

UNIVERSIDADE DE SÃO PAULO
ESCOLA DE ENGENHARIA DE SÃO CARLOS
DEPARTAMENTO DE ENGENHARIA MECÂNICA

LUCIANO CÁSSIO LULIO

*Técnicas de visão computacional aplicadas ao reconhecimento de
cenas naturais e locomoção autônoma em robôs agrícolas móveis*

São Carlos

2011

LUCIANO CÁSSIO LULIO

Técnicas de visão computacional aplicadas ao reconhecimento de cenas naturais e locomoção autônoma em robôs agrícolas móveis

Dissertação apresentada à Escola de Engenharia de São Carlos da Universidade de São Paulo, como parte dos requisitos para obtenção do título de Mestre em Engenharia Mecânica.

Área de concentração: Manufatura
Orientador: Mário Luiz Tronco

ESTE EXEMPLAR TRATA-SE
DA VERSÃO CORRIGIDA.
A VERSÃO ORIGINAL
ENCONTRA-SE DISPONÍVEL
JUNTO AO DEPARTAMENTO
DE ENGENHARIA
MECÂNICA DA EESC-SP

São Carlos

2011

AUTORIZO A REPRODUÇÃO E DIVULGAÇÃO TOTAL OU PARCIAL DESTE TRABALHO, POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO, PARA FINS DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

Ficha catalográfica preparada pela Seção de Tratamento da Informação do Serviço de Biblioteca – EESC/USP

L955t Lulio, Luciano Cássio.
Técnicas de visão computacional aplicadas ao reconhecimento de cenas naturais e locomoção autônoma em robôs agrícolas móveis. / Luciano Cássio Lulio ; orientador Mário Luiz Tronco. São Carlos, 2011.

Dissertação (Mestrado - Programa de Pós-Graduação em Engenharia Mecânica e Área de Concentração em Manufatura)— Escola de Engenharia de São Carlos da Universidade de São Paulo, 2011.

1. Manufatura. 2. Robôs móveis agrícolas. 3. Reconhecimento de cenas naturais. 4. Segmentação de imagens. 5. Classificação cognitiva e estatística. I. Título.

FOLHA DE JULGAMENTO

Candidato: Bacharel **LUCIANO CÁSSIO LULIO**

Título da dissertação: **Técnicas de visão computacional aplicadas ao reconhecimento de cenas naturais e locomoção autônoma em robôs agrícolas móveis.**

Data da defesa: 09/08/2011:

Comissão Julgadora:

Prof. Dr. **Mário Luiz Tronco (Orientador)**
(Escola de Engenharia de São Carlos/EESC)

Prof. Titular **Arthur José Vieira Porto**
(Escola de Engenharia de São Carlos/EESC)

Prof. Dr. **Emerson Carlos Pedrino**
(Universidade Federal de São Carlos/UFSCar)

Resultado:

APROVADO

APROVADO.

Aprovado.

Coordenador do Programa de Pós-Graduação em Engenharia Mecânica:
Prof. Associado **Marcelo Areias Trindade**

Presidente da Comissão de Pós-Graduação:
Prof. Associado **Paulo Cesar Lima Segantine**

Dedicatória

Aos meus pais.

Agradecimentos

Ao Prof. Dr. Mário Luiz Tronco, pela orientação nos propósitos desta pesquisa.

Ao Prof. Dr. Arthur José Vieira Porto, pela oportunidade de desenvolver este trabalho no Núcleo de Ensino e Pesquisa em Automação e Simulação (NEPAS/EESC/USP).

A CAPES pelo apoio financeiro, à Escola de Engenharia de São Carlos e ao Departamento de Engenharia Mecânica, pelo suporte no desenvolvimento do mestrado acadêmico.

Aos amigos, professores e funcionários do NEPAS, pela amizade e ideias discutidas.

Aos meus familiares, pelo apoio e incentivo.

A todos que, direta ou indiretamente, me apoiaram neste trabalho.

“Run, rabbit run. Dig that hole, forget the sun. And when at last the work is done.

Don't sit down; it's time to dig another one”.

Gilmour, Waters, Wright (Breathe/Pink Floyd)

Resumo

LULIO, L. C. **Técnicas de visão computacional aplicadas ao reconhecimento de cenas naturais e locomoção autônoma em robôs agrícolas móveis.** São Carlos, 353 p. Dissertação (Mestrado) – Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2011.

O emprego de sistemas computacionais na Agricultura de Precisão (AP) fomenta a automação de processos e tarefas aplicadas nesta área, precisamente voltadas à inspeção e análise de culturas agrícolas, e locomoção guiada/autônoma de robôs móveis. Neste contexto, no presente trabalho foi proposta a aplicação de técnicas de visão computacional nas tarefas citadas, desenvolvidas em abordagens distintas, a serem aplicadas em uma plataforma de robô móvel agrícola, em desenvolvimento no NEPAS/EESC/USP. Para o problema de locomoção do robô (primeira abordagem), foi desenvolvida uma arquitetura de aquisição, processamento e análise de imagens com o objetivo de segmentar, classificar e reconhecer padrões de navegação das linhas de plantio, como referências de guiagem do robô móvel, entre plantações de laranja, milho e cana. Na segunda abordagem, tais técnicas de processamento de imagens são aplicadas também na inspeção e localização das culturas laranja (primário) e milho (secundário), para análise de suas características naturais, localização e quantificação. Para as duas abordagens, a estratégia adotada nas etapas de processamento de imagens abrange: filtragem no domínio espacial das imagens adquiridas; pré-processamento nos espaços de cores RGB e HSV; segmentação não supervisionada JSEG customizada à quantização de cores em regiões não homogêneas nestes espaços de cores; normalização e extração de características dos histogramas das imagens pré-processadas para os conjuntos de treinamento e teste através da análise das componentes principais; reconhecimento de padrões e classificação cognitiva e estatística. A metodologia desenvolvida contemplou bases de dados para cada abordagem entre 700 e 900 imagens de cenas naturais sob condições distintas de aquisição, apresentando resultados significativos quanto ao algoritmo de segmentação nas duas abordagens, mas em menor grau em relação à localização de gramíneas, sendo que os milhos requerem outras técnicas de segmentação, que não aplicadas apenas em quantização de regiões não homogêneas. A classificação estatística, *Bayes* e *Bayes Ingênuo*, mostrou-se superior à cognitiva RNA e *Fuzzy* nas duas abordagens, e posterior construção dos mapas de classe no espaço de cores HSV. Neste mesmo espaço de cores, a quantificação e localização de frutos apresentaram melhores resultados que em RGB. Com isso, as cenas naturais nas duas abordagens foram devidamente processadas, de acordo com os materiais e métodos empregados na segmentação, classificação e reconhecimento de padrões, fornecendo características intrínsecas e distintas das técnicas de visão computacional propostas a cada abordagem.

Palavras chave: Robôs móveis agrícolas, reconhecimento de cenas naturais, segmentação de imagens, classificação cognitiva e estatística.

Abstract

LULIO, L. C. **Computer vision techniques applied to natural scenes recognition and autonomous locomotion of agricultural mobile robots.** São Carlos, 353 p. Dissertation (Master of Science) – School of Engineering of São Carlos, University of São Paulo, São Carlos, 2011.

The use of computer systems in Precision Agriculture (PA) promotes the processes' automation and its applied tasks, specifically the inspection and analysis of agricultural crops, and guided/autonomous locomotion of mobile robots. In this context, it was proposed in the present work the application of computer vision techniques on such mentioned tasks, developed in different approaches, to be applied in an agricultural mobile robot platform, under development at NEPAS/EESC/USP. For agricultural mobile robot locomotion, an architecture for the acquisition, image processing and analysis was built, in order to segment, classify and recognize patterns of planting rows, as references way points for guiding the mobile robot. In the second approach, such image processing techniques were applied also in the inspection and location of the orange crop (primary) and maize crop (secondary) aiming its natural features, location and quantification. For both mentioned approaches, the adopted image processing steps include: filtering in the spatial domain for acquired images; pre-processing in RGB and HSV color spaces; JSEG unsupervised segmentation algorithm, applied to color quantization in non-homogeneous regions; normalization and histograms feature extraction of preprocessed images for training and test sets, fulfilled by the principal components analysis (PCA); pattern recognition and cognitive and statistical classification. The developed methodology includes sets of 700 and 900 images' databases for each approach of natural scenes under different conditions of acquisition, providing great results on the segmentation algorithm, but not as appropriate as in the location of maize grass, considering other segmentation techniques, applied not only in the quantization of non-homogeneous regions. Statistical classification, *Bayes* and *Naive Bayes*, outperforms the cognitives *Fuzzy* and ANN on two approaches and subsequent class maps construction in HSV color space. Quantification and localization of fruits had more accurate results in HSV than RGB. Thus, natural scenes in two approaches were properly processed, according to the materials and methods employed in segmentation, classification and pattern recognition, providing intrinsic and different features of the proposed computer vision techniques to each approach.

Keywords: Agricultural mobile robots, natural scenes recognition, image segmentation, cognitive and statistical classifiers.

Lista de figuras

Figura 2.1: Estrutura sequencial de um sistema de visão computacional.	47
Figura 2.2: Modelo <i>pinhole</i> da câmera perspectiva com dois sistemas de coordenadas: externa (O_w) e interna (O_c) (CYGANNEK e SIEBERT, 2009).	49
Figura 2.3: (a) tensões aplicadas na camada de óxido e direção de transferência de carga;(b) estrutura da matriz de transferência dos registradores de <i>pixels</i> ; (c) diagrama elétrico simplificado com os transistores MOS, sensores de ativação e portas de transferência de carga para deslocamento analógico através da lógica <i>FDA</i> (Adaptado de EASTMAN KODAK, 2001).	52
Figura 2.4: Divisão dos <i>pixels</i> em uma imagem colorida para determinada parte da imagem. Cada <i>pixel</i> corresponde a um determinado quadrado, para um valor quantizado $f(x,y)$, cuja atribuição é a mesma para imagens monocromáticas.	54
Figura 2.5: Divisão de regiões categorizadas em subconjuntos da imagem R	59
Figura 2.6: Eixo-maior, eixo-menor, retângulo básico e diâmetro do eixo maior (d_{max}).	66
Figura 2.7: Código de cadeia (3211076755325) e direções dos segmentos baseada em região 8-conectada.	67
Figura 2.8: Número de formas, código de cadeias e diferenças para uma região de contorno.	68
Figura 2.9: Representação do contorno por momentos estatísticos, com números completos.	69
Figura 2.10: Área com regiões de elementos discretizados na imagem.	70
Figura 2.11: Classificação por exemplos (Adaptado de CAVANI, 2007).	75
Figura 2.12: Modelo de neurônio artificial <i>McCulloch e Pitts</i>	78
Figura 2.13: Diagrama de blocos de um controlador <i>Fuzzy</i> : (<i>EP</i> : entradas precisas; <i>SP</i> : saídas precisas).	81
Figura 2.14: Diagrama de blocos para o problema e solução da locomoção de robôs móveis (Adaptado de SIEGWART <i>et al.</i> , 2011).	89
Figura 2.15: Grade de ocupação, com as células representando o espaço ocupado das células.	97
Figura 2.16: Informações métricas adicionadas aos mapas topológicos (Adaptado de SILVA, 2008).	99
Figura 2.17: Processo de detecção e classificação de frutos.	104

Figura 2.18: (a) Aproximação poligonal de falha em superfície – cada nível de cinza representa uma borda diferente; (b-c) Área de segmentação para dois tipos de irregularidades nos frutos (LOPEZ <i>et al.</i> , 2009).	104
Figura 2.19: Detalhes do tamanho, forma e cor para as doenças estudadas (BLASCO <i>et al.</i> , 2009).	105
Figura 2.20: Segmentações de maçãs, tomates, laranjas e morangos (YIN <i>et al.</i> , 2008).	108
Figura 2.21: Imagem original, após filtragem e operação morfológica (WEI <i>et al.</i> , 2009)... ..	110
Figura 2.22: Caracterização da imagem (imagem original, imagem em tons de cinza e extração do caule (LIMING e TANCHAO, 2010).	111
Figura 2.23: Extração de características da imagem (a- imagem original, b- imagem segmentada, c- detecção de bordas, d- extração de características pelo método de distribuição de linhas) (LIMING e TANCHAO, 2010).	112
Figura 2.24: Trajetória da câmera e posições dos <i>landmarks</i> (WU <i>et al.</i> , 2009).	115
Figura 2.25: Algoritmo de detecção de linhas de guiagem, e teste real com o robô móvel na linha de plantio de vegetais (Adaptado de XUE e JU, 2010).	116
Figura 2.26: Aplicação do algoritmo proposto com as imagens originais, binarizadas e discriminadas após segmentação. (a) linha de plantio espaçada; (b) linha de plantio contínua; (c) ambas (XUE e JU, 2010).	117
Figura 2.27: Processamento de imagens: (a) imagem original; (b) imagem retificada; (c) imagem em tons de cinza pós cálculo da representação das componentes RGB; (d) seleção dos segmentos; (e) imagens combinadas limiarizadas; (f) imagens combinadas sobrepostas; (g) imagem original sobreposta;(h) acumulador <i>Hough</i> (BAKKER <i>et al.</i> , 2008).	119
Figura 2.28: Resultados do SLAM usando o algoritmo de otimização em lotes: os pontos vermelhos indicam a estimativa do mapa pelo algoritmo proposto;os pontos verdes indicam as estimativas do mapa pela visão em EKF; as elipses verdes e vermelhas representam as estimativas de covariância e o emprego do algoritmo, respectivamente (HERATH <i>et al.</i> , 2008).	121
Figura 2.29: Trajetória desejada do robô móvel por um círculo de raio 600mm: (a) posição inicial conhecida; (b) o robô mantém o localizador inativo e procura por <i>landmarks</i> a fim de estimar a posição global; (c) limite 3σ no erro em x , y e θ para a posição inicial conhecida; (d) melhoria do posicionamento, quando o algoritmo é aplicado (Adaptado de BAIS <i>et al.</i> , 2008).	123
Figura 2.30: Instalação da câmera e sistema de coordenadas; CG – centro de gravidade; RI – região de interesse do mapa de elevação; x_d (longitude), y_d (<i>offset</i> horizontal) e z_d (<i>offset</i> vertical) representam a localização da câmera no sistema de coordenadas; α , β e γ representam os ângulos de ação da câmera em <i>roll</i> , <i>tilt</i> e <i>pan</i> (KISE <i>et al.</i> , 2005).	124
Figura 2.31: (a) imagem original; (b) imagem com disparidade de <i>pixels</i> ; (c) mapa de elevação das linhas de plantio; (d) mapa de elevação suavizado - estéreo (Adaptado de KISE <i>et al.</i> , 2005).	125

Figura 2.32: Localização dos pontos de características após remodelagem do sensor óptico estéreo: (a) os 400 maiores gradientes de todos os <i>pixels</i> foram selecionados como pontos de características; (b) como resultado da localização de tais características, 89 pontos correspondentes foram encontrados na imagem adjacente, fornecendo maior exatidão na descrição das linhas de plantio (KISE e ZHANG, 2008).....	125
Figura 2.33: Resultado do processamento de imagens: (a) imagem original; (b) transformação da imagem em níveis de cinza; (c) binarização de OTSU; (d) projeção vertical para obter uma linha de plantio apenas; (e) detecção de bordas da linha de plantio; (f) transformada de <i>Hough</i> para delimitar a área navegável (JIANG e ZHAO, 2010).....	127
Figura 3.1: Estratégia modular para o processamento de imagens em cenas naturais.	130
Figura 3.2: Cenas das linhas de plantio de pomares.	133
Figura 3.3: Cenas de pomares individuais para toda a árvore e áreas locais de incidência de frutos.	134
Figura 3.4: (a) Milharal verde em período de colheita; (b) Milharal silagem em período de colheita; (c) Linhas de plantio de milho em período de safra; (d) Linhas de plantio de milho em período de colheita.	135
Figura 3.5: Linhas de plantio de canaviais em período de colheita (a) e (b) e em período de safra (c).....	136
Figura 3.6: Estrutura da primeira abordagem para o processamento e obtenção dos segmentos classificados.....	137
Figura 3.7: Estrutura da segunda abordagem para o processamento e obtenção dos segmentos classificados.....	138
Figura 3.8: Estratégia do algoritmo de segmentação não-supervisionado JSEG.	142
Figura 3.9: Fluxograma do algoritmo de distribuição espacial.	145
Figura 3.10: Histograma imagem quantizada 01 RGB.....	150
Figura 3.11: Histograma imagem quantizada 02 RGB.	151
Figura 3.12: Histograma imagem quantizada 03 RGB.....	152
Figura 3.13: Histograma imagem quantizada 01 HSV.	153
Figura 3.14: Histograma imagem quantizada 02 HSV.	154
Figura 3.15: Histograma imagem quantizada 03 HSV.	155
Figura 3.16: Etapas de extração e normalização de características pela ACP nos conjuntos de treinamento, teste e validação.....	161
Figura 3.17: Topologia RNA à primeira abordagem.	163
Figura 3.18: Topologia RNA à segunda abordagem.	163

Figura 3.19: Sistema de inferência <i>fuzzy</i> para a primeira abordagem.....	165
Figura 3.20: Sistema de inferência <i>fuzzy</i> para a segunda abordagem.	168
Figura 3.21: Algoritmo customizado à classificação pelo teorema de <i>Bayes</i>	174
Figura 4.1: Imagens adquiridas e processadas nos espaços de cores RGB e HSV: (a) linha de plantio de pomares; (b) linha de plantio de cana; (c) linha de plantio de milho; (d) plantação de milho; (e) plantação de laranja.	181
Figura 4.2: Imagem original à esquerda e quantizada em regiões de cores por níveis de cinza (direita). Observa-se que existem regiões na área navegável não desejados, posteriormente agrupados numa classe de padrões.	182
Figura 4.3: Imagens à esquerda mostram a segmentação (distribuição espacial) final, com os parâmetros inicialmente ajustados para encontrar a maior parte de irregularidades de <i>pixels</i> ; à direita, as mesmas cenas com reajuste dos parâmetros para agregar a maior parte das regiões em concentrações simplificadas aos classificadores.	184
Figura 4.4: Cenas de pomares originais (coluna 1); imagens com cores quantizadas (coluna 2); imagens segmentadas (coluna 3).	185
Figura 4.5: Plantação de milho em período de colheita: as duas primeiras linhas de imagens denotam a colheita em silagem e as duas últimas linhas denotam a colheita em milho verde.	186
Figura 4.6: Distribuição dos vetores de características à abordagem 1.	188
Figura 4.7: Distribuição dos vetores de características à abordagem 2.	189
Figura 4.8: Quantidade de dimensões a cada conjunto – pomares RGB.	190
Figura 4.9: Quantidade de dimensões a cada conjunto – pomares HSV.	190
Figura 4.10: Quantidade de dimensões a cada conjunto – canaviais RGB.	191
Figura 4.11: Quantidade de dimensões a cada conjunto – canaviais HSV.	192
Figura 4.12: Quantidade de dimensões a cada conjunto – milharais RGB.	192
Figura 4.13: Quantidade de dimensões a cada conjunto – milharais HSV.	193
Figura 4.14: Quantidade de dimensões a cada conjunto – laranja RGB.	194
Figura 4.15: Quantidade de dimensões a cada conjunto – laranja HSV.	194
Figura 4.16: Variação dos estados de inferência das entradas e saídas para as regras formuladas à primeira abordagem.	197
Figura 4.17: Variação dos estados de inferência das entradas e saídas para as regras formuladas à segunda abordagem.	198
Figura 4.18: Parâmetros à mistura do treinamento – Pomares RGB.	205

