser descrita por uma função bidimensional $F(x, y)$, onde (x, y) é a coordenada espacial e $F(x, y)$ o valor da característica em (x, y) . Dependendo do tipo da imagem, o valor da característica pode ser intensidade de luz, profundidade, intensidade de onda de rádio ou temperatura, etc. A imagem digital, por outro lado, é uma função discreta bidimensional $f(x, y)$, digitalizada pela coordenada espacial e pela magnitude do valor da característica. Uma imagem digital pode ser vista como uma matriz bidimensional, cujos índices linha e coluna identificam um ponto (pixel) na imagem e o valor do elemento da matriz correspondente identifica o nível de intensidade da característica.

Na literatura, encontram-se várias técnicas de segmentação, mas não existe um método simples que seja bom para todos os tipos de imagens. Nem todos os métodos são igualmente bons para um tipo particular de imagem. As metodologias baseiam-se na textura, cor, detecção de bordas, técnicas de limiarização, redes neurais, algoritmos genéticos, contornos ativos, lógica nebulosa e morfologia matemática, entre outros.

Muitas destas técnicas de segmentação utilizam informações de contexto entre regiões, algoritmos genéticos que são baseados em hipótese-verificação e métodos estocásticos que utilizam os modelos de Markov [PHI99]. Modelos para contornos ativos denominados *snakes* têm recebido destaque na segmentação de imagens por contorno ou borda. Técnicas de redes neurais associadas com as técnicas clássicas também aparecem na literatura [SON99].

A segmentação de imagens também pode ser feita através da morfologia matemática. As ferramentas morfológicas são muito úteis para segmentação de imagens com texturas e diferentes objetos.

Nesta seção, são apresentadas, de forma sucinta, algumas técnicas de segmentação e trabalhos correlatos.

3.3.1 Limiarização

A limiarização representa o processo mais simples de segmentação. A limiarização é a transformação da imagem de entrada $f(x,y)$ gerando uma imagem de saída $g(x,y)$ da seguinte forma: $g(x,y) = 1$ para $f(x,y) \geq T$ e $g(x,y) = 0$ para $f(x,y) < T$, onde T é o limiar escolhido para se estabelecer a segmentação. Como resultado tem-se uma imagem binária, onde, geralmente o objeto e/ou a região de interesse são classificados com valor “1” e o fundo ou o restante com valor “0”.

Quando um único limiar é aplicado à imagem toda o processo é chamado de limiarização global. Quando aplicado utilizando informação de partes da imagem é denominado de limiarização local. Em muitos casos a limiarização global resulta numa boa segmentação, porém geralmente a limiarização local apresenta resultados mais precisos.

A segmentação automática através do uso de limiar é feita com métodos de detecção automática do nível de limiarização. Se algumas propriedades da imagem são conhecidas previamente, a escolha do limiar para segmentação é simplificada. A detecção do limiar pode ser efetuada de diferentes formas: avaliando a forma do histograma e seus vales, determinando o limiar ótimo, etc.

Em histogramas com dois picos bem distintos, o limiar pode ser determinado como o mínimo entre os dois máximos locais. Estes histogramas mostram a presença de duas características bem distintas na imagem, como pode ser observado na figura 3.1. Os objetos da figura 3.1a estão bem evidenciados do fundo.

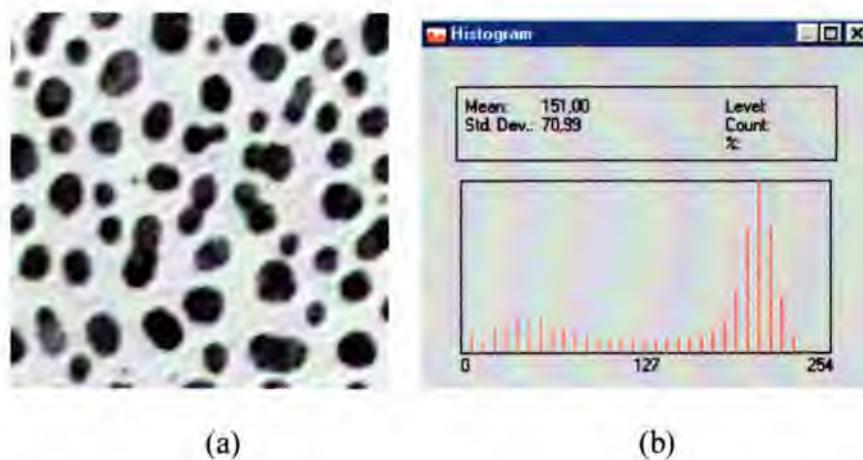


Figura 3.1: (a) Imagem original de manchas em um tecido e seu (b) histograma.

Outra forma de determinar o limiar com mais precisão pode ser feito levando-se em consideração os valores ponderados baseados na função de probabilidade de cada distribuição máxima, ou seja, define-se o limiar ótimo com base na probabilidade dos máximos nas distribuições, resultando no mínimo erro na segmentação.

Uma forma também utilizada para a segmentação por limiarização é a seleção de uma faixa para segmentação ao invés de apenas um único limiar. Na figura 3.2 pode ser visto um exemplo de limiarização na faixa de níveis de cinza de 5 a 129, ou seja, o resultado da segmentação está na faixa de tons de 5 a 129, representados por preto na figura 3.2a.

A utilização de uma ferramenta interativa permite erros maiores na segmentação pois depende do critério utilizado pelo operador. Assim, o ideal é utilizar limiarização automática.

Muitas aplicações utilizam imagens com várias bandas, como é o caso de imagens de satélite, bem como imagens coloridas, que utilizam sistemas de representação de cores em 3D, como o sistema RGB, HSV, etc. Nestes casos, utiliza-se a limiarização multi-espectral.

A limiarização multi-espectral nada mais é que a aplicação de um limiar T para cada banda ou cada componente de cor.

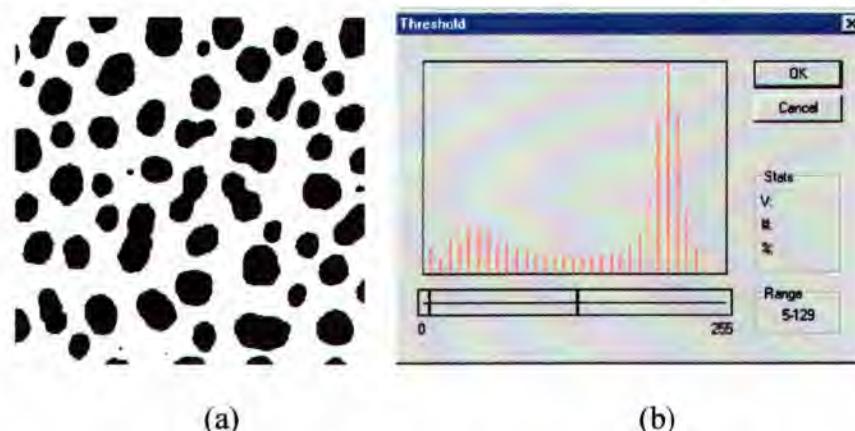


Figura 3.2: (a) Imagem segmentada com limiarização na faixa de 5-129 e (b) histograma com a faixa demarcada.

Um exemplo de segmentação por limiarização multi-espectral pode ser visto na figura 3.3. Na figura 3.3a observa-se a imagem de duas laranjas sobre um fundo escuro. Na figura 3.3b pode ser visto o resultado da segmentação do fundo a partir da limiarização por bandas em cada uma das componentes RGB da imagem. Na figura 3.4 apresenta-se os respectivos histogramas de cada uma das componentes R, G e B da imagem da figura 3.3a, com as faixas utilizadas para a segmentação da imagem devidamente selecionadas.

Da mesma forma pode-se utilizar outros sistemas de cor para a segmentação multi-espectral. Em alguns casos a utilização de apenas uma componente de cor para a limiarização produz excelentes resultados.

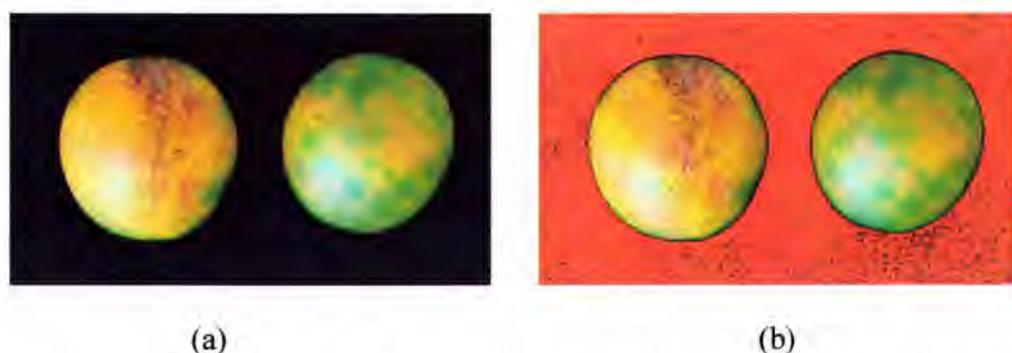


Figura 3.3: Imagem de duas laranjas: (a) original e (b) segmentada do fundo.

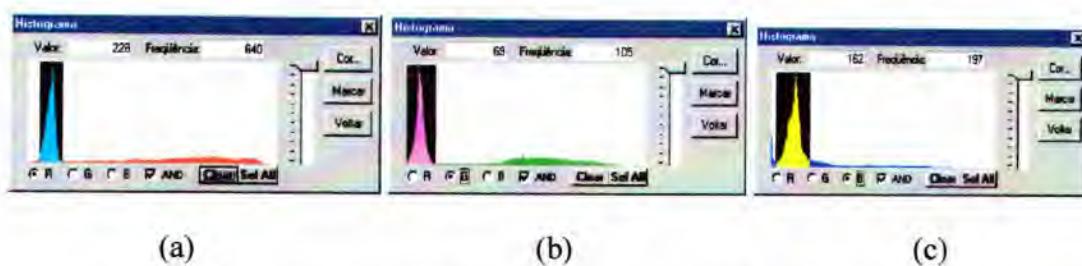


Figura 3.4: Histogramas (a) “R”, (b) “G” e (c) “B” da imagem da figura 3.3.

Todos esses métodos possuem uma desvantagem em comum: eles utilizam no cálculo somente a informação do histograma (ignorando os detalhes espaciais). Os métodos que binarizam o histograma utilizando detalhes espaciais geram resultados de segmentação mais significativos do que os métodos que usam somente a informação do histograma.

3.2.2 Detecção de Bordas

Segmentação baseada na detecção de bordas se resume em técnicas que identificam descontinuidades em tons de cinza, cores, textura, etc.

O principal problema, causado por ruídos ou informações desconhecidas na imagem, é a presença de arestas em locais onde não há uma borda real, ou o contrário, a ausência de arestas onde existe uma borda real.

A limiarização pode ser aplicada para se determinar descontinuidades através da escolha de um limiar apropriado.

Na tentativa do uso de um critério ótimo para a detecção de bordas, pode-se optar pela busca em grafos, que são extremamente eficientes neste tipo de segmentação. A detecção de borda é substituída pela busca do menor caminho em um grafo ponderado. Os custos são associados a cada nó do grafo e refletem a probabilidade da borda passar através de um nó em particular, ou pixel. O objetivo é encontrar o caminho ótimo que conecta dois nós específicos ou conjuntos de nós que representam o início e fim das bordas. [SON99].

Quando a forma dos objetos ou regiões a serem segmentados for conhecida pode-se proceder a segmentação através da transformada de *Hough* [GON87]. Esta transformada pode detectar linhas retas e curvas se suas equações analíticas forem conhecidas. Este é um método eficiente no reconhecimento de objetos com ruidos. Uma forma generalizada pode ser utilizada caso não sejam conhecidas as equações analíticas das formas procuradas. As equações paramétricas das curvas são baseadas em amostras e são determinadas em um estágio prévio de treinamento. [GON87].

A forma mais simples e largamente utilizada para a detecção de bordas é baseada em operadores derivativos locais, que pode ser feita buscando-se o máximo local através da primeira derivada ou o cruzamento com zero da segunda derivada. Phillip [PHI99] apresenta alguns algoritmos para extração de arestas a partir dos métodos derivativos mais utilizados. Muitos dos operadores são do tipo gradiente e podem ser aplicados através da operação de convolução de uma máscara com a imagem. A definição da máscara define o tipo de operador.

Há diferentes tipos de operadores diferenciais como:

- ✓ **Gradientes de Roberts, de Sobel e de Prewitt:** eles não respondem somente a bordas, mas também a pontos isolados. Estes operadores são chamados de operadores de derivada de primeira ordem.
- ✓ **Operador Laplaciano:** responde intensamente a ângulos, linhas e pontos isolados. Para imagens ruidosas, a menos que ela tenha um contraste baixo, o ruído produzirá valores de Laplaciano mais altos do que as bordas. Este operador é chamado de operador de segunda derivada.

De acordo com Philip [PHI99], um bom detetor de borda tem de ter as seguintes propriedades: probabilidade baixa de marcação de pontos que não pertençam à borda e de não marcação de pontos que pertençam à borda (boa detecção); pontos marcados como bordas devem estar tão próximos quanto possível do centro de bordas verdadeiras (boa localização); uma e somente uma resposta para um ponto de borda único (resposta única).

Nas figuras 3.5 e 3.6 podem ser observadas as segmentações por extração de bordas sobre a imagem da figura 3.1a, por limiarização e por operador derivativo local, respectivamente.

Enquanto a formação de regiões a partir de bordas completas é trivial, a determinação de regiões a partir de bordas parciais pode ser uma tarefa muito complexa. A construção de regiões pode ser baseada em probabilidades de que pixels estejam localizados dentro de uma região formada por uma borda parcial. Tais métodos quase sempre não apresentam regiões aceitáveis [SON99].

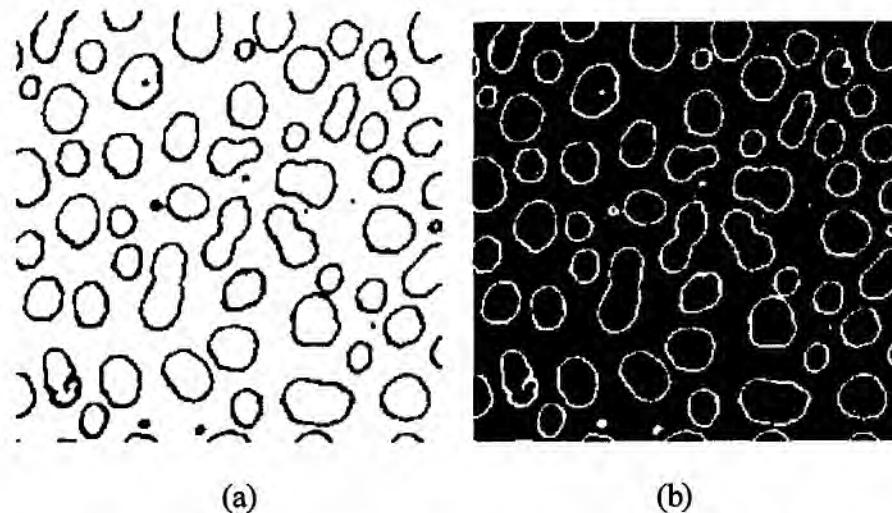


Figura 3.5: (a) Imagem resultante da detecção de bordas na figura 1a por limiarização e
(b) a segmentação das regiões a partir das bordas.

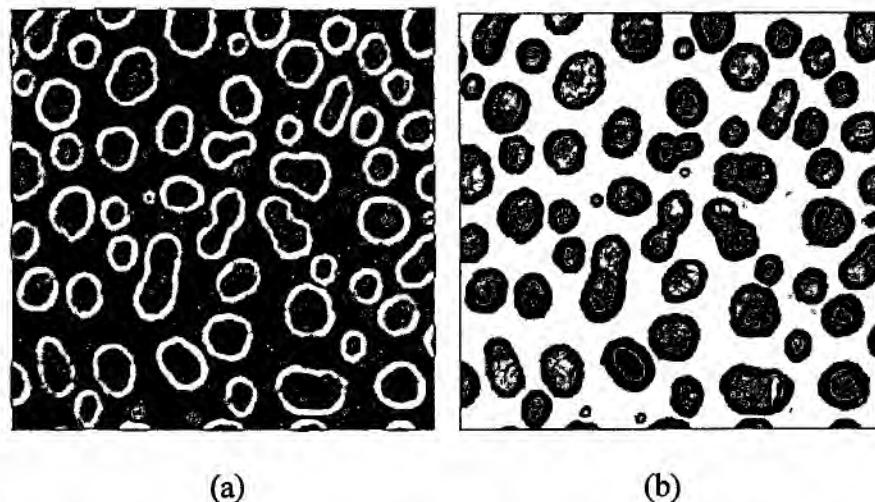


Figura 3.6: (a) Imagem resultante de um operador gradiente para detecção de bordas na
figura 1a e (b) a segmentação das regiões a partir das bordas

3.2.3 Segmentação por Regiões

A segmentação por regiões baseia-se em técnicas que procuram separar regiões com características homogêneas. A forma mais simples é o crescimento de regiões, que deve satisfazer a condição de que a imagem seja formada pela união de todas as regiões encontradas, que estas não tenham interseção e que sejam homogêneas. Também que representem as características distintas utilizadas para a segmentação.

Basicamente três abordagens aparecem no crescimento de regiões: união de regiões, divisão de regiões e união-divisão de regiões.

A união de regiões inicia-se com uma imagem segmentada. As regiões são unidas para satisfazer a condição que continuem formando um conjunto homogêneo. A divisão da imagem é o oposto, inicia-se com a imagem segmentada, ou não, porém com regiões não homogêneas. Começa-se a subdivisão da mesma de tal forma que as regiões formadas na divisão sejam homogêneas. Por fim, a combinação união-divisão, une os dois processamentos anteriores.

Um algoritmo muito utilizado para crescimento de regiões parte de uma semente, normalmente indicada interativamente, depois faz a busca nos vizinhos até que todas as condições de semelhança sejam testadas. Ao final deste procedimento, a região é formada. [PHI99].

Os métodos de crescimento de regiões sempre resultam em muitas regiões ou em poucas regiões. Para melhorar os resultados, uma grande quantidade de pós-processamentos tem sido feita.

Uma outra abordagem muito utilizada é a segmentação por *watershed* ou divisor de águas. Usando a definição de superfície de uma imagem e a terminologia de geografia, um mínimo regional é um máximo conjunto de pontos de nível constante, sem vizinhos menores. O *watershed* de uma região é o conjunto de pontos para os quais uma “gota de água” cai e flui para o mesmo mínimo regional. Para construir um *watershed*, o princípio de uma enchente progressiva de uma superfície é usado: imagine que as regiões de mínimo são preenchidas e que as superfícies são aumentadas. Quando duas superfícies tendem a se encontrar, uma barragem é construída conforme apresentado na figura 3.7. [SON99].

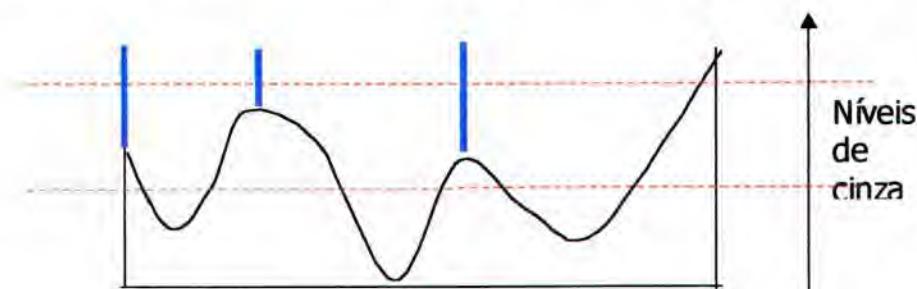


Figura 3.7: Construção da barragem no *Watershed*

3.2.4 Segmentação por Textura

Textura pode ser entendida como um ou mais padrões locais básicos que são repetidos periodicamente. Mas, não está claro exatamente qual é o padrão e como ele é repetido. E, além disso, não está claro se a textura é uma propriedade inerente de todas as coisas ou se

alguns objetos ou regiões são desprovidos totalmente de textura. Segmentar texturas, é partitionar uma imagem em regiões, onde cada região contém uma textura simples e distinta de seus vizinhos.

A textura pode ser representada por modelos que fornecem um meio para transformar uma janela de uma imagem em um conjunto de números. Esses modelos de textura podem ser classificados em três classes principais: estruturas de imagens piramidais, campos randômicos e métodos estatísticos.

O objetivo das estruturas de imagens piramidais é gerar um número de parâmetros homogêneos que representem a resposta de um banco de filtros em escalas variadas e possíveis orientações diferentes. Existem muitos tipos de filtros que podem ser usados para este propósito, como o filtro Laplaciano e o filtro *Gabor*.

Os modelos de campos randômicos assumem que os valores dos pixels são escolhidos por um processo estocástico bidimensional. As duas classes principais são: Campos Randômicos de Markov Gaussiano (*Gaussian Markov Random Fields* - GMRF's), utilizam a distribuição de probabilidade Gaussiana; Campos Randômicos de *Gibbs* (*Gibbs Random Fields* - GRF's), utilizam a distribuição de probabilidade exponencial. Ambas obedecem às propriedades de Markov, que declara o seguinte: “a probabilidade que o pixel tem de ter um certo nível de cinza dados todos os outros pixels da imagem é igual à probabilidade que o pixel tem de ter o mesmo nível de cinza dados somente os seus pixels vizinhos” [LIU99].

A classe dos métodos estatísticos agrupa várias técnicas:

- ✓ **Matriz de co-ocorrência:** representa a distância entre pixels e a dependência espacial dos níveis de cinza para diferentes orientações.
- ✓ **Transformada Discreta de Fourier e Transformada Discreta de Coseno:** são dois exemplos de uma classe geral de técnicas matemáticas chamadas transformadas ortogonais. Geralmente, um conjunto de funções ortogonais fundamentais é gerado, e, então, com cada função aplica-se a convolução com a imagem para produzir um conjunto de números para cada ponto.
- ✓ **Fractais:** são aplicados, às vezes, para a análise e segmentação de textura, porque eles possuem uma escala própria. Isto significa que uma análise fractal pode produzir a dimensão fractal da textura e indicar a escala. Desta forma, evita-se ter de adivinhar em que escala a análise deve proceder.

Na literatura, encontram-se vários trabalhos relacionados com as diversas técnicas de análise de texturas [BAT96], [HAR92].

3.2.5 Segmentação por Cor

Somente alguns animais possuem capacidade de visão colorida. Esta capacidade lhes oferece uma grande vantagem na luta pela sobrevivência em relação às outras espécies. Animais com visão colorida podem perceber detalhes de uma cena mais facilmente, o que lhes proporciona maior habilidade para se proteger e localizar comida. Estas espécies percebem diferentes faixas do espectro de luz visível diferentemente devido às suas próprias condições psicofísicas. Estas são respostas ao fenômeno fisiológico, percepção de cor.

O sistema da visão humana é caracterizado pela tricromatide, ou seja, possui três sensores na retina do olho chamados de cones. Cada um destes é mais sensível a uma porção específica do espectro de luz visível [SPI90]. A interação entre as respostas destes sensores é a causa primária da sensação de cor na espécie humana.

A cor é uma forma de percepção da luz, que por sua vez é um sinal eletromagnético. Quando uma fonte de luz que está iluminando um objeto, tal como o sol ou uma luz incandescente, possuindo todos os comprimentos de onda aos quais o sistema visual é sensível e o objeto reflete todos esses comprimentos de onda igualmente, o objeto é dito acromático, apresentando apenas o branco, preto ou tons intermediários de cinza [DUR87]. A luminosidade está associada à quantidade de luz refletida por um objeto (se o objeto reflete mais luz ele aparece mais branco, se ele absorve quase toda a luz ele aparece preto ou cinza escuro). Suponha três objetos, um branco, um cinza e outro preto. Cada um possui uma característica acromática que depende da quantidade de luz que é refletida. Se a iluminação sobre os mesmos aumenta, suas luminosidades permanecem as mesmas, porém o brilho é alterado. O objeto permanece branco, porém mais claro. Luminosidade é uma propriedade própria do objeto, e o brilho depende da iluminação sobre o mesmo. Sob certas condições, a luminosidade (preto-branco) e brilho (escuroclaro) são difíceis de serem distinguidos perceptivelmente. Usa-se o termo luminância para representar o efeito combinado de luminosidade e brilho.

Objetos que refletem ou emitem distribuições de comprimento de onda em diferentes proporções são chamados de cromáticos. As cores no objeto dependem dos comprimentos de onda dominantes, suas intensidades, do número e proporção de ondas refletidas.

Uma forma objetiva e quantitativa de especificação de uma cor é feita no ramo da física conhecido como colorimetria. Na colorimetria, a cor é definida por um sinal eletromagnético representado pelo seu comprimento de onda dominante. O comprimento de onda dominante de uma cor corresponde, do ponto de vista perceptivo, ao que chamamos de tonalidade ou matiz (em inglês *hue*) da cor e representa intuitivamente a cor propriamente dita (verde, amarelo, azul, etc). Uma fonte de luz monocromática pura que apresenta uma sensação de uma única tonalidade é rara de encontrar. Usualmente, a luz refletida pelo objeto é multi-cromática. Isto é, a sua energia está distribuída em um espectro de vários comprimentos de onda.

Uma vez determinada a tonalidade de uma cor, podemos alterá-la através da mistura com a cor branca. Quanto menos branco a cor contém mais saturada ela fica. Esse conceito perceptivo de saturação corresponde, em colorimetria, ao conceito de pureza de uma cor.

Para determinar uma cor necessitamos de um terceiro parâmetro: a luminância. Esta está ligada ao conceito perceptivo de intensidade da iluminação, conforme já apresentado. Na realidade, os modelos de interface utilizam o brilho que varia monotonicamente com a luminância.

Estes atributos da cor interagem entre si. Por exemplo, quando se aumenta o brilho de uma cor, esta se torna mais saturada, e com este aumento na saturação a luminância da luz emitida é menor, chegando a diferentes níveis para cada cor [DUR87]. Grandes aumentos na intensidade produzem níveis maiores de brilho, porém redução na saturação. Observa-se que a luminância e brilho são usados para descrever um efeito visual de variação de intensidade, mas os termos não são equivalentes. Brilho está ligado à percepção (dimensão psicofísica da cor), e a luminância é a medida de certas componentes de sensação de brilho (energia emitida).

A percepção da cor resulta da combinação de alguns fatores: as características físicas do próprio objeto; sua capacidade de refletir ou absorver certos comprimentos de onda da luz; as propriedades espectrais da fonte de luz que ilumina o objeto; as propriedades espectrais do sistema visual e seus receptores no momento do estímulo; uma experiência prévia com o objeto e a sensação de cor. Outros fatores, tais como, as propriedades da vizinhança do objeto ou a distância e meio através do qual a luz atravessa, podem também afetar a percepção da cor. O sistema visual humano responde a uma porção limitada do espectro eletromagnético, usualmente o intervalo de comprimentos de onda de 380 nm a 770 nm [HAL88].

A cor é uma característica com grande potencial discriminatório. Em diversas situações, o homem toma por base essa informação para diferenciar objetos ou estruturas. Há casos, onde a cor é adicionada ao processo artificialmente, através do uso de corantes. Isso é comum em estudos microscópicos, quando se torna necessário o destaque de estruturas de interesse específico.

Na análise automática de imagens, o processamento da cor é de grande importância na identificação e extração de características e objetos. Este fato deve-se principalmente ao aumento da capacidade de discernimento dos objetos em imagens coloridas. Entretanto, para aplicar a cor em processamento de imagens, é fundamental que se tenha conhecimento dos principais sistemas de representação desta informação.