Figura 4.19: Parâmetros à mistura do treinamento – Pomares HSV.	206
Figura 4.20: Parâmetros à mistura do treinamento – Canaviais RGB.	206
Figura 4.21: Parâmetros à mistura do treinamento – Canaviais HSV.	207
Figura 4.22: Parâmetros à mistura do treinamento – Milharais RGB.....	208
Figura 4.23: Parâmetros à mistura do treinamento – Milharais HSV.....	208
Figura 4.24: Parâmetros à mistura do treinamento – Laranja RGB.....	209
Figura 4.25: Parâmetros à mistura do treinamento – Laranja HSV.....	210
Figura 4.26: Taxa de acerto para os conjuntos de treinamento – abordagem 1 (pomares)...	212
Figura 4.27: Taxa de acerto para os conjuntos de treinamento – abordagem 1 (canaviais)...	213
Figura 4.28: Taxa de acerto para os conjuntos de treinamento – abordagem 1 (milharais)...	213
Figura 4.29: Taxa de acerto para os conjuntos de treinamento – abordagem 2 (laranja).	214
Figura 4.30: Imagem original RGB e HSV (laranja 01).	215
Figura 4.31: Mapa de classe 1 (Bayes RGB e HSV).	215
Figura 4.32: Mapa de classe 1 (Bayes ingênuo RGB e HSV).....	215
Figura 4.33: Mapa de classe 1 (Fuzzy preditivo RGB e HSV).	216
Figura 4.34: Mapa de classe 1 (RNA RGB e HSV).	216
Figura 4.35: Imagem original RGB e HSV (pomares 01).....	216
Figura 4.36: Mapa de classe 1 (Bayes RGB e HSV).....	216
Figura 4.37: Mapa de classe 1 (Bayes ingênuo RGB e HSV).....	217
Figura 4.38: Mapa de classe 1 (Fuzzy preditivo RGB e HSV).	217
Figura 4.39: Mapa de classe 1 (RNA RGB e HSV).	217
Figura 4.40: Imagem original RGB e HSV (milharais 01).	217
Figura 4.41: Mapa de classe 1 (Bayes RGB e HSV).	218
Figura 4.42: Mapa de classe 1 (Bayes ingênuo RGB e HSV).....	218
Figura 4.43: Mapa de classe 1 (Fuzzy preditivo RGB e HSV).	218
Figura 4.44: Mapa de classe 1 (RNA RGB e HSV).	219
Figura 4.45: Imagem original RGB e HSV (canaviais 01).	219

Figura 4.46: Mapa de classe 1 (Bayes RGB e HSV).	219
Figura 4.47: Mapa de classe 1 (Bayes ingênuo RGB e HSV).	219
Figura 4.48: Mapa de classe 1 (Fuzzy preditivo RGB e HSV).	220
Figura 4.49: Mapa de classe 1 (RNA RGB e HSV).	220
Figura 4.50: Mapas de classe RGB (esquerda) e HSV (direita) – cena 1.	222
Figura 4.51: Métricas próximas do limiar geométrico circular 1,0 para RGB (esquerda) e HSV (direita) – cena 1.	223
Figura 4.52: Representação da área e centróide dos objetos para RGB (esquerda) e HSV (direita) – cena 1.	223
Figura 4.53: Rótulo da associação do elemento fruto nos objetos encontrados aos dois casos – cena 1.	223
Figura 4.54: Mapas de classe RGB (esquerda) e HSV (direita) – cena 2.	223
Figura 4.55: Métricas próximas do limiar geométrico circular 1,0 para RGB (esquerda) e HSV (direita) – cena 2.	224
Figura 4.56: Representação da área e centróide dos objetos para RGB (esquerda) e HSV (direita) – cena 2.	224
Figura 4.57: Rótulo da associação do elemento fruto nos objetos encontrados aos dois casos – cena 2.	224
Figura 4.58: Mapas de classe RGB (esquerda) e HSV (direita) – cena 3.	225
Figura 4.59: Métricas próximas do limiar geométrico circular 1,0 para RGB (esquerda) e HSV (direita) – cena 3.	225
Figura 4.60: Representação da área e centróide dos objetos para RGB (esquerda) e HSV (direita) – cena 3.	225
Figura 4.61: Rótulo da associação do elemento fruto nos objetos encontrados aos dois casos – cena 3.	226
Figura A.1: Gráfico da entrada “RDDL” para os estados de inferência <i>acutângulo</i> , <i>retângulo</i> e <i>obtusângulo</i> , à primeira abordagem.	231
Figura A.2: Gráfico da entrada “RDEL” para os estados de inferência <i>acutângulo</i> , <i>retângulo</i> e <i>obtusângulo</i> , à primeira abordagem.	231
Figura A.3: Gráfico da entrada “RDDF” para os estados de inferência <i>escaleno</i> , <i>isósceles</i> e <i>equilátero</i> , à primeira abordagem.	232
Figura A.4: Gráfico da entrada “RDEF” para os estados de inferência <i>escaleno</i> , <i>isósceles</i> e <i>equilátero</i> , à primeira abordagem.	232

Figura A.5: Gráfico da saída “AN” para os estados de inferência <i>tex_P/amarelo_C</i> e <i>tex_P/marrom_C</i> , à primeira abordagem.	232
Figura A.6: Gráfico da saída “AnN” para os estados de inferência <i>plantação_E</i> , <i>plantação_H</i> e <i>plantação_D</i> , à primeira abordagem.	233
Figura A.7: Gráfico da entrada “CA100” para os estados de inferência <i>arco25</i> , <i>arco75</i> e <i>arco100</i> , à segunda abordagem.	233
Figura A.8: Gráfico da entrada “CA300” para os estados de inferência <i>arco25</i> , <i>arco75</i> e <i>arco100</i> , à segunda abordagem.	234
Figura A.9: Gráfico da entrada “AA300” para os estados de inferência <i>arco25</i> , <i>arco75</i> e <i>arco100</i> , à segunda abordagem.	234
Figura A.10: Gráfico da entrada “AIM” para os estados de inferência <i>arco25</i> , <i>arco75</i> e <i>arco100</i> , à segunda abordagem.	234
Figura A.11: Gráfico da entrada “AIB” para os estados de inferência <i>arco25</i> , <i>arco75</i> e <i>arco100</i> , à segunda abordagem.	235
Figura A.12: Gráfico da entrada “LA100” para os estados de inferência <i>arco25</i> , <i>arco75</i> e <i>arco100</i> , à segunda abordagem.	235
Figura A.13: Gráfico da entrada “LA300” para os estados de inferência <i>arco25</i> , <i>arco75</i> e <i>arco100</i> , à segunda abordagem.	235
Figura A.14: Gráfico da saída “FL” para os estados de inferência <i>arco25</i> , <i>arco75</i> e <i>arco100</i> , à segunda abordagem.	236
Figura A.15: Gráfico da saída “FRG” para os estados de inferência <i>arco25</i> , <i>arco75</i> e <i>arco100</i> , à segunda abordagem.	236
Figura A.16: Gráfico da saída “SVR” para os estados de inferência <i>arco25</i> , <i>arco75</i> e <i>arco100</i> , à segunda abordagem.	237
Figura A.17: Gráfico da saída “CF” para os estados de inferência <i>arco25</i> , <i>arco75</i> e <i>arco100</i> , à segunda abordagem.	237
Figura A.18: Gráfico da saída “GM” para os estados de inferência <i>arco25</i> , <i>arco75</i> e <i>arco100</i> , à segunda abordagem.	237
Figura A.19: Gráfico da saída “FC” para os estados de inferência <i>arco25</i> , <i>arco75</i> e <i>arco100</i> , à segunda abordagem.	238
Figura A.20: Gráfico da saída “CG” para os estados de inferência <i>arco25</i> , <i>arco75</i> e <i>arco100</i> , à segunda abordagem.	238
Figura B.1: Regressão das saídas alvo 10% para a segunda abordagem RGB.	240
Figura B.2: Regressão das saídas alvo 20% para a segunda abordagem RGB.	241
Figura B.3: Regressão das saídas alvo 30% para a segunda abordagem RGB.	242

Figura B.4: Regressão das saídas alvo 40% para a segunda abordagem RGB.	243
Figura B.5: Regressão das saídas alvo 50% para a segunda abordagem RGB.	244
Figura B.6: Regressão das saídas alvo 75% para a segunda abordagem RGB.	245
Figura B.7: Regressão das saídas alvo 100% para a segunda abordagem RGB.	246
Figura B.8: Regressão das saídas alvo 10% para a segunda abordagem RGB.	247
Figura B.9: Regressão das saídas alvo 20% para a segunda abordagem RGB.	248
Figura B.10: Regressão das saídas alvo 30% para a segunda abordagem RGB.	249
Figura B.11: Regressão das saídas alvo 40% para a segunda abordagem RGB.	250
Figura B.12: Regressão das saídas alvo 50% para a segunda abordagem RGB.	251
Figura B.13: Regressão das saídas alvo 75% para a segunda abordagem RGB.	252
Figura B.14: Regressão das saídas alvo 100% para a segunda abordagem RGB.....	253
Figura B.15: Regressão das saídas alvo 10% para a segunda abordagem RGB.	254
Figura B.16: Regressão das saídas alvo 20% para a segunda abordagem RGB.	255
Figura B.17: Regressão das saídas alvo 30% para a segunda abordagem RGB.	256
Figura B.18: Regressão das saídas alvo 40% para a segunda abordagem RGB.	257
Figura B.19: Regressão das saídas alvo 50% para a segunda abordagem RGB.	258
Figura B.20: Regressão das saídas alvo 75% para a segunda abordagem RGB.	259
Figura B.21: Regressão das saídas alvo 100% para a segunda abordagem RGB.....	260
Figura B.22: Regressão das saídas alvo 10% para a segunda abordagem RGB.	261
Figura B.23: Regressão das saídas alvo 20% para a segunda abordagem RGB.	262
Figura B.24: Regressão das saídas alvo 30% para a segunda abordagem RGB.	263
Figura B.25: Regressão das saídas alvo 40% para a segunda abordagem RGB.	264
Figura B.26: Regressão das saídas alvo 50% para a segunda abordagem RGB.	265
Figura B.27: Regressão das saídas alvo 75% para a segunda abordagem RGB.	266
Figura B.28: Regressão das saídas alvo 100% para a segunda abordagem RGB.....	267
Figura C.1: MSE 10% para a primeira abordagem RGB.....	269
Figura C.2: MSE 20% para a primeira abordagem RGB.....	270

Figura C.3: MSE 30% para a primeira abordagem RGB.....	270
Figura C.4: MSE 40% para a primeira abordagem RGB.....	271
Figura C.5: MSE 50% para a primeira abordagem RGB.....	271
Figura C.6: MSE 75% para a primeira abordagem RGB.....	272
Figura C.7: MSE 100% para a primeira abordagem RGB.....	272
Figura C.8: MSE 10% para a primeira abordagem HSV.....	273
Figura C.9: MSE 20% para a primeira abordagem HSV.....	273
Figura C.10: MSE 30% para a primeira abordagem HSV.....	274
Figura C.11: MSE 40% para a primeira abordagem HSV.....	274
Figura C.12: MSE 50% para a primeira abordagem HSV.....	275
Figura C.13: MSE 75% para a primeira abordagem HSV.....	275
Figura C.14: MSE 100% para a primeira abordagem HSV.....	276
Figura C.15: MSE 10% para a segunda abordagem RGB.....	276
Figura C.16: MSE 20% para a segunda abordagem RGB.....	277
Figura C.17: MSE 30% para a segunda abordagem RGB.....	277
Figura C.18: MSE 40% para a segunda abordagem RGB.....	278
Figura C.19: MSE 50% para a segunda abordagem RGB.....	278
Figura C.20: MSE 75% para a segunda abordagem RGB.....	279
Figura C.21: MSE 100% para a segunda abordagem RGB.....	279
Figura C.22: MSE 10% para a segunda abordagem HSV.....	280
Figura C.23: MSE 20% para a segunda abordagem HSV.....	280
Figura C.24: MSE 30% para a segunda abordagem HSV.....	281
Figura C.25: MSE 40% para a segunda abordagem HSV.....	281
Figura C.26: MSE 50% para a segunda abordagem HSV.....	282
Figura C.27: MSE 75% para a segunda abordagem HSV.....	282
Figura C.28: MSE 100% para a segunda abordagem HSV.....	283
Figura D.1: Gráfico de superfície para as combinações <i>RDDL/RDEL/AN</i> , <i>RDDL/RDEF/AN</i> , <i>RDDL/RDEF/AnN</i> , <i>RDEL/RDDL/AN</i> , <i>RDEL/RDDF/AN</i> , <i>RDEL/RDDF/AnN</i> ,	

<i>RDEL/RDEF/AN, RDDF/RDEL/AN, RDDF/RDEL/AnN, RDDF/RDEF/AnN, RDEF/RDDL/AN, RDEF/RDDL/AnN, RDEF/RDDF/AnN</i>	285
Figura D.2: Gráfico de superfície para as combinações <i>RDDL/RDEL/AnN</i> e <i>RDEL/RDDL/AnN</i>	286
Figura D.3: Gráfico de superfície para a combinação <i>RDDL/RDDF/AN</i>	286
Figura D.4: Gráfico de superfície para as combinações <i>RDDL/RDDF/AnN</i> e <i>RDDF/RDDL/AnN</i>	287
Figura D.5: Gráfico de superfície para as combinações <i>RDEL/RDEF/AnN, RDEF/RDEL/AnN</i>	287
Figura D.6: Gráfico de superfície para a combinação <i>RDDF/RDDL/NA</i>	288
Figura D.7: Gráfico de superfície para a combinação <i>RDDF/RDEF/AN</i>	288
Figura D.8: Gráfico de superfície para a combinação <i>RDEF/RDDF/AN</i>	289
Figura D.9: Gráfico de superfície para as combinações <i>CA100/CA200/FL, CA100/CA200/CF, CA200/CA100/FL</i> e <i>CA200/CA100/CF</i>	289
Figura D.10: Gráfico de superfície para as combinações <i>CA100/AA300/FL, CA100/AA300/CF, CA200/AA300/FL</i> e <i>CA200/AA300/CF</i>	290
Figura D.11: Gráfico de superfície para as combinações <i>AA300/CA100/FL, AA300/CA100/CF, AA300/CA200/FL</i> e <i>AA300/CA200/CF</i>	290
Figura D.12: Gráfico de superfície para as combinações <i>LA100/LA300/GM</i> e <i>LA100/LA300/CG</i>	291
Figura D.13: Gráfico de superfície para as combinações <i>LA300/LA100/GM</i> e <i>LA300/LA100/CG</i>	291
Figura D.14: Gráfico de superfície para a combinação <i>AIM/AIB/FRG</i>	292
Figura D.15: Gráfico de superfície para a combinação <i>AIM/AIB/SVR</i>	292
Figura D.16: Gráfico de superfície para a combinação <i>AIM/AIB/FC</i>	293
Figura D.17: Gráfico de superfície para a combinação <i>AIB/AIM/FRG</i>	293
Figura D.18: Gráfico de superfície para a combinação <i>AIB/AIM/FRG</i>	294
Figura D.19: Gráfico de superfície para a combinação <i>AIB/AIM/FRG</i>	294
Figura E.1: Imagem original RGB e HSV (laranja 02).....	295
Figura E.2: Mapa de classe 2 (Bayes RGB e HSV).....	295
Figura E.3: Mapa de classe 2 (Bayes ingênuo RGB e HSV).....	295

Figura E.4: Mapa de classe 2 (Fuzzy preditivo RGB e HSV).....	296
Figura E.5: Mapa de classe 2 (RNA RGB e HSV).....	296
Figura E.6: Imagem original RGB e HSV (laranja 03).....	297
Figura E.7: Mapa de classe 3 (Bayes RGB e HSV).....	297
Figura E.8: Mapa de classe 3 (Bayes ingênuo RGB e HSV).	297
Figura E.9: Mapa de classe 3 (Fuzzy preditivo RGB e HSV).....	298
Figura E.10: Mapa de classe 3 (RNA RGB e HSV).....	298
Figura E.11: Imagem original RGB e HSV (laranja 04).....	299
Figura E.12: Mapa de classe 4 (Bayes RGB e HSV).....	299
Figura E.13: Mapa de classe 4 (Bayes ingênuo RGB e HSV).	299
Figura E.14: Mapa de classe 4 (Fuzzy preditivo RGB e HSV).....	300
Figura E.15: Mapa de classe 4 (RNA RGB e HSV).....	300
Figura E.16: Imagem original RGB e HSV (laranja 05).....	301
Figura E.17: Mapa de classe 5 (Bayes RGB e HSV).....	301
Figura E.18: Mapa de classe 5 (Bayes ingênuo RGB e HSV).	301
Figura E.19: Mapa de classe 5 (Fuzzy preditivo RGB e HSV).....	302
Figura E.20: Mapa de classe 5 (RNA RGB e HSV).....	302
Figura E.21: Imagem original RGB e HSV (pomares 02).	303
Figura E.22: Mapa de classe 2 (Bayes RGB e HSV).....	303
Figura E.23: Mapa de classe 2 (Bayes ingênuo RGB e HSV).	303
Figura E.24: Mapa de classe 2 (Fuzzy preditivo RGB e HSV).....	303
Figura E.25: Mapa de classe 2 (RNA RGB e HSV).....	304
Figura E.26: Imagem original RGB e HSV (pomares 03).	305
Figura E.27: Mapa de classe 3 (Bayes RGB e HSV).....	305
Figura E.28: Mapa de classe 3 (Bayes ingênuo RGB e HSV).	305
Figura E.29: Mapa de classe 3 (Fuzzy preditivo RGB e HSV).....	305
Figura E.30: Mapa de classe 3 (RNA RGB e HSV).....	306