Os sistemas de representação de cor, de uma forma geral, fornecem uma padronização para a cor, permitindo a sua quantificação, independente das condições psicofísicas dos observadores. Vários são os sistemas para a representação da cor encontrados na literatura; as duas principais abordagens são:

- ✓ **RGB:** consiste na utilização de um espaço tridimensional onde cada cor é obtida a partir das cores primárias vermelha (R), verde (G) e azul (B), respectivamente baixa, média e alta freqüências do espectro percebido pelo olho humano.
- ✓ **HSI:** H é a matiz (descreve a cor pura: vermelha, azul, amarela); S a saturação da cor (grau de pureza); I é a componente que representa a intensidade de luz (brilho) da cor.

Estes dois sistemas de cores são mais bem descritos no apêndice A.

Ohta et al. [OHT80] realizaram um estudo comparativo entre diversos sistemas de representação de cor, buscando as componentes mais significativas para segmentação de imagens coloridas. No processo de comparação foram considerados a qualidade da segmentação e o cálculo envolvido na transformação para obterem-se as novas componentes da cor a partir do sistema RGB. Além de diversas componentes de sistemas de representação de cor conhecidos, tais como RGB e HSI, foram propostas três novas componentes, dadas por: $I1 = (R + G + B)/3$, $I2 = R - B$ e $I3 = (2G - R - B)/2$. Essas componentes foram consideradas significativas nesta ordem e mostraram ser as mais eficientes durante os experimentos realizados. Um dos motivos foi a transformação linear simples necessária para obterem-se as novas componentes do sistema RGB. Uma abordagem parecida foi apresentada por Ohlander et al. [OHL78].

Mais recentemente, Carevic e Caelli [CAR97], utilizaram o mesmo sistema de componentes introduzidos por Ohta et al. [OHT80] e a aplicação da Transformada de “Hotelling” (TH) numa técnica de agrupamento, ou *clustering*, que se mostrou bastante eficiente na identificação de regiões auto-similares da imagem.

Brito Júnior [BRI95] mostra também as vantagens de descorrelacionar as componentes de cor RGB através da TH no processo de segmentação de imagens coloridas, comparando-o com sistemas que utilizam formas de representação mais próximas do sistema de visão humana, por exemplo o sistema HSI. A partir do sistema HSI foi aplicada a TH e efetuada a segmentação a partir de amostras supervisionadas em Brito Júnior [BRI96] e testada por Weber [WEB97].

A segmentação de regiões baseada na classificação por meio da técnica supervisionada do *k*-ésimo vizinho mais próximo, obtida a partir de amostras coletadas pelo operador, foi apresentada por Uchiyama e Arbib [UCH94] e Moreira e Costa [MOR95]. Foram também utilizados conceitos de redes neurais na segmentação por regiões [MOR96] e por agrupamento [VER97]; [KEH98]. Em Phillips e Jain [PHI98] foi efetuada a segmentação local através da decomposição de valores singulares de uma amostra da imagem.

Algumas aplicações para recuperação de imagens coloridas similares em bancos de dados aparecem mais recentemente e os critérios utilizados também são úteis para a segmentação de regiões em imagens coloridas. Hafner et al. [HAF95], Funt e Finlayson [FUN95], Mehltre et al. [MEH95] e Kankanhalli et al. [KAN96] fazem a recuperação baseando-se na similaridade de histogramas medida a partir da distância Euclidiana entre cada um, assim como Jain e Vailaya [JAI96], porém acrescentando, além do atributo cor, a forma. Saber et al. [SAB96] fazem a recuperação baseada na classificação de cores supervisionada através de uma função Gaussiana bidimensional, cujo vetor principal e matriz de covariância para cada classe são estimados por treinamento apropriado e verificados com a avaliação a posteriori. Gong e Sakauchi [GON95] fazem a recuperação através da avaliação no sistema de coordenadas HVC (*Hue*, *Value*, *Croma*), considerando uma função polinomial específica. Mathias e Conci [MAT98] comparam diferentes métricas e espaços de cores na recuperação de imagens coloridas em bancos de imagens.

Também através de histogramas normalizados baseados na proposta de Ohta et al. [OHT80], Healey [HEA92] apresenta uma forma simples de fazer a segmentação por cores normalizadas.

Liu e Yang [LIU94] partitionam as imagens coloridas em regiões homogêneas usando modelos randômicos de Markov e utilizam a avaliação a posteriori para estimar a segmentação. Com uma abordagem baseada nos histogramas das componentes de cor, identificam *clusters* no espaço de medida. Huang et al. [HUA92] também utilizam Markov aliado às outras técnicas para segmentação de imagens coloridas.

Histogramas de cores locais também são utilizados como proposta para segmentar objetos em imagens coloridas em Ennesser e Medioni [ENN95].

Outras propostas adicionam a avaliação dos atributos de textura e forma como complemento na segmentação de imagens coloridas [CAE93], [SCH94], [TAN94].

Do ponto de vista computacional, o processamento da informação cor implica em considerável aumento dos custos, além de necessitar de um espaço de armazenamento três vezes maior. Em algumas aplicações, observa-se a transformação de imagens originalmente coloridas para níveis de cinza. Mas nem sempre o contraste da imagem monocromática é suficiente para a obtenção de bons resultados no processamento de imagens, principalmente quando o objetivo é a segmentação.

3.2.6 Segmentação por Redes Neurais

Uma rede neural (RN) é uma coleção de unidades de processamento, chamadas de neurônios, dispostos em uma estrutura de uma ou várias camadas e conectados uns aos outros de forma paralela. Reagem coletivamente a um dado estímulo e produzem uma única resposta global. Uma rede neural tem a habilidade de aprender um dado tipo de conhecimento, armazenar este conhecimento e usá-lo mais tarde para executar uma tarefa. Algumas das suas vantagens são: paralelismo, memória distribuída, capacidade de adaptação, tolerância a falhas e robustez diante de ruídos.

O conhecimento de uma rede neural artificial (RNA) está na estrutura das conexões entre as unidades (neurônios) que a compõe. A cada conexão associa-se um peso (valor numérico) que caracteriza a força da conexão entre dois neurônios. O aprendizado em uma RNA é realizado por um processo de adaptação dos seus pesos.

Os conhecimentos das redes se resumem a um conjunto de valores numéricos descrevendo as conexões, e por consequência, estes valores descrevem também o comportamento da rede. A escolha do tipo de neurônio artificial utilizado é fundamental para se definir como será a representação interna dos conhecimentos da rede.

Cada unidade da rede é dotada de um estado interno, denominado estado de ativação. As unidades podem propagar seu estado de ativação para as outras unidades através das conexões existentes. A regra que determina a ativação de um neurônio em função das entradas, ponderadas pelos respectivos pesos, se chama regra de ativação ou função de ativação.

Na figura 3.8 pode ser visto um neurônio artificial típico. A entrada fixa é utilizada para referência, mas pode ser omitida.

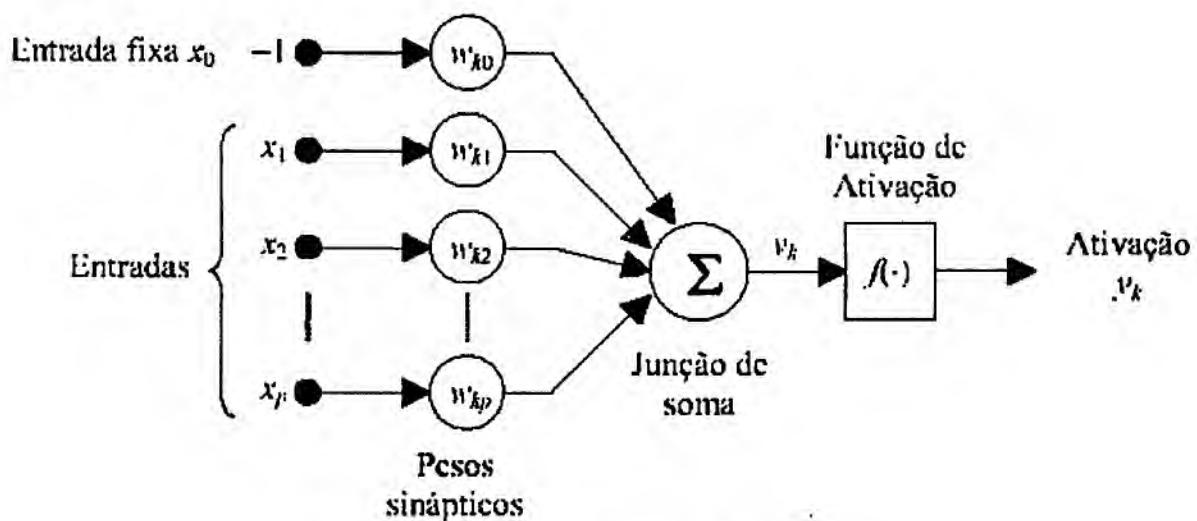


Figura 3.8: Modelo típico de um neurônio

Alterações nas conexões e na estrutura da rede é o que permite realizar o aprendizado de um novo comportamento. Desta maneira, pode-se modificar o estado de ativação na saída da rede em resposta a uma certa configuração de entradas. Portanto, a rede é capaz de estabelecer associações de entrada-saída (estímulo e resposta) para se adaptar a uma situação proposta. No caso de uma rede com aprendizado supervisionado, a rede deve adaptar os seus pesos de maneira que possa responder de acordo com o exemplo dado, ou seja, gerando na sua saída um estado de ativação compatível para com o esperado. O método utilizado para modificar o comportamento de uma rede é denominado de regra de aprendizado.

A forma de aprendizado é em geral um processo gradual e iterativo, onde os pesos são modificados várias vezes, pouco a pouco, seguindo-se uma regra de aprendizado que estabelece a forma como estes pesos são alterados. O aprendizado é realizado utilizando-se um conjunto de dados disponível para aprendizado (base de exemplos). Pode ser classificado como:

- ✓ **Aprendizado supervisionado:** o usuário dispõe de um comportamento de referência preciso que ele deseja ensinar para a rede. A rede deve ser capaz de medir a diferença entre seu comportamento atual e o comportamento de referência e corrigir os pesos de maneira a reduzir este erro (desvio de comportamento em relação aos exemplos de referência). Nos casos de problemas de classificação, a saída é a classe à qual cada exemplo está associado.
- ✓ **Aprendizado não-supervisionado:** os pesos da rede são modificados em função de critérios internos, tais como, por exemplo, a repetição de padrões de ativação em paralelo de vários neurônios. O comportamento resultante deste tipo de aprendizado é usualmente comparado com técnicas de análise de dados empregadas na estatística (*clustering*).

A maneira pela qual a rede é alimentada pelos dados pode ser:

- ✓ **Aprendizado por pacotes:** o conjunto de dados de aprendizado é apresentado à rede várias vezes de modo que é possível otimizar a resposta da rede, reduzindo os erros e minimizando o erro obtido na sua saída. Este modo de aprendizado é caracterizado por trabalhar com uma alteração dos pesos para cada época, ou seja, para cada passagem completa de todos os exemplos da base de aprendizado.
- ✓ **Aprendizado contínuo:** o algoritmo de aprendizado leva em consideração continuamente os exemplos que lhe são repassados. Se o conjunto de dados é bem delimitado, denomina-se este método de aprendizado on-line. Uma rede com uma grande facilidade de adaptação pode “esquecer” rapidamente os conhecimentos anteriormente adquiridos e uma rede com uma grande estabilidade pode ser incapaz de incorporar novos conhecimentos.

A adaptação/otimização dos pesos também pode ser implementada por diferentes métodos, segundo o tipo de regra de aprendizado que for empregado:

- ✓ **Métodos de correção do erro:** Exemplos de modelos deste tipo são: *Adaline*, *Perceptron*, *Back-Propagation*, *Cascade-Correlation*;
- ✓ **Métodos por reforço:** Exemplos: *Driver-Reinforcement Learning*;
- ✓ **Métodos por competição ou por auto-organização:** Exemplos: *Kohonen Self-Organizing Feature Maps (SOM)*, *Adaptive Resonance Theory -ART*;
- ✓ **Métodos através da criação de protótipos ou clusters:** Exemplos: ART1, ART2; ·

- ✓ **Métodos baseados em memórias associativas:** (auto-associativas ou hetero-associativas). Exemplos: Modelo de Hopfield;
- ✓ **Métodos de aprendizado de seqüências temporais (redes recorrentes):** Exemplos: BPTT (*back propagation temporal*).

Neste trabalho são apresentados os métodos de correção de erro, por auto-organização e por criação de protótipos.

Outro elemento importante diz respeito a forma como os neurônios armazenam as informações: unidades baseadas em protótipos e unidades do tipo *Perceptron*.

- ✓ **Redes à base de protótipos:** utiliza neurônios que servem para representar protótipos dos exemplos aprendidos. As unidades têm uma representação interna que agrupa as características comuns e típicas de um grupo de exemplos. As redes baseadas em protótipos têm normalmente um aprendizado não supervisionado (com um ou mais protótipos associados a cada classe). Uma das vantagens deste tipo de rede é a possibilidade de fazer um aprendizado contínuo e incremental, uma vez que não é muito difícil de conceber um algoritmo capaz de aumentar a rede neural através da adição de novos protótipos. Os protótipos são também denominados de *clusters*;
- ✓ **Redes à base de Perceptrons:** as unidades do tipo *Perceptron* foram criadas por Frank Rosenblatt em 1950. Este é um dos modelos de neurônios mais utilizados na atualidade. Ele é a base de diversos tipos de RNA com aprendizado supervisionado utilizando uma adaptação por correção de erros (usualmente baseada na descida da superfície de erro usando o gradiente). O modelo do Perceptron de múltiplas camadas (MLP – *Multi-Layer Perceptron*) tornou-se muito conhecido e aplicado, sendo na maior parte das vezes associado a regra de aprendizado *BackPropagation*, apresentada no próximo capítulo, na descrição do sistema implementado.

Vários autores estudaram a segmentação de imagens usando redes neurais do tipo MLP de três camadas, onde o número de neurônios na camada de entrada depende do número de características de entrada para cada pixel e o número de neurônios na camada de saída é igual ao número de classes. Em uma rede multi-camada para binarizar imagens, a entrada da rede pode ser o histograma, enquanto a saída seria o limiar desejado. Para a fase de treinamento, é preciso um conjunto grande de amostras de imagens com limiares

conhecidos que produzam saídas adequadas. Na prática, é muito difícil obter um número grande de amostras de imagens com limiares conhecidos [JES99].

A abordagem neural também é aplicada na segmentação de texturas e na segmentação de imagens coloridas. Yin et al. [YIN00] utilizaram um tipo variante de rede neural com três camadas para o problema de segmentação em cenas naturais, cuja entrada são os valores RGB dos pixels das imagens. Moreira e Costa [MOR96] aplicaram uma rede neural do tipo mapas auto-organizáveis (SOM) na segmentação de imagens coloridas. Vehtari & Lampinen (2000) utilizaram uma rede do tipo MLP para a segmentação e a classificação de imagens de uma floresta. Da mesma forma, Andrade e Martins [AND01], testaram diferentes arquiteturas no processamento de imagens de satélite. Para a classificação supervisionada de imagens de sensoriamento remoto foram também utilizadas combinações de redes neurais e algoritmos estatísticos [GIA00].

Redes neurais probabilísticas (PNN) foram utilizadas para análise de imagens de sementes de milho visando a sua classificação [STE99]. Na segmentação de imagens em profundidade foi feita a aplicação de redes neurais oscilatórias [LIU99b] e de uma MLP [VAR99].

Na determinação de parâmetros para segmentação e identificação de caminhos em sistemas agrícolas autônomos, foi utilizada uma rede MLP em Pinto et al. [PIN99]. Na aplicação em tempo real de herbicidas através da identificação de texturas nas imagens de cobertura do solo foi utilizada uma rede MLP de três camadas [TAN99]. A textura e a cor foram utilizadas numa rede SOM para segmentar imagens de cenas naturais [CAM95a], [CAM95b], [CAM96].

3.2.7 Algoritmos Genéticos (AG)

Os algoritmos genéticos são inspirados na teoria da evolução de Darwin. Eles pertencem à classe dos métodos probabilísticos de busca e otimização, embora não sejam aleatórios. Utiliza-se o conceito de probabilidade, mas os AG's não são buscas simples. Pelo contrário, eles tentam dirigir as buscas para regiões do espaço onde é provável que os pontos ótimos estejam.

De modo geral, os AG's têm as seguintes características [SON99]:

- ✓ Operam em uma população (conjunto) de pontos e não a partir de um ponto isolado;

- ✓ Operam em um espaço de soluções codificadas e não em um espaço de busca diretamente;
- ✓ Necessitam somente de informação sobre o valor de uma função objetivo para cada membro da população e não requerem derivadas ou qualquer outro tipo de conhecimento;
- ✓ Utilizam transições probabilísticas e não regras determinísticas.

Um dos problemas dos AG's é a determinação de parâmetros. No algoritmo básico, o usuário deve definir o tamanho da população, além das probabilidades de recombinação (um processo aleatório que ocorre com probabilidade fixa) e mutação (equivalente à busca aleatória). Em AG's mais sofisticados, há ainda mais parâmetros, comprometendo parte da robustez dos algoritmos. Infelizmente, não há regras claras para a escolha desses parâmetros.

3.2.8 Modelos de Contornos Ativos – *Snakes*

Em 1987, divulgou-se para a comunidade científica o modelo de contornos ativos. O principal objetivo era uma representação de contornos da imagem que permitisse a interação com processos de alto nível. Foi introduzido um modelo baseado em funções energéticas, cuja minimização traria um conjunto de soluções possíveis devido aos vários mínimos locais encontrados. Os processos de alto nível são utilizados para escolher uma solução dentro de um conjunto encontrado. Esses processos são introduzidos através da adição de novos termos energéticos ao modelo, fazendo com que este se adapte às características das imagens desejadas. Durante o processo de minimização, o contorno ativo tem um comportamento parecido com uma cobra rastejando, por isso a denominação de *Snake*.

A técnica *Snake* possibilita avaliar os dados da imagem, considerando uma estimativa inicial do contorno e propriedades do contorno desejado. O contorno ativo é modelado diretamente como uma curva, mantendo-se como tal durante todo o processo iterativo. As características da curva (comprimento, curvatura e adequação) podem ser modificadas e calculadas no algoritmo. A solução ótima baseia-se na minimização de energia do contorno ativo com base nas características do contorno ativo e da imagem. A técnica pode ser utilizada em aplicações 2D e estendida para aplicações 3D.

Existem algumas desvantagens. Como o processo é iterativo, ao encontrar um mínimo local, o processo toma essa solução não ótima como ótima, não determinando o mínimo

pretendido. O modelo dos contornos ativos precisa de uma iniciação próxima do contorno a extrair, não resolvendo de forma independente o problema, podendo ficar obstruído em mínimos locais durante a minimização da sua energia. O contorno ativo deriva na forma de curvatura mínima quando não estão presentes forças exteriores: em um ponto, no caso de um contorno ativo fechado e em uma linha, no caso de um contorno ativo aberto. É necessário englobar no modelo forças externas que influenciem o contorno ativo para que haja atração para as características da imagem que se pretende extraír.

Desde que foi introduzido o modelo dos contornos ativos, várias foram as aplicações do método e também vários os tipos de imagens utilizadas. O maior domínio de aplicação é o campo da medicina, sendo a cardiovascular a preferida. Na literatura, encontram-se vários trabalhos que aplicam o modelo dos contornos ativos para extraír/segmentar as características desejadas [SON99].

3.2.9 Métodos Baseados na Teoria *Fuzzy*

O conceito de conjuntos *fuzzy* é utilizado onde o conhecimento impreciso pode ser usado para definir um evento. A relevância da teoria dos conjuntos *fuzzy*, em problemas de reconhecimento de padrões, tem sido adequadamente tratada na literatura. Em vários estágios de um sistema de reconhecimento de padrões, a teoria dos conjuntos *fuzzy* pode ser incorporada para manipular incertezas. Estas incertezas originam-se de deficiências de informação, a qual pode ser incompleta, imprecisa, vaga, contraditória e mal definida. A incerteza em um padrão de imagem pode ser interpretada em termos de ambigüidade do conceito de nível de cinza ou ambigüidade espacial (geométrica) ou ambas. A ambigüidade de níveis de cinza significa a incerteza na decisão se um pixel é branco ou preto. A ambigüidade geométrica refere-se à incerteza do formato e geometria de uma região dentro da imagem.

As técnicas convencionais de análise e reconhecimento de imagens consistem na segmentação da imagem em regiões significativas, extraíndo suas bordas e esqueletos, processando várias características/propriedades (por exemplo: área, perímetro, centróide, etc.) e primitivas (por exemplo: linha, curva, cantos), e o relacionamento entre as regiões. Finalmente, são desenvolvidas regras/gramáticas de decisão para descrever, interpretar e/ou classificar a imagem e suas sub-regiões. No sistema convencional, cada uma dessas operações envolve decisões claras (isto é, sim ou não, branco ou preto, 0 ou 1) sobre regiões, características, primitivas, propriedades, relacionamentos e interpretações. Como as regiões de uma imagem nem sempre são definidas claramente, incertezas podem surgir

em todas as tarefas citadas acima. Uma decisão em um nível particular terá um impacto em todas as atividades dos níveis mais altos. Por exemplo, o problema de extrair objetos em uma cena. A questão é: “Como definir exatamente o objeto ou a região do objeto na cena quando sua borda está mal definida?” Uma binarização para a extração do objeto propagará a incerteza associada para estágios subsequentes (por exemplo, afinamento, extração do esqueleto, seleção de primitivas) e isto pode, sucessivamente, afetar a análise e reconhecimento de características. Portanto, é conveniente permitir que segmentos ou contornos ou esqueletos sejam subconjuntos *fuzzy* da imagem; os subconjuntos são caracterizados pela possibilidade (grau) de cada pixel pertencer a eles.

Similarmente, para descrever e interpretar informações estruturais mal definidas em um padrão é natural definir linha, ângulo, curva, etc. e o relacionamento entre eles, usando rótulos de conjuntos *fuzzy*. A produção de regras de uma gramática pode similarmente ser nebulosa para calcular a incerteza no relacionamento entre as primitivas. Deste modo, aumenta o poder produtivo de uma gramática para reconhecimento sintático de um padrão.

Na literatura, encontram-se muitos métodos de segmentação *fuzzy*, alguns deles são: binarização *fuzzy*, agrupamento *fuzzy* e detecção de bordas *fuzzy*. O Algoritmo de Agrupamento *Fuzzy* também foi bastante usado na segmentação de imagens [SON99]. Esse algoritmo usa uma otimização iterativa de uma função objetivo baseada na medida de similaridade ponderada entre os pixels da imagem e cada um dos centros dos agrupamentos. Um extremo local desta função objetivo indica um agrupamento ótimo dos dados de entrada.

3.2.10 Segmentação por Análise Multivariada

A segmentação da imagem, de tal forma fazer a extração automática de características e classificação dos pixels também pode ser feita pela análise multivariada.

A análise multivariada é uma análise estatística responsável por examinar como variáveis dependentes relacionam-se dentro de um determinado conjunto de dados. Estes métodos de análise têm como objetivo a redução e simplificação da complexidade do problema. Na prática, esses métodos podem reduzir o número de variáveis a serem computadas ou eliminar aquelas difíceis de serem observadas ou medidas, garantindo, ainda assim, que informações relevantes não sejam perdidas.

Dentre os objetivos da análise multivariada pode-se destacar: o agrupamento de dados, ou seja, como selecionar variáveis similares a partir de medidas de características; a

dependência entre variáveis, avaliando se são mutuamente independentes ou se existem uma ou mais variáveis que dependem de outras; e a previsão, determinando o relacionamento entre as varáveis com a finalidade de prever os valores de uma ou mais variáveis de interesse.

Os métodos utilizados para neste trabalho são:

- ✓ Transformada de Hotelling;
- ✓ Discriminante Linear de Mahalanobis.

Em todos estes métodos deseja-se fazer a extração de características, processo através do qual um espaço de dados é transformado em um espaço de características que, teoricamente, tem exatamente a mesma dimensão do espaço de dados original. Todavia, a transformação pode ser projetada de modo que o conjunto de dados possa ser representado por um número reduzido de características relevantes.

✓ **Transformada de “Hotelling”**

A Transformada de “Hotelling” (TH), também conhecida por Transformada de “Karhunen-Loëve”, ou ainda Transformada das Componentes Principais [PRA91] baseia-se nas propriedades estatísticas de um conjunto de dados, neste trabalho, no conjunto de cores da imagem.

Esta transformação introduzida por Hotelling [HOT33], faz a análise de componentes principais, uma das técnicas mais conhecidas de análise multivariada. É um dos métodos estatísticos multivariados mais simples.

A análise das componentes principais, como o próprio nome sugere, toma p variáveis e procura por combinações lineares dessas que sejam não correlacionadas. A falta de correlação é uma propriedade útil, pois significa que os índices estão medindo diferentes “dimensões” dos dados.

Ao realizar uma análise de componentes principais, espera-se sempre que as variâncias da maioria dos índices sejam tão pequenas que possam ser desprezadas. Nesse caso, a variação do conjunto de dados pode ser descrita adequadamente pelas poucas variáveis cujas variâncias são consideradas não desprezíveis. Assim, consegue-se reduzir o número de variáveis.

Considere um conjunto de M vetores aleatórios da forma:

$$\mathbf{x}_i = [x_{i1} \ x_{i2} \ \dots \ x_{in}]^T \quad (3.1)$$

O vetor médio \mathbf{m}_x do conjunto de M vetores pode ser descrito por:

$$\mathbf{m}_x = E\{\mathbf{x}\} \cong \frac{1}{M} \sum_{i=1}^M \mathbf{x}_i \quad (3.2)$$

Sendo $E\{\text{variável}\}$ o valor esperado da variável.

A matriz de covariância C_x desta população de vetores será dada por:

$$C_x = \begin{bmatrix} c_{11} & c_{12} & \dots & c_{1n} \\ c_{21} & c_{22} & \dots & c_{2n} \\ \vdots & \vdots & & \vdots \\ c_{n1} & c_{n2} & \dots & c_{nn} \end{bmatrix} \quad (3.3)$$

Onde o elemento c_{jk} de C_x é dado por:

$$c_{jk} = E[(x_j - m_j)(x_k - m_k)^T] \quad (3.4)$$

Sendo x_j , x_k e m_j , m_k o j -ésimo e o k -ésimo componentes de \mathbf{x}_i e \mathbf{m}_x , respectivamente.

Então a matriz de covariância desta população de vetores será dada por:

$$C_x = E[(\mathbf{x} - \mathbf{m})(\mathbf{x} - \mathbf{m})^T] \quad (3.5)$$

Sendo \mathbf{x} um vetor n-dimensional, a matriz de covariância C_x tem dimensão $n \times n$.

Se \mathbf{x} é um vetor real e $x_i x_j = x_j x_i$ a matriz C_x será real e simétrica ($C_{ij} = C_{ji}$). Logo a população de vetores terá seus elementos x_i e x_j descorrelacionados se os elementos $C_{ij} = C_{ji}$ forem iguais a zero.

Expandindo a equação 3.5 tem-se:

$$C_x = E\{\mathbf{x}\mathbf{x}^T - \mathbf{x}\mathbf{m}_x^T - \mathbf{m}_x\mathbf{x}^T + \mathbf{m}_x\mathbf{m}_x^T\} = E\{\mathbf{x}\mathbf{x}^T\} - \mathbf{m}_x\mathbf{m}_x^T \quad (3.6)$$

Logo:

$$C_x = \frac{1}{M} \sum_{i=1}^M \mathbf{x}_i \mathbf{x}_i^T - \mathbf{m}_x \mathbf{m}_x^T \quad (3.7)$$

Como C_x é real e simétrica, pode-se encontrar um conjunto de autovetores ortonormais e dispô-los nas linhas de uma matriz na ordem decrescente dos autovalores λ a que cada autovetor está associado, ou seja, $\lambda_i \geq \lambda_{i+1}$, onde $i=1, 2, \dots, n-1$ é o índice da linha da matriz.

A transformação da forma:

$$y = A(x - m_x) \quad (3.8)$$

é conhecida como Transformada de “Hotelling” (TH), onde A é a matriz de autovetores citada anteriormente.

O novo sistema resultante da TH com três novas componentes HOT pode ser obtido por:

$$\begin{bmatrix} H \\ O \\ T \end{bmatrix} = \begin{bmatrix} E_{11} & E_{12} & E_{13} \\ E_{21} & E_{22} & E_{23} \\ E_{31} & E_{32} & E_{33} \end{bmatrix} \cdot \begin{bmatrix} R - m_R \\ G - m_G \\ B - m_B \end{bmatrix} \quad (3.9)$$

Onde $[E_{11}E_{12}E_{13}]$, $[E_{21}E_{22}E_{23}]$ e $[E_{31}E_{32}E_{33}]$ são os autovetores da matriz de covariância.

Propriedades da Transformada de “Hotelling”:

A partir da equação 3.8 pode-se verificar que:

$$E\{y\} = 0 \quad (3.10)$$

$$C_y = AC_x A^T \quad (3.11)$$

$$C_y = \begin{bmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ 0 & & \ddots & \\ & & & \lambda_n \end{bmatrix} \quad (3.12)$$

Uma característica importante da TH é o fato das componentes associadas aos maiores autovalores conterem a maior parte da informação presente no conjunto de M vetores.

Alguns sistemas de cores possuem suas componentes totalmente correlacionadas, como é o caso do sistema RGB, mostrando-se pouco indicados quando se deseja fazer a segmentação de cores que se torna muito complicada.

A TH pode ser utilizada para descorrelacionar as componentes dos sistemas estudados. As novas componentes H, O e T são descorrelacionadas e apresentam as características presentes na imagem. Pode-se utilizar as componentes de maior variância nos algoritmos de segmentação, reduzindo a dimensão do sistema. Neste caso, a primeira componente, H, é a mais representativa, uma vez que é aquela formada pelo maior autovalor.

✓ Análise Discriminante Linear

Este tipo de análise visa separar dois ou mais grupos de indivíduos tendo como dados de entrada as características desses indivíduos. É estabelecida uma função discriminante capaz de revelar as diferenças entre os grupos.

Dentre os vários discriminantes existentes destaca-se uma abordagem para discriminação baseada na distância de Mahalanobis. Nessa abordagem, os vetores média para as m amostras podem ser considerados como estimativas dos verdadeiros vetores média para as classes. As distâncias de Mahalanobis entre os indivíduos e os centros de cada classe podem então ser calculados e cada indivíduo pode ser alocado na classe cuja distância for a menor. Essa classe pode ou não ser a classe à qual o indivíduo realmente pertence. A porcentagem de alocações corretas é uma indicação de quão bem as classes podem ser separadas levando-se em conta as variáveis disponíveis.

O procedimento é definido como segue. Sejam $x_{im} = (x_{1m}, x_{2m}, \dots, x_{pm})$ o vetor média para as observações do i -ésimo grupo, C_i a matriz de covariância para a mesma amostra e C a matriz de covariância para o total de amostras. A distância de Mahalanobis de uma observação $x_i = (x_1, x_2, \dots, x_p)$ à média (centro) do grupo i é estimada como:

$$D_i^2 = (x - \bar{x}_i)^T C_i^{-1} (x - \bar{x}_i) = \sum_{r=1}^p \sum_{s=1}^p (x_r - \bar{x}_{ri}) C^{rs} (x_s - \bar{x}_{si}) \quad (3.13)$$

onde C^{rs} é o elemento da r -ésima linha e s -ésima coluna de C^T . A observação x é alocada no grupo para o qual D_i^2 tem o menor valor.

Normalmente, este discriminante é indicado para resolver problemas causados por:

- ✓ Características mal dimensionadas;
- ✓ Características altamente correlacionadas.

É importante não confundir esta análise discriminante com análise de *cluster*. A primeira requer conhecimento *a priori* sobre as classes, geralmente na forma de uma amostra para cada classe. Já na segunda, os dados disponíveis não incluem informações sobre os membros das classes. Classificação e *Clustering* são, portanto, métodos distintos. O objetivo operacional da classificação é, dadas m classes, sendo m conhecido, associar cada nova observação a uma dessas classes. Já a análise de *Cluster* é uma técnica na qual não é feita nenhuma suposição envolvendo o número ou estrutura das classes. Os agrupamentos são feitos com base em similaridades ou em distâncias. As entradas para a

análise por *cluster* são, portanto, medidas de similaridade ou dados a partir dos quais possam ser computadas tais medidas. Nesse caso, a medida de similaridade mais conhecida e usada é a distância euclidiana.

3.4 Considerações Finais

Diante da revisão dos métodos de segmentação de imagens na literatura, pode-se observar que existem muitos métodos, desde os mais simples até os mais complexos utilizando técnicas não convencionais. A maioria, porém, não pode ter seu uso generalizado para qualquer aplicação.

Por outro lado, avaliar a cobertura do solo por imagens aéreas não é uma tarefa simples, principalmente pelo tamanho das áreas envolvidas e pelas características das imagens. São imagens com variações de luz, sem muita padronização em termos de intensidade e cor. Também são necessárias avaliações ao longo do tempo, demandando uma grande quantidade de dados. Dentre os trabalhos similares que apresentam imagens com todas estas variações, destacam-se os que utilizam as técnicas estatísticas e as não convencionais.

Na avaliação da cobertura do solo a cor é uma das características mais importantes a serem observadas. Pela revisão feita na literatura fica evidente que o sistema de representação de cores é relevante. Utilizar modelos baseados na percepção humana torna mais fácil a automatização do processo. A correlação das componentes de cores também é outro fator determinante para a obtenção de bons resultados, justificando a grande utilização da Transformada de *Hotelling* na descorrelação.

A utilização de uma técnica isolada pode resultar em uma boa segmentação, no entanto, fica evidente que testar mais de uma técnica é necessário antes da escolha da mais adequada. Em alguns casos, torna-se imprescindível a combinação de mais de uma técnica.

Com a utilização de técnicas não convencionais como as Redes Neurais Artificiais e a Lógica *Fuzzy*, dentre outras, têm-se conseguido romper a dificuldade de definir regras rígidas ou equações matemáticas complexas para a segmentação. Também, com o avanço de sistemas dedicados de hardware, como os circuitos integrados neurais, possibilita-se o aumento da velocidade de processamento, ainda limitante em muitas aplicações.

No estudo de cobertura do solo em sistemas embarcados nas aeronaves de fase III do projeto ARARA, visando a Agricultura de Precisão, é necessária uma maior rapidez na

segmentação das imagens. Avaliar técnicas não convencionais para esta aplicação se torna imprescindível.

No próximo capítulo são apresentadas as implementações dos métodos selecionados, utilizando a cor como característica discriminante nas imagens de cobertura do solo, e os testes realizados para a sua validação.

Capítulo 4

Teste e Implementação

4.1 Considerações Iniciais

No capítulo anterior foi feita uma revisão das principais técnicas existentes para segmentação e classificação de imagens. Neste capítulo são apresentadas as técnicas implementadas e testadas para a avaliação da cobertura do solo em imagens aéreas. Foram utilizadas técnicas supervisionadas e não supervisionadas.

Na seção 4.2 é especificado o sistema e apresentado o diagrama de fluxo de dados proposto para este trabalho, bem como o detalhamento dos processos do diagrama de fluxo de dados e as respectivas implementações e testes.

O sistema desenvolvido deverá ser futuramente integrado no sistema SIARCS, apresentado no capítulo 2, com o intuito de torná-lo automático.

4.2 Especificação do Sistema

Este trabalho tem por objetivo apresentar uma solução para o SIARCS no estudo automático de cobertura vegetal. Na versão anterior, foi utilizada uma ferramenta para segmentação manual através de limiarização utilizando a cor como discriminante.

As imagens de cobertura são imagens complexas, com padrões difíceis de serem definidos. O uso da cor como único atributo para diferenciar o solo, plantas ou resíduos pode ser bastante limitante em alguns casos. Entretanto, um dos objetivos deste trabalho é o de comparar os resultados da técnica manual com a automática, utilizando o mesmo atributo.

Os padrões variam para cada tipo de solo, dependem das condições de iluminação ambiente, da resolução da imagem, do tipo de planta e resíduos sobre o solo, dentre outros fatores. Apesar disso, a informação espacial, ou seja, a textura não foi considerada. Para

minimizar possíveis erros optou-se por um pré-processamento na imagem, de tal forma tornar as regiões mais homogêneas e os padrões de textura na imagem, mais suaves.

4.2.1 Diagrama de Fluxo de Dados

O diagrama de fluxo de dados apresentado na figura 4.1 reflete as alterações feitas no diagrama de fluxo de dados de nível 1 do SIARCS e é constituído dos seguintes processos:

- ✓ **Pré-processar a imagem:** Este processo é responsável pela filtragem da imagem com um filtro de suavização da imagem de tal forma minimizar os efeitos de ruídos e texturas;
- ✓ **Selecionar Classes:** Esse processo é responsável pela definição e seleção dos padrões de cada classe a ser utilizada na classificação supervisionada;
- ✓ **Calcular Transformada de Hotelling:** Esse processo é utilizado para calcular a Transformada de Hotelling ou Transformada das Componentes Principais, utilizada no classificador estatístico e nas redes neurais;
- ✓ **Classificar com Mahalanobis:** Esse processo é utilizado como discriminante linear no classificador estatístico com a distância de Mahalanobis;
- ✓ **Treinar Rede MLP:** Esse processo é utilizado de duas formas: como um método supervisionado para a segmentação da imagem e como classificador para o mapa de Kohonen;
- ✓ **Classificar com MLP:** Esse processo implementa a fase de teste do método supervisionado de classificação dos pixels através da rede Perceptron Multicamadas
- ✓ **Treinar Rede de Kohonen:** Esse processo implementa o método não supervisionado proposto por Kohonen, cuja função é encontrar os possíveis agrupamentos pelas características discriminantes (as componentes do sistema de cor, neste trabalho). É a fase de treinamento da rede, onde todos os pesos são ajustados;
- ✓ **Gerar Mapa Auto-Organizado:** Esse processo é parte da Rede de Kohonen, porém é apenas a fase de classificação depois dos ajustes dos pesos. É responsável pela geração dos mapas de agrupamentos dos pixels com características semelhantes, também conhecidos como mapas de Kohonen. Gera uma matriz de saída com os pixels agrupados segundo suas características;

- ✓ **Classificar com ART2:** Esse processo implementa a rede ART2, cuja entrada é a imagem e a saída é a imagem segmentada. É um método totalmente não supervisionado.

Após a abertura da imagem, pode-se utilizar um filtro de realce do tipo média para suavizar as regiões da imagem, resultando assim, numa melhor segmentação. Este filtro de média é um filtro convencional no domínio espacial por convolução com uma máscara que varia de 3x3, 5x5 a 7x7. Quanto maior a máscara maior a suavização da imagem e, portanto, maior a perda de nitidez. Exemplos deste filtro serão apresentados no capítulo de resultados.

Com a imagem processada calcula-se a Transformada de Hotelling. Como resultado desta transformada obtém-se um novo sistema de coordenadas de cores, aqui denominado de HOT, sendo H a primeira componente, de valor mais significativo, O a segunda componente e T a terceira. Com a matriz HOT é possível aplicar o classificador estatístico e as redes neurais utilizando componentes não correlacionadas entre si.

Para as técnicas supervisionadas, o classificador estatístico e a rede MLP, foi implementada uma ferramenta interativa para seleção de padrões sobre a imagem para o treinamento *a priori*. No estudo de cobertura, restringiu-se às seguintes classes de treinamento: solo nu, cobertura vegetal e resíduos. Em alguns casos diferenciam-se os tipos de cobertura vegetal, por exemplo, plantas rasteiras de árvores.

Através do discriminante linear de Mahalanobis segmenta-se a imagem através da classificação dos pixels. Este método é automático, apesar de supervisionado. Os padrões devem ser selecionados apenas uma vez. Depois disso, a análise para outras imagens é feita sem proceder ao treinamento novamente, a não ser que, os padrões de cobertura ou as áreas sob análise sejam outras.

Ainda de forma supervisionada pode-se utilizar a rede MLP. A entrada para esta rede são os valores RGB ou os valores HOT, dependendo do sistema utilizado. Da mesma forma, seleciona-se padrões que são usados para a fase de treinamento da rede. Ou seja, treina-se a rede apresentando várias amostras dos respectivos padrões. Após o treinamento, apresenta-se a imagem toda à rede e esta procede com a fase de validação, segmentando a imagem.

De forma não supervisionada pode-se executar as rotinas para análise da rede ART2. Para a entrada RGB ou a imagem HOT, obtém-se diretamente na saída da rede os

resultados. Dependendo dos parâmetros de ajuste da rede esta rede fica mais sensível a variações aumentando muito o número de padrões encontrados. Depois de alguns testes, chega-se aos valores de ajuste como os quais a rede chega perto de três padrões de cobertura de solo.

De forma intermediária tem-se a Rede de Kohonen. Apesar de ser um modelo não supervisionado, esta rede gera na saída um mapa de agrupamento de pixels segundo suas próprias características. Este mapa auto-organizado deve então ser rotulado ou classificado. Várias são as opções de rotular o mapa, neste trabalho utilizou-se uma rede MLP, ou simplesmente pela definição manual de uma semente para cada classe.

O Diagrama de Fluxo de Dados da figura 4.1 detalha o fluxo a ser seguido na ferramenta implementada. A seguir, são detalhados os processos que constituem este diagrama.

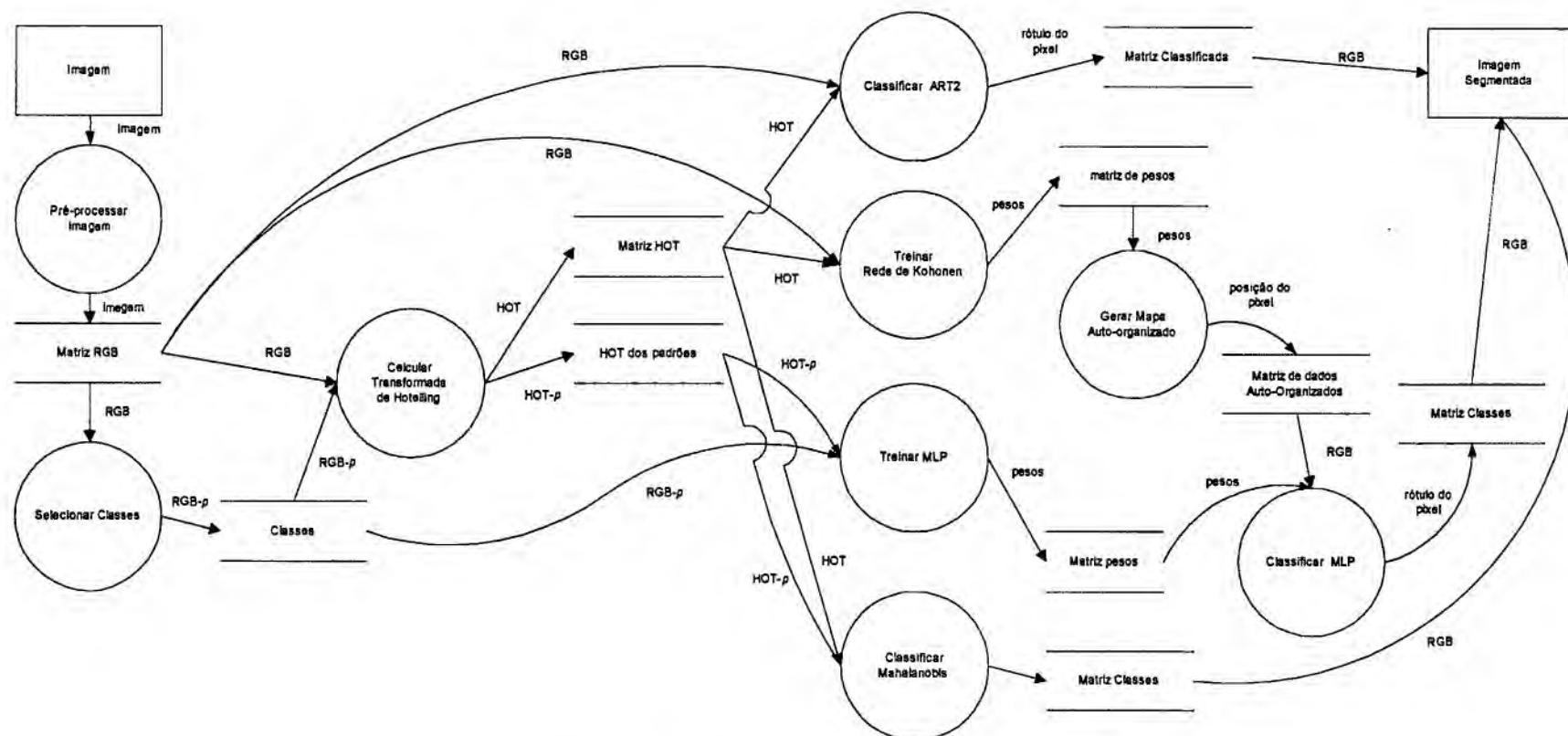


Figura 4.1: Diagrama de Fluxo de Dados

4.2.2 Pré-processar a imagem

Neste processo é realizado o processamento da imagem através de uma operação de convolução, aplicando-se uma máscara sobre a imagem. A máscara é uma pequena matriz de pesos, geralmente quadrada, que é deslocada sobre a imagem original, a fim de obter cada pixel da imagem convoluída. A operação de convolução, no caso discreto, é dada por:

$$y(m, n) = \sum_i \sum_j h(m - i, n - j)x(i, j) \quad (4.1)$$

onde x é a imagem original, representada como uma matriz de ordem $[M \times N]$, h é a máscara que desejamos aplicar sobre a imagem, representada como uma matriz de ordem $[I \times J]$ e y é a imagem resultante da convolução, que será uma matriz de ordem $[(I+M-1) \times (J+N-1)]$.

Na implementação da convolução, uma das principais preocupações foi possibilitar o usuário editar suas próprias máscaras, de forma que este não fique limitado a um pequeno número de máscaras preestabelecidas, apesar de ser utilizada apenas a máscara do filtro de média. Desta forma, foi criada uma janela onde o usuário pode abrir ou criar, editar e salvar suas próprias máscaras, conforme mostra a figura 4.2.

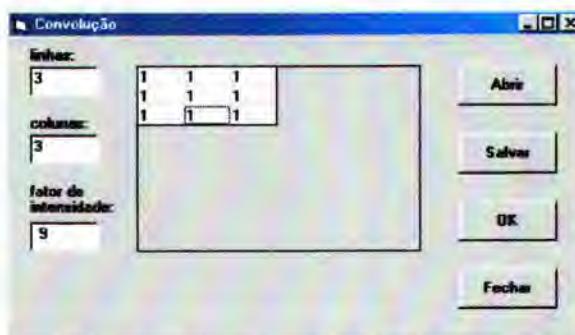


Figura 4.2: Interface para operação de filtragem da imagem

A aplicação de diferentes máscaras resultará em diferentes filtragens sobre uma imagem. O filtro utilizado para suavizar a imagem foi um filtro de média. Este filtro reduz os ruídos da imagem.

Na figura 4.3 pode ser visto o resultado da filtragem em uma das imagens de cobertura.



Figura 4.3: (a) Imagem original e (b) imagem suavizada.

4.2.3 Selecionar Classes

Os métodos supervisionados, como o classificador por Mahalanobis e a rede MLP, necessitam de uma fase de treinamento *a priori*, através da qual são selecionados as classes e os respectivos padrões para cada uma. Este processo permite a criação das classes e seleção dos padrões através de uma interface interativa, onde é possível selecionar amostras de padrões, com o *mouse* diretamente sobre a imagem. Para cada amostra associa-se uma classe. Para cada classe, pode-se selecionar mais de uma amostra padrão.

Na figura 4.4 pode ser vista a interface desenvolvida, aqui denominada de Agro Imagem v.01, onde já foram criadas as classes e selecionados os respectivos padrões. A cada uma associa-se uma cor que será utilizada no resultado final da segmentação.

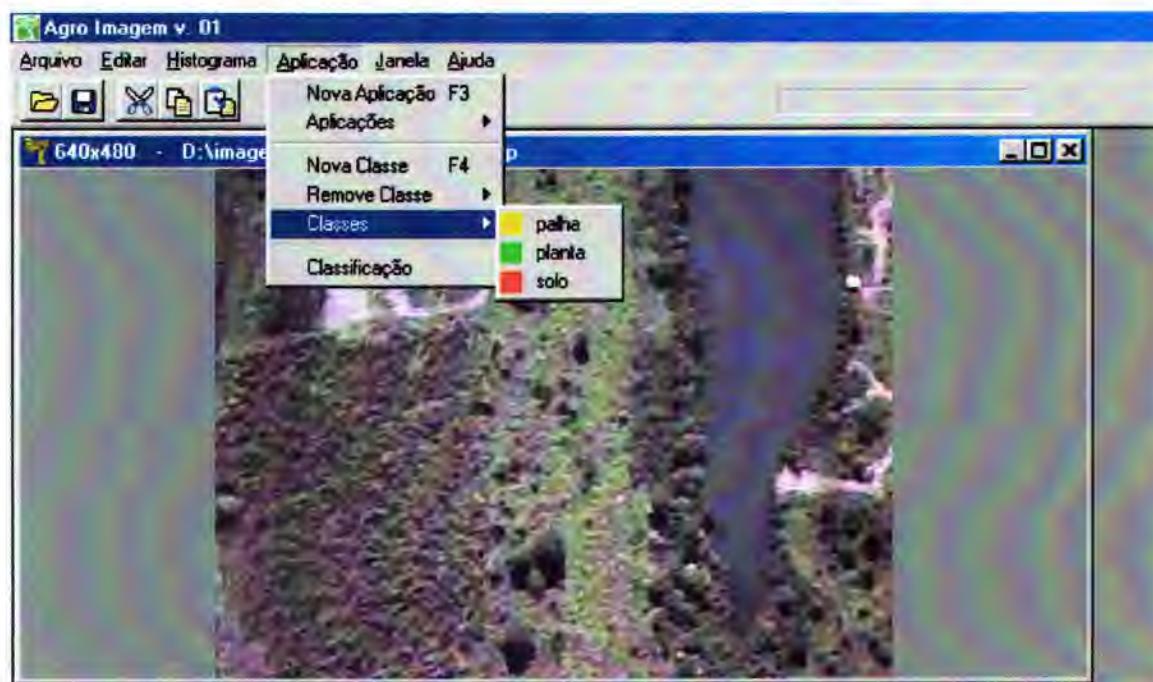


Figura 4.4: Interface para Seleção de Classes

Para cada amostra selecionada, calculam-se os valores médios de RGB e HOT. Estes são armazenados em um banco de dados para utilização posterior.

4.2.4 Calcular Transformada de Hotelling

Este processo é o responsável pela transformação da matriz RGB na matriz HOT. Ou seja, pelo cálculo das novas componentes não correlacionadas.

Executado logo no início quando uma imagem é aberta e quando as classes são selecionadas.

O cálculo segue a transformação descrita anteriormente quando apresentada a TH.

Para determinar a TH é encontrada a matriz de covariância como:

$$\mathbf{C}_{RGB} = \begin{bmatrix} C_{RR} & C_{RG} & C_{RB} \\ C_{GR} & C_{GG} & C_{GB} \\ C_{BR} & C_{BG} & C_{GG} \end{bmatrix} \quad (4.2)$$

onde C_{RR} representa a auto-covariância calculada a partir da componente R de cada pixel da imagem de acordo com a expressão:

$$C_{RR} = \frac{1}{N} \sum_{i=1}^N R_i^2 - \mu_R^2 \quad (4.3)$$

onde N é o número de pixels da imagem, R_i é o valor R de cada pixel and μ_R é o valor médio de todas componentes R da imagem. O valor C_{GG} é a auto-covariância, para a componente G, substituindo R por G na equação 4.3. O mesmo pode ser feito para o cálculo de C_{BB} , correspondendo à componente B.

As covariâncias entre R e G (C_{RG} e C_{GR}), tanto quanto entre R e B (C_{RB} and C_{BR}) e entre G e B (C_{GB} and C_{BG}), podem ser obtidas da mesma forma:

$$C_{xy} = \frac{1}{N} \sum_{i=1}^N X_i Y_i - \mu_x \mu_y \quad (4.4)$$

onde X e Y representam R, G ou B . Os valores μ_X e μ_Y representam os valores médios de X e Y, respectivamente.

Uma vez obtida a matriz de covariância, calculam-se os autovetores através do método de rotações sucessivas de Jacobi para uma matriz simétrica [PRE92]. Depois de encontrados os autovetores, as componentes HOT são encontradas como na expressão 3.9.

4.2.5 Classificar com Mahalanobis

Este processo simplesmente aplica o cálculo da distância de Mahalanobis, conforme equação 3.13. Uma vez calculados os valores médios de HOT para cada padrão de cada classe, calculam-se as distâncias destes a cada pixel. O pixel é classificado com a cor da classe cuja distância ao seu padrão for a menor.

Neste caso pode-se utilizar as três componentes principais ou apenas a primeira.

4.2.6 Treinar Rede MLP

Neste processo, executa-se a fase de treinamento de uma rede MLP com o algoritmo *backpropagation* (BP), ou de *retro-propagação de erro*, ou simplesmente *retro-propagação*, utilizado no treinamento de redes neurais multi-camadas com uma ou mais camadas escondidas. Devido a grande difusão da rede MLP com esta regra de aprendizagem, é comum referir-se a ela com o nome da própria regra de aprendizagem, ou seja, rede BP.

O algoritmo de retro-propagação consiste em dois passos de computação: o processamento direto e o processamento reverso. No processamento direto, uma entrada é aplicada à rede neural e seu efeito é propagado pela rede, camada a camada.

Durante o processamento direto, os pesos da rede permanecem fixos. No processamento reverso, um sinal de erro calculado na saída da rede é propagado no sentido reverso, camada a camada, e ao final deste processo os pesos são ajustados de acordo com uma regra de correção de erro. O treinamento da rede através de retro-propagação é realizado de forma supervisionada, ou seja, é apresentada à rede os padrões selecionados interativamente e também é disponibilizada a resposta desejada para cada padrão de entrada.

O algoritmo *backpropagation* foi criado a partir da generalização da regra de aprendizado *Widrow-Hoff*, que foi introduzida por Bernard Widrow & Marcian Hoff em 1960-1962 para redes do tipo *feedforward perceptron* [CAR98].

A regra de aprendizado *Widrow-Hoff* também é conhecida como “Regra Delta” (LMS - minimização do erro médio quadrático), que ajusta os pesos das conexões entre os neurônios da rede de acordo com o erro, ou seja, esta regra tem como objetivo encontrar um conjunto de pesos e polarizações que minimizem a função erro:

$$E = \frac{1}{2} \sum_{j=1}^R \sum_{i=1}^S (y_{d,j,i} - y_{j,i})^2 \quad (4.5)$$

onde R = número de padrões ou vetores de entrada; S = número de neurônios de saída (dimensão do vetor de saída); $y_{d,j,i}$ = saída desejada no i -ésimo neurônio, quando o j -ésimo padrão é apresentado; $y_{j,i}$ = saída obtida pela rede no i -ésimo neurônio, quando o j -ésimo padrão é apresentado.