Figura E.31: Imagem original RGB e HSV (pomares 04).	307
Figura E.32: Mapa de classe 4 (Bayes RGB e HSV).	307
Figura E.33: Mapa de classe 4 (Bayes ingênuo RGB e HSV).	307
Figura E.34: Mapa de classe 4 (Fuzzy preditivo RGB e HSV).	307
Figura E.35: Mapa de classe 4 (RNA RGB e HSV).	308
Figura E.36: Imagem original RGB e HSV (pomares 05).	309
Figura E.37: Mapa de classe 5 (Bayes RGB e HSV).	309
Figura E.38: Mapa de classe 5 (Bayes ingênuo RGB e HSV).	309
Figura E.39: Mapa de classe 5 (Fuzzy preditivo RGB e HSV).	309
Figura E.40: Mapa de classe 5 (RNA RGB e HSV).	310
Figura E.41: Imagem original RGB e HSV (milharais 02).	311
Figura E.42: Mapa de classe 2 (Bayes RGB e HSV).	311
Figura E.43: Mapa de classe 2 (Bayes ingênuo RGB e HSV).	311
Figura E.44: Mapa de classe 2 (Fuzzy preditivo RGB e HSV).	312
Figura E.45: Mapa de classe 2 (RNA RGB e HSV).	312
Figura E.46: Imagem original RGB e HSV (milharais 03).	313
Figura E.47: Mapa de classe 3 (Bayes RGB e HSV).	313
Figura E.48: Mapa de classe 3 (Bayes ingênuo RGB e HSV).	313
Figura E.49: Mapa de classe 3 (Fuzzy preditivo RGB e HSV).	314
Figura E.50: Mapa de classe 3 (RNA RGB e HSV).	314
Figura E.51: Imagem original RGB e HSV (milharais 04).	315
Figura E.52: Mapa de classe 4 (Bayes RGB e HSV).	315
Figura E.53: Mapa de classe 4 (Bayes ingênuo RGB e HSV).	315
Figura E.54: Mapa de classe 4 (Fuzzy preditivo RGB e HSV).	316
Figura E.55: Mapa de classe 4 (RNA RGB e HSV).	316
Figura E.56: Imagem original RGB e HSV (milharais 05).	317
Figura E.57: Mapa de classe 5 (Bayes RGB e HSV).	317

Figura E.58: Mapa de classe 5 (Bayes ingênuo RGB e HSV).	317
Figura E.59: Mapa de classe 5 (Fuzzy preditivo RGB e HSV).	318
Figura E.60: Mapa de classe 5 (RNA RGB e HSV).	318
Figura E.61: Imagem original RGB e HSV (canaviais 02).	319
Figura E.62: Mapa de classe 2 (Bayes RGB e HSV).	319
Figura E.63: Mapa de classe 2 (Bayes ingênuo RGB e HSV).	319
Figura E.64: Mapa de classe 2 (Fuzzy preditivo RGB e HSV).	320
Figura E.65: Mapa de classe 2 (RNA RGB e HSV).	320
Figura E.66: Imagem original RGB e HSV (canaviais 03).	321
Figura E.67: Mapa de classe 3 (Bayes RGB e HSV).	321
Figura E.68: Mapa de classe 3 (Bayes ingênuo RGB e HSV).	321
Figura E.69: Mapa de classe 3 (Fuzzy preditivo RGB e HSV).	321
Figura E.70: Mapa de classe 3 (RNA RGB e HSV).	322
Figura E.71: Imagem original RGB e HSV (canaviais 04).	323
Figura E.72: Mapa de classe 4 (Bayes RGB e HSV).	323
Figura E.73: Mapa de classe 4 (Bayes ingênuo RGB e HSV).	323
Figura E.74: Mapa de classe 4 (Fuzzy preditivo RGB e HSV).	323
Figura E.75: Mapa de classe 4 (RNA RGB e HSV).	324
Figura E.76: Imagem original RGB e HSV (canaviais 05).	325
Figura E.77: Mapa de classe 5 (Bayes RGB e HSV).	325
Figura E.78: Mapa de classe 5 (Bayes ingênuo RGB e HSV).	325
Figura E.79: Mapa de classe 5 (Fuzzy preditivo RGB e HSV).	326
Figura E.80: Mapa de classe 5 (RNA RGB e HSV).	326
Figura F.1: Quantificação 03 – mapas de classe, métricas circulares, centróides e localização de frutos (RGB – esquerda; HSV – direita).	328
Figura F.2: Quantificação 04 – mapas de classe, métricas circulares, centróides e localização de frutos (RGB – esquerda; HSV – direita).	329

Figura F.3: Quantificação 05 – mapas de classe, métricas circulares, centróides e localização de frutos (RGB – esquerda; HSV – direita).....	330
Figura F.4: Quantificação 06 – mapas de classe, métricas circulares, centróides e localização de frutos (RGB – esquerda; HSV – direita).....	331
Figura F.5: Quantificação 07 – mapas de classe, métricas circulares, centróides e localização de frutos (RGB – esquerda; HSV – direita).....	332
Figura F.6: Quantificação 08 – mapas de classe, métricas circulares, centróides e localização de frutos (RGB – esquerda; HSV – direita).....	333
Figura F.7: Quantificação 09 – mapas de classe, métricas circulares, centróides e localização de frutos (RGB – esquerda; HSV – direita).....	334
Figura F.8: Quantificação 10 – mapas de classe, métricas circulares, centróides e localização de frutos (RGB – esquerda; HSV – direita).....	335

Lista de tabelas

Tabela 3.1: Distribuição dos vetores de características às classes de Pomares em RGB.	156
Tabela 3.2: Distribuição dos vetores de características às classes de Pomares em HSV.	157
Tabela 3.3: Distribuição dos vetores de características às classes de Canaviais em RGB. ...	157
Tabela 3.4: Distribuição dos vetores de características às classes de Canaviais em HSV.....	157
Tabela 3.5: Distribuição dos vetores de características às classes de Milharais em RGB.	158
Tabela 3.6: Distribuição dos vetores de características às classes de Milharais em HSV.	158
Tabela 3.7: Distribuição dos vetores de características às classes na localização de frutos em RGB.	159
Tabela 3.8: Distribuição dos vetores de características às classes na localização de frutos em HSV.....	159
Tabela 3.9: Distribuição dos vetores de características às classes na localização de milhos em RGB.	160
Tabela 3.10: Distribuição dos vetores de características às classes na localização de milhos em HSV.....	160
Tabela 3.11: Base de regras para as entradas e saídas à primeira abordagem.....	171
Tabela 3.12: Base de regras para as entradas e saídas à segunda abordagem (A).....	172
Tabela 3.13: Base de regras para as entradas e saídas à segunda abordagem (B).....	172
Tabela 4.1: Probabilidade <i>a priori</i> estimadas dos conjuntos de treinamento de pomares.	201
Tabela 4.2: Probabilidade <i>a priori</i> estimadas dos conjuntos de treinamento de canaviais. ...	202
Tabela 4.3: Probabilidade <i>a priori</i> estimadas dos conjuntos de treinamento de milharais. ...	202
Tabela 4.4: Probabilidade <i>a priori</i> estimadas dos conjuntos de treinamento de laranjas.	203

Lista de abreviaturas e siglas

AP	Agricultura de Precisão
ACP	Análise das Componentes Principais
AGV	<i>Autonomous Guided Vehicle</i> (veículo guiado autonomamente)
ANN	<i>Artificial Neural Network</i>
AG	Algoritmo Genético
ANFIS	<i>Adaptive Neuro-Fuzzy Inference System</i>
CAN	<i>Controller Area Network</i>
CDA	Centro de Área (centróide)
CDD	<i>Charge-Coupled Device</i>
CIE	<i>Commission International de l'Eclairage</i>
CML	<i>Concurrent Mapping and Localization</i>
CMOS	<i>Complementary Metal-Oxide Semiconductor</i>
CRLB	<i>Cramer-Rao Lower Bound</i>
DSP	<i>Digital Signal Processor</i> (processador de sinais digitais)
ECU	<i>Electronic Control Unit</i> (unidade de controle eletrônico)
EIF SLAM	<i>Extended Information Form SLAM</i>
EKF	<i>Extended Kalman Filter</i> (filtro de Kalman estendido)
EM	<i>Expectation-Maximization</i>
EQM	Erro Quadrático Médio
EESC	Escola de Engenharia de São Carlos
FANN	<i>Fast Artificial Neural Network</i>
GLA	<i>General Lloyd Algorithm</i>

GPS	<i>Global Positioning System</i> (sistema de posicionamento global)
HDD	<i>Hard Disk Drive</i> (disco rígido)
HSB	Matiz (<i>hue</i>), saturação e brilho
HSI	Matiz (<i>hue</i>), saturação e intensidade
HSL	Matiz (<i>hue</i>), saturação e luminosidade
HSV	Matiz (<i>hue</i>), saturação e valor
IA	Inteligência Artificial
ICP	<i>Interactive Closest Point</i>
IHPF	<i>Ideal Highpass Filter</i> (filtro ideal passa-alta)
ILPF	<i>Ideal Lowpass Filter</i> (filtro ideal passa-baixa)
iRPROP	<i>Improved Resilient Propagation</i>
JM	<i>Jeffries-Matusita</i>
JPEG	<i>Joint Photographic Experts Group</i>
KF	<i>Kalman Filter</i> (filtro de Kalman)
LSM	<i>Least-Squares Method</i> (método dos mínimos quadrados)
LVQ	<i>Learning Vector Quantization</i> (aprendizagem por quantização vetorial)
MCP	<i>Model Predictive Control</i> (modelo controle preditivo)
ML	<i>Maximum-Likelihood</i> (máxima verossimilhança)
MLP	<i>Multilayer Perceptron</i> (Perceptron multicamadas)
MOS	<i>Metal-Oxide Semiconductor</i>
MSE	<i>Mean Square Error</i>
Nrgb	RGB normalizado
NEPAS	Núcleo de Ensino e Pesquisa em Automação e Simulação
PC	Computador Pessoal
PCA	<i>Principal Component Analysis</i>

PDF	<i>Density Probability Function</i> (função densidade de probabilidade)
PDFc	PDF conjunta
PDI	Processamento Digital de Imagens
PGF	<i>Peer Group Filtering</i> (filtragem por grupo de pares)
RAM	<i>Random Access Memory</i> (memória de acesso aleatório)
RBPf	<i>Rao-Blackwellised Particle Filter</i>
RGB	<i>Red, Green, Blue</i>
RM	Robótica Móvel
RNA	Redes neurais artificiais
RSS	<i>Root Sum Squares</i>
SEIF	<i>Sparse Extended Information Form</i>
SFM	<i>Structure from Motion</i> (estrutura por movimento)
SIFT	<i>Scale Invariant Feature Transformation</i>
SLAM	<i>Simultaneous Localization and Mapping</i>
SLAMB	<i>Simultaneous Localization and Map Building</i>
SVM	<i>Support Vector Machine</i>
UKF	<i>Unscented Kalman Filter</i> (filtro de Kalman imprevisível)
USP	Universidade de São Paulo
VC	Visão Computacional
VRT	<i>Variable Rate Technology</i> (taxa de variação tecnológica)

Lista de símbolos

O_W	Sistema de coordenadas externa da câmera <i>pinhole</i>
O_C	Sistema de coordenadas interna na câmera <i>pinhole</i>
T	Matriz de translação
R	Matriz de rotação
H	Plano da imagem por grade de <i>pixels</i>
O_c	Centro de origem (ponto focal)
P	Equação linear da câmera <i>pinhole</i>
D	Matriz de projeção de imagem
Ψ	Matriz de calibração da câmera <i>pinhole</i>
b	Distância focal
b	Imagem discretizada
R	Comprimento de onda da componente R do espaço de cores RGB
G	Comprimento de onda da componente G do espaço de cores RGB
B	Comprimento de onda da componente B do espaço de cores RGB
S_R	Filtro da componente R (red/vermelho)
S_G	Filtro da componente G (green/verde)
S_B	Filtro da componente B (blue/azul)
λ	Comprimento de onda
$E(\lambda)$	Energia luminosa
r	Componente R normalizada
g	Componente G normalizada
b	Componente B normalizada

H	Componente H (hue/matiz) do espaço de cores HSI
S	Componente S (saturação) do espaço de cores HSI
I	Componente I (intensidade) do espaço de cores HSI
B	Componente B (brilho) do espaço de cores HSB
L	Componente L (luminosidade) do espaço de cores HSL
V	Componente V (valor) do espaço de cores HSV
Y	Componente Y (luminosidade) do espaço de cores YIQ
I	Diferença entre alaranjado e ciano do espaço de cores YIQ
Q	Diferença entre verde e magenta do espaço de cores YIQ
P_g	Probabilidade do k-ésimo nível de cinza da imagem
N_k	Número de <i>pixels</i>
$d_i(n)$	Distância euclidiana intra- <i>pixels</i>
$J(i)$	Estimação de <i>pixels</i> para a região de bordas
$f_i(n)$	Grau de ruído da filtragem por grupo de pares da imagem
w_i	Pesos gaussianos da filtragem por grupo de pares da imagem
$P(n)$	Soma ponderada de <i>pixels</i>
G_x	Gradiente de detecção de bordas vertical
G_y	Gradiente de detecção de bordas horizontal
R	Imagem segmentada
$P(R_i)$	Critério de homogeneidade
$s_{(k)}$	Descritor de <i>Fourier</i>
A	Autovetores de matriz de correlação em transformação linear
k	Número total de classes em probabilidade condicional
y_k	Saída desejada para a rede neural artificial
$\varphi(.)$	Função de ativação de rede neural artificial

$E(n)$	Função erro da regra Delta na retro-propagação de RNA
B_{ij}	Distância de <i>Bhattacharyya</i>
M_{ij}	Vetor de média para distância Mahalanobis
C_{ij}	Matriz de covariância das médias para distância Mahalanobis
J_{ij}	Distância de Jeffries-Matusita
D	Distorção global de regiões segmentadas
J	Valor J da quantização de cores da segmentação JSEG
S_T	Soma dos pontos quantizados na imagem
S_W	Medida de distâncias entre os mapas de classe
T_J	Limiar de segmentação das regiões homogêneas

Sumário

1. Introdução	41
1.1. Considerações iniciais	41
1.2. Objetivos	42
1.3. Estrutura do Trabalho	42
2. Visão Computacional e Robótica Móvel Agrícola.....	45
2.1. Visão computacional	47
2.1.1. Sistemas de aquisição de imagens	49
2.1.2. Representação gráfica monocromática e colorida	53
2.1.3. Pré-processamento em realce e restauração de imagens.....	57
2.1.4. Segmentação de imagens	58
2.1.5. Reconhecimento de Classes e Padrões	63
2.2. Robótica móvel	87
2.2.1. Percepção e incertezas na estimativa de dados.....	89
2.2.2. Localização por mapas, navegação e planejamento de trajetórias	92
2.3. Estado da arte em visão computacional aplicada à robótica móvel agrícola.....	101
2.3.1. Tecnologia e Agricultura de Precisão	101
2.3.2. Análise, inspeção e localização de culturas agrícolas.....	103
2.3.3. Locomoção guiada e autônoma de robôs móveis agrícolas	113
2.4. Considerações finais	127
3. Metodologia proposta.....	129
3.1. Características de aquisição e análise das imagens	131
3.2. Processamento de imagens	136
3.2.1. Pré-processamento e representações gráficas.....	140
3.2.2. Segmentação de imagens em cenas naturais	140
3.2.3. Vetores, normalização e extração de características.....	148
3.2.4. Classificadores cognitivos e estatísticos	161
3.2.5. Modelos de validação dos classificadores.....	176
3.3. Considerações finais	177
4. Resultados.....	179
4.1. Pré-processamento em realce de imagens	179
4.2. Segmentação das cenas de culturas agrícolas	181
4.3. Segmentação das cenas de linhas de plantio.....	184
4.4. Normalização e extração de características	188
4.5. Classificação dos segmentos para as duas abordagens.....	195
4.5.1. Classificador baseado em redes neurais artificiais	195

4.5.2.	Classificador baseado em lógica nebulosa preditiva	196
4.5.3.	Classificador baseado no teorema Bayes e Bayes ingênuo.....	200
4.5.4.	Avaliação dos classificadores.....	210
4.6.	Mapas de classes para reconhecimento de padrões.....	214
4.7.	Pós-processamento na quantificação de objetos	220
4.8.	Considerações finais	226
5.	Conclusões	227
5.1.	Considerações finais e discussões	227
5.2.	Propostas de trabalhos futuros.....	228
APÊNDICE A.	<i>Gráficos das variáveis de entrada e saída do controlador Fuzzy preditivo às duas abordagens</i>	231
APÊNDICE B.	<i>Gráficos de regressão das saídas alvo para o classificador RNA</i>	239
APÊNDICE C.	<i>EQM/MSE do classificador RNA para as duas abordagens</i>	269
APÊNDICE D.	<i>Gráficos de superfície dos estados de inferência Fuzzy</i>	285
APÊNDICE E.	<i>Mapas de classes para as duas abordagens</i>	295
APÊNDICE F.	<i>Gráficos de quantificação e localização de objetos</i>	327
APÊNDICE G.	<i>Publicações científicas</i>	337
6.	Referências Bibliográficas.....	339

1. Introdução

1.1. Considerações iniciais

O recente crescimento de novas tecnologias em sistemas eletrônicos embarcados e em automação vem promovendo o desenvolvimento de atividades de pesquisa científica direcionadas para a área de Agricultura de Precisão (AP), com foco em aplicações de máquinas e implementos agrícolas de alto desempenho (PEREIRA, 2008). Tal desenvolvimento abrange a aplicação de procedimentos experimentais (com base em teorias científicas consagradas na literatura), a determinadas tarefas voltadas para a aplicação em robótica móvel agrícola, entre as quais:

- fusão de dados opto-sensoriais por visão computacional na aplicação de técnicas de processamento digital de imagens para a localização e análise de culturas agrícolas;
- navegação e planejamento de trajetórias para locomoção;
- métodos heurísticos analíticos, agentes inteligentes artificiais e algoritmos probabilísticos no processamento de informações gráficas e no reconhecimento de padrões;

Neste contexto, no presente trabalho são descritas técnicas de processamento de imagens digitais para duas situações em cenas naturais: a primeira propõe uma abordagem para a análise das áreas de linha de plantio em pomares, milharais e canaviais, cuja finalidade é delimitar a área de tráfego de um robô móvel, em áreas denominadas passíveis de navegação, baseadas no ângulo de guiagem e deslocamento lateral. A segunda situação lida com a inspeção, localização e o reconhecimento de culturas agrícolas como laranja e milho (experimental), e suas respectivas características naturais, além dos elementos intermediários como árvores, folhas, galhos, ramos, solo e céu.

1.2. Objetivos

Devido à complexidade da discriminação dos elementos para o processamento gráfico, os objetivos aplicados às duas situações são sintetizados a seguir:

- Pré-processamento das imagens adquiridas em filtragem no domínio espacial e identificação de regiões de interesse nos espaços de cores RGB e HSV;

- Aplicação do algoritmo de segmentação não supervisionado proposto em quantização de cores para regiões não homogêneas e avaliação do comportamento com outras técnicas clássicas nas imagens das culturas citadas;

- Validação dos vetores de características na identificação dos segmentos das imagens coletadas e relacionadas, para a formação das regiões de interesse e sua classificação.