A alteração dos pesos $W_{i,j}$ da regra de Widrow-Hoff é calculada da seguinte maneira:

$$\Delta W_{i,j} = -\eta \frac{\partial E}{\partial W_{i,j}} \quad (4.6)$$

onde, η = taxa de aprendizado e $\frac{\partial E}{\partial W_{i,j}}$ é a derivada parcial do erro em relação ao peso da respectiva conexão (gradiente).

O algoritmo BP utilizado para treinamento incremental é descrito pelos seguintes passos:

- ✓ **Passo 1:** Inicializar os pesos, as polarizações e os demais parâmetros de treinamento;
- ✓ **Passo 2:** Apresentar à rede um padrão de entrada do conjunto de treinamento e computar a sua saída; O conjunto de treinamento é a matriz que representa a janela da qual foi selecionada um padrão da classe;
- ✓ **Passo 3:** Calcular o erro para os neurônios da camada de saída, subtraindo a saída desejada da saída calculada;

Para um dado vetor de entrada o nível de ativação $v_j^{(l)}(n)$ para o j -ésimo neurônio na camada l é:

$$v_j^{(l)}(n) = \sum W_{ji}^{(l)}(n) y_i^{(l-1)}(n) \quad (4.7)$$

onde $y_i^{(l-1)}(n)$ é o sinal de saída do i -ésimo neurônio na camada $l-1$ na iteração n e $w_{ji}^{(l)}(n)$ é o peso do neurônio i na camada $l-1$ para o neurônio j na camada l .

Assumindo uma função sigmoidal como função de ativação definida por:

$$y(n) = \frac{1}{1+e^{-(vb)}} \quad (4.8)$$

onde b é responsável pela inclinação da sigmóide e v o nível de ativação.

Assim o erro é calculado por:

$$e(n) = (y_d(n) - y(n)) \quad (4.9)$$

onde: y_d é a saída desejada e y a saída real (saída gerada pela rede).

- ✓ **Passo 4:** Calcular os gradientes locais, ou deltas, para ajuste dos pesos. Retropropagar o erro para as camadas escondidas. Como não existe uma saída desejada para os neurônios das camadas escondidas, deve-se calcular o erro destes a partir do erro dos neurônios pertencentes à camada de saída e das conexões que os interligam.

$$\delta_j^{(L)}(n) = e_j^{(L)}(n) \cdot y_j(n) [1 - y_j(n)] \quad (4.10)$$

para o neurônio j na camada de saída L , ou ainda,

$$\delta_j^{(l)}(n) = y_j^{(l)}(n) [1 - y_j(n)] \sum \delta_k^{(l+1)}(n) w_{kj}^{(l+1)}(n) \quad (4.11)$$

para o neurônio j na camada intermediária l .

- ✓ **Passo 5:** Têm-se assim, a seguinte equação para calcular o ajuste dos pesos na camada l de acordo com a regra delta generalizada:

$$w_{jl}^{(l)}(n+1) = w_{jl}^{(l)}(n) + \alpha [w_{jl}^{(l)}(n) - w_{jl}^{(l)}(n-1)] + \eta \delta_j^{(l)}(n) y_j^{(l-1)}(n) \quad (4.12)$$

onde η é o parâmetro de velocidade de aprendizado e α é a constante momento.

Esta constante é responsável por aumentar a velocidade de aprendizado e evitar o perigo de instabilidade da rede.

- ✓ **Passo 6:** Calcular o erro acumulado da rede. Nesta etapa, deve ser verificado se o erro total sobre todos os padrões de entrada pode ser considerado desprezível, isto é, se caiu abaixo de um limiar de aceitação. Se assim for o caso, o algoritmo deve parar, caso contrário, deve-se voltar ao passo 2.

Múltiplas camadas de neurônios não-lineares permitem à rede BP aprender relações lineares e não-lineares existentes entre os valores de entrada e saída. Nesse caso, se a camada de saída for composta por neurônios lineares, a rede pode produzir valores fora do

intervalo $[0,1]$. Neste trabalho, restringiu-se a saída da rede a valores no intervalo $[0,1]$ fazendo a camada de saída também consistir de neurônios sigmoidais.

A Figura 4.5 apresenta a configuração da rede de 3 camadas (1 de entrada, 1 escondida e 1 de saída) utilizada no reconhecimento das três classes de cobertura do solo, onde a entrada é o valor RGB ou HOT.

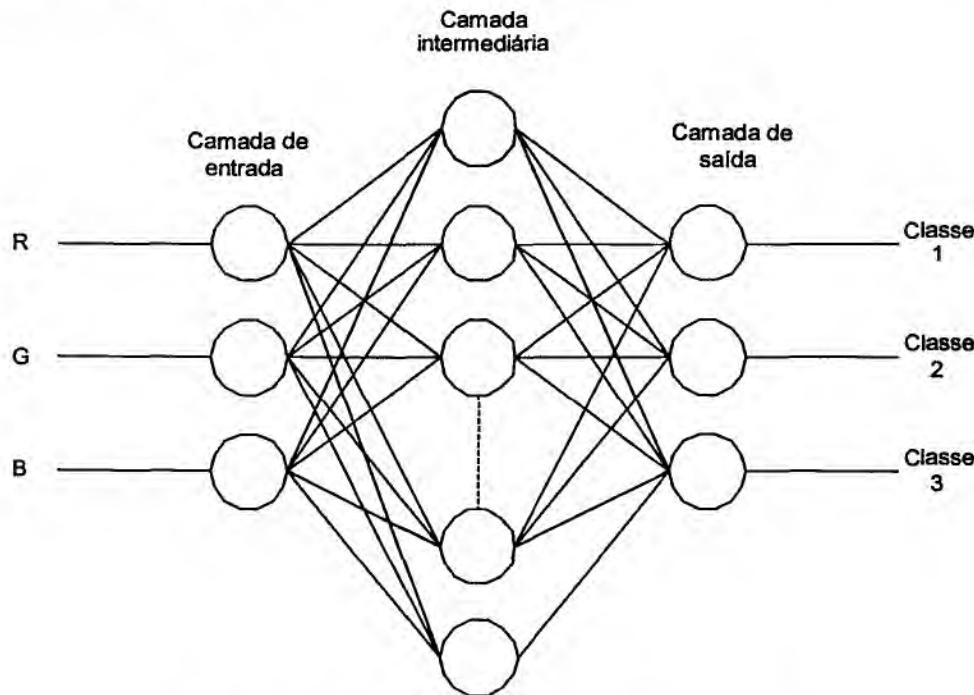


Figura 4.5: Topologia da rede MLP utilizada

A verificação da fase de treinamento da rede ocorre através da verificação do erro. Neste trabalho estudou-se a variação do erro variando-se η e α e a topologia da rede, com uma imagem teste formada por uma imagem de cobertura de tamanho reduzido.

Na figura 4.6 podem ser observados os resultados de erro para uma rede com duas camadas intermediárias e com diferentes valores de η e α . Neste caso, observa-se que com $\eta=0,7$ e $\alpha=0,3$ obtém-se a melhor convergência da rede.

Na figura 4.7 são apresentados os resultados da variação de erro da rede com uma camada intermediária e com diferentes valores de η e α . Neste caso, observa-se que com $\eta=0,9$ e $\alpha=0,1$ obtém-se a melhor convergência da rede. Além disto, como apenas uma camada os resultados foram melhores que com duas camadas intermediárias.

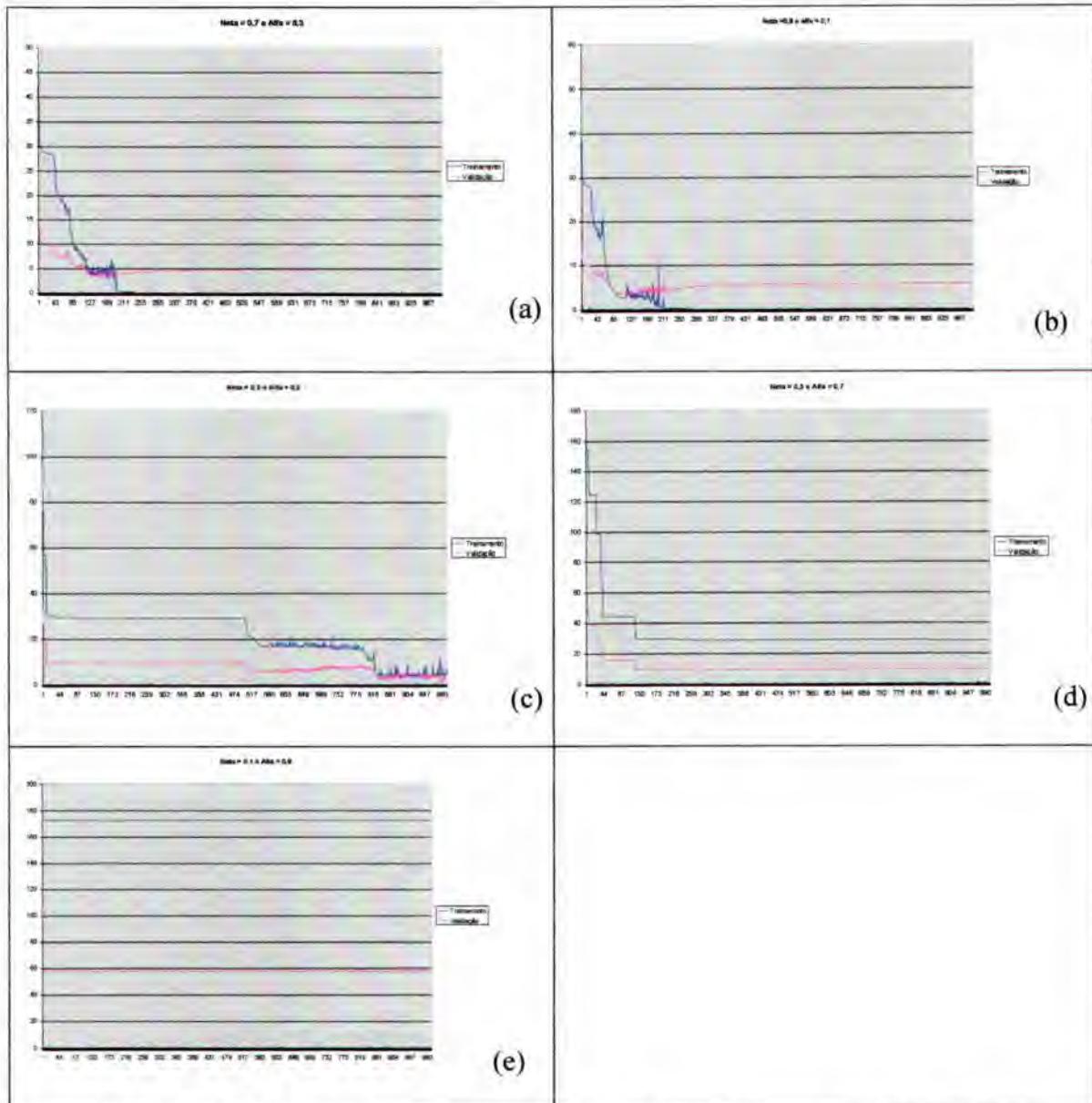


Figura 4.6: Variação de erro com diferentes valores de η e α , com duas camadas intermediárias.

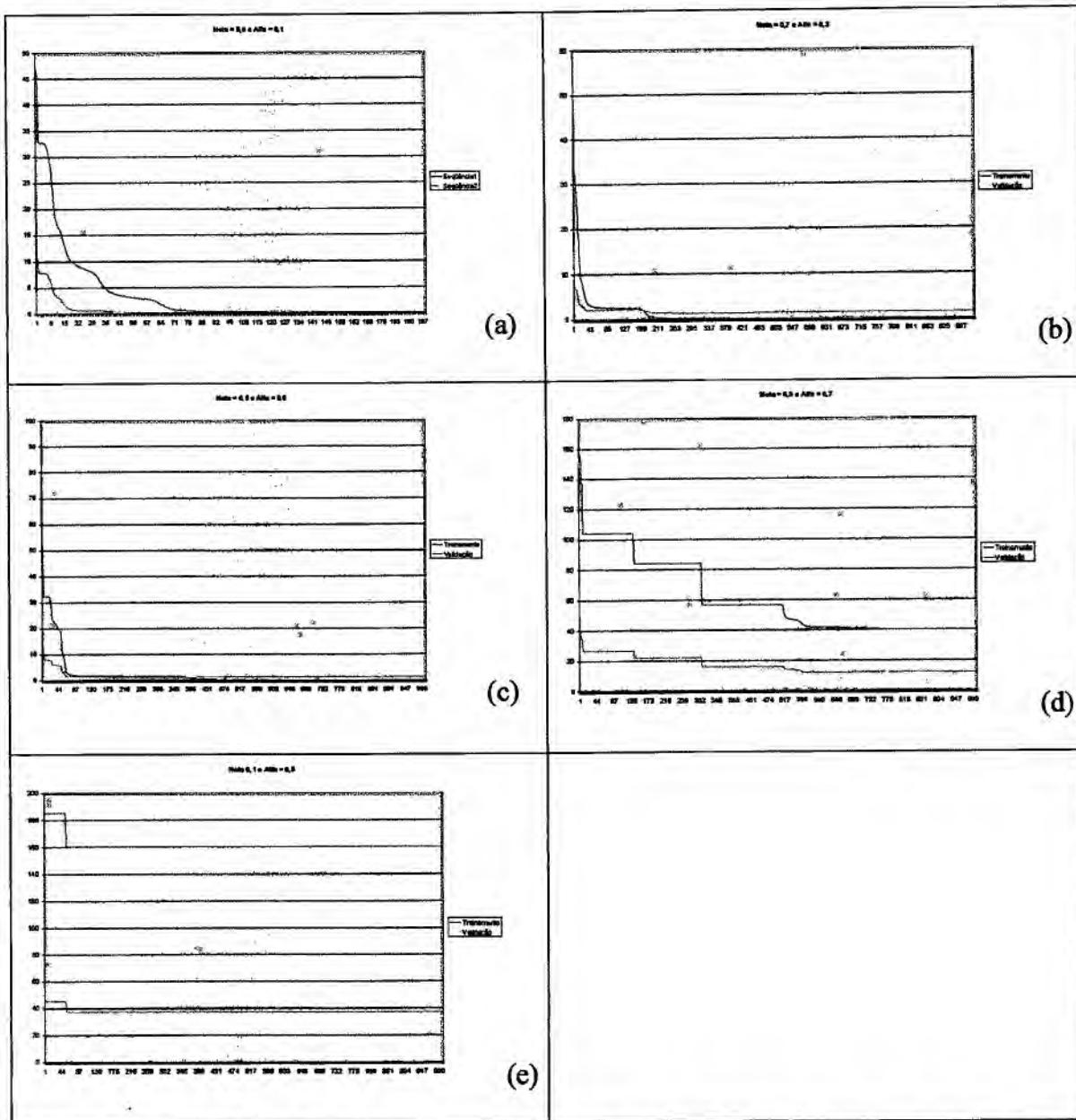


Figura 4.7: Variação de erro com diferentes valores de η e α , com uma camada intermediária.

4.2.7 Classificar com MLP

Uma vez treinada a rede MLP, o processo classificar com MLP nada mais é que a fase de validação onde são apresentadas as imagens a serem classificadas.

Na figura 4.8 pode ser observada a variação do erro quadrático para a fase de teste, ou validação, para 4 diferentes imagens de testes nas mesmas condições, mantendo-se a mesma topologia e valores de η e α .

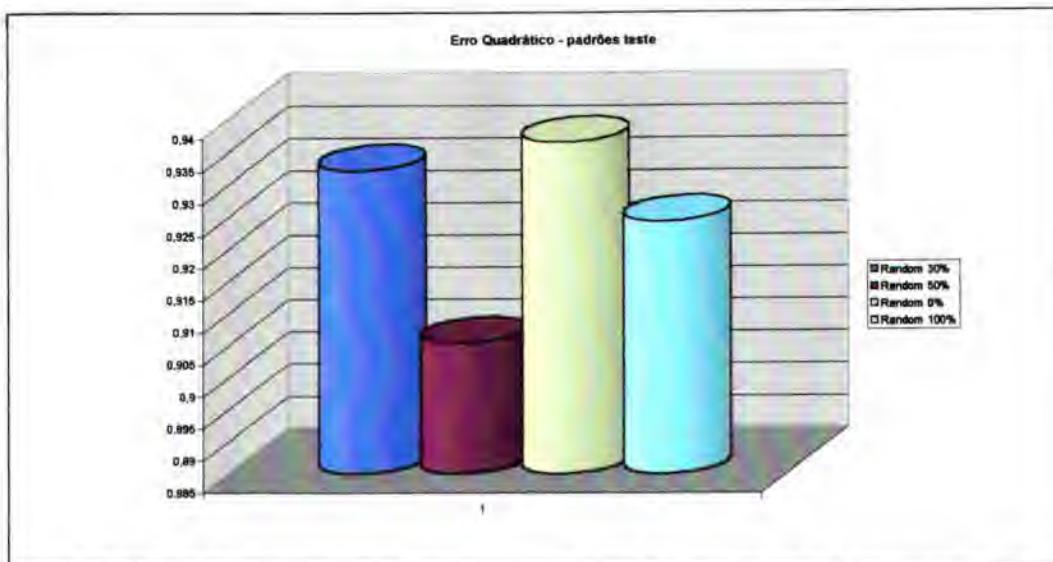


Figura 4.8: Variação do erro quadrático para a fase de teste para 4 diferentes arquivos, na mesma topologia e valores de η e α .

Pode-se observar que existe uma pequena variação do erro, indicando que dependendo da imagem de entrada o comportamento da rede, apesar de já ter sido treinada, pode variar de imagem para imagem.

Este processo também é utilizado para classificar o mapa gerado pela rede de Kohonen, com o treinamento para os respectivos padrões. Desta forma consegue-se rotular cada agrupamento do mapa segundo as cores de cada classe: solo, planta ou resíduo.

4.2.8 Treinar Rede de Kohonen e Gerar Mapa Auto-Organizado

A rede de Kohonen foi proposta por Teuvo Kohonen da Universidade de Tecnologia de Helsinki, na Finlândia e surgiu devido ao interesse que o grupo de Reconhecimento de Padrões dessa universidade encontrou na tarefa de extrair características de um conjunto de fonemas observados, usando elementos da Teoria de Probabilidades e da Teoria das Decisões Clássicas. Estes modelos de rede conhecidos como Mapas Auto-organizáveis (SOM) foram estudados por Kohonen [VAR00]. Não é necessária a apresentação da saída desejada para realizar a correção dos erros, ou seja, as redes são não supervisionadas.

O modelo de rede neural proposto por Kohonen [CAR98] baseia-se no método de Classificação de Vizinho mais próximo. Este método interpreta um conjunto de dados como padrões e usa estatísticas tradicionais para formar classes C_i de padrões, onde cada classe é representada por um pequeno conjunto de padrões P_i . Cada novo padrão p pode ser classificado através da identificação da classe C_p para a qual p é mais próximo com

respeito a alguma métrica já definida. Neste caso utiliza-se a medida de distância Euclidiana.

Possui uma arquitetura muito simples. Apenas duas camadas compõem a rede: uma de entrada (que na verdade não é uma camada e sim um filtro que é utilizado para apresentar os padrões) e uma de saída, de forma similar à rede apresentada na figura 4.9. Nesta figura pode-se observar uma entrada conectada com o mapa de saída.

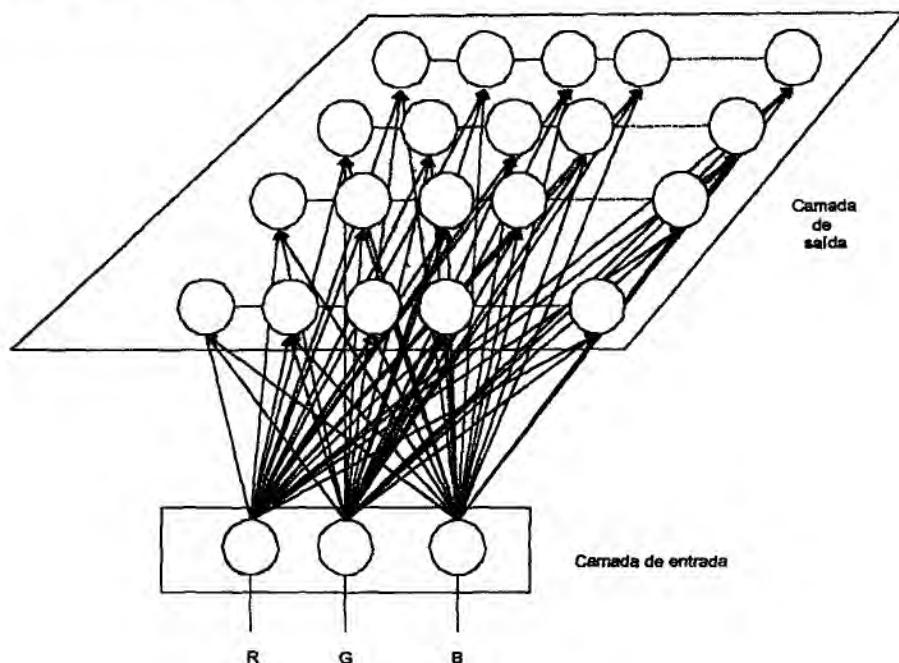


Figura 4.9: Topologia da Rede de Kohonen

Cada neurônio da camada de saída possui conexões com cada uma das entradas e existe um neurônio associado a cada uma dessas conexões. Quando um padrão é apresentado à rede, cada neurônio calcula a diferença do seu valor de pesos em relação ao padrão que está sendo apresentado.

Depois de cada unidade ter calculado essa diferença, o neurônio ganhador é o que possui a menor diferença. Neste contexto, uma vez escolhida a unidade ganhadora, somente ela tem o direito de atualizar suas conexões e as dos seus vizinhos. Um critério de vizinhança pode ser utilizado, por exemplo, escolher regiões retangulares em volta da unidade ganhadora. No caso deste trabalho, utiliza-se uma região circular, conforme mostra a figura 4.10.

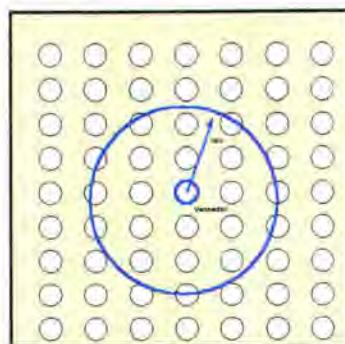


Figura 4.10: Critério de vizinhança para atualização de pesos

Após a atualização dos pesos em volta do neurônio ganhador, a rede passa a apresentar grupos ou *clusters*. Cada *cluster* corresponde a um grupo de padrões similares.

O aprendizado da rede de Kohonen é um método para classificação do tipo vizinho mais próximo, no qual os vetores-peso iniciais (número fixo) são progressivamente modificados para mapear o espaço de entradas.

Durante o aprendizado da rede de Kohonen, todos os seus neurônios, que formam o chamado mapa de preservação de topologia, recebem o vetor de entrada x *n-dimensional*. Após a atualização do neurônio vencedor e da vizinhança de raio progressivamente decrescente, os pesos dos neurônios pertencentes a esta vizinhança são modificados. Esse processo é repetido para todas as entradas apresentadas à rede e por fim, o mapa bidimensional de saída representará o espaço de características das entradas *n-dimensionais*.

O algoritmo Kohonen funciona como classificador devido ao fato do treinamento da rede estimular a competição pelas entradas num conjunto limitado de neurônios. Deste modo, regiões de neurônios com vetores-peso similares são criadas. Estas regiões tendem a ser ordenadas topologicamente do mesmo modo que os correspondentes vetores de entrada. Portanto, a rede pode ser vista como um classificador de vetores entrada em classes de similaridade topológicas.

✓ Algoritmo da Rede de Kohonen:

Cria-se uma matriz de r linhas por c colunas, que representa a chamada camada de Kohonen. Cada elemento desta camada tem a mesma dimensão do vetor de entrada. Deve-se observar que a camada de Kohonen deve ser suficientemente grande em relação ao número de entradas. No caso, esta camada deve ser maior que a imagem de entrada.

Para cada vetor de entrada $x(n)$ ([RGB] ou [HOT] ou somente H), executa-se os seguintes passos:

- 1) Identifica-se um neurônio z com a entrada apresentada, isto é, o neurônio cujo vetor-peso tiver a menor distância euclidiana em relação à entrada apresentada.
- 2) Atualizar gradualmente este neurônio e sua vizinhança N_z . Isto é, atualizar desde um raio inicial que é igual à largura da rede até um raio igual a zero, que corresponde a atualizar somente o neurônio escolhido.

A atualização do vetor-peso associado ao neurônio j ocorre na direção de x pela equação:

$$\Delta w_j = \begin{cases} \alpha(x - w_j), & j \in N_z \\ 0, & \text{c.c.} \end{cases} \quad (4.13)$$

onde α é uma função do tempo e determina o quanto w_j deve caminhar na direção de x .

Assumindo, também, uma discretização do tempo tem-se que a fórmula acima pode ser reescrita na seguinte forma:

$$w_j(n+1) = \begin{cases} w_j(n) + \alpha(n)[x(n) - w_j(n)], & j \in N_z(n) \\ w_j(n), & \text{c.c.} \end{cases} \quad (4.14)$$

onde n aumenta de uma unidade a cada variação de vizinhança e a cada nova entrada apresentada.

Para assegurar a convergência dos vetores-peso, dois requisitos devem ser satisfeitos:

- 1) α deve satisfazer as seguintes condições:

$$\begin{cases} 0 < \alpha(n) < 1 & \rightarrow \forall n \\ \alpha(n) = \frac{1}{2} \alpha(n-1) & (\text{deve ser uma função decrescente}), \end{cases}$$

Quando α atinge certo limite, diz-se que o aprendizado ocorreu. Neste caso, estabelece-se um valor bem baixo.

- 2) O tamanho da vizinhança N_z deve satisfazer duas condições:

$$\begin{cases} 1 < N_z(n) < \max(\text{linha}, \text{coluna}) & \rightarrow \forall n \\ N_z(n) = N_z(n-1)-1 & (\text{deve ser uma função decrescente}), \end{cases}$$

onde $\max(\text{linha}, \text{coluna})$ é considerado o diâmetro da rede.

Como $\alpha(n)$ deve ser decrescente durante o processo de aprendizado, esta função deve ser atualizada a cada novo padrão (entrada) apresentado à rede. Depois que todas as entradas foram apresentadas, verifica-se um certo limite, que deve ser um valor bem próximo de zero, atingido pela função decrescente $\alpha(n)$. Em caso positivo, a etapa de treinamento está concluída; caso contrário, as entradas devem ser reapresentadas à rede, até que o limite especificado seja alcançado.

Finalmente, a cada nova entrada é associado o neurônio que melhor a represente de acordo com o critério da menor distância euclidiana. Cada neurônio no mapa topológico deve reapresentar apenas uma entrada. Assim, o resultado obtido na rede pode ser visualizado na matriz escolhida para representar a topologia da rede.

Na figura 4.11 pode-se observar a entrada de dados para a rede. Para simplificar foi testada numa interface feita no sistema operacional DOS. Na figura 4.12 apresenta-se uma das imagens utilizadas para teste. Nesta imagem existem 4 classes de cores, nítidamente identificadas.

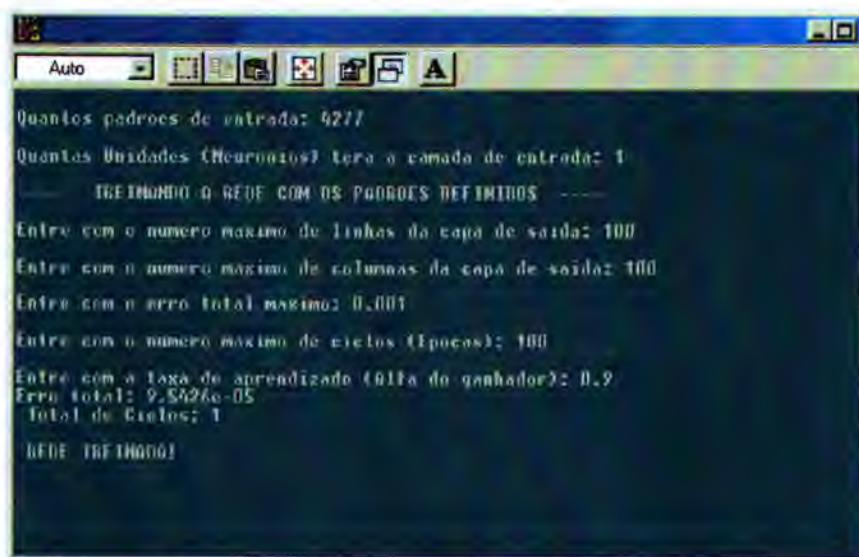


Figura 4.11: Entrada de parâmetros na Rede de Kohonen

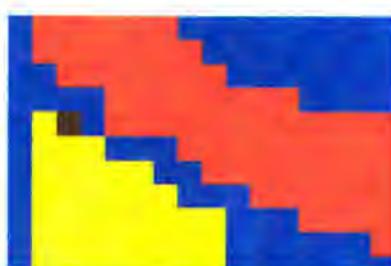


Figura 4.12: Imagem teste

0	1	4	9	16	25	36	49	64	81	100	121	144	169	-1	-1	-1	-1
2	3	5	10	17	26	37	50	65	82	101	122	145	170	-1	-1	-1	-1
6	7	8	11	18	27	38	51	66	83	102	123	146	171	-1	-1	-1	-1
12	13	14	15	19	28	39	52	67	84	103	124	147	172	-1	-1	-1	-1
20	21	22	23	24	29	40	53	68	85	104	125	148	173	-1	-1	-1	-1
30	31	32	33	34	35	41	54	69	86	105	126	149	174	-1	-1	-1	-1
42	43	44	45	46	47	48	55	70	87	106	127	150	175	-1	-1	-1	-1
56	57	58	59	60	61	62	63	71	88	107	128	151	-1	-1	-1	-1	-1
72	73	74	75	76	77	78	79	80	89	108	129	152	-1	-1	-1	-1	-1
90	91	92	93	94	95	96	97	98	99	109	130	153	-1	-1	-1	-1	-1
110	111	112	113	114	115	116	117	118	119	120	131	154	-1	-1	-1	-1	-1
132	133	134	135	136	137	138	139	140	141	142	143	155	-1	-1	-1	-1	-1
156	157	158	159	160	161	162	163	164	165	166	167	168	-1	-1	-1	-1	-1
-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1

Figura 4.14: Mapa de saída com entradas RGB da figura 4.13, com $\alpha = 0,7$, erro=0,001 e ciclos=1000

Avaliando-se a imagem 4.12, através do histograma das componentes HOT da figura 4.15, ou mais precisamente, na componente H, verifica-se a existência de 4 cores predominantes, sendo uma delas com pouca ocorrência na imagem.

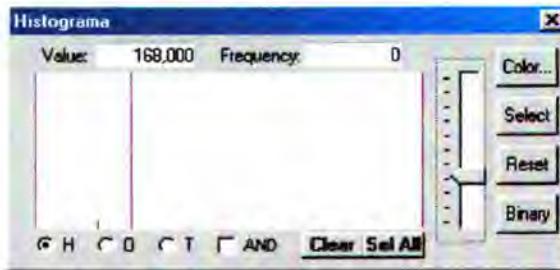


Figura 4.15 : Histograma de cores no sistema HOT

Na figura 4.16 pode-se verificar a matriz H gerada para a imagem da figura 4.12.

0.960	0.240	0.240	0.240	0.240	0.240	0.240	0.240	0.240	0.960	0.960	0.960	0.960	0.960	0.960	0.960	0.960	0.960
0.960	0.240	0.240	0.240	0.240	0.240	0.240	0.240	0.240	0.960	0.960	0.960	0.960	0.960	0.960	0.960	0.960	0.000
0.960	0.240	0.240	0.240	0.240	0.240	0.240	0.240	0.240	0.960	0.960	0.960	0.960	0.960	0.000	0.000	0.000	0.000
0.960	0.240	0.240	0.240	0.240	0.240	0.240	0.240	0.240	0.960	0.960	0.960	0.960	0.960	0.000	0.000	0.000	0.000
0.960	0.240	0.240	0.240	0.240	0.240	0.240	0.240	0.240	0.960	0.960	0.960	0.960	0.960	0.000	0.000	0.000	0.000
0.960	0.240	0.240	0.240	0.240	0.240	0.240	0.240	0.240	0.960	0.960	0.960	0.960	0.960	0.000	0.000	0.000	0.000
0.960	0.240	0.240	0.240	0.240	0.240	0.240	0.240	0.240	0.960	0.960	0.960	0.960	0.960	0.000	0.000	0.000	0.000
0.960	0.240	0.240	0.240	0.240	0.240	0.240	0.240	0.240	0.960	0.960	0.960	0.960	0.960	0.000	0.000	0.000	0.000
0.960	0.240	0.240	0.240	0.240	0.240	0.240	0.240	0.240	0.960	0.960	0.960	0.960	0.960	0.000	0.000	0.000	0.000
0.960	0.240	0.240	0.240	0.240	0.240	0.240	0.240	0.240	0.960	0.960	0.960	0.960	0.960	0.000	0.000	0.000	0.000
0.960	0.240	0.240	0.240	0.240	0.240	0.240	0.240	0.240	0.960	0.960	0.960	0.960	0.960	0.000	0.000	0.000	0.000
0.960	0.240	0.240	0.240	0.240	0.240	0.240	0.240	0.240	0.960	0.960	0.960	0.960	0.960	0.000	0.000	0.000	0.000
0.960	0.240	0.240	0.240	0.240	0.240	0.240	0.240	0.240	0.960	0.960	0.960	0.960	0.960	0.000	0.000	0.000	0.000
0.960	0.240	0.240	0.240	0.240	0.240	0.240	0.240	0.240	0.960	0.960	0.960	0.960	0.960	0.000	0.000	0.000	0.000
0.960	0.240	0.240	0.240	0.240	0.240	0.240	0.240	0.240	0.960	0.960	0.960	0.960	0.960	0.000	0.000	0.000	0.000
0.960	0.240	0.240	0.240	0.240	0.240	0.240	0.240	0.240	0.960	0.960	0.960	0.960	0.960	0.000	0.000	0.000	0.000
0.960	0.240	0.240	0.240	0.240	0.240	0.240	0.240	0.240	0.960	0.960	0.960	0.960	0.960	0.000	0.000	0.000	0.000
0.960	0.240	0.240	0.240	0.240	0.240	0.240	0.240	0.240	0.960	0.960	0.960	0.960	0.960	0.000	0.000	0.000	0.000
0.960	0.240	0.240	0.240	0.240	0.240	0.240	0.240	0.240	0.960	0.960	0.960	0.960	0.960	0.000	0.000	0.000	0.000
0.960	0.240	0.240	0.240	0.240	0.240	0.240	0.240	0.240	0.960	0.960	0.960	0.960	0.960	0.000	0.000	0.000	0.000
0.960	0.240	0.240	0.240	0.240	0.240	0.240	0.240	0.240	0.960	0.960	0.960	0.960	0.960	0.000	0.000	0.000	0.000
0.960	0.240	0.240	0.240	0.240	0.240	0.240	0.240	0.240	0.960	0.960	0.960	0.960	0.960	0.000	0.000	0.000	0.000
0.960	0.240	0.240	0.240	0.240	0.240	0.240	0.240	0.240	0.960	0.960	0.960	0.960	0.960	0.000	0.000	0.000	0.000

Figura 4.16: Matriz H da imagem da figura 4.14

Especificando-se um mapa de saída de 20x20, pode-se observar na figura 4.17 a matriz gerada por Kohonen, com $\alpha=0,9$, depois de 100 ciclos e para um limite de erro 0,01.

31	44	47	62	77	93	108	130	162	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
45	46	58	63	78	94	109	131	163	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
59	60	61	72	79	95	110	132	164	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
73	74	75	76	87	100	111	133	165	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
88	89	90	91	92	101	116	134	166	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
102	103	104	105	106	107	117	135	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
118	98	119	120	121	122	123	136	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
145	146	147	148	149	150	151	161	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
172	173	174	175	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
144	152	153	154	155	156	157	158	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
124	125	126	127	128	129	137	159	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
71	80	84	85	86	96	138	160	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
41	42	43	48	55	99	139	167	-1	-1	-1	-1	53	54	65	66	67	68	69	81			
25	26	27	28	56	112	140	168	-1	-1	-1	-1	82	24	33	34	35	36	37	38			
12	13	14	29	57	113	141	169	-1	-1	-1	-1	83	39	7	8	17	18	19	40			
9	10	15	30	64	114	142	170	-1	-1	-1	-1	97	49	20	2	3	4	21	50			
0	11	16	32	70	115	143	171	-1	-1	-1	-1	51	22	5	1	6	23	52				

Figura 4.17: Mapa de Kohonen para $\alpha = 0.9$, ciclos=100 e erro= 0.01

Observa-se que neste caso foram encontrados 3 grupos ou *clusters* na imagem. Em cada grupo observa-se a posição do pixel que o compõe. Aquele pixel cuja cor é diferente e que aparece no histograma como uma única ocorrência, não foi discriminado com estes ajustes apresentados. Sendo assim, variando-se o α para 0,7 e aumentando-se o número de ciclos para 1000 e diminuindo o limiar de erro para 0,001, verifica-se na matriz gerada, apresentada na figura 4.18 que é possível discriminar este pixel de cor diferente. Isto mostra que a rede é sensível aos valores de ajuste.

31	44	47	62	77	93	108	131	163	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	
45	46	58	63	78	94	109	132	164	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	
59	60	61	72	79	95	110	133	165	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	
73	74	75	76	87	100	111	134	166	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	
88	89	90	91	92	101	116	135	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	
102	103	104	105	106	107	117	136	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	
118	119	120	121	122	123	130	145	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	98	
146	147	148	149	150	151	161	162	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	
-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
53	54	65	66	67	68	69	81	82	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	
83	24	33	34	35	36	37	38	97	-1	-1	-1	167	139	99	32	41	42	43	48	55			
-1	39	7	8	17	18	19	40	-1	-1	-1	168	140	112	56	14	15	16	25	26				
-1	49	20	2	3	4	21	50	-1	-1	-1	169	141	113	57	27	9	10	11	28				
-1	51	22	5	1	6	23	52	-1	-1	-1	170	142	114	64	29	12	0	13	30				

Figura 4.18: Mapa de Kohonen $\alpha = 0.7$, ciclos= 1000, erro= 0.001

Uma imagem de cobertura obtida com uma torre de filmagem (figura 4.19), porém com dimensões reduzidas, também foi utilizada aqui para demonstrar o funcionamento da rede. Na figura 4.20 pode-se observar o mapa gerado de tamanho 100x100, com os valores de

$\alpha=0,7$, 1000 ciclos e limiar de erro 0,001. Cada *cluster* é visto em cor cheia. Foram encontrados 5 agrupamentos, dois de maior freqüência, que representam o solo e as plantas. Os outros mostram a grade colocada sobre o solo. Na figura 4.21, pode-se observar detalhes de parte do mapa gerado.

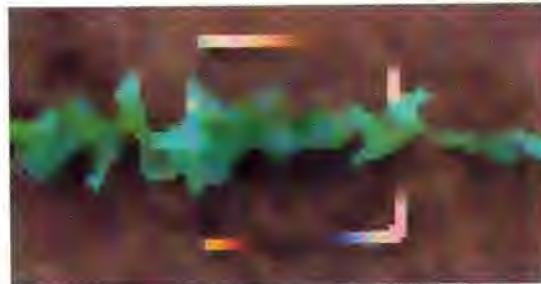


Figura 4.19: Imagem de cobertura depois do filtro de suavização

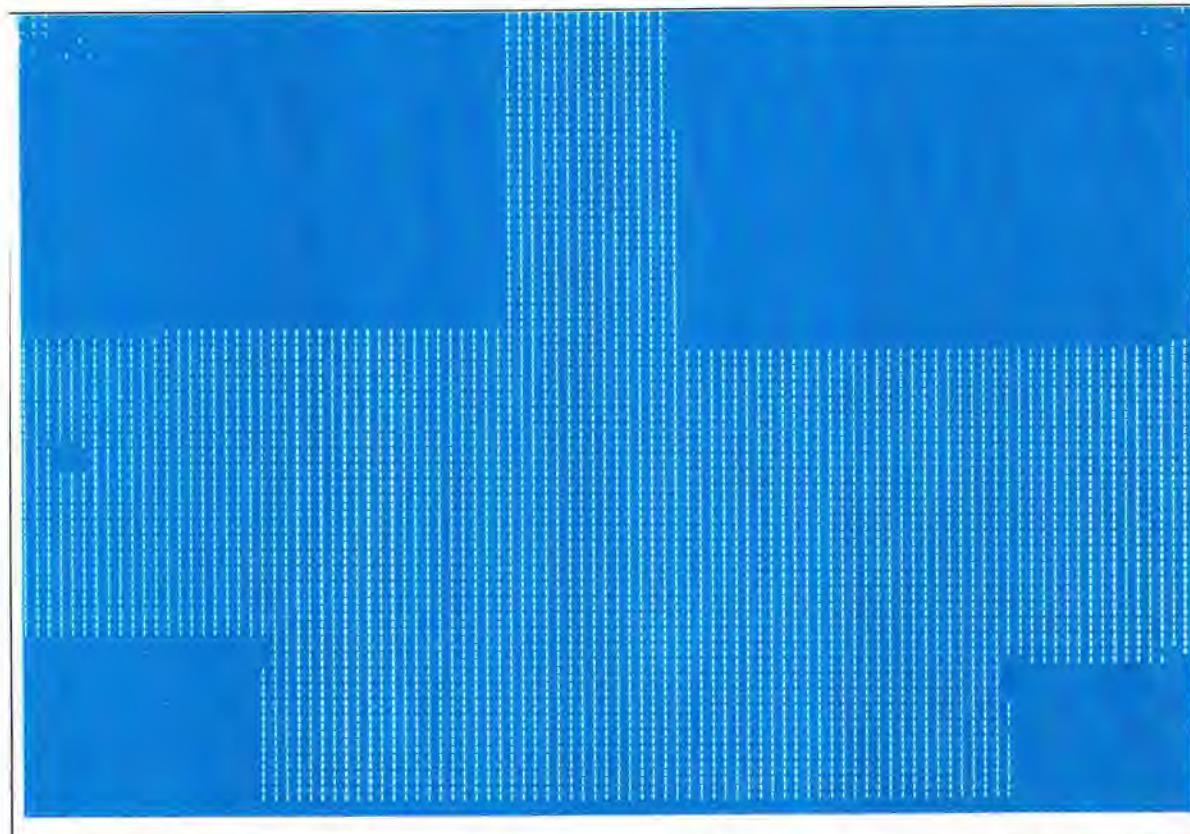


Figura 4.20: Mapa de Kohonen para a imagem da figura 4.17, com $\alpha=0,7$, erro= 0,001
e ciclos= 1000

tempo por padrões de classes diferentes. Uma alternativa para resolver este problema é a adição de um mecanismo de realimentação localizado entre a camada de entrada e a camada competitiva. Este mecanismo facilita o aprendizado de novos padrões sem esquecer a informação já armazenada, a alternância entre os modos.

Existem dois tipos de arquiteturas de redes ART adequadas para a classificação de padrões, a ART1 e a ART2. A diferença está no tipo dos dados com que elas trabalham. A rede ART1 é apropriada para entradas binárias e a ART2 para entradas analógicas.

A arquitetura básica ART envolve duas camadas de neurônios. A camada de entrada, também conhecida como F1, faz o processamento dos dados de entrada. A camada de saída, também conhecida como F2, contém os neurônios que representam os *clusters* (Figura 4.22).

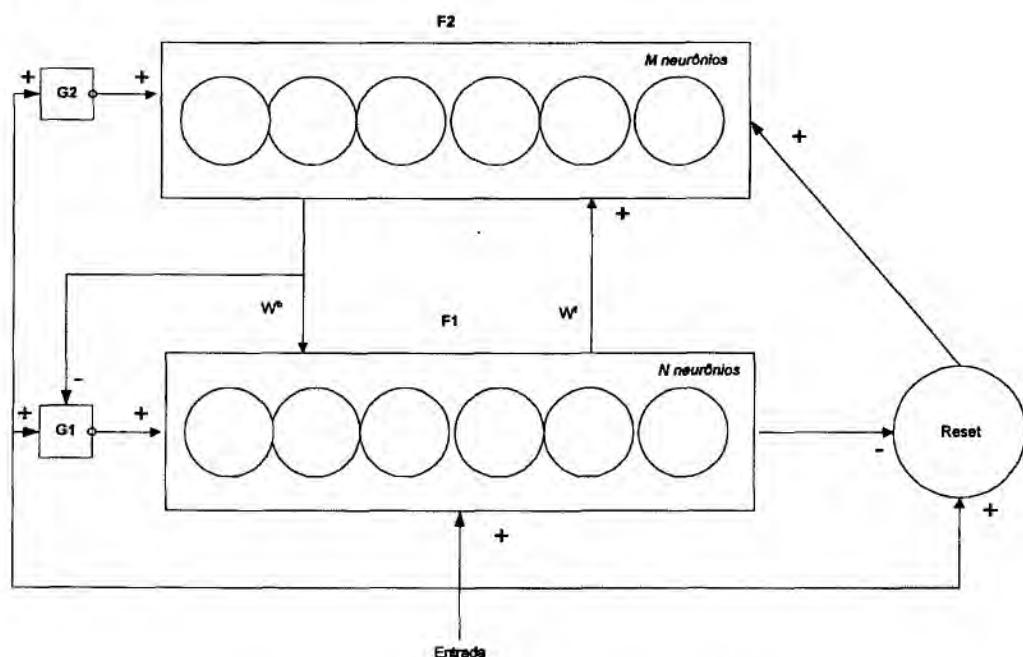


Figura 4.22: Arquitetura básica ART

Estas duas camadas estão conectadas através de dois vetores de pesos. Um vetor de pesos *feedforward*, representado por W_f , assume valores reais e segue da camada de entrada para a camada de saída. Outro vetor de pesos *feedback*, representado por W_b , assume valores binários e vai da camada de saída para a camada de entrada. Estes pesos são também conhecidos como filtros adaptativos porque mudam dinamicamente para possibilitar o aprendizado de novos padrões de entrada. O vetor de pesos *feedforward*, associado a cada neurônio da camada de saída, pode ser visto como um protótipo ou exemplar dos padrões que fazem parte do *cluster* representado por este neurônio.

Para cada camada da rede ART há uma unidade externa de controle. Essas unidades controlam o fluxo de dados através das camadas em cada fase de um ciclo de treinamento. Elas estão representadas por G1 e G2 na figura 4.22.

Entre a camada de entrada e a camada de saída há também uma unidade de *reset*, que é responsável por comparar as entradas da rede com um limiar de vigilância ρ . Este limiar de vigilância determina se uma nova classe de padrões poderá ser criada para um determinado padrão de entrada, ou seja, controla o grau de similaridade.

✓ Algoritmo ART2

Uma arquitetura típica da rede ART2 pode ser vista com mais detalhes na figura 4.23. Nesta, a camada F2 é conectada à camada F1 através dos pesos W_b , aqui denominados de b_{ij} . A camada F2 é conectada à camada F1 através dos pesos W_t , aqui denominados por t_{ji} . Este duplo sentido de conexões é utilizado para controlar a similaridade dos padrões de um mesmo grupo. A camada F1 consiste de seis tipos de unidades (W, X, U, V, P e Q). Existem n unidades destas dependendo da dimensão do vetor de entrada. As conexões entre estas unidades indicam transformações que ocorrem com o sinal de uma unidade para a outra. Somente na conexão entre as duas camadas é que ocorre a multiplicação por pesos dos sinais.

O modelo ART2 é mais complexo que o modelo ART1, pois possui mecanismo de normalização de ruídos incluídos na camada F1. A eliminação de ruído é controlada pelo parâmetro θ (Theta) que ajuda a rede se tornar estável.

Os modos de aprendizado deste tipo de rede podem ser: rápido ou lento. O modo de aprendizagem afeta, além da velocidade, as características dos grupos e dos vetores de pesos a serem formados.

No modo rápido os pesos são atualizados durante a ressonância, e o tempo consumido é equivalente ao tempo de apresentação de um padrão. O modo rápido na ART2 fornece vetores de pesos equivalentes aos vetores obtidos na ART1. A aprendizagem rápida requer poucos ciclos de execução, mas é necessário realizar várias iterações do algoritmo, na parte que se refere à alteração dos pesos da camada F1, para cada padrão de entrada.

No modo lento, as alterações dos pesos ocorrem durante o tempo equivalente ao tempo de aprendizagem e, são necessárias muitas apresentações dos padrões para se obter a estabilidade da rede.

A aprendizagem rápida ou lenta na ART2 pode ser controlada pelo parâmetro de aprendizagem α . Valores pequenos para este parâmetro podem reduzir a velocidade de aprendizagem, mas garante que os pesos alcançam o equilíbrio e a distribuição dos padrões nos grupos.

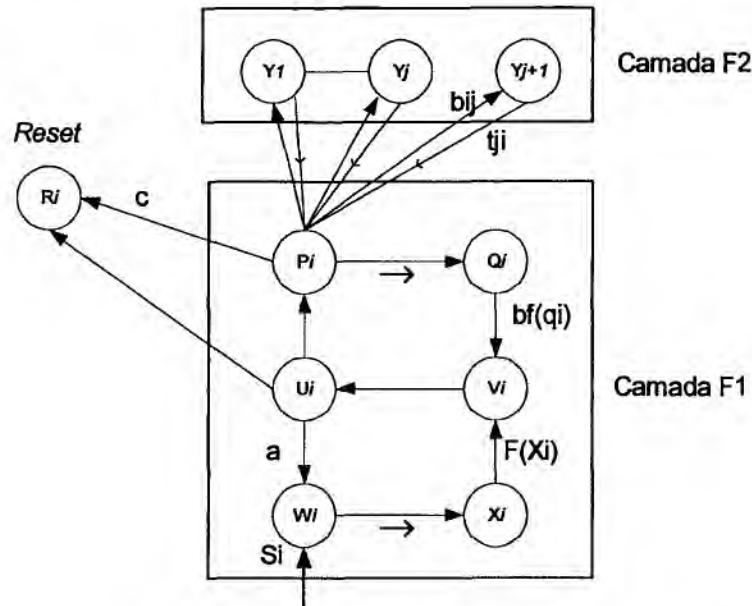


Figura 4.23: Detalhe da arquitetura ART2

O algoritmo do modelo ART 2 pode ser resumido como segue:

1. Repita (N vezes ou N_ciclos (Épocas))
 - a. Ativar a camada $F1$
 - b. Atualizar $F1$
 - c. Calcular sinais para enviar a $F2$
 - d. Enquanto ($RESET = Verdade$)
 - i. Encontrar o neurônio vencedor Yj
 - ii. Calcular $RESET$
 - iii. Se $RESET$ é verdadeiro {Vencedor Rejeitado} Então INIBIR Yj
Fim Enquanto passo(d)
 - e. Atualizar $F1$
 - f. Repita (i) , (ii) , (iii) até que os pesos do neurônio ganhador sejam atualizados
 - i. Atualiza os pesos do ganhador
 - ii. Atualiza $F1$
 - iii. Fim repita (f)
 - g. Repita o passo (a) para o próximo padrão
2. Repita passo (1) para o próximo número de ciclos

Para apresentar o comportamento da rede para a imagem de cobertura da figura 4.19, onde existem predominantemente 3 classes, solo planta e o reticulado, foram utilizados θ (Theta) = 0,0; ρ (Parâmetro Vigilância) = 0,9 e α (Taxa Aprendizagem) = 0,6. Os ciclos são apresentados na figura 4.24.