- Avaliação de classificadores e descritores de padrões nas imagens processadas, baseadas em redes neurais artificiais (RNA), lógica *Fuzzy* e métodos estatísticos como filtros de *Bayes* e *Bayes Ingênuo*, comparando seus respectivos desempenhos na classificação.

Ainda, na segunda abordagem há a etapa de pós-processamento para a identificação, localização e quantificação de objetos (frutos) nas cenas classificadas e com padrões já mapeados (mapas de classe).

Dois sistemas monoculares foram usados como recurso sensorial na aquisição de imagens em perspectiva plana, frontal e lateral na locomoção do veículo, e em foco direto na localização de objetos.

1.3. Estrutura do Trabalho

A organização da dissertação foi estruturada inicialmente na abordagem teórica dos conceitos de visão computacional, e os trabalhos entre esta e a robótica móvel para cenas naturais agrícolas. Os materiais e métodos, como a metodologia de trabalho, são apresentados, assim como as discussões para os resultados obtidos, e as conclusões.

O conteúdo de cada capítulo é detalhado a seguir.

No capítulo dois são apresentados os conceitos teóricos de visão computacional, relacionando as teorias e aplicações em processamento de imagens digitais, as etapas de pré-processamento, técnicas de segmentação de imagens e reconhecimento de classes e padrões, detalhando os procedimentos para a descrição de elementos base, a normalização e extração de características, avaliação e classificação de padrões em cenas naturais, e validação dos classificadores analíticos e estatísticos. Em seguida, é realizada uma revisão dos conceitos teóricos de robótica móvel, priorizando as técnicas e teorias aplicadas na localização, mapeamento, navegação e planejamento de trajetórias. Por fim, neste capítulo, a tecnologia em agricultura de precisão, com abordagem dos implementos agrícolas aplicados a cada área de interesse, fusão de sensores e etapas de processamento computacional são apresentadas, bem como os trabalhos em estado da arte relacionando visão computacional e robótica móvel agrícola para aplicações em inspeção, localização e análise de culturas, e técnicas para navegação e planejamento de trajetórias, entre as áreas de linhas de plantio, de forma autônoma e robusta, com aplicações da teoria probabilística e cognitiva.

No capítulo três são apresentadas e detalhadas a metodologia proposta para as etapas de processamento e segmentação de imagens, para a caracterização das cenas naturais, bem como as ferramentas computacionais, técnicas, algoritmos e modelos propostos para tal. Os métodos são apresentados em seguida, relacionando os tópicos sequenciais para os objetivos deste trabalho: segmentação de imagens, extração e normalização de características, classificação em abordagens conjuntas analíticas e estatísticas dos segmentos de interesse, e a avaliação/validação destes classificadores segundo parâmetros de desempenho, flexibilidade, robustez e eficácia.

No capítulo quatro são apresentados e discutidos os resultados obtidos pelo uso de algoritmos de segmentação de imagens, os resultados da avaliação na extração dos vetores de

características, além de comparações com os classificadores cognitivos e estatísticos. Ao final do capítulo são apresentadas as imagens segmentadas e processadas com as regiões de interesse para a discriminação dos segmentos, na formação do mapa de classe de cada cena, bem como o pós-processamento na quantificação de frutos.

As considerações finais, conclusões e trabalhos futuros são comentados no capítulo cinco, destacando os temas relacionados na abordagem proposta, nos resultados esperados e obtidos.

2. Visão Computacional e Robótica

Móvel Agrícola

Define-se Visão Computacional como a ciência que desenvolve as bases matemáticas e algorítmicas pelas quais informações de alto nível do ambiente são extraídas de imagens (entrada de dados) para análise e processamento (descrição ou interpretação da saída de dados), com uso de sensores ópticos, técnicas de processamento digital de imagens, reconhecimento de padrões por métodos analíticos e estatísticos, e aplicações de Inteligência Artificial (IA) (FORSYTH e PONCE, 2002; SEBE *et al.*, 2005).

Na fusão sensorial óptica para a aquisição de dados em visão computacional, há considerável quantidade de informações na imagem adquirida, entre as quais: dimensão e representação espacial/geométrica, cor, textura e informação temporal para a aquisição dinâmica de imagens. As técnicas e operações aritméticas dedicadas ao estudo e análise das informações de tais imagens contemplam a área de Processamento Digital de Imagens (PDI), cuja aplicação busca a melhoria e realce de determinados aspectos da imagem tratada.

O processamento de imagens digitais é direcionado para funções e finalidades específicas, uma vez que o tempo de processamento, características e informações de entrada, e base de armazenamento devem ser considerados, para que possam ser aplicados em tempo-real ou manipulados de acordo com a necessidade temporal.

No processamento em tempo-real, para imagens complexas, que apresentam informações como elementos gráficos de forma, tamanho, cor e textura variadas, em cenas de distintas intensidades luminosas, e dependentes dos locais de aquisição, são utilizados algoritmos e técnicas sofisticadas de síntese, processamento, análise, classificação e reconhecimento,

envolvendo métodos analíticos e estatísticos para tais, em conjunto com a abordagem de Inteligência Artificial (IA). IA é definida como o estudo de agentes que percebem características do ambiente e executam ações específicas à percepção (RUSSELL e NORVIG, 2009). Em cada ação, é definida uma sequência de funções de percepção sensorial de um dado agente. Assim, estas funções de percepção são representadas por modelagens intrínsecas ao problema proposto, seja um sistema de produção, sistema de controle, sistema reativo e que desempenham a interpretação do ambiente estudado, por técnicas de aprendizagem. Tais técnicas são aplicadas nos módulos de visão computacional voltados à caracterização, classificação e reconhecimento de padrões.

Durante o aprendizado, há várias informações que apresentam ruídos e incertezas qualitativas, causadas pelas adversidades naturais, enquanto são adquiridas as imagens e após estas mesmas aquisições pelas limitações dos sensores ópticos usados. As incertezas são corrigidas com técnicas probabilísticas de estimação de estados recursivos, com aplicação de filtros não-paramétricos, gaussianos e robustos (THRUN *et al.*, 2005). Mesmo com tais métodos matemáticos, deve-se considerar a natureza ambígua do processo de visão, cuja interpretação destas é distinta ao modo perceptivo de cada modelo de aquisição. Desta forma, a construção de sistemas computacionais adequados e análogos à visão humana repele as linearidades de processamento, cuja complexidade natural de percepção obriga ao uso de modelos mais simplificados.

A visão computacional, quando aplicada à Agricultura, expande-se em várias etapas e funções desta área, desde coleta e análise de dados de determinada lavoura, inspeção e localização de culturas agrícolas, até a navegação guiada e/ou autônoma de veículos, localização, mapeamento e planejamento de trajetórias por fusão sensorial, que associa o emprego de robôs móveis para tal finalidade. Com isso, são aplicados critérios para a efetiva Automação entre robôs móveis e visão computacional na agricultura de precisão (SOUZA,

2007): *aquisição e análise de dados*: informações de entidades de interesse do agricultor são controláveis e programáveis, com o auxílio de algoritmos dedicados em visão computacional; *redução de custos*: racionalização controlada na aplicação de herbicidas contra plantas invasoras, localizando e identificando somente os locais caracterizados pela invasão, reduzindo os custos de aplicação uniforme na plantação; *automação e produtividade*: automação de tarefas (substituição de mão de obra) para a colheita automática de culturas, com auxílio de fusão de sistemas computacionais, opto-sensores e manipuladores robóticos, através da guiagem automática de robôs móveis.

2.1. Visão computacional

Visão computacional (VC) é a ciência que caracteriza informações de interesse numa imagem, a qual é estruturada na definição do problema visual a ser analisado e solucionado, organizado nas etapas sequenciais de aplicação: aquisição, pré-processamento, segmentação, extração de características e reconhecimento. Trata-se da integração de processos e representações gráficas usadas para a percepção visual (GONZAGA, 2010).

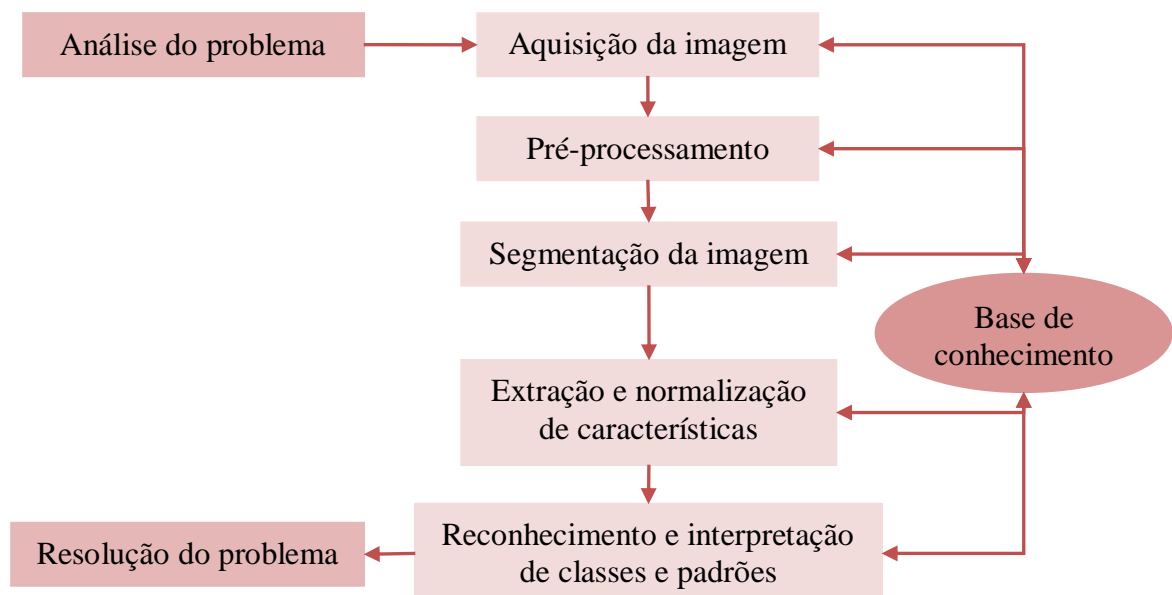


Figura 2.1: Estrutura sequencial de um sistema de visão computacional.

Como mostrado na Figura 2.1, a estrutura de um sistema de visão computacional contempla as etapas sequenciais, mediante a descrição de um problema visual: aquisição, pré-processamento, segmentação, normalização das características de interesse, classificação e reconhecimento de padrões, à resolução do problema. Todas interligadas a uma base de conhecimento dos resultados obtidos.

A estrutura sequencial das etapas presentes no sistema de VC se divide ainda em três estágios: *visão em baixo nível* (aquisição e pré-processamento); *visão em nível intermediário* (segmentação, extração e normalização de características); *visão em alto nível* (classificação, reconhecimento e correspondência) (FORSYTH e PONCE, 2002).

Para todos estes estágios, a tarefa final de um sistema de VC é modelar ou representar a imagem de entrada com o correspondente na base de conhecimento, através de interpretação, decisão ou descrição dos classificadores. A representação das imagens, então, pode ser definida pela área de aplicação:

- a) *Imagens generalizadas*: são icônicas e atendem pelas propriedades físicas da cena;
- b) *Imagens segmentadas*: são classificadas pelos segmentos ou regiões de interesse na imagem;
- c) *Imagens geométricas*: atendem pela disposição geométrica da cena, para simulação de efeitos entre as componentes envolvidas no formato bidimensional ou tridimensional;
- d) *Imagens relacionais*: são estruturadas em interpolação de componentes dentro da cena, cujo interesse do problema a ser solucionado baseia-se na sua inter-relação.

Com a definição destes conceitos é possível abordar as teorias envolvidas no processamento de imagens digitais para o propósito deste trabalho, tais como a representação de imagens monocromáticas e coloridas, o pré-processamento em realce de imagens, segmentação de imagens e reconhecimento de classes e padrões.

2.1.1. Sistemas de aquisição de imagens

Sistemas de aquisição de imagens podem ser passivos, cuja representação da cena é feita por símbolos no processo de identificação dos objetos, uma vez que a observação da cena dá-se pela imobilidade dos objetos; e ativos, onde há interação entre o observador e o ambiente, tornando dinâmica a reconstrução da cena - ajustes de aproximação da área de visão (*zoom*), da posição da câmera, da regulagem da intensidade luminosa (diafragma fotográfico) e os modos de interação da aquisição, tais como imagens perspectivas, panorâmicas, lineares e omnidirecionais (FLORCZYK, 2005; SEBE *et al.*, 2005).

Modelagem física de câmeras

O modelo mais simples da concepção física de aquisição de imagens é a câmera *pinhole* (micro-orifício). Devido ao *pinhole* estar situado entre o plano de imagem e a cena do ambiente tridimensional, qualquer raio de luz que é emitido ou refletido de determinada superfície passa por ele antes de ser projetado no plano de imagem (CYGANNEK e SIEBERT, 2009). A modelagem matemática da câmera *pinhole* é ilustrada na Figura 2.2, abaixo.

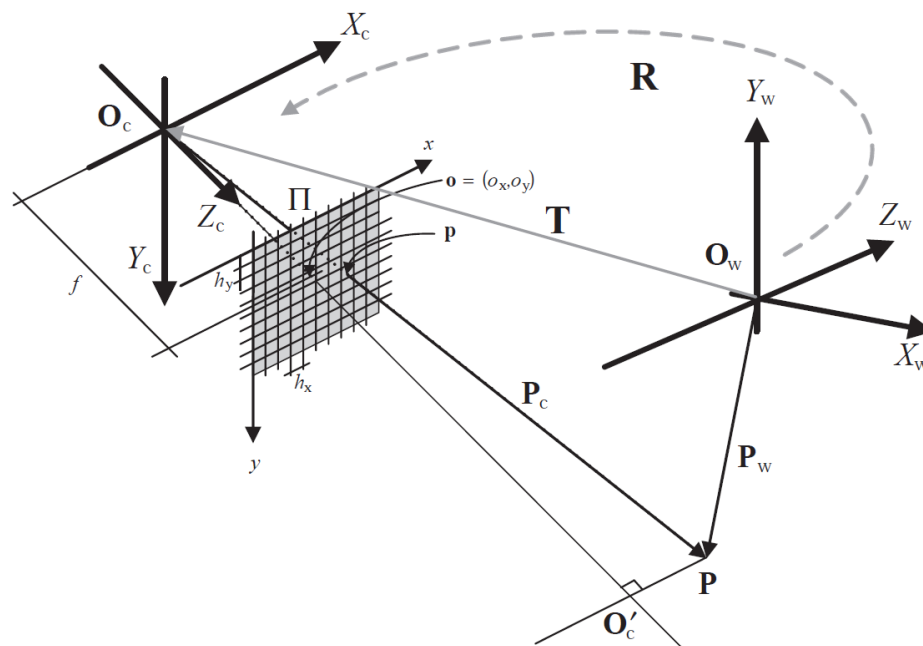


Figura 2.2: Modelo *pinhole* da câmera perspectiva com dois sistemas de coordenadas: externa (O_w) e interna (O_c) (CYGANNEK e SIEBERT, 2009).

Os pontos de projeção da cena tridimensional são relacionados matematicamente ao plano de imagem por dois sistemas de coordenadas: o sistema de coordenadas externo (W - *world*), o qual é independente da localização e dos parâmetros da câmera, e o sistema de coordenadas interno (C - câmera).

Os dois sistemas de coordenadas são relacionados pela matriz de translação T , e a matriz de rotação R . A origem O_c , denominado *centro* ou *ponto focal*, em conjunto com os eixos X_c , Y_c e Z_c , determina o sistema de coordenadas da câmera. Tal modelo assume o plano de imagem Π , que é quadriculado, e auxilia na representação de regiões em locais fotossensíveis discretos que amostram qualquer imagem projetada no plano. Cada quadrado é denominado *pixel* (*picture element*) e é indexado por um par de coordenadas expresso por valores escalares. Ainda, o plano de imagem Π representa uma grade de *pixels* discreta, cuja projeção da origem O_c na direção de Z_c determina o ponto principal das coordenadas locais (o_x, o_y) . O eixo principal é um segmento de reta entre os pontos O_c e O'_c . A distância do plano de imagem ao ponto principal é chamada de distância focal. Por fim, h_x e h_y determinam as dimensões físicas de um único *pixel*.

A localização de um dado ponto \mathbf{P} do espaço tridimensional depende do sistema de coordenadas escolhido: para O_c , tal ponto é representado por um vetor coluna \mathbf{P}_c ; para O_w , tal ponto é representado por um vetor coluna \mathbf{P}_w . O ponto \mathbf{p} é uma imagem do ponto \mathbf{P} sob a projeção com centro na origem O_c , no plano Π . Ambos os pontos são representados abaixo por suas coordenadas, na equação 2.1.

$$\begin{aligned}\mathbf{P} &= [X, Y, Z]^T \\ \mathbf{p} &= [x, y, z]^T\end{aligned}\tag{2.1}$$

Logo, como os eixos ópticos são perpendiculares ao plano de imagem, considerando que os triângulos $\Delta O_c \mathbf{p} o$ e $\Delta O_c \mathbf{P} O'_c$ são idênticos, e igualando $z = f$, obtém-se a base da modelagem física da câmera *pinhole* (equação 2.2).

$$\begin{aligned}
 x &= f \frac{X}{Z}, \\
 y &= f \frac{Y}{Z}, \\
 z &= f
 \end{aligned}
 \tag{2.2}$$

Câmeras CCD (*Charged Coupled Device*)

Câmeras CCD (dispositivos de carga acoplada) provêm elementos sensoriais de conversão digital, onde os sinais elétricos referentes aos *pixels* são transferidos a linhas e colunas de projeção, acoplados a carga de cada ponto da matriz fotoelétrica. Tais elementos são superiores aos dispositivos CMOS (*Complementary Metal-Oxide Semiconductor*) devido ao baixo nível de ruído durante a aquisição e tratamento da imagem, maior sensibilidade aos pontos de aquisição da cena, apresentando considerável nitidez, resolução e contraste de objetos na imagem (EASTMAN KODAK, 2001; JANESICK, 2001; JANESICK, 2002).

A estrutura básica CCD é dotada de transistores MOS posicionados em uma linha sequencial e conectados a uma fonte de tensão de três fases. Na primeira etapa, há a fase de armazenamento com a tensão V_a aplicada a um conjunto de eletrodos e a outra tensão V_b (baixa) aplicada aos outros dois conjuntos. Numa extremidade há carga elétrica que será fornecida pela luz incidente na estrutura, ao contrário da outra extremidade alimentada por V_b . Na segunda etapa, na fase de transferência, uma fonte de tensão de maior valor V_p é aplicada às armaduras dos capacitores para a transferência de carga elétrica acoplada sequencialmente. Nas duas últimas etapas, as fontes de tensão primárias V_a e V_b são reiniciadas ao estado de armazenamento com a informação da carga elétrica deslocada pela última extremidade. Tal informação é novamente processada pelas etapas anteriores até ser possível ler o conteúdo armazenado da conversão fotoelétrica (BOYLE e SMITH, 1970).