Ciclo 1	Ciclo 2	Ciclo 3
1 1 1 1 1 2 1 1 1 1 1 1 1 1 2 2 2 1 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 3 2 2 3 1 2 1 1 3 1 1 1 2 2 1 2 1 3 3 3 3 3 3 1 3 3 1 3 2 3 2 2 3 2 3 3 3 3 2 3 3 1 3 3 3 3 3 3 3 2 2 1 3 3 2 3 2 3 3 3 3 3 3 3 2 1 3 3 3 3 3 3 3 3 3 2 3 3 1 3 3 3 3 3 2 1 3	1 1 1 1 1 2 1 1 2 1 1 1 1 1 1 2 1 1 1 1 2 1 1 2 1 2 2 2 1 1 1 1 1 2 2 1 2 2 2 1 1 2 1 2 1 2 1 1 1 1 1 1 2 1 1 1 2 1 2 3 3 2 2 2 2 2 2 1 3 2 2 2 2 2 2 1 2 2 1 2 3 1 2 2 2 2 2 3 3 2 2 1 2 2 1 1 2 2 2 2 3 2 3 2 1 1 2 2 2 2 2 2 2 2 2 2 2 2 1 2 1 2 2 2 2 2 3 2 3 3 3 3 3 3 3 2 3 3 3 3 3 3 3 2 3 3 3 3 3 3 2 3 3 3 3 3 3	1 1 1 1 1 1 2 1 1 1 1 1 1 1 2 1 1 1 1 1 1 2 1 2 2 2 1 1 1 1 1 2 2 1 2 2 2 1 2 1 1 2 1 1 1 1 1 1 2 1 1 1 2 1 2 2 2 1 2 1 1 2 2 2 3 2 2 2 2 2 1 2 2 2 1 2 3 1 2 2 2 2 2 3 1 2 2 2 1 1 1 2 2 2 2 3 2 3 1 1 2 2 2 2 2 2 2 2 1 2 2 1 1 2 1 2 2 2 2 2 3 2 3 3 3 3 3 3 3 2 3 3 3 3 3 3 3 2 3 3 3 3 3 3 2 3 3 3 3 3 3
Ciclo 4	Ciclo 5	Ciclo 6
1 1 1 1 1 1 2 1 1 1 1 1 1 2 1 1 1 1 1 1 2 1 2 2 2 1 1 1 1 1 2 2 1 2 2 2 1 1 2 1 1 1 2 1 1 1 1 1 1 2 1 1 1 2 1 2 2 2 2 1 2 1 1 2 2 2 3 2 2 2 2 2 1 2 2 2 1 2 3 1 2 2 2 2 2 2 3 1 2 2 2 1 1 1 2 2 2 2 3 2 3 1 1 1 2 2 2 2 2 2 2 2 2 1 2 2 1 1 2 1 2 2 2 2 2 3 2 3 3 3 3 3 3 3 2 3 3 3 3 3 3 3 2 3 3 3 3 3 3 2 3 3 3 3 3 3	1 1 1 1 1 1 2 1 1 1 1 1 1 2 1 1 1 1 1 1 2 1 2 2 2 1 1 1 1 1 2 2 1 2 2 2 1 1 2 1 1 1 2 1 1 1 1 1 1 2 1 1 1 2 1 2 2 2 2 1 2 1 1 2 2 2 3 2 2 2 2 2 1 2 2 2 1 2 3 1 2 2 2 2 2 2 3 1 2 2 2 1 1 1 2 2 2 2 3 2 3 1 1 1 2 2 2 2 2 2 2 2 1 2 2 1 1 2 1 2 2 2 2 2 3 2 3 3 3 3 3 3 3 2 3 3 3 3 3 3 3 2 3 3 3 3 3 3 2 3 3 3 3 3 3	1 1 1 1 1 1 2 1 1 1 1 1 1 2 1 1 1 1 1 1 1 2 1 2 2 2 1 1 1 1 1 2 2 1 2 2 2 1 1 2 1 1 1 2 1 1 1 1 1 1 2 1 1 1 2 1 2 2 2 2 1 2 1 1 2 2 2 3 2 2 2 2 2 1 2 2 2 1 2 3 1 2 2 2 2 2 2 3 1 2 2 2 1 1 1 2 2 2 2 3 2 3 1 1 1 2 2 2 2 2 2 2 2 1 2 2 1 1 2 1 2 2 2 2 2 3 2 3 3 3 3 3 3 3 2 3 3 3 3 3 3 3 2 3 3 3 3 3 3 2 3 3 3 3 3 3
Ciclo 7	Ciclo 8	Ciclo 9
1 1 1 1 1 1 2 1 1 1 1 1 1 2 1 1 1 1 1 1 2 1 2 2 2 1 1 1 1 2 2 1 2 2 2 1 1 2 1 1 1 2 1 1 1 1 1 1 2 1 1 1 2 1 2 2 2 2 1 2 1 1 2 2 2 3 2 2 2 2 2 1 2 2 2 1 2 3 1 2 2 2 2 2 2 3 1 2 2 2 1 1 1 2 2 2 2 3 2 3 1 1 1 2 2 2 2 2 2 2 2 1 2 2 1 1 2 1 2 2 2 2 2 3 2 3 3 3 3 3 3 3 2 3 3 3 3 3 3 3 2 3 3 3 3 3 3 2 3 3 3 3 3 3	1 1 1 1 1 1 2 1 1 1 1 1 1 2 1 1 1 1 1 1 2 1 2 2 2 1 1 1 1 2 2 1 2 2 2 1 1 2 1 1 1 2 1 1 1 1 1 1 2 1 1 1 2 1 2 2 2 2 1 2 1 1 2 2 2 3 2 2 2 2 2 1 2 2 2 1 2 3 1 2 2 2 2 2 2 3 1 2 2 2 1 1 1 2 2 2 2 3 2 3 1 1 1 2 2 2 2 2 2 2 2 1 2 2 1 1 2 1 2 2 2 2 2 3 2 3 3 3 3 3 3 3 2 3 3 3 3 3 3 3 2 3 3 3 3 3 3 2 3 3 3 3 3 3	1 1 1 1 1 1 2 1 1 1 1 1 1 2 1 1 1 1 1 1 1 2 1 2 2 2 1 1 1 1 2 2 1 2 2 2 1 1 2 1 1 1 2 1 1 1 1 1 1 2 1 1 1 2 1 2 2 2 2 1 2 1 1 2 2 2 3 2 2 2 2 2 1 2 2 2 1 2 3 1 2 2 2 2 2 2 3 1 2 2 2 1 1 1 2 2 2 2 3 2 3 1 1 1 2 2 2 2 2 2 2 2 1 2 2 1 1 2 1 2 2 2 2 2 3 2 3 3 3 3 3 3 3 2 3 3 3 3 3 3 3 2 3 3 3 3 3 3 2 3 3 3 3 3 3

Figura 4.24: Saída da rede ART2 para a imagem da figura 4.19

Os números mostrados representam as classes encontradas. A cada número de classe está associada uma outra matriz com a posição dos respectivos pixels na imagem para então ser gerada a imagem de saída.

Os valores correspondentes à supressão de ruído (θ) têm que ser valores baixos para que se tenha uma melhor clusterização, caso contrário os valores altos acabam por inibir o parâmetro de vigilância ρ , tendo como consequência um menor número de *clusters*. Caso o valor de θ não seja bem escolhido, a rede ficará instável e oscilará entre *clusters* diferentes no lugar de convergir.

4.3 Considerações Finais

Depois de testar cada uma das técnicas utilizadas pode-se afirmar que tanto os métodos supervisionados como os não supervisionados possibilitam a análise das imagens de forma bem mais independente do usuário quando comparados com o método originalmente utilizado no SIARCS. Os métodos supervisionados dependem ainda das amostras selecionadas *a priori* pelo operador, porém a seleção de uma área representando cada padrão ou classe é muito mais simples do que selecionar cor a cor na tabela, como era feito antes.

Além dos resultados da segmentação, alguns outros estudos puderam ser feitos, tais como: verificação da possibilidade de redução da dimensão dos dados de entrada mantendo o mesmo desempenho na segmentação com a utilização das componentes principais; melhor entendimento dos parâmetros de ajuste de cada uma das redes neurais implementadas, tais como atualização pesos do ganhador e seus vizinhos em Kohonen, o parâmetro de vigilância da rede ART2 e a influência do termo momento na MLP com relação à estabilização.

Em todos os métodos existem detalhes de ajustes que variam com o tipo da aplicação, devendo ser feitos ajustes para cada uma. Algumas dificuldades encontradas foram os tempos de processamento, objetivando principalmente a utilização em arquiteturas embarcadas nas aeronaves do projeto ARARA. Para isto, torna-se necessário hardware dedicado.

No próximo capítulo, são apresentados os resultados obtidos com os diferentes métodos quando aplicados nas imagens do projeto ARARA. Não foram processadas todas as imagens obtidas em virtude da baixa velocidade de processamento e da quantidade de testes efetuados. Foram escolhidos os parâmetros de cada um dos métodos com base nos testes efetuados neste capítulo. Não foram feitos novos testes variando-se todos os parâmetros, devido o elevado tempo de processamento.

Capítulo 5

Resultados Obtidos

5.1 Considerações Iniciais

No capítulo anterior foram apresentados a implementação e testes dos diversos métodos considerados para a segmentação de imagens de cobertura de solo. Neste capítulo são apresentados os resultados obtidos com imagens capturadas por diferentes formas: com um helicóptero do ARARA e câmera fotográfica química; com um avião do ARARA e câmera fotográfica digital; com um avião do ARARA e câmera de vídeo e com vídeo obtido em torre de filmagem.

As imagens utilizadas nos testes são apresentadas no item 5.2. No item 5.3 são apresentados os resultados com os diferentes métodos. No item 5.4 é feita uma discussão com as medidas das respectivas áreas segmentadas. As considerações finais sobre os testes realizados são apresentadas no item 5.5.

5.2 Imagens Utilizadas

As figuras 5.1 até 5.4 apresentam imagens típicas selecionadas entre o conjunto capturado durante a fase de testes. Cabe ressaltar que todas foram obtidas dentro do projeto ARARA, com apoio financeiro da EMBRAPA. No entanto, para simplificar a apresentação e minimizar o tempo de processamento, foram escolhidos apenas alguns casos para serem processados.

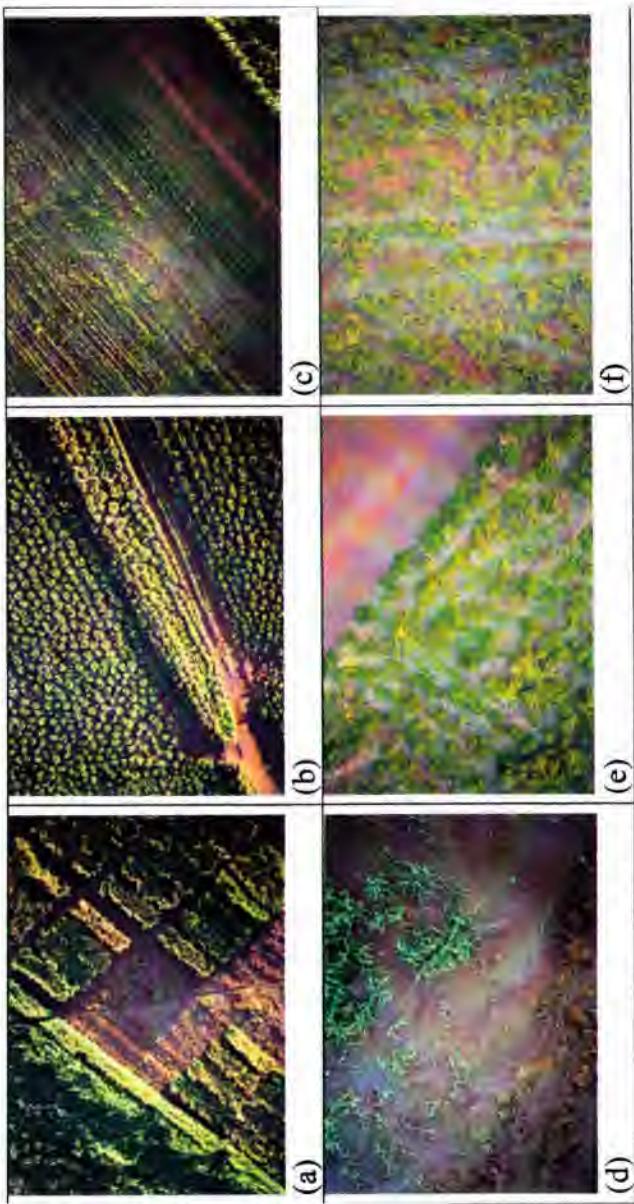


Figura 5.1: Imagens obtidas com fotos químicas no helicóptero

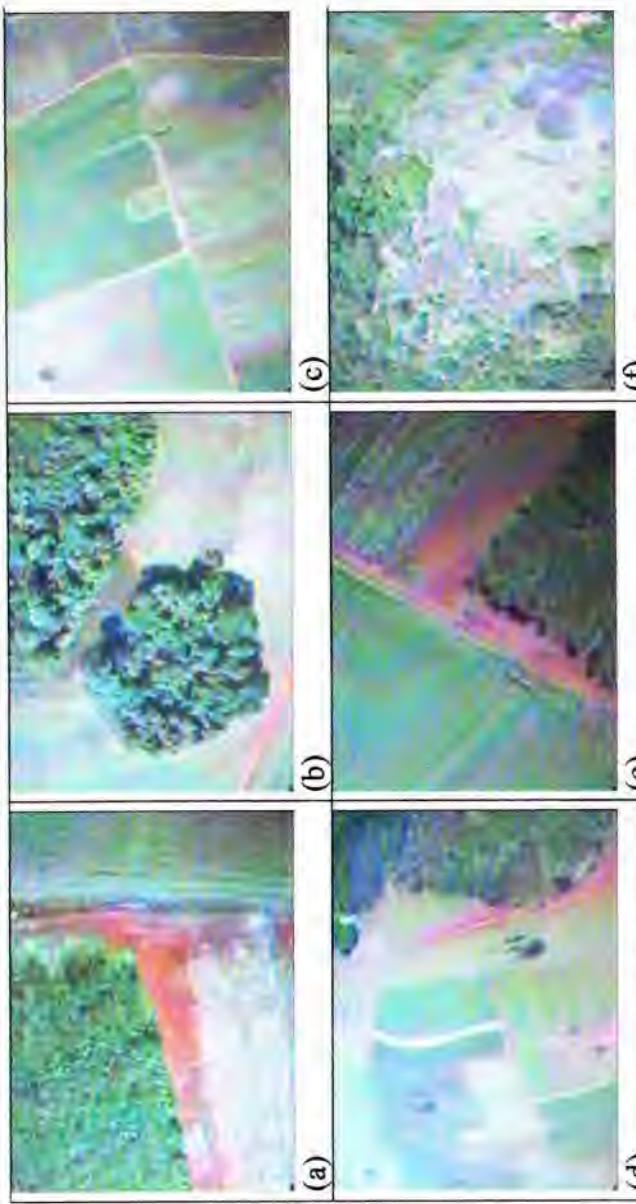


Figura 5.2: Imagens obtidas acima de 10m do chão, em voo no avião.

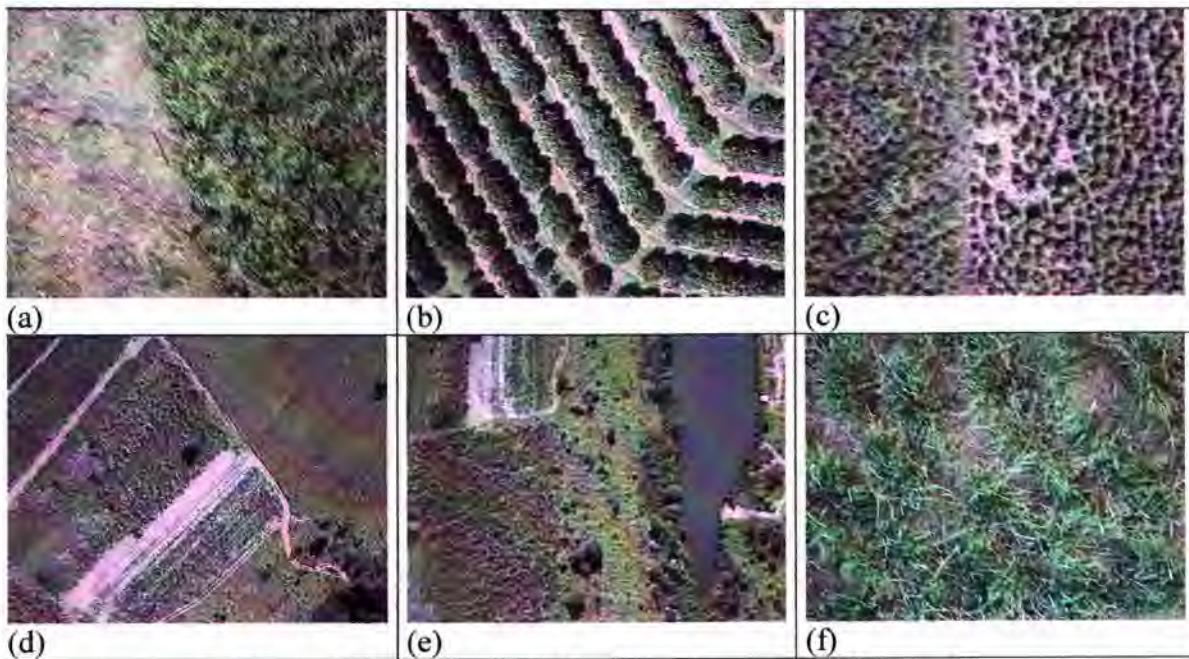


Figura 5.3: Imagens obtidas com câmera digital no avião.

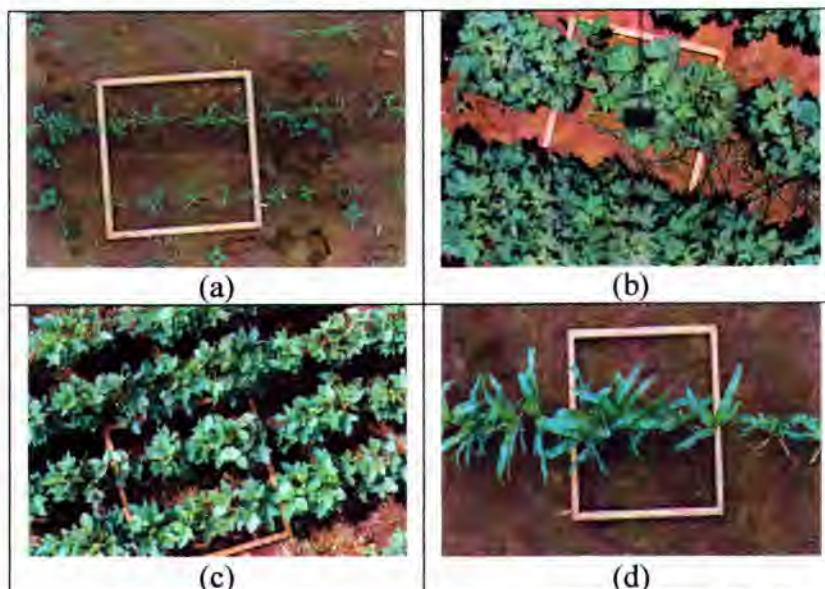


Figura 5.4: Imagens com filmagem em vídeo pela torre de filmagem

Deve-se notar que a qualidade das imagens obtidas está mais relacionada com a qualidade da câmera utilizada para a sua captura do que com o dispositivo utilizado para transporte/sustentação da câmera.

Um fato bastante marcante foi a dificuldade de generalização da Rede ART2. Para duas imagens com mesmas características visuais, às vezes o resultado era totalmente diferente. Este tipo de rede neural se mostrou muito sensível aos ajustes. Os resultados apresentados foram os melhores desempenhos conseguidos.

Chama-se de generalização a capacidade de um modelo de aprendizado responder corretamente aos exemplos que lhe são apresentados, sendo que estes exemplos não devem estar presentes na base de aprendizado. Um modelo que tem uma boa generalização é aquele modelo que responde corretamente não somente aos exemplos contidos na base de aprendizado, mas também a outros exemplos diferentes daqueles da base de aprendizado, e que estão contidos em uma base de teste. A capacidade de generalizar é a principal capacidade desejada nas tarefas que envolvem aprendizado. Uma rede pode se especializar demasiadamente em relação aos exemplos contidos na base de aprendizado. Este tipo de comportamento leva ao problema de aprendizado conhecido como super-aprendizado (*over-training / over-fitting*). Normalmente o *over-fitting* pode ser detectado/evitado através do uso de um teste de generalização variando-se o conjunto de validação.

5.3 Avaliação qualitativa da cobertura do solo pelos diferentes métodos

Nas figuras 5.5 a 5.9 podem ser vistos os resultados dos processamentos feitos nas imagens selecionadas. Em todos os casos são apresentados os resultados obtidos com a utilização da primeira componente de Hotelling, a componente H. Nos métodos supervisionados foram utilizadas 3 amostras para cada classe detectada na imagem. As saídas foram pintadas seguindo-se uma escala de cores pré-determinada: verde para planta, vermelho para solo, amarelo para resíduos e azul claro para classe água e palha. Procurou-se também estabelecer um limiar para que a classificação fosse feita, ou seja, caso a melhor classificação obtida apresentasse erro maior que esse limiar, o pixel não era classificado.

Na rede de Kohonen procurou-se utilizar 1000 ciclos, limiar de erro 0,001 e $\alpha = 0,7$. Na rede ART2 procurou-se utilizar $\theta = 0$; $\rho = 0,9$ e $\alpha = 0,6$. Também foi utilizado $\rho = 0,7$. Na rede MLP utilizou-se $\eta=0,9$ e $\alpha=0,1$, e uma arquitetura com 3 camadas, sendo a intermediária com 5 neurônios.

Nas figuras 5.5 e 5.6 são apresentados os resultados para imagens obtidas com o helicóptero. Nas figuras 5.7 e 5.9 aparecem os resultados com as imagens obtidas com o avião. A figura 5.8 apresenta os resultados para as imagens obtidas com a torre de filmagem. Neste caso pode-se observar que, de uma forma geral, foram separadas as regiões de solo nu e de cobertura verde. Em alguns casos houveram pixels não classificados por terem fugido do limiar preestabelecido de erro. O pior desempenho pode ser atribuído para a rede ART2, ajustada com $\rho = 0,7$.

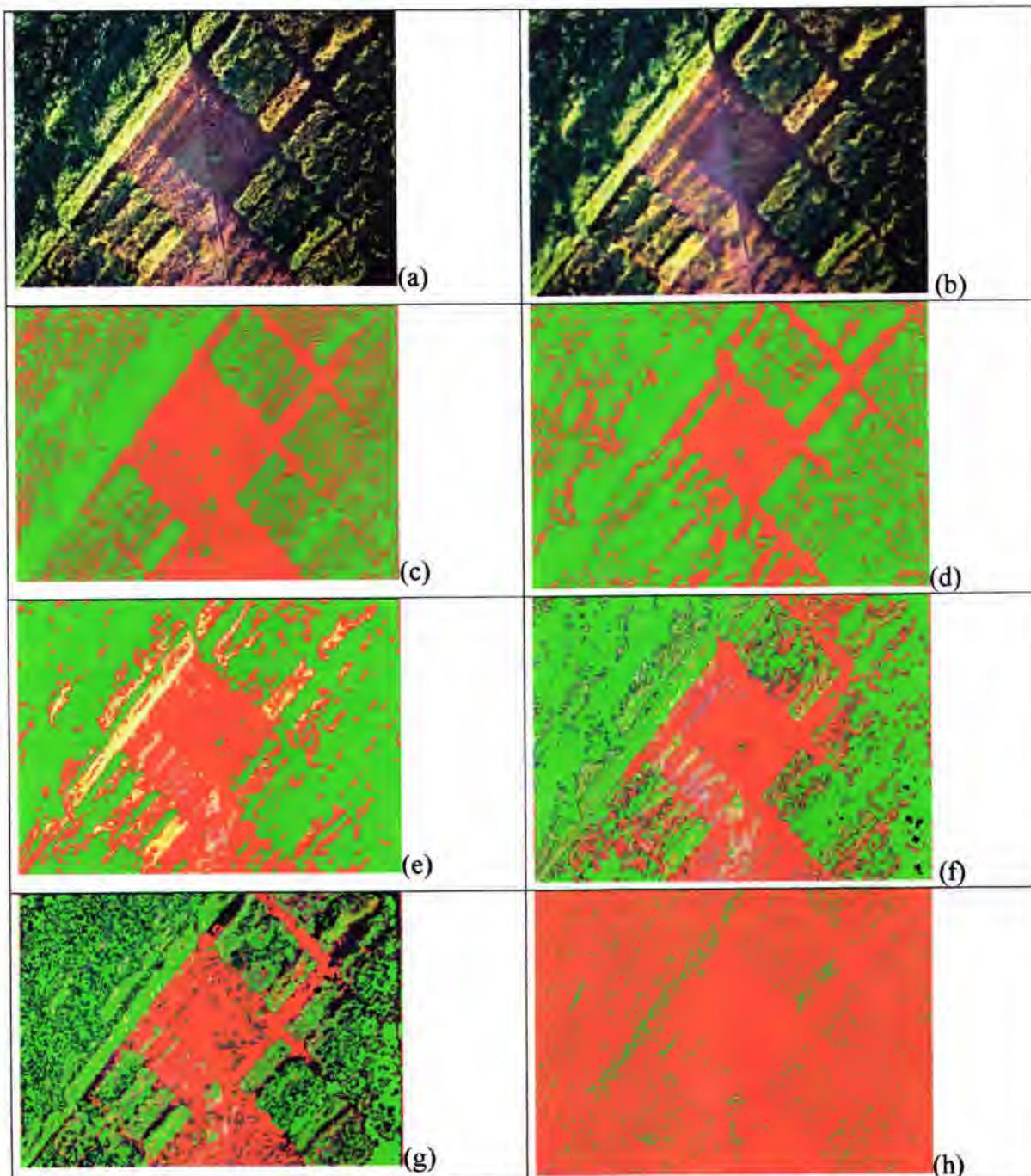


Figura 5.5: (a) Imagem obtida com helicóptero (b) com a suavização, (c) a segmentação manual com o SIARCS, a segmentação automática (d) com o discriminador linear, (e) com rede MLP, treinada com três amostras de cada padrão, (f) com rede de Kohonen e rotulada pela mesma MLP, (g) com ART2, $\rho = 0,9$ e (h) a mesma ART2, $\rho = 0,7$.

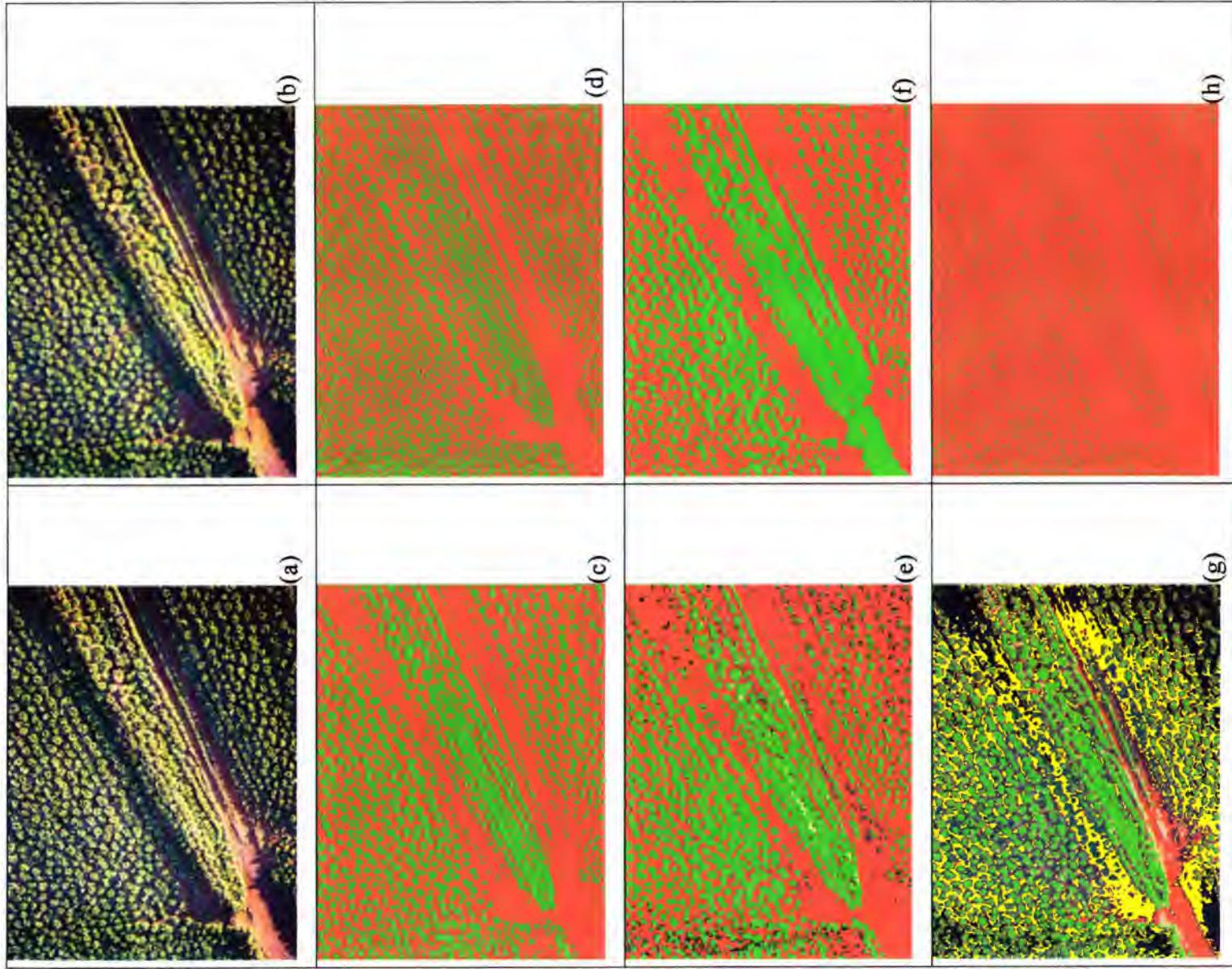


Figura 5.6: (a) Imagem obtida com helicóptero (b) com a suavização, (c) a segmentação manual com o SIARCS, a segmentação automática (d) com o discriminador linear, (e) com rede MLP, treinada com três amostras de cada padrão, (f) com rede de Kohonen e rotulada pela mesma MLP, (g) com ART2, $\rho = 0,9$ e (h) a mesma ART2, $\rho = 0,7$.

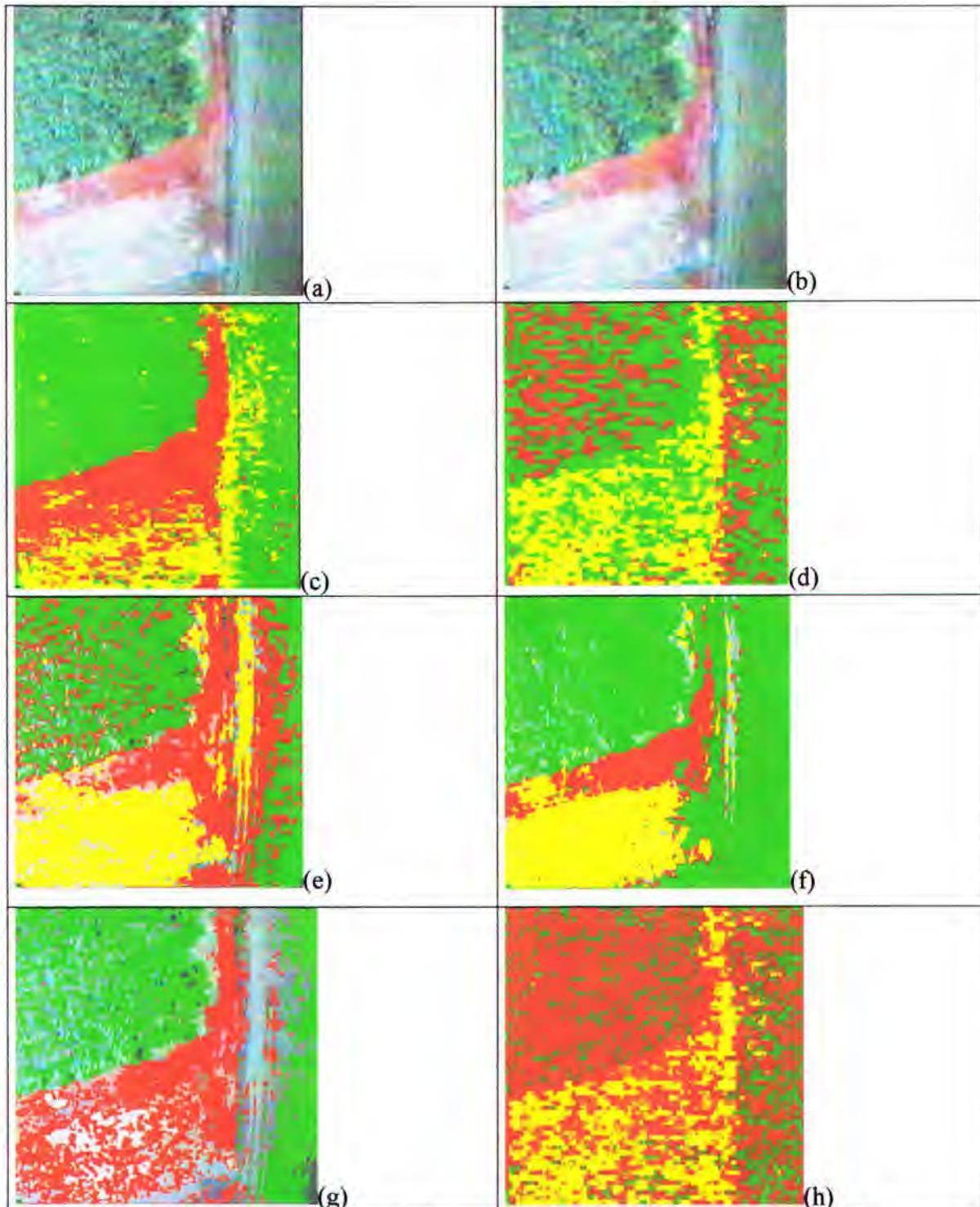


Figura 5.7: (a) Imagem obtida com avião (b) com a suavização, (c) a segmentação manual com o SIARCS, a segmentação automática (d) com o discriminador linear, (e) com rede MLP, treinada com três amostras de cada padrão, (f) com rede de Kohonen e rotulada pela mesma MLP, (g) com ART2, $\rho = 0,9$ e (h) a mesma ART2, $\rho = 0,7$.

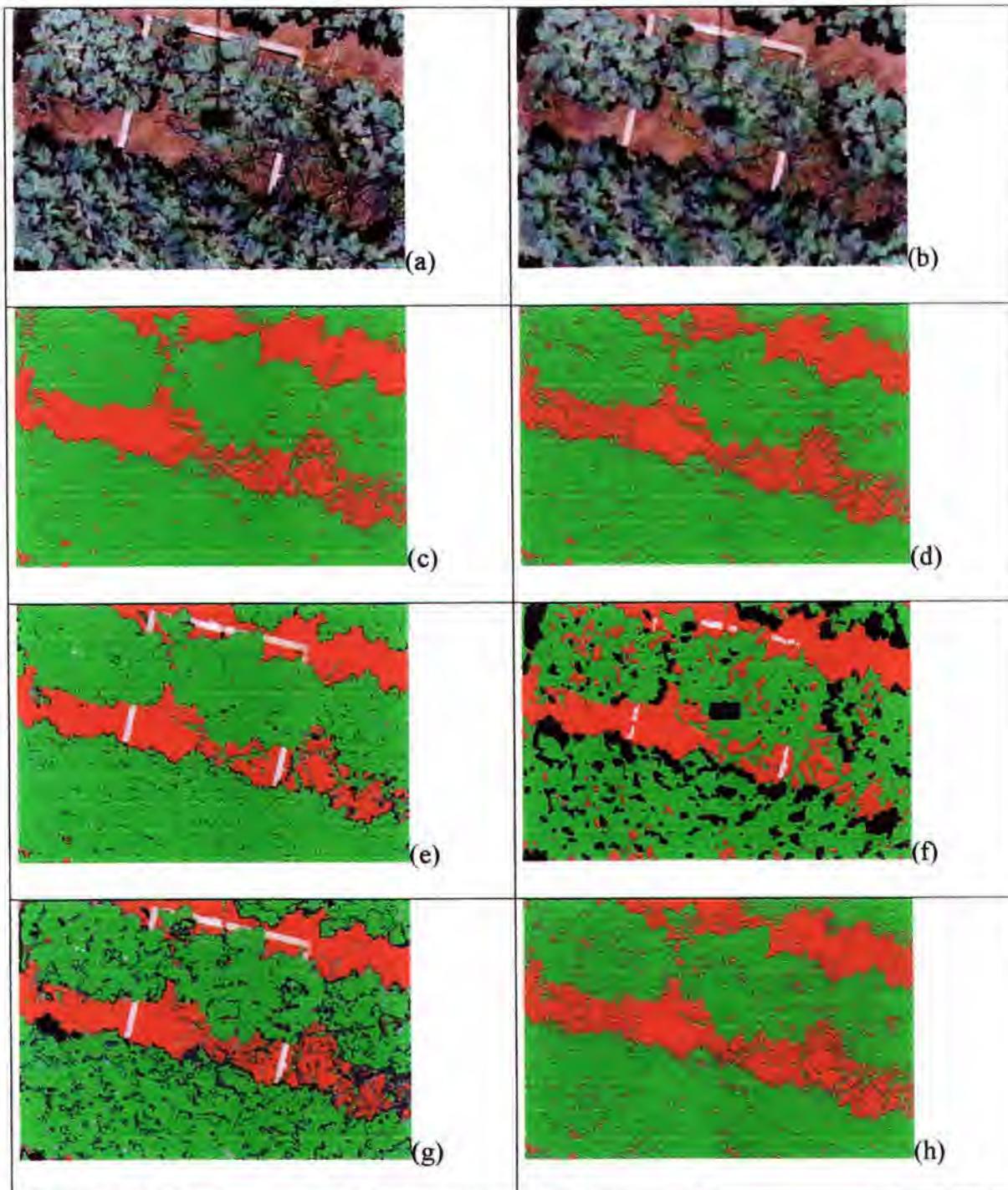


Figura 5.8: (a) Imagem obtida com torre de filmagem (b) com a suavização, (c) a segmentação manual com o SIARCS, a segmentação automática (d) com o discriminador linear, (e) com rede MLP, treinada com três amostras de cada padrão, (f) com rede de Kohonen e rotulada pela mesma MLP, (g) com ART2, $\rho = 0,9$ e (h) a mesma ART2, $\rho = 0,7$.

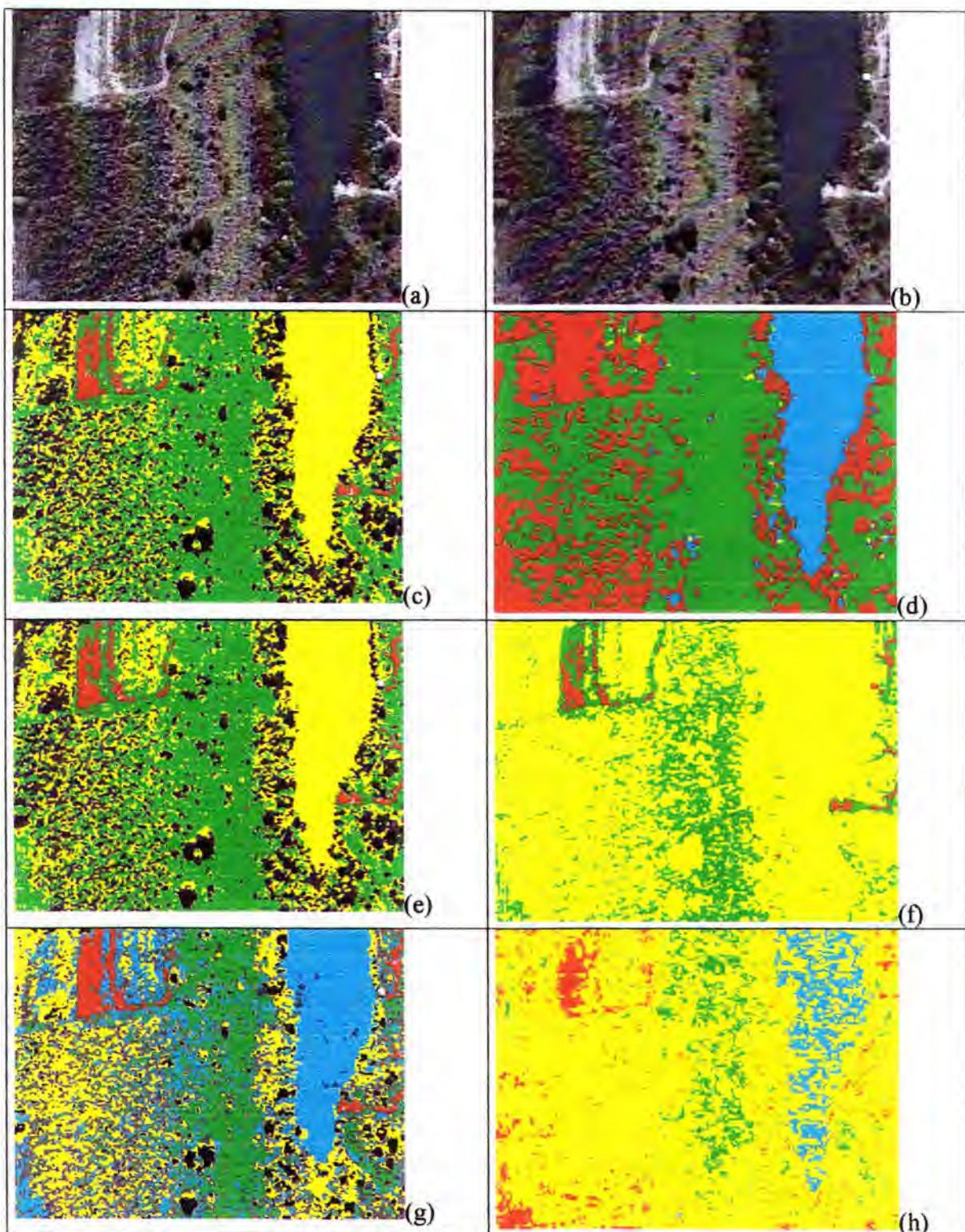


Figura 5.9: (a) Imagem obtida com avião (b) com a suavização, (c) a segmentação manual com o SIARCS, a segmentação automática (d) com o discriminador linear, (e) com rede MLP, treinada com três amostras de cada padrão, (f) com rede de Kohonen e rotulada pela mesma MLP, (g) com ART2, $\rho = 0,9$ e (h) a mesma ART2, $\rho = 0,7$.

Observando-se as imagens resultantes, pode-se afirmar que as mais suaves, como as da figura 5.8 obtidas com a torre de filmagem, apresentaram resultados bons em todos os

métodos. Ressalta-se, porém, que os ajustes na rede ART2 não foram simples de realizar. Uma característica especial nestas imagens foi o sombreamento feito (luz difusa) quando foram tomadas.

As imagens obtidas com o avião, com filmagem, foram as mais difíceis de analisar e ajustar os parâmetros, uma vez que a resolução das imagens é bem ruim. Um caso um pouco melhor foram as imagens também de avião, mas da câmera digital. Apesar de pouca definição para certos tipos de trabalhos e levando-se em conta a qualidade da câmera digital utilizada, algumas redes e técnicas conseguiram uma segmentação próxima da aceitável.

Uma análise quantitativa das segmentações efetuadas é apresentada no próximo item.

5.4 Avaliação quantitativa da cobertura do solo pelos diferentes métodos

Com o objetivo de avaliar de forma quantitativa os resultados obtidos, foram determinadas e comparadas as áreas das regiões segmentadas pelos diversos métodos.

Pode-se observar na tabela 5.1 os valores obtidos em porcentagem de área da imagem representando cada classe. Foram apenas contabilizadas as classificações como solo, planta e resíduos. Os pixels não classificados não foram considerados.

Na análise das imagens obtidas com a torre de filmagem, os resultados obtidos pelos diversos métodos podem ser considerados equivalentes. Isto comprova que o problema da condição de luz é relevante neste tipo de análise.

De uma forma geral, os métodos supervisionados se mostraram ligeiramente melhores, ou seja, se aproximaram um pouco melhor do resultado com o SIARCS, considerado como padrão de referência. Considerando os métodos não supervisionados, pode-se afirmar que a rede de kohonen apresentou resultados excelentes. A rede ART2, apesar de apresentar, em alguns casos, resultados razoáveis, apresenta dificuldades na escolha dos parâmetros de ajuste.

Tabela 5.1: Área de solo nu, cobertura vegetal e resíduos para as imagens das figuras 5.5 a 5.9, com os diferentes métodos.

Figura	SIARCS (%)	Mahalanobis (%)	MLP (%)	Kohonen (%)	ART2 ($\rho=0,9$) (%)	ART2 ($\rho=0,7$) (%)
5.5	Solo	42,47	30,11	26,53	33,06	20,43
	Planta	57,43	69,89	47,86	66,94	41,68
	Resíduo	0,0	0,0	0,0	0,0	0,0
5.6	Solo	60,99	63,68	58,53	55,65	3,81
	Planta	39,01	36,68	41,35	35,48	16,22
	Resíduo	0,0	0,0	0,0	24,14	81,61
5.7	Solo	23,76	22,57	10,97	29,15	25,68
	Planta	57,11	53,93	62,55	34,33	32,67
	Resíduo	18,97	23,50	20,71	24,00	0,0
5.8	Solo	27,98	27,89	24,74	20,12	18,84
	Planta	72,02	72,11	74,03	62,78	62,78
	Resíduo	0,0	0,0	0,0	0,01	0,0
5.9	Solo	2,62	31,64	1,53	2,62	3,84
	Planta	38,59	53,38	21,71	38,59	16,34
	Resíduo	35,18	0,39	76,75	35,18	55,04
						81,61

5.5 Considerações Finais

Todos os métodos de uma forma ou outra apresentaram resultados bons. Os mais difíceis de implementar foram os mapas de Kohonen e a rede ART2.

Os métodos não supervisionados são os mais difíceis de serem ajustados e tomam mais tempo no início da análise de cada conjunto de imagens. Os métodos supervisionados, apesar da fase inicial de treinamento, foram mais práticos para uso em conjuntos de imagens com características diferentes ou distintas.

No caso do modelo ART2, percebe-se que o mesmo apresenta pouca generalização, necessitando de pequenos ajustes para as diversas imagens. Se os padrões de cor da imagem se alteram muito, o modelo necessita de ajustes finos.

Através dos experimentos realizados neste trabalho, tornou-se possível perceber vantagens e desvantagens de cada método abordado para realizar a segmentação de imagens de cobertura do solo.

Alguns problemas inerentes à aquisição das imagens não foram tratados tais como, a influência da luminosidade, distorções e condições de obtenção das imagens, de tal forma viabilizar o uso em Agricultura de Precisão.

Pode-se destacar que:

- A abordagem que utiliza RNA apresenta cálculos bem menos complexos do que aqueles executados nos métodos estatísticos;
- Apesar da dificuldade inicial de entendimento dos parâmetros de ajuste de cada rede, esse processo é mais simples que em algumas das técnicas estatísticas;
- Uma vez que os resultados foram equivalentes aos métodos convencionais testados, e existe a possibilidade de implementação em hardware embarcado, as RNA demonstraram ser de grande interesse prático;
- Comparando-se os modelos de redes utilizados, o mais simples é o MLP, e apesar de supervisionado é de grande utilidade. O método de Kohonen, apesar de não ser supervisionado, retorna como saída um mapa com agrupamentos de pixels que de alguma forma devem ser rotulados nos padrões existentes. Esta é uma tarefa supervisionada, a não ser que sejam utilizadas técnicas de *clustering* convencionais.
- Outra questão está relacionada com a rede ART2. Apesar de algumas vezes esta rede apresentar excelentes resultados, sua generalização não é boa. Segundo artigos pesquisados existem outras variações que utilizam lógica Fuzzy e organização multi-camadas que podem resolver o problema.

No próximo capítulo são apresentadas as conclusões finais do trabalho e são sugeridos novos trabalhos em continuidade ao até aqui desenvolvido.

Capítulo 6

Conclusões

6.1 Considerações Iniciais

Neste trabalho foram apresentados alguns métodos de segmentação aplicados no estudo da cobertura de solo em imagens aéreas. Foram comparadas abordagens clássicas, que utilizam técnicas estatísticas bem conhecidas e abordagens alternativas, utilizando redes neurais artificiais.

Neste capítulo são apresentadas as conclusões finais do trabalho realizado. Também são apresentadas diversas sugestões para a sua continuidade. Entre essas sugestões destacam-se aquelas relacionadas com a implementação em hardware dedicado de um método selecionado para utilização embarcada no projeto ARARA.

6.2 Conclusões

As abordagens clássicas de análise das componentes principais, através da transformada de Hotelling e a análise de discriminantes lineares, através de Mahalanobis, foram implementadas e testadas em diversos conjuntos de imagens.

Os modelos de redes neurais avaliados demonstraram que as arquiteturas de RNA estudadas podem apresentar resultados próximos aos obtidos pelos métodos clássicos.

Os resultados de ambos os métodos no estudo da cobertura vegetal são bastante animadores, indicando que, apesar dos desafios, é possível incorporá-los em pacotes de tecnologia para a Agricultura de Precisão, tornado a Roça Cibernética uma realidade cada vez mais concreta.

6.1 Proposta de trabalhos futuros

Através dos experimentos e estudos comparativos realizados é possível apontar algumas vantagens e desvantagens de cada método abordado nesta dissertação. E principalmente poder vislumbrar propostas futuras para continuidade deste trabalho:

- ✓ Utilizar um número maior de imagens, assim como imagens em diferentes condições de aquisição com o objetivo de avaliar mais exaustivamente o desempenho das técnicas utilizadas;
- ✓ Proceder a implementação em hardware específico para cada rede, uma vez que existem várias opções comerciais de custo bem acessível, permitindo avaliar os ganhos de desempenho possíveis;
- ✓ Implementar variações dos modelos ART, tais como: FuzzyART, FuzzyARTMAP, ART3 comparando seu desempenho na segmentação de imagens com o dos métodos já estudados;
- ✓ Testar o modelo de Kohonen com marcadores morfológicos na segmentação de imagens coloridas;
- ✓ Utilizar o fato da transformada de Hotelling possibilitar uma redução na dimensão das imagens para ganhar tempo de processamento, ou mesmo, para envio das imagens processadas pelos enlaces de rádio que estão sendo usados no projeto ARARA;
- ✓ Implementar em hardware o módulo embarcado de segmentação de imagens do projeto ARARA.

Referências Bibliográficas

- [AND00] ANDRADE, A.O.; MARTINS, W. Redes neurais aplicadas ao processamento de imagens de satélite. <http://www.eee.ufg.br/paper.html> (dia 01/04/2001).
- [BAR00] BARBOSA, L.C.P. SINACOM - *Sistema de Navegação e Controle de Manobras do Projeto ARARA*; mini-dissertação para qualificação ao mestrado defendida no ICMC-USP, 55 pgs, São Carlos, Abril de 2000.
- [BAT96] BATISTA NETO, J.E.S. Techniques for computer-based anatomical segmentation of the Brain using MRI. *Thesis...* University of London. March 1996.
- [BEV96] BEVERLY, R.B. Video Image Analysis as a nondestructive measure of plant vigor for precision agriculture. *Commun. Soil Sci. Plant Anal.* v.27, n.3/4, p.607-614. 1996.
- [BRI95] BRITO JÚNIOR, A.M. *Separação de cores em uma imagem via transformada de Hotelling*. Projeto final de Computação Gráfica, UNICAMP-Campinas. 13p. 1995.
- [BRI96] BRITO JÚNIOR, A. S. *Segmentação por cor utilizando a transformada de Karhunen-Loève*. Dissertação de Mestrado, CEFET-Curitiba-PR. 1996.
- [CAE93] CAELLI, T.; REYE, D. On the classification of image regions by colour, texture and shape. *Pattern Recognition*. v.26, n.4, p.461-470, 1993.
- [CAL92] CALDAS, L.S.; BRAVO, C.; PICCOLO, H.; FARIA, C.R.S.M. Measurement of leaf area with a hand-scanner linked to a microcomputer. *Revista Brasileira de Fisiologia Vegetal*, Brasília, v.4, n.1, p.17-20, 1992.
- [CAM95a] CAMPBELL, N.W.; MACKEOWN, W.P.J.; THOMAS, B.T.; TROSCIANKO, T. The automatic classification of outdoor images. IN: BRITISH MACHINE VISION CONFERENCE. *Proceedings...* Birmingham, UK, Sept. 1995. P.297-306. 1995.

- [CAM95b] CAMPBELL, N.W.; THOMAS, B.T.; TROSCIANKO, T. Segmentation of Natural Images using Self-Organizing Feature Maps. IN: BRITISH MACHINE VISION CONFERENCE. *Proceedings...* Birmingham, UK, Sept. 1995.
- [CAM96] CAMPBELL, N.W.; THOMAS, B.T.; TROSCIANKO, T. Neural networks for the segmentation of outdoor images. IN: ENGINEERING APPLICATIONS OF NEURAL NETWORKS. *Proceedings...* London, 1996.
- [CAR97] CAREVIC, D.; CELLI, T. Region-based coding of color images using Karhunen-Loeve transform. *Graphical Models and Image Processing*, v.59, n.1, p.27-38, Jan. 1997.
- [CAR98] CARVALHO, A.C.P.L.F. Notas de aula de redes neurais. *Notas didáticas...*ICMC-USP, n.32. 1998
- [CRE94] CRESTANA, S.; GUIMARÃES, M.F.; JORGE, L.A.C.; RALISCH, R.; TOZZI, C.L.; TORRE-NETO, A.; VAZ, C.M.P. Avaliação da distribuição de raízes no solo auxiliada por processamento de imagens digitais. *Revista Brasileira de Ciência do Solo*, Campinas, v.18, n.3, p.365-371, set./dez. 1994.
- [CRE96] CRESTANA, S.; CRUVINEL, P.E.; MASCARENHAS, S.; BISCEGLI, C.I.; MARTIN-NETO, L.; COLNAGO, L.A., ed. *Instrumentação agropecuária: contribuições no limiar do novo século*. Brasília: EMBRAPA-SPI, 1996. 291p.
- [DUR87] DURRET, H.J. *Color and the computer*. Academic Press, 1987.
- [ELW76] ELWELL, H.A.; GARDNER, S. *Comparison of two techniques for measuring percent crop canopy of row crops in erosion research programmes*. Harare: Department of Conservation and Extention, Institute of Agricultural Engineering, 1976. 20p. (Research Bulletin, 19).
- [ELW77] ELWELL, H.A.; WENDELAAR, F.E. *To initiate a vegetal cover data bank for soil loss estimation*. Harare: Department of Conservation and Extention, Institute of Agricultural Engineering, 1977. 42p. (Research Bulletin, 23).

- [EMB96] (EMBRAPA, 1996) JORGE, L. A. C.; CRESTANA, S.; GUIMARÃES, M. F.; MARIA, I. C. *Sistema Integrado para Análise de Raízes e Cobertura do Solo - SIARCS*. 1996. Patente: Modelo de Utilidade. n. PI9604721-6, "SIARCS". 13 de abr. de 1998 (Depósito); 04 de set. de 1998 (Exame); 06 de nov. de 1999 (Concessão).
- [ENN95] ENNESER, F.; MEDIONI, G. Finding waldo, or focus of attention using local color information. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. V.17, n.8, p.805-809. Aug. 1995.
- [FRA95] FRANZ, E. GEBHARDT, M.R.; UNKLESBAY, K.B. Algorithms for extracting leaf boundary information from digital images of plant foliage. *Transactions of the ASAE*, v.38, n.2, p.625-633. 1995.
- [FUN95] FUNT, B.V.; FINLAYSON, G.D. Color constant color indexing. *IEEE Transactions on Pattern Analysis and machine Intelligence*. V.17, n.5, p.522-529. May 1995.
- [GIA00] GIACINTO, G.; ROLI, F.; BRUZZONE, L. Combination of neural and statistical algorithms for supervised classification of remote-sensing images. *Pattern Recognition Letters*. v.21, p.385-397. 2000.
- [GOM94] GOMES, J.; VELHO, L. Computação gráfica: imagem. IMPA-SBM, 1994.
- [GON95] GONG, Y.; SAKAUCHI, M. Detection of regions matching specified chromatic features. *Computer Vision and image Understanding*. V.61, n.2, p.263-269. Mac. 1995.
- [GON87] GONZALEZ, R.C. & WINTZ, P. *Digital Image Processing*. Addison Wesley, 1987.
- [HAF95] HAFNER, J.; SAWHNEY, H.S.; EQUITZ, W.; FLICKNER, M.; NIBLACK, W. Efficient color histogram indexing for quadratic form distance functions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v.17, n.7, p.729-736. Jul. 1995.
- [HAG90] HAGERUP, M.; SONDERGARD, I.; NIELSEN, J.K. measurements of areas consumed from leaf discs by chewing phytophagousinsects: Description of a new method onvolving image processing. *Entomologia Experimentalis et Applicata*. V.57, p.105-113, 1990.