Na Figura 2.3, abaixo, são ilustradas a estrutura básica das tensões aplicadas na camada de óxido e a direção de transferência de carga em (a), a estrutura da matriz de transferência dos

registradores de *pixels* em (b), e a simplificação do diagrama elétrico com os transistores MOS, o sensor de ativação fotodiodo e as portas de transferência de carga para o sinal de saída, através da *FDA* (*Floating Diffusion Amplifier* – amplificador de difusão flutuante) em (c).

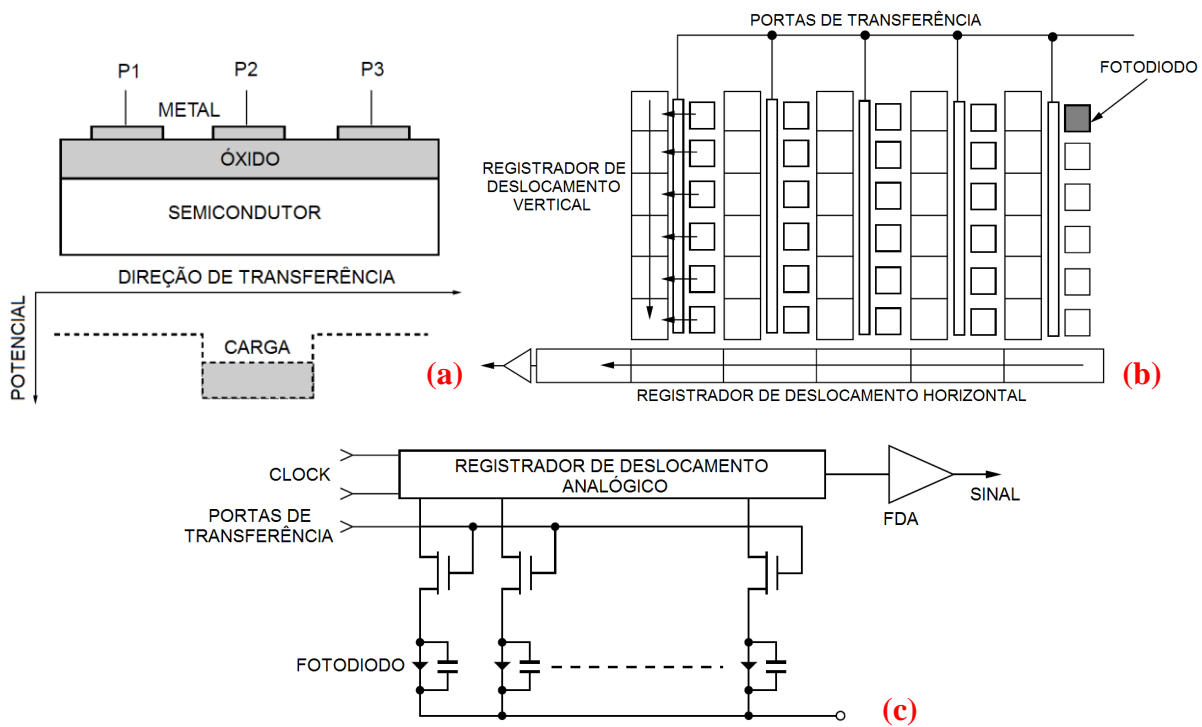


Figura 2.3: (a) tensões aplicadas na camada de óxido e direção de transferência de carga; (b) estrutura da matriz de transferência dos registradores de *pixels*; (c) diagrama elétrico simplificado com os transistores MOS, sensores de ativação e portas de transferência de carga para deslocamento analógico através da lógica *FDA* (Adaptado de EASTMAN KODAK, 2001).

O modelo matemático das câmeras CCD se assemelha ao *pinhole*. O eixo óptico da câmera é representado pelo eixo Z do sistema de coordenadas O_C . Um ponto X da cena tridimensional é projetado do ponto de incidência até a matriz CCD C , a qual é simbolizada pelo sistema de coordenadas do sensor bidimensional O_S . Tal sistema de coordenadas é determinado pelo ponto de intersecção entre o eixo óptico e o sensor C . A distância focal b é determinada pela distância entre a origem O_C e a origem O_S . Por fim, um determinado ponto $(x, y, z)_C$ no sistema de coordenadas da câmera é mapeado ao sistema de coordenadas do sensor pela equação 2.3 (BURT, 1974; ZORPETTE, 1998; FLORCZYK, 2005).

$$\begin{bmatrix} x_s \\ y_s \end{bmatrix} = \frac{b}{z_c} \begin{bmatrix} x_c \\ y_c \end{bmatrix} \quad (2.3)$$

2.1.2. Representação gráfica monocromática e colorida

A representação gráfica digital de uma imagem é realizada através da discretização de seu correspondente do modelo real observado. Em sua concepção matemática, uma imagem é uma função bidimensional $f(x,y)$ cujas componentes (x,y) são as coordenadas no plano onde a imagem é projetada, e o valor de tal função representa a intensidade da luz neste ponto. A discretização/digitalização da imagem é realizada através da amostragem espacial ou pela quantização da amplitude (FORSYTH e PONCE, 2002; GONZALEZ e WOODS, 2007).

Na amostragem espacial, as coordenadas que formam o plano cartesiano de projeção da imagem (x,y) são representados por um número finito de pontos (*pixels*), formando uma estrutura bidimensional do modelo gráfico real. Os *pixels* apresentam variação de distribuição e quantidade, sendo diretamente proporcionais à quantidade de informação espacial obtida. Cada *pixel* é associado ao respectivo valor quantizado de $f(x,y)$ simbolizado por pequenos quadrados na imagem, como pode ser observado na Figura 2.4.

Como uma imagem discreta apresenta uma matriz de *pixels* ($M \times N$), representando a largura e comprimento, respectivamente, cada *pixel* corresponde a 2^l tons de cinza – l denota a escala de tons de cinza. Assim, o número de *bits* (b) necessários para armazenar a imagem digital é dado pela equação 2.4:

$$b = N \times M \times 2^l \quad (2.4)$$

Em visão computacional, como imagens monocromáticas possuem apenas um comprimento de onda, o processamento de suas informações gráficas é mais rápido em comparação com imagens coloridas, porém se a análise dos dados depender de características

de cores diferentes, outros modelos de processamento para diversos comprimentos de onda devem ser considerados.



Figura 2.4: Divisão dos *pixels* em uma imagem colorida para determinada parte da imagem. Cada *pixel* corresponde a um determinado quadrado, para um valor quantizado $f(x,y)$, cuja atribuição é a mesma para imagens monocromáticas.

A recepção de cores no sistema visual humano é dotada de três tipos diferentes, chamados cones, cujos picos de receptividade em comprimentos de onda são: 650nm para vermelho ($R - red$), 530nm para verde ($G - green$) e 430nm para azul ($B - blue$). Tais comprimentos de onda emitem uma intensidade específica de luz em cada componente, calculados por filtros ideais, apresentados pela equação 2.5:

$$\begin{aligned}
 R &= \int_{\lambda} E(\lambda)S_R(\lambda)d\lambda \\
 G &= \int_{\lambda} E(\lambda)S_G(\lambda)d\lambda \\
 B &= \int_{\lambda} E(\lambda)S_B(\lambda)d\lambda
 \end{aligned} \tag{2.5}$$

Para cada componente, S_R , S_G e S_B são os filtros vermelho, verde e azul, respectivamente, que absorvem os pigmentos presentes nos cones do olho humano. $E(\lambda)$ representa a energia luminosa e λ o comprimento de onda (LEVINE, 1985; CAVANI, 2007). Assim, a partir da equação 2.5 é possível definir o espaço de cores RGB, cuja representação geométrica é associada a um cubo com os três eixos espaciais correspondendo às três componentes. Logo, uma imagem de função $f(x,y)$ será o conjunto de vetores tridimensionais associado ao espaço de cores RGB, e a quantidade de *bits* necessária para seu armazenamento digital é dada por:

$$b = 3 \times N \times M \times 2^l \quad (2.6)$$

Analogamente à equação 2.6 a imagem colorida em RGB é três vezes o tamanho da sua similar monocromática (em *bits* – b), relacionando o produto com a largura (N), o comprimento (M) e os tons de cores em RGB (2^l).

Trabalhos que demonstram a relação dos canais RGB para com o desempenho de algoritmos de segmentação em cenas de culturas naturais de ambientes agrícolas são relatados em (JING *et al.*, 2003; DAINESE *et al.*, 2004; JAFARI *et al.*, 2004; STEWARD *et al.*, 2004; HAN *et al.*, 2010; STORY *et al.*, 2010).

Nestes trabalhos, destaca-se a dificuldade do desempenho na aquisição e processamento de imagens em RGB, devido à correlação existente nos três canais de cores, além de ser impossível medir similaridades de cores usando distância euclidiana. Todavia, tais problemas do espaço de cores RGB podem ser resolvidos por meio de transformações lineares ou não-lineares, que geram outros espaços de cores com características intrínsecas. Para contornar o problema da correlação, é realizada uma transformação não-linear, originando o espaço de cores RGB normalizado ($Nrgb$ ou $nRGB$), cujas componentes são dadas pela equação 2.7. Neste espaço de cores, as componentes do $Nrgb$ representam apenas a informação cromática (CHENG *et al.*, 2010; CHENG *et al.*, 2010).

$$\begin{aligned} r &= \frac{R}{R+G+B}; \\ g &= \frac{G}{R+G+B}; \\ b &= \frac{B}{R+G+B}; \end{aligned} \quad (2.7)$$

O espaço de cores $Nrgb$, por ser normalizado e adequado para aplicações triviais em tempo real, com operações aritméticas simples, não possui comportamento estável em ambientes de baixa iluminação. Assim, para sistemas de visão computacional robustos, com aplicações *outdoor* estruturados ou não, é necessária a junção de espaços de cores distintos

para cada etapa do processamento das informações gráficas de interesse, com derivações providas da separação cromática e acromática da luz, tais como: HSI (matiz/*hue*, saturação e intensidade), e HSV (matiz/*hue*, saturação e valor).

Estes espaços de cores apresentam a vantagem da discriminação dos canais compostos, sendo dois canais caracterizando a cor, e um canal a luminosidade. Com isso este último canal fornece baixa influência na cor, admitindo informações mais gradativas sobre sua tonalidade na imagem analisada. A matiz/*hue* (H) do espaço de cores HSI define o comprimento de onda dominante da cor. O canal de saturação (S) verifica a quantidade da cor branca misturada na cor analisada (pureza), e o canal (I) verifica a intensidade da luz. A relação algébrica deste espaço de cores com o RGB é definida por Cheng (CHENG *et al.*, 2001), representada pelas equações 2.8, 2.9 e 2.10:

$$H = \arctan\left(\frac{\sqrt{3}(G-B)}{(R-G)(R-B)}\right) \quad (2.8)$$

$$I = \frac{(R+G+B)}{3} \quad (2.9)$$

$$S = 1 - \frac{\min(R, G, B)}{I} \quad (2.10)$$

Há algumas variações do HSI, como HSV, HSB (*B* – brilho) e HSL (*L* – luminosidade). Como o HSI é resultado de uma divisão de informações cromáticas e acromáticas da luz, este não é estável quando a saturação é baixa. Logo, os algoritmos de segmentação não atribuem os *pixels* para os segmentos corretos, devido à instabilidade numérica nesta região de saturação, já que em situações de baixa luminosidade, a informação cromática é irrelevante, resultando em representações de cores inadequadas pelos canais H e S, não comprometendo o pré-processamento em cenas naturais.

2.1.3. Pré-processamento em realce e restauração de imagens

Responsável pelas operações de realce e restauração de imagens, a etapa de pré-processamento é fundamental como técnica de análise na obtenção de imagens dedicadas a aplicações específicas que necessitam de informações de quadrantes ou regiões de interesse ao processamento, sendo realizadas como transformações de intensidade em filtros no domínio do espaço ou no domínio da frequência (GONZALEZ e WOODS, 2007).

No domínio espacial do pré-processamento, são realizadas operações globais em transformada de Laplace, que consideram determinada característica intrínseca de toda imagem, descaracterizando computacionalmente regiões menos significativas ao algoritmo. Neste sentido, operações envolvendo alteração de histograma consideram a distribuição estatística dos valores dos *pixels*, com o objetivo de melhorar a qualidade da imagem equalizando as regiões distintas de contraste. Assim, pequenas regiões são mapeadas em informações locais, com algoritmos dedicados ao processamento paralelo de outras localidades da imagem.

Processamento de histograma

O histograma de uma imagem em níveis de cinza identifica a percentagem de representação dos *pixels* neste espaço de cores. A qualidade ao nível de contraste e a intensidade de brilho são verificadas através do modelo gráfico de barras (GONZALEZ e WOODS, 2007).

A percentagem é dada em conjuntos para cada intensidade de *pixels*, logo o histograma é definido como: seja uma imagem g de $M \times N$, a probabilidade do k -ésimo nível de cinza da imagem é dada pela equação 2.11:

$$p_g(g_k) = \frac{n_k}{M \cdot N} \quad (2.11)$$

em que n_k é número de *pixels* correspondente ao nível de cinza k , M e N representam o produto do número total de *pixels* na imagem, ou sua resolução espacial, k corresponde ao número de níveis de cinza da imagem digitalizada, e $0 \leq r_k \leq 1$.

2.1.4. Segmentação de imagens

A segmentação de imagens é a etapa que une operações de processamento, como a redução de ruídos e extração de bordas das regiões de interesse dentro da imagem, e o reconhecimento de padrões para a interpretação de cenas (EGMONT-PETERSEN *et al.*, 2002; FORSYTH e PONCE, 2002; GONZALEZ e WOODS, 2007). Os elementos de interesse são separados do restante da imagem, para a extração de informações que serão usadas em sua caracterização.

Desta forma, os elementos são extraídos pelas discontinuidades das propriedades de *pixels* com o restante da cena na imagem, e também podem ser particionados em regiões homogêneas com propriedades internas semelhantes ou distintas entre si.

Define-se matematicamente, então, a segmentação de imagens (CHENG *et al.*, 2001; YOO, 2004; ZHANG, 2006). Seja $P()$ o critério de homogeneidade definido em determinados *pixels* interconectados. O mesmo critério é particionado em um novo conjunto, ou imagem, R sobre outros subconjuntos de *pixels* ($R_1, R_2, R_3, \dots, R_n$), dado pela equação 2.12.

$$R = \bigcup_{i=1}^n R_i \rightarrow R_i \cap R_j = \emptyset, (i \neq j) \quad (2.12)$$

A segmentação completa da imagem resulta em limiares concêntricos em cenas simples (equação 2.13), os quais representam a transformação de uma imagem de entrada f à outra imagem de saída h (segmentada), por binarização em imagens em tons de cinza ou imagens coloridas, cujo objetivo é evidenciar regiões de contorno.

$$\begin{aligned} h(i, j) = 1 &\rightarrow f(i, j) \geq T, \\ h(i, j) = 0 &\rightarrow f(i, j) < T \end{aligned} \quad (2.13)$$

onde T é o limiar, $h(i,j) = 1$ para elementos de objetos na imagem, e $h(i,j) = 0$ para elementos do plano de fundo.

Logo, o critério de homogeneidade $P(R_i)$ será verdadeiro para todas as regiões e $P(R_i \cup R_j)$ quando houver vizinhança entre R_i e R_j , e $i \neq j$.

A Figura 2.5, abaixo, mostra uma imagem R segmentada em regiões de interesse.

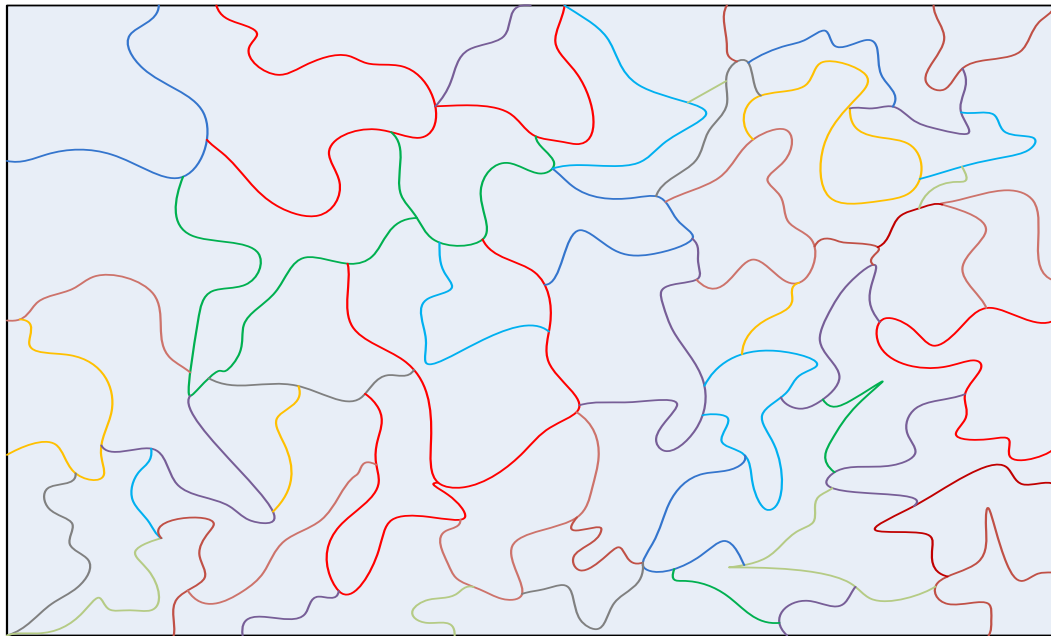


Figura 2.5: Divisão de regiões categorizadas em subconjuntos da imagem R .

Como demonstrado acima, as regiões obtidas através da segmentação devem manter relação condizente com as regiões de interesse ao critério de homogeneidade $P(R_i)$, uma vez que nem todos os algoritmos de segmentação não se aplicam à mesma imagem, tampouco um algoritmo apenas é aplicado a várias imagens adequadamente. Por isso, determinados sistemas de visão computacional operam em domínios distintos de imagens, relacionando-as de forma mais simples em estruturas definidas, ou complexas, devido às características do ambiente desconhecido.

Desta forma, deve-se considerar a aplicação de algoritmos de segmentação de imagens para imagens em tons de cinza, ou coloridas. No primeiro caso, apenas o valor do *pixel* é

usado na distinção de máscaras para o critério de homogeneidade, e em alguns casos a textura formada pelos *pixels* da região de interesse. Nas imagens coloridas, a variação da cor é discriminada para a obtenção de $P(R_i)$, bem como a textura analisada.

Conforme abordado por Cavani (CAVANI, 2007), a técnica de segmentação deve dividir regiões ou segmentos na imagem, cujo valor de $P(R_i)$ seja verdadeiro. Assim, o critério de homogeneidade escolhido deve corresponder ao algoritmo implementado para a segmentação, dependente da aplicação realizada, o que torna tal algoritmo específico a determinada tarefa.