- [HAR92] HARALICK, R.M.; SHAPIRO, L.G. *Computer and Robot Vision*, Addison-Wesley Publishing Company, Reading, Massachusetts. 1992.
- [HAL88] HALL, R. *Illumination and color in computer generated imagery*. Springer-Verlag, 1988.
- [HAR88] HARGROV, W.W.; CROSSLEY JR., D.A. Video digitizer for the rapid measurement of leaf area lost to herbivorous insects. *Annals of the Entomological Society of America*. V.81, p.593-598, 1988.
- [HAR78] HARWIG, R.O.; LAFLEN, J.M. A meterstick method for measuring crop residue cover. *Journal of Soil and Water Conservation*, v.33, n.2, p.90-91. 1978.
- [HEA92] HEALEY, G. Segmentation images using normalized color. *IEEE Transactions on Systems, Man and Cybernetics*. V.22, n.1, p.64-73. Jan./Fev. 1992.
- [HUA92] HUANG, C.; CHENG, T.; CHIEN, C. Color images segmentation using scale space filter and markov random field. *Pattern Recognition*, v. 25, n.10, p.1217-1229. 1992.
- [JAI96] JAIN, A.K.; VAILAYA, A. Image retrieval using color and shape. *Pattern Recognition*, v.29, n.8, p.1233-1244. 1996.
- [JAI89] JAIN, A.K. Fundamentals of digital image processing. *Prentice-Hall International Editions*, 1989.
- [JES99] JESUS A. *Uma abordagem morfológica para segmentação de logotipos em cheques bancários brasileiros sem conhecimento a priori*. Dissertação de Mestrado, Puc-Curitiba. 1999.
- [JON93] JONG VAN LIER, Q; SPAROVEK, G.; VASQUES FILHO, J. Análise de imagens utilizando um “scanner” manual: aplicações em agropecuária. *R. bras. Ci. Solo*, Campinas, 17:479-482, 1993.
- [JOR93] JORGE, L.A.C.; De MARIA, T.C.; LOMBERDI NETTO, F.; CRESTANA. S. Avaliação da cobertura do solo por culturas através de processamento de imagens. In: CONGRESSO BRASILEIRO DE CIÊNCIA DO SOLO. Cerrados: fronteira agrícola no século XXI, 24. V.3, Goiânia, 1993. *Resumo...* Goiânia, SBCS, 1993. P.139-140.

- [JOR99] JORGE, L.A.C.; SOUZA, N.P.; TRINDADE JR., O.; TOZZI, C.L; NETO, A.T. Studies of Soil Conservation for Precision Farming by Digital Color Image Analysis. IN: ASAE/CSAE-SCGR ANNUAL INTERNATIONAL MEETING, 1999. *Proceedings...* Toronto, Ontario Canadá. (CD-ROM).
- [KAM94] KAMPMANN, H.H.; HANSEN, O.B. Using colour image analysis for quantitative assessment of powdery mildew on cucumber. *Euphytica*. V.79, p.19-27, 1994.
- [KAN96] KANKANHALLI, M.S.; MEHTRE, B.M.; WU, J.K. Cluster-based color matching for image retrieval. *Pattern Recognition*, c.29, n.4, p.701-708. 1996.
- [KEH98] KEHTARNAVAZ, N.; MONACO, J.; NIMTSCHEK, J.; WEEKS, A. Color image segmentation using multi-scale clustering. In: IEEE SOUTHWEST SYMPOSIUM ON IMAGE ANALYSIS AND INTERPRETATION. April 1998, Tucson, Arizona, USA. *Proceedings. IEEE*. p.142-145. 1998. 1998.
- [KOH95] KOH, J.; SUK, M.; BHANDARKAR, S.M. A multilayer self-organizing feature map for range image segmentation. *Neural Networks*. V.8, n.1, p.67-86. 1995.
- [LI98] LI, F.; CHAPLIN, J. Evaluation of large area field sampling methods for crop residue coverage measurement. *Transactions of the ASAE*, v.41, n.3, p.645-651. 1998.
- [LIN83] LINDOW, S.E.; WEBB, R.R. Quantification of foliar plant disease symptoms by microcomputer-digitized video analysis. *Phytopathology*, v.73, p.520-524. 1983.
- [LIU94] LIU, J.; YANG, Y. Multiresolution color image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. V.16, n.7, p.689-700. Jul. 1994.
- [LIU99a] LIU, J.; WANG,L. LI, S.Z. MRMRF texture classification and MCMC parameter estimation. IN: VISION INTERFACE, *Proceedings...* Trois-Rivières, Canada, May. 1999.

- [LIU99b] LIU, X.; WANG, D.L. Range image segmentation using a relaxation oscillator network. *IEEE Transactions on Neural Networks*, v.10, n.3, may. 1999.
- [MAN69] MANNERING, J.V.; JOHNSON, C.B. Effect of crop row spacing on erosion and infiltration. *Agronomy Journal*, v.61, p.902-905. 1969.
- [MAR00] MARQUES, A.; SANTOS, G.R.; MARTINCI, A.; ARAUJO, M. Análise da cobertura vegetal da reserva biológica de UNA através de técnicas de geoprocessamento. IN: CONGRESSO BRASILEIRO DE CIÊNCIA DO SOLO, *Resumos...* 2000.
- [MAT98] MATHIAS, E.; CONCI, A. Comparing the influence of color space and metrics in content-based image retrieval. In: In: SIBGRAPI'98. Oct. 1998. *Proceedings...* SBC-Rio de Janeiro. p.371-378. 1998.
- [MEH95] MEHTRE, B.M.; KANKANHALLI, M.S.; NARASIMHALU, A.D. Color matching for image retrieval. *Pattern Recognition Letters*, v.16, p.325-331. 1995.
- [MER00] MERON, M.; COHEN, S.; MELMAN, G. Tree shape and volume measurement by light interception and aerial photogrammetry. *Transactions of ASAE*. v.43, n.2, p.475-481. 2000.
- [MOL97] MOLIN, J.P. Agricultura de precisão. PARTE I: O que é e estado da arte em sensoriamento. *Engenharia Agrícola*, v.17, n.2, p.97-107, dez. 1997
- [MOR97] MORAN, M.S.; INOUE, Y.; BARNES, E.M. Opportunities and limitations for image-based remote sensing in precision crop management. *Remote Sensing Environment*, v.61, p.319-346, 1997.
- [MOR95] MOREIRA, J.; COSTA, L.F. Multispectral image segmentation by chromaticity classification. IN: SIBGRAPI, VIII. *Anais...* SBC-Rio de Janeiro. p.119-125. 1995.
- [MOR96] MOREIRA, J.; COSTA, L.F. Neural-based color image segmentation and classification using self-organizing maps. In: SIBGRAPI, IX. Out. 1996. *Anais...* SBC-Rio de Janeiro. p.47-54. 1996.

- [NEV86] NEVATIA, R. Chapter 9: Image segmentation. In: YOUNG T.Y.; FU, K.S. *Handbook of pattern recognition and image processing*. Academic Press, p.215-232. 1986.
- [OHL78] OHLANDER, R.; PRICE, K.; REDDY, D.R. Picture segmentation using a recursive splitting method. *Comp. Graph. Image Process.*, v.8, p.313-333. 1978.
- [OHT80] OHTA, Y.; KANADE, T. SAKAI, T, Color Information for region segmentation. *Computer Graphics and Image Processing*. V.13, pp.222-241, 1980.
- [PAL99] PALMA, E.C; TRINDADE, JR., O., Uma Aeronave Rádio Controlada para Aplicação na Agricultura – Instrumentação de Bordo. IN: SIMPÓSIO DE INICIAÇÃO CIENTÍFICA DA USP, 7. *Anais Eletrônicos...* São Paulo, 1999.
- [PER86] PEREIRA, V.P.; MALHEIROS, E.B.; RAGAZZI, W.L.; ZANON, J.C. Comparação entre alguns métodos fotográficos e o método meterstick, para avaliação de cobertura de resíduos de culturas no solo, com modelos simulados. In: CONGRESSO BRASILEIRO DE CONSERVAÇÃO DO SOLO, 6; Campo Grande, 1986. *Resumo...* Campinas: Fund. Cargill, 1986. P.99.
- [PHI98] PHILLIPS, C.B.; JAIN, R.C. Local color image segmentation using singular value decomposition. In: IEEE SOUTHWEST SYMPOSIUM ON IMAGE ANALYSIS AND INTERPRETATION. April 1998, Tucson, Arizona, USA. *Proceedings. IEEE.* p.148-153. 1998.
- [PHI99] PHILIPP, S. Image segmentation: the state of the art In: Brazilian symposium on computer graphics and image processing, 12. UNICAMP, Campinas. *Minicursos/Tutorials ... Oct.1999.* 7p.
- [PIN99] PINTO, F.A.C.; REID, J.F.; ZHANG, Q.; NOGUCHI, N. Guidance parameter determination using artificial neural network classifier. IN: ASAE/CSAE-SCGR ANNUAL INTERNATIONAL MEETING, 1999. *Proceedings...* Toronto, Ontario Canadá. (CD-ROM).

- [PRA91] PRATT, W.K. *Digital Image processing*. Wiley-Interscience. New York, 1991.
- [PRE92] PRESS, W.H.; TEUKOLSKY, S.A.; VETTERLING, W.T.; TLANNERY, B.P. *Numerical recipes in C: The Art of Scientific Computing*. Cambridge University Press, Cambridge, 1992.
- [REZ00] REZENDE, S.B.; LANI, J.L.; MOREAU, M.S.; RIBEIRO, J.C. Aerofotos não convencionais de baixo custo – uma ferramenta no planejamento do uso do solo. IN: CONGRESSO BRASILEIRO DE CIÊNCIA DO SOLO. *Anais...* 2000. (Cd-ROM).
- [RIB00] RIBEIRO, R.E. *Um Sistema para Telemetria e Telecomando de Aeronaves do Projeto ARARA*; dissertação de mestrado defendida no ICMC-USP, 136 pgs, São Carlos, Maio de 2000.
- [ROG87] ROGERS, D.F. & ERNSHAW, R.A. *Techniques for computer graphics*. Spring-Verlag, 1987.
- [SAB96] SABER, E.; TEKALP, M.; ESCHIBACH, R.; KNOX, K. Automatic image annotation using adaptative color classification. *Graphical Models and Image Processing*. V.58, n.2, p.115-126. Mar. 1996.
- [SAL88] SALLAWAY, M.M.; LAWSON, D.; YULE, D.F. Ground cover during fallow from wheat, sorghum and sunflower stubble under three tillage practices in Central Queensland. *Soil & Tillage Research*, v.12, p.347-364, 1988.
- [SCH94] SCHETTINI, R. Multicolored object recognition and location. *Pattern Recognition Letters*, v.15, p.1089-1097. 1994.
- [SLO77] SLONEKER, L.L.; MOLDENHAUER, W.C. Measuring the amounts of crop residue remaining after tillage. *Journal of Soil and Water Conservation*, v.32, n.5, p.231-236. 1977.
- [SMI78] SMITH, A.R. Color gamut transform pairs. In: SIGGRAPH 78. *Proceedings*. p.12-19. 1978.
- [SMI91] SMITH, S.M.; SCHEREIER, H.E.; BROWN, S. Spatial analysis of forage parameters using geographic information system and image-analysis techniques. *Grass and Forage Science*, v.46, p.183-189, 1991.

- [SON99] SONKA, M.; HLAVAC, V.; BOYLE, R. *Image processing, analysis, and machine vision*. PWS Publishing, 1999.
- [SOU99] SOUZA, N.P. *Sistemas Alternativos para a Obtenção de Imagens Aéreas e sua Aplicação na Agricultura*; dissertação de mestrado defendida no ICMC-USP, 124 pgs, São Carlos, Março de 1999.
- [SOU98] SOUZA, N.P.; TRINDADE JR., O.; JORGE, L.A.C. Uso de Aeronaves Rádio Controladas na Obtenção de Fotografias Aéreas - Resultados Iniciais. IN: SIMPÓSIO NACIONAL DE INSTRUMENTAÇÃO AGROPECUÁRIA, 2. *Anais...* São Carlos, SP, 1998.
- [SPI90] SPILLMANN, L. & WERNER, J.S. *Visual perception: the neurophysiological foundations*. Academic Press, 1990.
- [STE99] STEENHOEK, L.; MISRA, M.; BATCHELOR, W.; DAVIDSON, J. Probabilistic neural networks for segmentation of features in corn kernel images. IN: ASAE/CSAE-SCGR ANNUAL INTERNATIONAL MEETING, 1999. *Proceedings...* Toronto, Ontario Canadá. (CD-ROM).
- [TAN94] TAN, T.S.C.; KITTLER, J. Colour texture analysis using colour histogram. IEE Proc. Vis. *Image Signal Process.* V.141, n.6, p.403-412. Dec. 1994.
- [TAN99] TANG, L.; TIAN, L.F.; STEWARD, B.L.; REID, J.F. Texture-based weed classification using Gabor Wavelets and Neural Network for real-time selective herbicide applications. IN: ASAE/CSAE-SCGR ANNUAL INTERNATIONAL MEETING, 1999. *Proceedings...* Toronto, Ontario Canadá. (CD-ROM).
- [TRE00] TREVIZANI, K.M. *Sistema Embocado para Transmissão de Dados e Vídeo Digital no Projeto ARARA*; mini-dissertação para qualificação ao mestrado defendida no ICMC-USP, 41 pgs, São Carlos, Abril de 2000.
- [TRI00] TRINDADE JR., O.; CARMO, M.S.; RIBEIRO, R.E. Small Aircraft for Application in Agriculture, IN: 2000 ASAE ANNUAL INTERNATIONAL MEETING. *Proceedings...* Milwaukee, Wisconsin, USA, 2000. (electronic format)

- [UCH94] UCHIYAMA, T.; ARBIB, M.A. Color image segmentation using competitive learning. *IEEE Transactions on Pattern Analysis and machine Intelligence*, v. 16, n.12, p.1197-1206. Dec. 1994.
- [VAR00] VARGAS, E.C.; ROMERO, R.A.F. Estudo de modelos de redes neurais para a indexação e recuperação de dados. *Notas didáticas...* ICMC-USP, São Carlos-SP, Ago. 2000. n.44, 32p.
- [VEH00] VEHTARI, A.; LAMPINEN, J. Bayesian MLP neural network for image analysis. *Pattern Recognition Letters*. Submitted 14th Sept. 2000.
- [VER97] VERRIKAS, A.; MALMQVIST, K.; BERGMAN, L. Colour image segmentation by modular neural network. *Pattern Recognition Letters*, v.18, p.173-185. 1997.
- [WEB97] WEBER, A. *Avaliação da proliferação celular baseada em PCNA-CICLINA por análise morfológica aplicada a linfomas malignos não-hodgkin*. Dissertação de Mestrado, CEFET-Curitiba-PR. 1997.
- [YIN00] YIN, X.; GUO, D.; XIE, M. Hand image segmentation using color and RCE neural network. *Robotics and Autonomous Systems*, v.903, p.1-17, 2001.
- [YON96] YONEKAWA, S; SAKAI, N.; KITANI, O. Identification of idealized leaf type using simple dimensionless shape factors by image analysis. *Transactions of the ASAE*, v.39, n.4, p.1525-1533. 1996.

Apêndice A

Modelos de Representação das Cores

O modelo espacial do sinal de cor associa para cada comprimento de onda uma medida da cor dada através de alguma das grandezas de energia radiante. Esse modelo espacial do sinal de cor é chamado usualmente em colorimetria de distribuição espectral. A distribuição espectral é de extrema importância no estudo de cor.

O uso do modelo tricromático para representar uma cor é largamente aceito, principalmente porque ele se basea no sistema visual humano. Cada um dos estímulos do modelo tricromático é traduzido em uma distribuição espectral representando a resposta visual do sistema para baixas, médias e altas freqüências do espectro visível. Nesta abordagem, as cores são representadas por um vetor no espaço tridimensional, chamado de espaço tricromático, que é definido pelas três componentes cromáticas primárias e a cor é formada pela combinação dessas componentes. A componente cromática primária R (*red*) corresponde às freqüências baixas, a primária G (*green*) às médias e a primária B (*blue*) às freqüências altas. A resposta espectral para cada um dos sensores (cones) da retina humana apresenta os picos de resposta ao estímulo exatamente nos comprimentos de onda das cores azul, verde e vermelha do espectro visível, de onde surgiu a teoria tricromática baseada nos três estímulos.

Um estímulo de cor é unicamente definido pela sua distribuição espectral, a qual descreve a cor em termos da distribuição de sua energia radiante dentro do espectro visível. Os valores de cada uma das componentes (estímulos) são obtidos pelo produto da função de sensibilidade espectral do sistema visual pela distribuição espectral do estímulo de cor, para todo o espectro visível.

A CIE - Comissão Internacional de Iluminação é o órgão responsável pela padronização na área de fotometria e colorimetria. Existem vários padrões de cor estabelecidos pela CIE, grande parte na década de 30.

Em 1931, a CIE, definiu um conjunto de funções de sensibilidade espectral para o sistema visual humano, que representam as funções de reconstrução de cores primárias para um observador. Estas ficaram conhecidas como sistema CIE-XYZ. Neste, uma cor visível é representada pela mistura aditiva das cores primárias do sistema, sendo estas cores primárias abstrações matemáticas, não existindo na realidade [ROG87]. Estas funções de reconstrução das cores foram obtidas através de experimentos, sendo estes repetidos com mais precisão em 1964 [HAL88].

O conjunto XYZ é constituído por um espaço tridimensional de cores que pode ser representado graficamente através de um sólido de cor, no qual define-se um sistema de coordenadas.

Pode-se utilizar diferentes sistemas de cor, atendendo às peculiaridades das diversas aplicações. Esse fato é um caso particular de um princípio geral: para cada problema procura-se o sistema de coordenadas mais adequado para o seu equacionamento e solução.

É de grande importância para a comparação de cor entre diferentes sistemas o uso de sistemas padrões que sejam definidos de forma independente de qualquer aplicação. O objetivo desses sistemas é estabelecer paradigmas que possam ser utilizados como base para definir sistemas de cor adequados às diversas aplicações.

Um aspecto importante da existência de diversos sistemas de cor é a possibilidade de fazer uma conversão das coordenadas de cor entre eles, o que se traduz matematicamente em uma mudança do sistema de coordenadas.

A definição de sistemas padrões tem por objetivo estabelecer sistemas universais de cor que permitam a especificação de cor de uma forma independente do dispositivo ou das dificuldades específicas de determinadas aplicações. Deste modo eles se tornam extremamente úteis nos problemas de comparação de cor, na definição de novos sistemas e no armazenamento da informação de cor associada a um determinado problema.

No sistema CIE-XYZ as cores podem ser representadas como vetores com direção e magnitude, passando pela origem (Figura A-1). Em outras palavras, a direção está associada à tonalidade e a magnitude à luminância. É interessante notar que XYZ são comumente associadas às altas, médias e baixas freqüências, também associadas com as tonalidades azul, verde e vermelho. Em geral, o espaço tricromático é também chamado de espaço RGB, por causa da associação das cores primárias com as faixas de freqüências.

Esse sólido é chamado de cubo RGB. A origem do cubo (0,0,0) corresponde à cor preta, e ponto de coordenadas (1,1,1) corresponde ao branco de luminância máxima (Figura A-2).

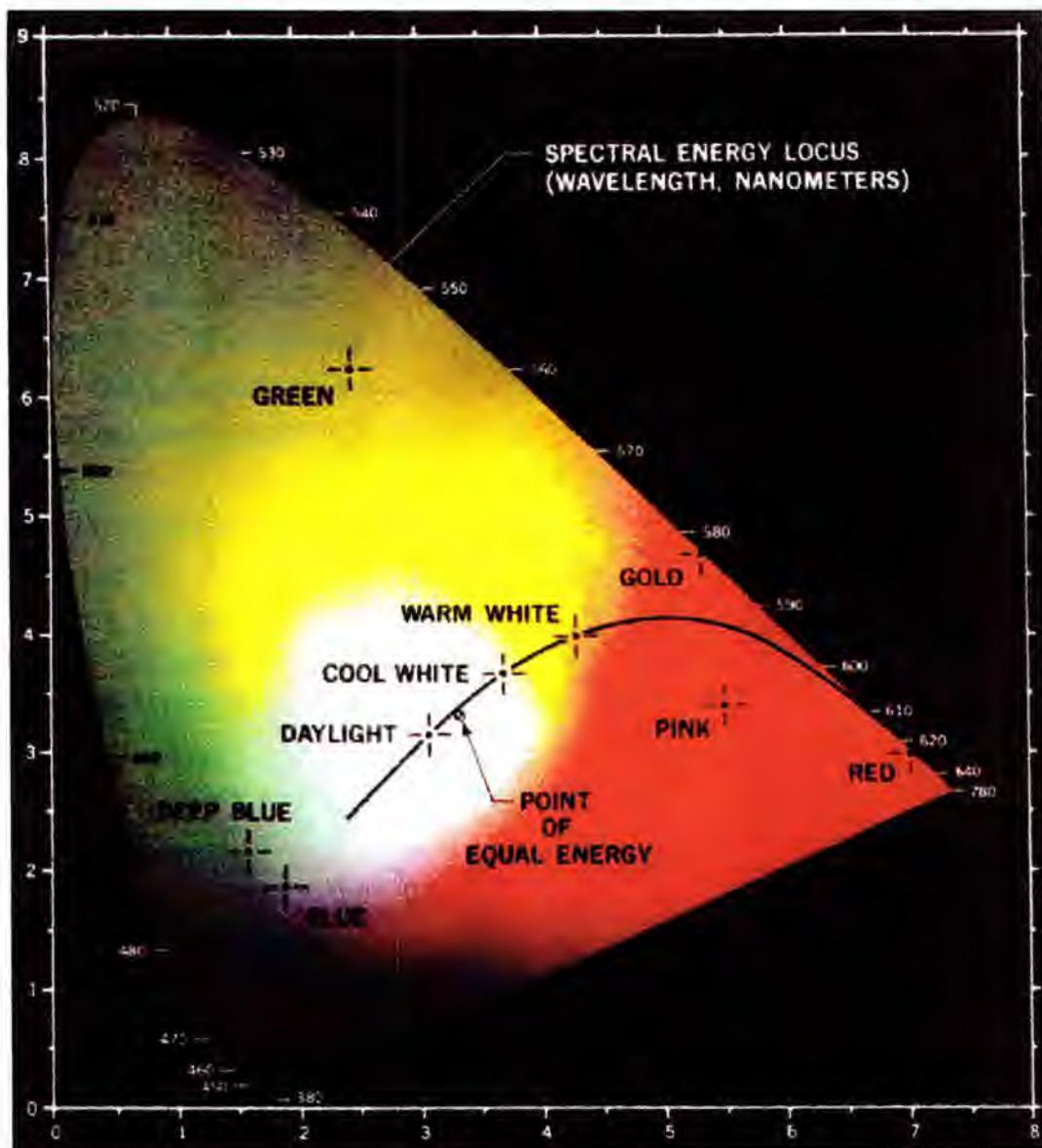


Figura A-1: Sistema CIE-XYZ

Os modelos de especificação de cor baseados no modelo de espaço vetorial, como o modelo RGB, apesar de práticos do ponto de vista computacional, são inadequados do ponto de vista de comunicação com o usuário, com relação à especificação simples e intuitiva de uma determinada cor. Um exemplo típico da dificuldade de se especificar cor no sistema RGB ocorre quando desejamos determinar uma cor de tonalidade e luminância conhecidas, e precisamos apenas a saturação, isto é, diminuir ou aumentar a quantidade de branco na cor. Para realizar essa alteração as três componentes de cor não têm qualquer relação intuitiva com os parâmetros perceptivos de tonalidade, luminância e saturação.

Esses três parâmetros estão mais próximos do modo como efetuamos a escolha de cor, e devemos procurar modelos nos quais a especificação de um determinado estímulo de cor está mais próxima desse paradigma.

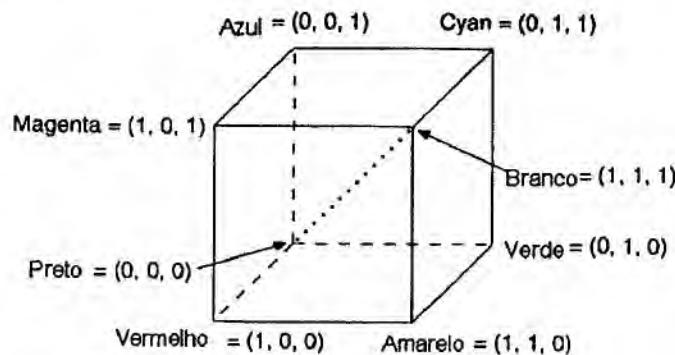


Figura A-2: Cubo RGB

Vários modelos de representação de cor baseiam-se nos conceitos perceptivos de tonalidade, saturação e brilho. De um ponto de vista puramente descritivo, o sólido de cor desses modelos é um subconjunto tridimensional, parametrizado de modo que suas coordenadas, que correspondem à tonalidade, saturação e brilho, variam como em um sistema de coordenadas cilíndricas. A tonalidade varia ao longo dos círculos horizontais, a saturação varia na direção radial e o brilho varia na direção ortogonal ao plano tonalidade-saturação (Figura A-3).

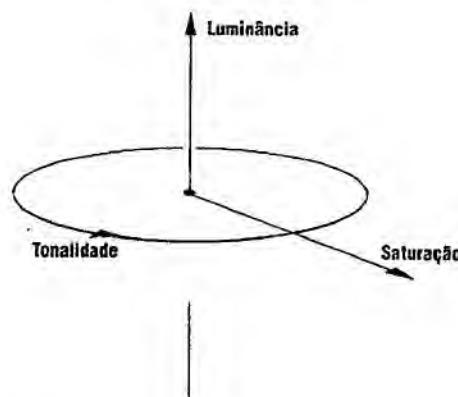


Figura A-3: Modelo típico perceptivo

Várias propostas de modelos perceptivos baseados em coordenadas cilíndricas foram sugeridas, e todos eles são modelos pseudo-perceptivos. Em muitos desses modelos a

componente acromática ($R+G$) representa a luminância ou brilho e a razão das componentes cromáticas ($R-G$, $Y-B$), definem a matiz. A combinação das componentes cromáticas com as acromáticas define a saturação.

Sistemas perceptivos padrão também foram propostos pela CIE e são aceitos por muitos cientistas. Os sistemas mais utilizados são o CIE-LUV e o CIE-LAB, ambos derivados do CIE-XYZ. Estes sistemas são chamados de sistemas uniformes, pois apresentam uniformidade no espaço de cor criado.

Diversos outros sistemas perceptivos foram criados a partir do sistema RGB. Dentre os mais difundidos pode-se destacar aqueles que usam como representação um sólido de cor do tipo cone com base hexagonal, de acordo com a variação da tonalidade, saturação e do brilho (Figura A-4). Uma variação com duplo cone pode ser visto na figura A-5.

O sistema proposto por Smith (1978), o HSV (*hue, saturation, value*), também chamado de HSB, sendo B para o brilho, é um exemplo deste tipo de representação. Em Smith [SMI78] são apresentados os algoritmos utilizados neste trabalho para conversão do sistema RGB para HSV e do HSV para o RGB.

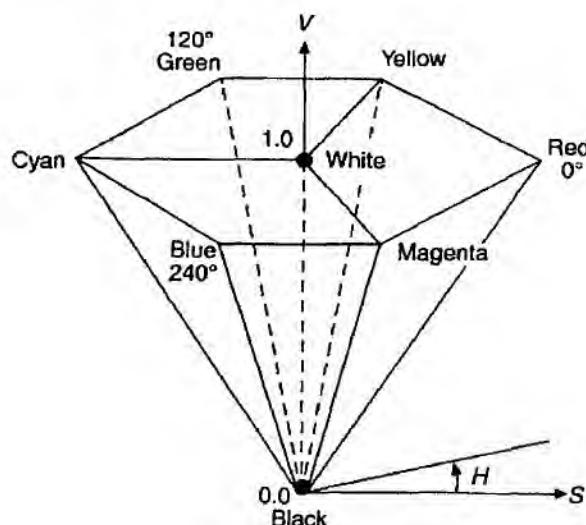


Figura A-4: Sistema HSV



Figura A-5: Duplo cone para sistema HSB