Neste sentido, a classificação dos algoritmos de segmentação obedece à técnica usada para criar os segmentos de interesse na imagem, e são relatadas a seguir (CHENG *et al.*, 2001; CHEN e CHEN, 2002; CHEN *et al.*, 2005; ZHANG, 2006), com enfoque mais detalhado em (LUCCHESE e MITRA, 2001; LUCCHESE e MITRA, 2004; GONZALEZ e WOODS, 2007). Uma revisão sobre os trabalhos em estado da arte que lidam com as técnicas de segmentação abaixo e as aplicações na área agrícola é relatada na seção 2.3:

Segmentação por detecção de pontos, linhas e bordas: Baseado na detecção de mudanças locais abruptas de intensidade, cujas características principais são os pontos isolados, as linhas e as bordas. Nas bordas, os *pixels* são extraídos pela mudança abrupta da intensidade de uma função imagem, e seus segmentos são conjuntos destes pixels convexos. Determina as fronteiras (bordas) das regiões de interesse. Se aplicada em detecção paralela de bordas, considera um conjunto de *pixels* pertencentes a uma borda. Se aplicada em detecção sequencial de bordas, o conjunto de *pixels* pertencentes a uma borda influencia os demais *pixels* vizinhos, sendo implementada por buscas heurísticas de estados.

Segmentação por limiarização de histogramas: Os algoritmos implementados verificam os vales e picos em histogramas para a relação de limiares que definam as regiões das imagens analisadas, sendo estas majoritariamente em tons de cinza, quando regiões de interesse são apropriadamente separadas, uma vez que o grau computacional aplicado é baixo.

Estendendo tal técnica em imagens coloridas, apenas é viável quando o histograma tridimensional é calculado por partes, segundo as componentes analisadas.

Segmentação por regiões: Nesta segmentação, a característica da imagem usada como informação de entrada é a disposição espacial dos *pixels*, cujas manipulações são feitas em cores não uniformes e em texturas, baseando-se no crescimento de regiões (seleção de um *pixel* ou pequena região que crescerá agrupando os *pixels* que satisfazem um critério customizado), divisão de regiões (partição de regiões não homogêneas em regiões homogêneas com o decréscimo de área dividida) e fusão de regiões (combinação de crescimento e divisão de regiões).

Segmentação por formação de agrupamentos no espaço de características: Como o espaço de características usado é o espaço de cores, e este apenas provê a informação cromática, a classe de algoritmos exige que a imagem formada por regiões detenha somente cores uniformes. Trata-se de uma técnica para algoritmos não supervisionados (*k-means* e *fuzzy c-means*), e para algoritmos supervisionados (MLP e *k-vizinhos mais próximos*) cujo cálculo do número de agrupamentos é relativamente complexo. A análise destes agrupamentos é feita através de concentrações de cores semelhantes em determinados grupos, então os *pixels* de um mesmo grupo recebem o mesmo rótulo, caracterizando-o como um segmento.

Segmentação por watersheds (divisão de águas) morfológicas: Baseada na interpretação topográfica da imagem, cujas três dimensões gráficas caracterizam as coordenadas espaciais e a intensidade. Neste tipo de segmentação, são considerados três tipos de pontos: os que pertencem a um mínimo regional, os que identificam *watermarks* (marcas d'água) em um mínimo simples e os pontos que identificam a mesma probabilidade de *watermarks* em um mínimo simples.

Segmentação por lógica difusa (nebulosa/fuzzy): Faz uso da teoria dos conjuntos difusos para a manipulação de incertezas e ambigüidades, no tratamento de informações em conjuntos de características de regiões similares em várias imagens.

Segmentação por Redes Neurais Artificiais (RNA): Baseado no conhecimento por treinamento de dados na rede neural, os segmentos são classificados e relacionados a determinados neurônios ativos, e estes dependem de características dos dados armazenados, como a tolerância a ruídos e desempenho de processamento computacional. Como as redes neurais artificiais são estruturadas em diferentes arquiteturas, podem-se realizar agrupamentos de cores homogêneas por rede de *Kohonen*, como também aprimorar o critério de homogeneidade por rede de *Hopfield*. RNA por Perceptron Multicamadas (MLP) são direcionadas a aplicações envolvendo segmentação supervisionada, que são relatadas em trabalhos iniciais de Eegmont-Petersen (EGMONT-PETERSEN *et al.*, 2002).

Dentre os métodos de segmentação apresentados acima, neste trabalho é adotado uma arquitetura modular de segmentação e classificação de imagens, abrangendo as técnicas de extração de características em subitens: algoritmos para agrupamento de características (abordagem estatística), algoritmos baseados em regiões que consideram a disposição espacial dos pixels, por metodologias não supervisionadas. Estas técnicas são aplicadas no contexto das imagens agrícolas, pois precisam ser projetadas e customizadas à segmentação de cenas naturais.

Neste contexto, é necessário avaliar tais algoritmos e validar seus respectivos resultados. Para a avaliação dos algoritmos são relatadas duas abordagens (CAVANI, 2007): subjetiva e objetiva. A abordagem subjetiva, usada em alguns trabalhos (JING *et al.*, 2003; CHEN *et al.*, 2005; WEI *et al.*, 2009), avalia se os segmentos são coerentes com a imagem, através da inspeção visual da imagem segmentada, o que a torna suscetível a erros. Na abordagem objetiva, dois modelos são aplicados (ZHANG, 2008) – modelo analítico e modelo empírico.

O modelo analítico considera a complexidade e propriedades do algoritmo, porém não de forma precisa na segmentação. O modelo empírico avalia indiretamente o algoritmo através das imagens segmentadas e, a partir destes, são calculados os índices de qualidade da segmentação em relação a algum critério estabelecido. Ainda, neste modelo, são considerados dois tipos de avaliação: adequabilidade e discrepância.

Na adequabilidade, são adotados critérios que medem a uniformidade dos segmentos, o contraste entre os mesmos ou a região geométrica de divisão, prescrevendo sua auto-avaliação de forma subjetiva. Na discrepância, são adotados critérios para avaliar os segmentos reais da segmentação ideal, através de imagens artificiais, o que os tornam objetivos e quantitativos, pois estabelecem a relação do número de pixels atribuídos incorreta ou corretamente para determinado segmento.

2.1.5. Reconhecimento de Classes e Padrões

Norbert Wiener (WIENER, 1942) *apud* Gonzalez e Woods (GONZALEZ e WOODS, 2007) fez a seguinte afirmação sobre reconhecimento de padrões:

“Um padrão é essencialmente um arranjo. Ele é caracterizado pela ordem dos elementos que o compõe, e não pela natureza intrínseca desses elementos.”

As metodologias de reconhecimento de padrões se dividem em duas categorias: a primeira, denominada de *decisão teórica*, abrange os padrões de descritores quantitativos; a segunda, denominada de *decisão estrutural*, abrange os padrões de descritores qualitativos. Ambas são usadas para extrair informações semânticas segundo algum critério.

Um padrão então é um arranjo de descritores, sendo estes referenciados por conjuntos de características, e uma classe de padrões é uma família ou agrupamento de padrões que possuem propriedades comuns, indicadas por $\omega_1, \omega_2, \dots, \omega_n$, sendo ω o número de classes.

Na visão computacional, a identificação de imagens é feita por meio dos descritores quantitativos, baseada no conteúdo registrado, como chave de busca para que o sistema

retorne imagens semelhantes (BUENO, 2001; CASTANÓN e TRAINA, 2003; PARK *et al.*, 2004; CHANG *et al.*, 2010; NANDEDKAR, 2010; RATLE *et al.*, 2010). A identificação dos elementos em uma imagem é usada para a manipulação física de peças ou objetos reais, através de robôs industriais articulados (TRONCO, 1999; PEDRO, 2008a; PEDRO *et al.*, 2008b; PEDRO e CAURIN, 2008c), bem como reconhecimento de culturas agrícolas, na identificação de irregularidades em frutos, plantas e obtenção de ângulos de guiagem pelas linhas de plantio (PLEBE e GRASSO, 2001; TANG *et al.*, 2003; ANNAMALAI *et al.*, 2004; BULANON *et al.*, 2004; THRUN *et al.*, 2004a; BLASCO *et al.*, 2007-a; BULANON *et al.*, 2009; SLAUGHTER *et al.*, 2009; WANG *et al.*, 2009; WEI *et al.*, 2009).

A classificação de padrões ou classes de padrões enfatiza a categorização das qualidades dos objetos de interesse, seja em uma imagem, som, ou qualquer taxonomia física na geração de uma referência. Os objetos com qualidades comuns são agrupados de forma que os grupos tenham significado/semântica. Assim, tais objetos podem ser denominados casos, unidade de dados, observações, eventos, indivíduos, entidades, relações e taxonomias operacionais (COSTA e CESAR JÚNIOR, 2001; CAVANI, 2007). Desta forma, cria-se uma hierarquia de informações organizadas, relacionando a quantidade em *bits* para sua representação e de sua classe.

Assim, conceituam-se as características dos objetos representados por valores numéricos, cujos agrupamentos formam vetores de características (*feature vectors*) para descrições quantitativas, *strings* e árvores para descrições qualitativas estruturais. A implementação do algoritmo de classificação tem como objetivo a divisão do espaço formado pelos vetores de características em regiões associadas às classes dos objetos, que são *n*-dimensionais, denominados espaços de características. A quantidade de características é calculada por *n*.

Para a discriminação dos aspectos relacionados com as etapas de desenvolvimento a um classificador de padrões e classes de padrões aplicados em imagens, são abordados os

seguintes procedimentos: composição das características dos elementos das imagens; normalização, extração e classificação de tais características; seleção, avaliação e validação do classificador; proporção da redução da dimensionalidade; erro e regra da decisão de *Bayes* em modelos estatísticos para classes gaussianas de padrões (SANTOS, 2006).

Composição das características dos elementos em uma imagem

Durante a escolha das características para a representação dos objetos de interesse, deve-se considerar como os vetores formados por estas estarão dispersos no espaço de características. Assim, são relacionadas as etapas para a caracterização dos objetos:

- a) Relacionar características distintas entre as classes, evitando ambigüidades, de forma objetiva com o mínimo de parâmetros;
- b) Evitar características correlacionadas, uma vez que são redundantes, exigindo maior processamento computacional e espaço para armazenamento;
- c) Evitar discrepâncias de padrão em objetos pertencentes às mesmas classes (*outliers*);
- d) Adotar quantidade mínima de características para a redução do custo associado ao processo de obtenção e classificação das mesmas;
- e) Relacionar características invariantes, com descritores quantitativos.

Para organizar os dados, respeitando as etapas acima, devem-se elaborar tabelas e gráficos de dispersão com informações das projeções dos vetores de características, pertencentes a um espaço n -dimensional. Para tal, como são inúmeras as características existentes na literatura, são considerados os objetos com relação intrínseca à visão computacional, pela forma de representação e descrição dos elementos. A imagem, ou o segmento que deve ser classificado é representado por uma matriz numérica com valores de intensidade da luz e cores, no espaço de cores analisado. Outras representações da imagem fornecem menores quantidades de

informações e auxiliam no cálculo de determinados descritores, que caracterizam os elementos.

Nessa representação da imagem ou do segmento, considera-se o contorno ou os *pixels* de preenchimento dos elementos. Com isso, para a abordagem da decisão teórica dos descritores, são detalhados a seguir (GONZALEZ e WOODS, 2007):

Excentricidade: Razão entre o comprimento do eixo-maior e eixo-menor. Este último é a reta perpendicular ao eixo-maior com comprimento igual a um dos lados do retângulo que envolve todo elemento. O outro lado do retângulo é do mesmo tamanho do eixo-maior, e chamado de retângulo básico, como mostrado na Figura 2.6.

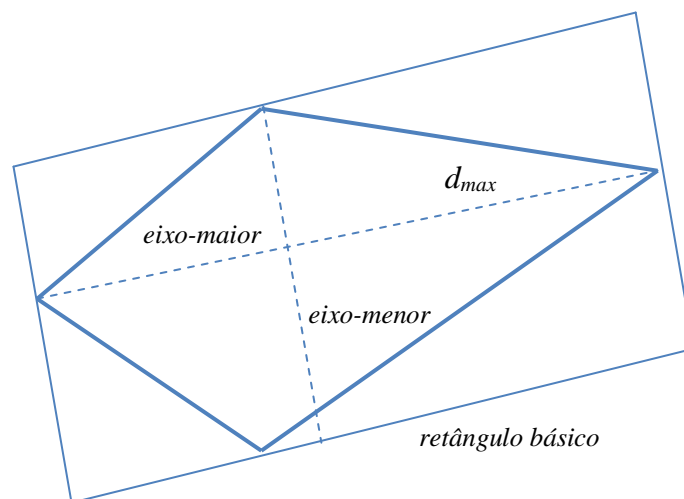


Figura 2.6: Eixo-maior, eixo-menor, retângulo básico e diâmetro do eixo maior (d_{max}).

Curvatura: Taxa de variação de inclinação da curva definida pelo contorno. Como, os contornos são localmente ondulados, são realizadas medidas diferentes de curvatura, através da diferença da inclinação entre dois segmentos adjacentes do contorno. Este descritor depende da relação entre o comprimento total e o comprimento do trecho a ser analisado.

Na Figura 2.7 abaixo, o segmento do contorno que representa uma linha reta, pode ser inclinado para várias direções, conectando pontos da região de contorno para uma dada ondulação, como na conexão de 3-2, iniciando o código de cadeia.

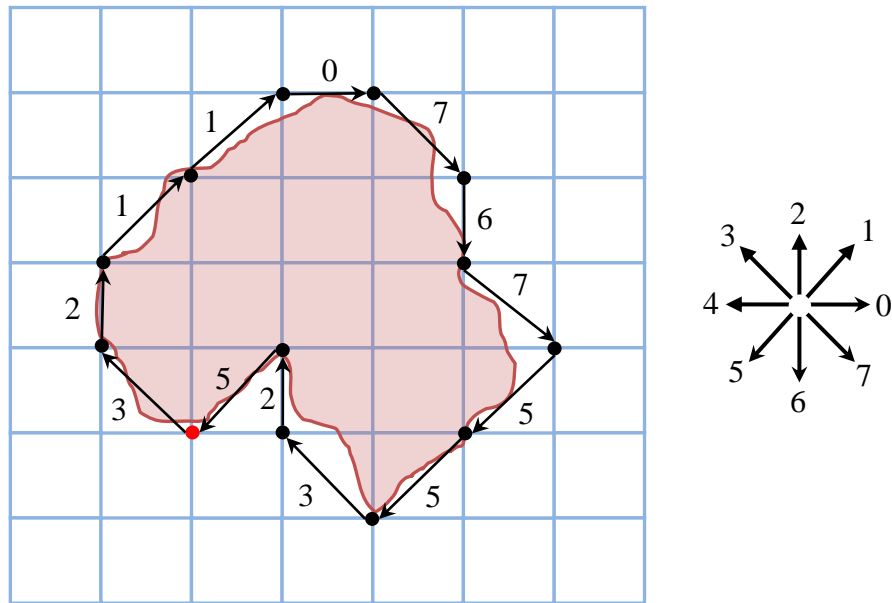


Figura 2.7: Código de cadeia (3211076755325) e direções dos segmentos baseada em região 8-conectada.

Número de formas: São calculados a partir da primeira diferença do código de cadeia para que seja a primeira posição da sequência, de forma circular. Os números de formas são independentes da rotação, mas não da orientação da grade formada pelos *pixels*, que pode ser alinhada com o retângulo básico, para a normalização da orientação. A quantidade de divisões da grade de *pixels* é determinada pela ordem do número de formas, representando os valores na cadeia. O procedimento se baseia inicialmente na escolha do valor para ordem, então o retângulo básico é dividido em uma grade, cujo perímetro é igual ao valor da ordem, a quantidade de divisões em cada um dos lados é proporcional aos eixos maior e menor. Caso resulte em uma cadeia com ordem maior que a escolhida, deve-se escolher uma grade com perímetro menor que a ordem.

Para a Figura 2.8, tem-se o código de cadeia em: 21110100033033232212. Supondo que inicialmente o valor da ordem seja 25, então, o número de formas será: 33003130030303130313, cujas diferenças são: 30031300303031303133.

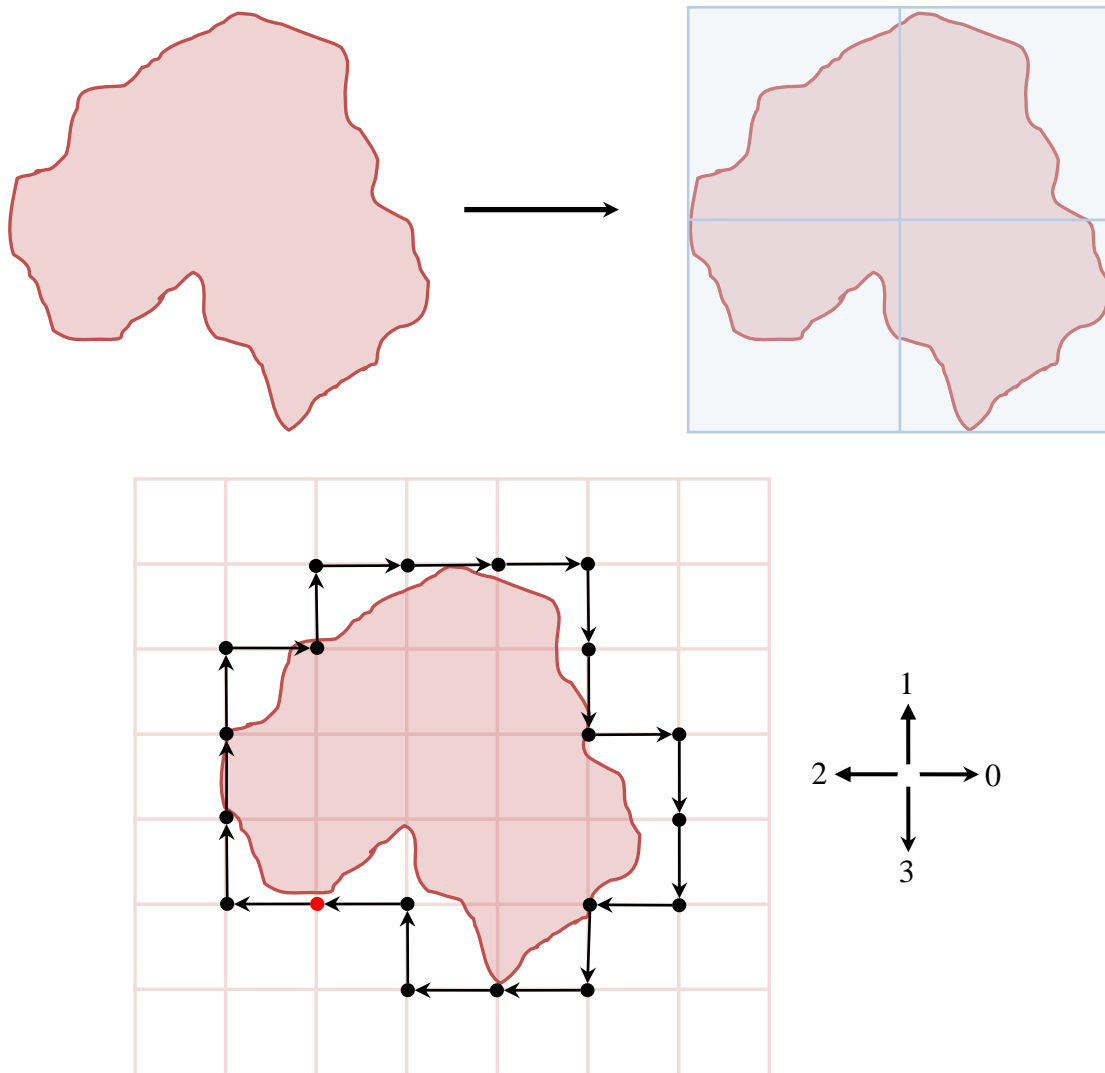


Figura 2.8: Número de formas, código de cadeias e diferenças para uma região de contorno.

Comprimento ou perímetro: Número de *pixels* do contorno da imagem ou código de cadeia, que é representado através de segmentos com tamanho fixo e direção variável. São usados para aproximar o contorno original, e cada segmento associa-se a um valor representativo de direção. O código de cadeia delimita a sequência destes valores. Assim, o comprimento da região de contorno é calculado pela soma entre quantidade de segmentos na vertical, na horizontal, e pelo produto da quantidade de segmentos na diagonal por $\sqrt{2}$.

Diâmetro: Distância mínima entre dois pontos pertencentes à borda de dado elemento, cuja reta que interliga tais pontos é denominada de eixo-maior, considerada como característica.

Momentos estatísticos: Dados pela média, variância e outros momentos, descrevem contornos representados por assinaturas. Estas simbolizam uma função com imagem unidimensional. A amplitude da função de assinatura é dada como variável aleatória, para o cálculo dos momentos. São descritores insensíveis a rotação e são normalizados com relação ao tamanho. Na Figura 2.9, observa-se tal decisão teórica de reconhecimento de padrões.

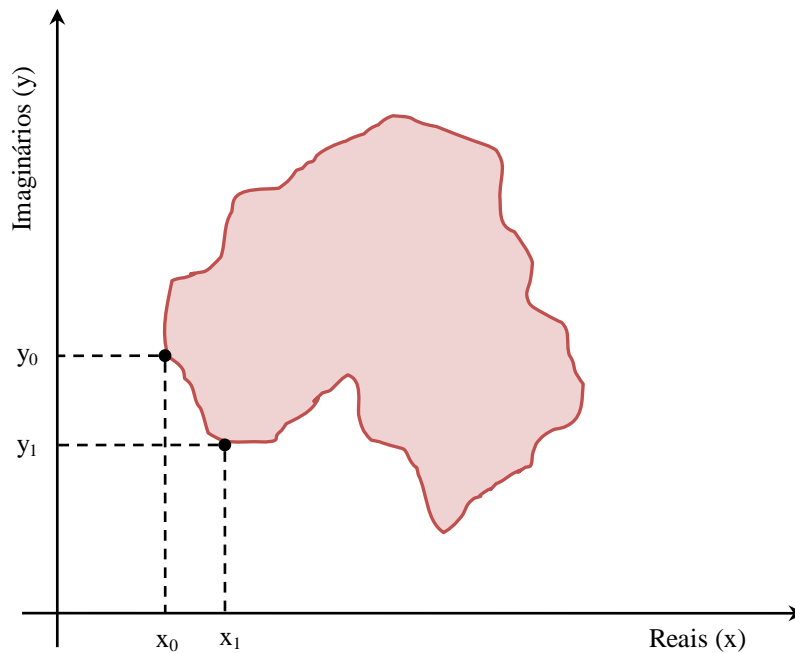


Figura 2.9: Representação do contorno por momentos estatísticos, com números completos.

Descritores de Fourier: A interpretação do contorno de um elemento é dada por uma função com imagem complexa, onde os valores reais e imaginários representam as coordenadas dos *pixels* da borda.

$$s(k) = x(k) + iy(k) \quad (2.14)$$

Na equação 2.14, acima, k é um pixel de borda, $x(k)$ e $y(k)$ são as coordenadas nos eixos x e y , respectivamente. Os coeficientes da transformada direta de Fourier são calculados para a função $s(k)$, para a denominação dos descritores, mas com quantidade predeterminada de coeficientes. Como o descritor de Fourier é insensível ao início da função que descreve o contorno, mas sensível a transformações geométricas, são adicionados parâmetros para correção desta limitação.

Para a relação dos descritores que utilizam representação interna, são listados abaixo:

Área: Aproximação do número de *pixels* no preenchimento da região da imagem (Figura 2.10). Os elementos avaliados por este descritor não sofrem transformações de escala, evitando distorções. Os elementos são discretizados pela relação da grade envoltória, e sua área é dada pelo preenchimento de cada quadrado, representando os *pixels*.

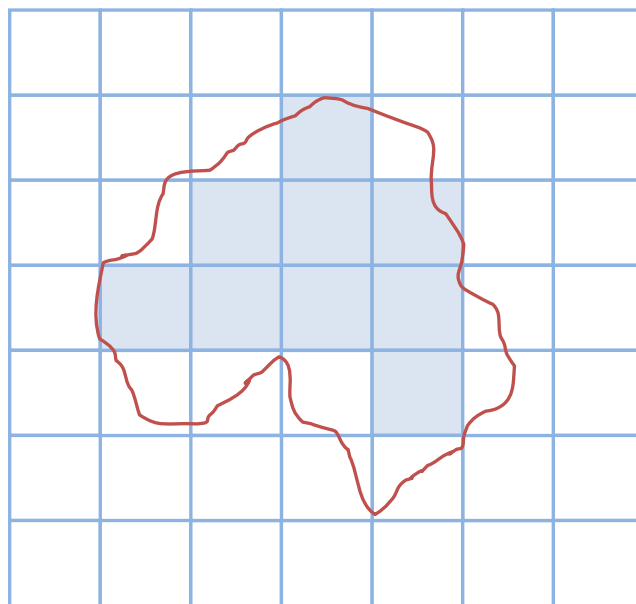


Figura 2.10: Área com regiões de elementos discretizados na imagem.

Compactação: O descritor de compactação é dado pela relação dos descritores da área e do perímetro (comprimento). A razão entre o quadrado dos valores destes últimos fornece o valor do descritor de compactação – mínimo para elementos com formato circular, sendo adimensional e insensível a mudanças de escala de orientação.

Topologia: Estudo das propriedades das imagens sem deformações, para a representação global das características, independentes de medidas de distâncias, tais como: buracos de elementos, número de elementos conectados e número de *Euler* (dado pela diferença entre o número de componentes e o número de buracos).

Textura e análise multi-resolução: Descritores caracterizados pela suavidade, regularização e direção. São extraídos de forma estatística, estrutural e espectral. Na primeira

abordagem, são calculados os momentos estatísticos a partir dos histogramas dos tons de cinza, fornecendo texturas suaves ou granulares, sem considerar a informação espacial. Por outro lado, considerando tal informação, os momentos estatísticos são calculados através das matrizes de co-ocorrência de tons de cinza, para obtenção de parâmetros como máxima probabilidade, uniformidade e entropia. Na abordagem estrutural, são usadas regras sintáticas para a criação de texturas com características mais complexas. Na abordagem espectral, é realizada a transformada de Fourier para analisar a textura à direção principal e ao período fundamental, representado pelo espectro polar.

Para a análise em multi-resolução, consideram-se as imagens com detalhes para maiores resoluções espaciais e imagens que apresentam padrões de texturas grosseiras, sem considerável resolução geométrica. Em ambos os casos, são usadas duas técnicas (transformada *Wavelet* e filtros de Gabor) para diferentes resoluções ou escalas numa mesma imagem, para a extração de seus descritores, que analisam o espectro em diversas escalas das imagens, relacionando as resoluções espacial e espectral, obtendo também, a orientação da textura. A transformada de *wavelet* analisa a direção das texturas e invariante a deslocamentos. Os filtros de Gabor tratam da análise em frequência, orientações e espaços de aplicação das texturas (FORSYTH e PONCE, 2002; HILL, 2002; GONZALEZ e WOODS, 2007).

Redução da Dimensionalidade: Normalização e extração de características

As características que compõem cada padrão, seja por textura ou cor, são interdependentes e podem não discriminar as classes entre si, sendo necessária a redução do conjunto de dados de entrada do sistema de reconhecimento de padrões para um conjunto mínimo pré-estabelecido por critérios ótimos. Considerando não apenas o fator computacional, como característica desfavorável ao trabalho com espaço de características de dimensionalidade

muito alta, há ainda a “maldição/mal da dimensionalidade”, que torna o espaço de características esparso e proporcional para o número de padrões em treinamento e número de dimensões (KORN *et al.*, 2001; SANTOS, 2006).

Em dada situação, padrões podem não possuir vizinhos conhecidos para o conhecimento de uma classe. A redução da dimensionalidade se adéqua a um tipo de pré-processamento aos padrões do classificador. Considerando algumas medidas invariantes a rotação e mudanças de escala, suas informações facilitam a escolha das melhores características e das transformações adequadas ao pré-processamento, uma vez que são conhecidos *a priori*.

Caso esta informação não esteja disponível, recorre-se ao redutor automático de dimensionalidade que pode transformar ou invalidar as características. Para a transformação, é realizado o processo de extração e normalização de características, estabelecendo a mesma arquitetura para o reconhecimento de padrões, do aprendizado até o classificador escolhido. Na invalidação das características, é realizado o processo de seleção ou classificação de características, como modelagem ao reconhecimento de padrões para as características mais relevantes.

Para a construção de um classificador, deve-se atentar para a etapa de extração de características no reconhecimento de padrões. Estas podem não originar espaços onde a dispersão dos vetores não seja proporcional a classificação, o que torna necessária a transformação do espaço de características para a discriminação das classes. Assim, são usadas técnicas para a redução do número de dimensões, projetando vetores n -dimensionais para o conjunto de características sobre os elementos, evidenciando o espaço de características mais apropriado. Tais técnicas podem afetar o espaço da forma e influenciar na dispersão das classes, assim deve-se avaliar o classificador final para a produção de resultados satisfatórios (COSTA e CESAR JÚNIOR, 2001; CAVANI, 2007).

Na normalização, as unidades das características influenciam a dispersão dos vetores no espaço, calculado em relação a uma referência, como maior valor dos dados. Formalizando algebricamente, a normalização das características ($F(i,j)$) é feita através da equação 2.15 abaixo:

$$F(i, j) = \frac{f(i, j) - \mu_j}{\sigma_j} \quad (2.15)$$

Com isso, o conjunto de dados terá média (μ_j) igual a zero e desvio padrão (σ_j) igual a um, sendo $f(i,j)$ o valor da característica j do vetor i .

Uma vez realizada a normalização, é feita a análise de componentes principais (ACP), com a transformação linear na minimização da covariância e maximização da variância. Segundo Haykin (2008), as características encontradas pela ACP são não-correlacionadas, evitando a redundância entre as características. Com isso, as características ou componentes principais são relacionadas pelos maiores tamanhos de informações contidas nos dados, reduzindo assim, o número de dimensões.

Para uma matriz $Z_{m \times n}$, cada coluna representa um vetor m -dimensional. A ACP é calculada, apresentando a matriz A com a projeção de Z sobre as componentes principais e Q é a matriz de transformação linear, composta pelos autovetores da matriz de correlação Z , que são organizados em ordem decrescente (equação 2.16).

$$A = Z^T Q \quad (2.16)$$

Os autovetores q_j e autovalores λ_j são encontrados pela equação 2.17 abaixo, sendo $R = E[ZZ^T]$, como a matriz de correlação de Z e $j = 1, \dots, m$. Já a matriz Q é obtida pela reescrita da equação de correlação, sendo $Q = [q_1, q_2, \dots, q_m]$ e $A = \text{diag}[\lambda_1, \lambda_2, \dots, \lambda_m]$.

$$Rq_j = \lambda_j q_j \rightarrow RQ = \Lambda Q \quad (2.17)$$

A matriz Q é reescrita apenas com os primeiros l autovetores, para que a projeção do espaço m -dimensional em um espaço l -dimensional seja possível, considerando $l < m$. O erro

quadrático médio desta projeção será o valor da soma das variâncias das componentes eliminadas, para mínimos valores de autovetores, onde $l + 1 \leq j \leq m$.

Redução da Dimensionalidade: Classificação das características

A classificação de padrões busca a divisão espacial de características em regiões, com vetores pertencentes a determinadas classes. Assim, as fronteiras das classes são definidas, através de pesos intrínsecos a cada característica de acordo com um critério imposto. O método e o critério são modelos que descrevem os dados e suas classes, como medidas de similaridade e dissimilaridade, estatísticas e limiares, critérios perto ou longe, probabilidade e níveis pressupostos.

Logo, tanto o método quanto o critério são conhecidos através de algoritmos que aproximem seus valores, o que leva a divisão de três técnicas: critério imposto, classificação por exemplos e critério aberto.

- a) *Critério imposto*: As classes e as qualidades dos objetos para a discriminação de uma classe são conhecidas e, conseqüentemente, o critério e o método. Desta forma, somente a extração de características é realizada.
- b) *Classificação por exemplos*: Denominada de classificação supervisionada, é usada quando há incertezas sobre o condicionamento dos objetos para uma classe determinada, mesmo que esta seja conhecida. Os algoritmos de classificação supervisionada de padrões são implementados de duas formas distintas: primeiramente ocorre a análise do conjunto de exemplos ou conjunto de treinamento de dados – o método e/ou critério e parâmetros associados a um modelo são denominados; logo o algoritmo realiza o reconhecimento de padrões através de uma comparação com os modelos de classificação existentes na base de dados (Figura 2.11).

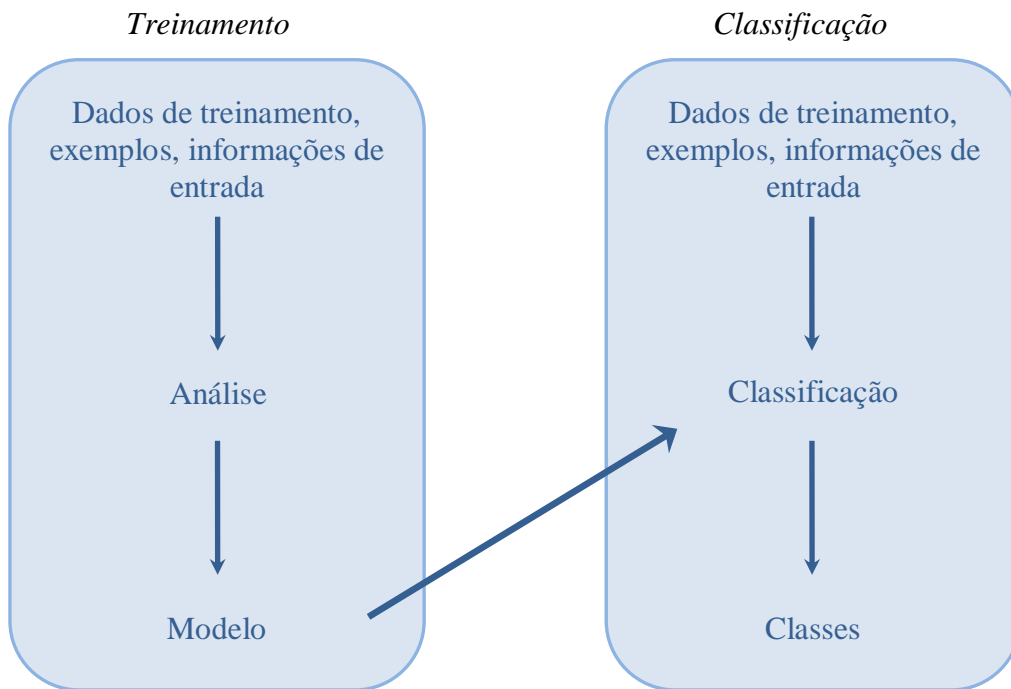


Figura 2.11: Classificação por exemplos (Adaptado de CAVANI, 2007).

Alguns classificadores supervisionados usam métodos específicos ou generalizados para sua abstração, da mesma forma para critérios específicos. Exemplos destes tipos de classificadores são os estatísticos bayesianos e diferentes topologias de Redes Neurais Artificiais (RNA) (HAYKIN, 2008), bem como modelos relacionais em Lógica *Fuzzy* (difuso-nebulosa) (ZADEH, 1965). No primeiro caso, adota-se o critério de maior probabilidade para estimar modelos dos dados (função distribuição/densidade de probabilidade – PDF) a partir das informações de treinamento. Nesse caso, o cálculo da probabilidade é realizado *a posteriori*. No segundo caso, classificadores analíticos conexionistas como as RNA, aprendem o método e o critério a partir de dados durante a etapa de treinamento. Estes dados são ponderados e reajustados em pesos sinápticos aos neurônios na camada de entrada na rede (vetor de características de um objeto), recalculados e propagados em direção a camada de saída da rede, aproximando a classificação em funções não-lineares, dividindo o espaço de características nas classes desejadas ou de referências.

i) *Teorema de Bayes e Bayes Ingênuo (filtragem e inferência)* (BAYES e PRICE, 1763; GNEDENKO, 1982):

O teorema de Bayes na classificação de padrões é tido como o modelo estatístico mais usualmente empregado, principalmente em inferências de dados relacionais que necessitam de alto grau de processamento na discriminação de características. Na equação abaixo 2.18, é possível identificar a probabilidade do evento B pertencer à classe A , em que $P(B)$ é definida como a constante normalizadora (COSTA e CESAR JÚNIOR, 2001).

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad (2.18)$$

Considerando que, conhecendo a probabilidade *a priori* de A ocorrer, e do evento B ocorrer para a classe A , tal equação denota as probabilidades em todas as classes pretendidas, e que resultar na maior probabilidade será escolhida. Devido à probabilidade condicional abaixo (equação 2.19), o problema na resolução consiste na determinação de k (classe de B) e do conhecimento prévio de C_i (evento relacionado à classe i), $P(C_i)$ (probabilidade da classe C_i ocorrer), $P(B/C_i)$ (probabilidade de B ocorrer para a classe C_i) e K (número total de classes), para que o problema seja reformulado (equação 2.20).

$$P(A|B) = \frac{P(A \cap B)}{P(A)} \quad (2.19)$$

$$k = \arg \max_i (P(B|C_i)P(C_i)), i = 1, 2, \dots, K \quad (2.20)$$

Assim, como a probabilidade condicional é indisponível nos casos em que a probabilidade *a priori* seja usada para a estimação de conjuntos de treinamento, o teorema de Bayes é reformulado novamente para que seja possível usar a função densidade de probabilidade (PDF – *probability density function*).

$$P(C_i | y) = \frac{p(y|C_i)P(C_i)}{\sum_{j=1}^K p(y|C_j)P(C_j)} \quad (2.21)$$

Na equação 2.21, acima, $p(y/C_i)$ é a PDF condicional da classe i , y é um valor escalar ou vetorial, já que neste último é viável o manuseio de várias características, cujas variáveis aleatórias são denominadas de multi-variadas com PDF conjunta (PDFc).

Para o classificador *Bayes Ingênuo* (*naive Bayes*), as características são independentes, o que caracteriza uma suposição improvável/incerta. Logo, para cada classe ou conjunto de classes, a PDF é supostamente denotada pela equação 2.22.

$$p(y | C_i) = \prod_{j=1}^n p(y_j | C_i) \quad (2.22)$$

Nesta formulação, $p(y/C_i)$ é a função de densidade marginal condicional, estimada como a PDFc, n denota o número de características. Como no teorema de Bayes, o treinamento e a validação são realizados da mesma forma, porém considera-se que a PDF admite várias características, já que a PDF marginal relaciona-se com apenas uma dimensão, sendo possível a junção de PDF marginais contínuas e discretas.

ii) Redes Neurais Artificiais (RNA) (HAYKIN, 2008):

Redes Neurais Artificiais (RNA) são modelos lógicos computacionais que processam de forma paralela, dados não-lineares e complexos, inspiradas no cérebro humano, precisamente pelas estruturas celulares independentes, os neurônios. Sua estrutura baseia-se na complexidade de processamento paralelo, taxa de aplicação e redução do tempo de execução. Uma RNA deve apresentar os conceitos naturais do cérebro humano, como a aprendizagem, a qual fomenta a obtenção do conhecimento a partir de uma aplicação, e as operações sinápticas ponderadas, que são as conexões intra-neurônios para o treinamento das informações.

Assim, um neurônio artificial possui dendritos como as informações de entrada, o corpo celular para a função de ativação, onde são realizados os cálculos ponderados, e os axônios, que denotam as informações de saída. O estímulo de entrada de informações é representado de forma algébrica por x_1, x_2, \dots, x_m , cujos pesos sinápticos são representados por w_{kl} ,

w_{k2}, \dots, w_{km} , e que resultam numa soma limiarizada por b_k , responsável pela ativação do neurônio em resposta ao estímulo inicial. Em seguida, a função de ativação $\varphi(\cdot)$ delimita o sinal de saída (axônio) y_k a um intervalo finito e definido. Tal modelo é equacionado a seguir (equação 2.23).

$$y_k = \varphi \left(\sum_{i=1}^m w_{ki} x_i + b_k \right) = \sum_{i=1}^m w_{ki} x_i \geq \varphi \quad (2.23)$$

A Figura 2.12, a seguir, mostra esquematicamente o modelo de neurônio artificial de *McCulloch e Pitts*.

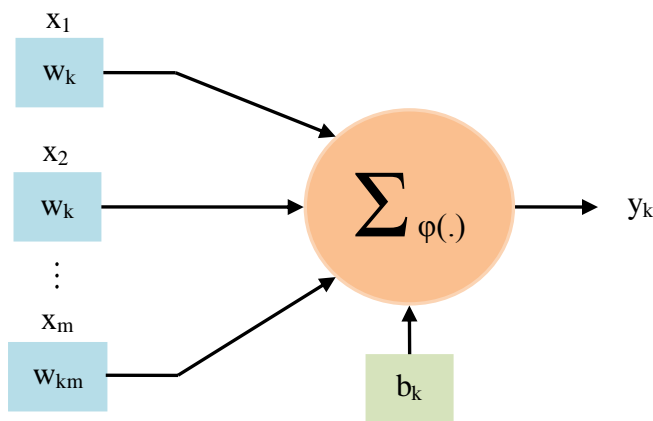


Figura 2.12: Modelo de neurônio artificial *McCulloch e Pitts*.

No modelo do neurônio artificial *McCulloch e Pitts* (McCULLOCH e PITTS, 1943), as entradas são disparadas simultaneamente a um estímulo e não dependem dos estados anteriores. Para tal modelo, os neurônios são organizados em camadas, cujas entradas são conectadas na primeira camada, e os cálculos ponderados na camada de processamento. Esta estrutura é denominada de *feedforward*, cujo processamento é sequencial nas camadas necessárias à RNA. Os pesos sinápticos e o limiar são ajustados com um algoritmo de treinamento supervisionado, baseado no erro da rede (diferença entre os valores de saída e os de referência), e são dedicadas apenas a funções lineares, como Adaline e Perceptron de camada única (HAYKIN, 2008).

Para camadas intermediárias (escondidas), os cálculos ponderados são direcionados à maior complexidade e tendem a resultar em maior precisão após a limiarização do intervalo pretendido. Assim, tal arquitetura é denominada de Perceptron Multicamadas (MLP – *multilayer perceptron*), e dedicada a problemas matemáticos não-lineares. A sua função de ativação deve ser contínua, não linear e diferenciável em qualquer ponto de convergência no cálculo do gradiente e no ajuste dos pesos sinápticos, para possibilitar a minimização do erro.

Tal erro depende intrinsecamente ao modo de aprendizagem (treinamento) da RNA. Esta, então, é definida como o processo de adaptação dos parâmetros livres (pesos sinápticos dos neurônios e os limiares de ativação) na estimulação pelo ambiente, a qual a RNA é inserida, e seu modo é determinado pela modificação de tais parâmetros. A aprendizagem pode ser não-supervisionada ou supervisionada. A aprendizagem não supervisionada não provê a existência de um professor para a discriminação do estímulo na resposta da função de ativação, o que relaciona apenas os padrões de entrada como a proporção e identificação dos dados, mediante uma regularidade estatística durante a classificação.

A aprendizagem supervisionada provê a existência de um professor, que conhece o ambiente de representação da RNA, submetido à avaliação dos resultados pela diferença entre os dados esperados e os dados obtidos (erro). Tal erro é usado para o ajuste dos parâmetros livres da RNA, a fim de torná-lo mínimo ou tão próximo de zero. Esse ajuste é conhecido como regra *Delta* generalizada (*Widrow-Hoff*). Na equação 2.24 abaixo, e representa o erro, k o neurônio após ajuste de peso e n a medida de tempo discreta.

$$E(n) = \frac{1}{2} e_k^2(n) \quad (2.24)$$

A regra *Delta* é generalizada no algoritmo *Backpropagation*, cujo erro da RNA é atribuído no ajuste dos pesos sinápticos da camada de saída (*forward*), e nas demais camadas ajustados pela propagação inversa (*backward*). Tal ajuste é proporcional ao gradiente da função erro (gradiente de descida) e pode não ser ótimo se considerar o mínimo local e não global da

mesma função. Para correção deste cálculo, o algoritmo *backpropagation* foi adaptado em propagação resiliente aprimorada, iRPROP (IGEL e HÜSKEN, 2000). Tal derivação provê ao algoritmo, adaptação ao gradiente de descida, maximizando-o e no ajuste da taxa de treinamento, minimizando as oscilações, tornando robusto e preciso. Além disso, o algoritmo iRPROP apresenta número reduzido de iterações, o que diminui o tempo de convergência.

iii) Lógica Fuzzy (difuso-nebulosa) (ZADEH, 1965):

Conjuntos difusos são uma generalização dos conjuntos clássicos. A diferença entre ambos consiste na variação de pertinência (0 ou 1) a determinado elemento para o modelo clássico, e $\mu = [0,1]$ como intervalo aos graus de pertinência no modelo nebuloso. Sua pertinência neste intervalo é relacionada à própria normalização dos dados (LULIO *et al.*, 2010b).

As funções de pertinência assumem várias formas geométricas (sino, triangulares, trapezoidais). Para a expressão da função trapezoidal utilizam-se quatro parâmetros ($\alpha, \beta, \gamma, \delta$) que correspondem aos vértices do trapézio. Os parâmetros α e δ representam os valores mínimo e máximo da variável analisada, que têm pertinência dentro do conjunto. São definidas as operações de união e interseção para estes conjuntos.

Para tais valores são atribuídos números difusos. Um número difuso é um valor numérico que representa um conjunto. As variáveis linguísticas representam vários conjuntos formados pelos números difusos, que assumem valores dentro de um conjunto de termos linguísticos, ou seja, palavras ou frases, assim elas assumem instâncias linguísticas.

As regras aplicadas para as instâncias linguísticas neste caso são do tipo condição-ação, nestas um conjunto de condições de entrada é ligada a ações que fazem com o processo estudado chegue às respostas esperadas. A idéia central é representar o conhecimento através de um conjunto de regras, nas quais as condições dadas a partir de um conjunto de termos

linguísticos são associadas às variáveis de saída/entrada do processo. A relação entre as entradas e as regras de saída é conhecida como inferência ou estado de inferência.

Assim, um controlador de inferências *Fuzzy* trabalha as variáveis da seguinte forma: as variáveis de entrada são *fuzzyficadas*, processo de escalonamento para normalização dos dados e transformação dos números difusos em conjuntos ou variáveis linguísticas. Criando a base de regras o controlador já possui uma base de conhecimento, a qual gera a estratégia de controle e os resultados. O processo de inferência, por sua vez, lida com as ações de controle para a saída, conhecidas como *defuzzyficação*, que transformam as variáveis linguísticas para números difusos. Um dos métodos mais utilizados em *defuzzyficação* é o método do centro de área (CDA), que retorna o centro da área do conjunto difuso de saída, ou conjunto inferido (GOMIDE e GUDWIN, 1994). Abaixo, na Figura 2.13, observa-se o digrama de blocos para um controlador *Fuzzy* com as etapas acima citadas.

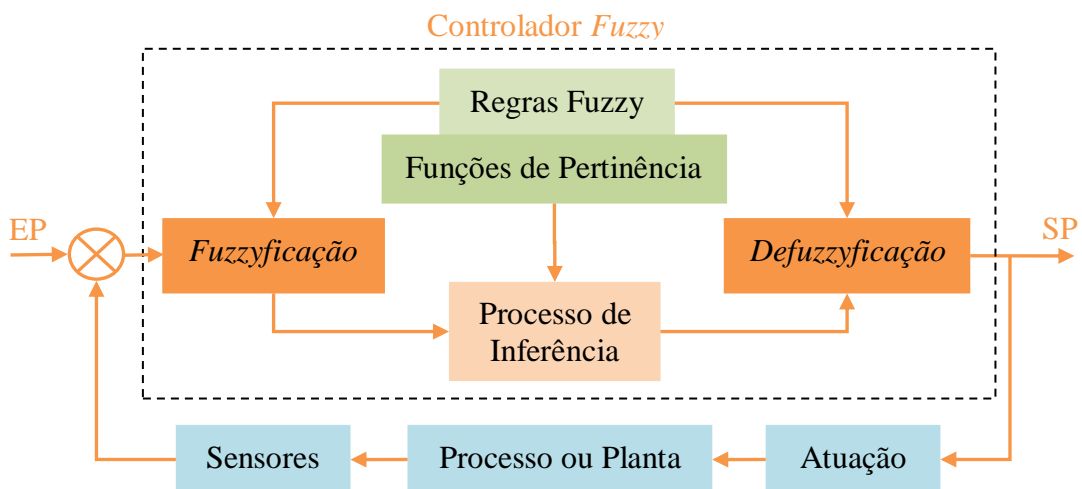


Figura 2.13: Diagrama de blocos de um controlador *Fuzzy*: (*EP*: entradas precisas; *SP*: saídas precisas).

c) *Critério aberto*:

Estes classificadores não possuem um conjunto de treinamento como etapa de aprendizagem, dividido em classes pré-estabelecidas, e tampouco são conhecidos os dados que serão classificados, tornando o método não-supervisionado.

Os algoritmos desenvolvidos para este critério buscam padrões evidentes nos objetos, agrupando-os em características semelhantes e com diferenças maximizadas entre os agrupamentos. São aplicados em métodos de função critério, por buscas recursivas e modelo *Branch and Bound* (LAND e DOIG, 1960).

No primeiro método o espaço de características é dividido em um número definido de agrupamentos, mediante os critérios estabelecidos previamente, que pode dispersar objetos em agrupamentos. Uma função critério $J(.)$ qualifica subconjuntos de características, recebendo como entrada um conjunto de números inteiros representando as características, e retornando um valor real para o subconjunto (SANTOS, 2006; RONCATTI, 2008).

As características selecionadas com uma função critério dependente são específicas para o algoritmo de reconhecimento de padrões correspondente, fornecendo resultados mais significativos. A dependência ocorre pela avaliação do algoritmo, e são detalhadas abaixo:

i) *Distância de Bhattacharyya:*

Define a distribuição normal para as classes i e j , como:

$$B_{ij} = \frac{1}{8} (M_j - M_i)^T \cdot \left(\frac{C_i + C_j}{2} \right)^{-1} \cdot (M_j - M_i) + \frac{1}{2} \ln \left(\frac{\left| \frac{C_i + C_j}{2} \right|}{\sqrt{|C_j| |C_i|}} \right) \quad (2.25)$$

Na equação 2.25, acima, é realizada a soma de dois termos, sendo o primeiro a distância de *Mahalanobis* entre os centróides das duas classes, em que estes são equivalentes aos valores de médias. M_i e M_j são vetores de médias e C_i e C_j são as matrizes de covariância das classes i e j , respectivamente. A diferença entre as classes é medida em relação às diferenças das médias e das covariâncias, se $C_i = C_j$, o segundo termo será reduzido a zero. Se $M_i = M_j$, o primeiro termo é reduzido a zero, e M_j corresponderá à diferença entre as covariâncias das classes.

ii) *Distância de Jeffries-Matusita*

Representado pela seguinte equação 2.26:

$$J_{ij} = 2\left(1 - e^{-B_{ij}}\right) \quad (2.26)$$

A distância de *Jeffries-Matusita* (JM) entre as classes i e j para a distribuição normal, estabelece que o valor de J_{ij} varie entre 0 e 2, sendo que a maior corresponde a separação completa das classes em questão. Caso a probabilidade *a priori* seja a mesma para todas as classes, a função critério é relacionada como a média entre a distância JM e todos os pares de classes.

iii) *Máxima verossimilhança*

Esta função critério baseia-se nas medidas de média e variância de cada classe dos padrões de treinamento como entrada de dados, dado pela distribuição normal na estimação da PDF de cada classe, que usa o critério bayesiano na classificação dos padrões para cada classe.

iv) *Distância mínima*

Esta função critério usa a classificação dos centróides de cada classe do conjunto de treinamento na entrada de dados, que equivale ao vetor de médias para as classes i e j . Primeiramente, usa-se a distância euclidiana, com variações para normalizada e *Mahalanobis*. O treinamento é feito através do cálculo dos vetores de médias e representa o quadrado da distância euclidiana para um dado padrão Z e o centróide M_i da classe i (equação 2.27)

$$e(Z, M_i)^2 = (Z - M_i)^T \cdot (Z - M_i) \quad (2.27)$$

Para a determinação da classe c (equação 2.28) do padrão Z , considerando o número total de classes como N :

$$c = \arg \min_{i=1, \dots, N} \left(e(Z, M_i)^2 \right) \quad (2.28)$$

Para a determinação da distância euclidiana normalizada (e'), considerando P_i como a matriz diagonal com as variâncias das características dos dados de entrada para o treinamento da classe i , a equação 2.29 denota a normalização.

$$e'(Z, M_i)^2 = (Z - M_i)^T \cdot P_i^{-1} \cdot (Z - M_i) \quad (2.29)$$

Pode-se também, minimizar a soma das distâncias, ou a dissimilaridade dos elementos de um mesmo agrupamento e seus membros, conhecido como algoritmo de busca recursiva, *k-means* (k-médias). Este é processado na seguinte sequência: a quantidade k de *clusters* (agrupamentos de padrões no espaço de características) é definida com um referencial de estimação; as coordenadas de um padrão são arbitrariamente atribuídas para cada *cluster*; cada padrão é identificado ao seu *cluster* mais próximo; o centro de todos os *clusters* é recalculado, baseando-se na média dos padrões que lhes pertencem. Esta sequência é repetida até a estabilização dos centros de todos os *clusters*, e o resultado depende da ordem da pesquisa dos padrões, do número de classes e da escolha inicial dos *clusters*, para total convergência.

Por último, na classificação por critério aberto, tem-se a seleção por *Branch and Bound*, que usa critérios de similaridade interclasses para direcionar a busca no espaço de soluções. Tal similaridade é calculada pelas distâncias *Mahalanobis* ou *Bhattacharyya*, usadas como critérios ótimos na determinação do melhor subconjunto de características por busca exaustiva. Assim, se a medida obtiver um aumento linear proporcional à dimensionalidade, usa-se o argumento de que o acréscimo de uma característica a um subconjunto sempre aumenta a precisão do sistema, e consideram-se as combinações possíveis de características numa estrutura de dados em árvore partindo-se dos subconjuntos maiores, que são os nós pais, para os menores, que são os nós filhos.

Seleção e avaliação/validação dos classificadores

Na classificação de padrões, modelos dos dados são criados em classes ou agrupamentos de interesse, apresentando incertezas intrínsecas ao método de classificação, o que torna necessário a validação dos resultados, para garantir a eficácia e eficiência do classificador. Diz-se que um modelo adequado generaliza, se é capaz de estimar coerentemente a classe de um vetor não-supervisionado, e sua capacidade é associada a complexidade das características envolvidas no modelo, como o número de parâmetros ajustados durante o treinamento. A complexidade do modelo fomenta a adaptação de dados organizados, mas não generaliza, o que certifica a validação de modelos equilibrados, para tanto generalizar, quanto aprender novos dados estruturalmente elaborados.

As técnicas para a seleção e validação dos modelos de classificação, são definidas inicialmente pelo erro esperado do teste e o erro real de treinamento. Considerando $X \in RP$ um vetor aleatório e $Y \in R$ uma variável aleatória de saída, o modelo estima Y , tal que X seja dado pela função $g(X)$. A função $\hat{g}(X)$ é a predição de $g(X)$ encontrada através do aprendizado. O erro entre Y e $\hat{g}(X)$, é denotada por $L(Y, \hat{g}(X))$ (CAVANI, 2007).

$$L(Y, \hat{g}(X)) = (Y - \hat{g}(X))^2 = |Y - \hat{g}(X)| \quad (2.30)$$

O modelo que generaliza apresenta menor erro do teste, disponível após o treinamento, pois quanto mais complexo o modelo, menor será o erro do treinamento e maior será o erro do teste (*overfitting*).

O erro de treinamento e o erro de teste (estimado de generalização) são dados abaixo, respectivamente, pelas equações 2.31 e 2.32.

$$\overline{erro} = \frac{1}{N} \sum_{i=1}^N L(y_i, \hat{g}(x_i)) \quad (2.31)$$

$$Erro = E[L(Y, \hat{g}(X))] \quad (2.32)$$

O erro de teste define a complexidade do modelo em sua seleção, analiticamente. O modelo com menor erro de teste é validado com o conjunto de testes, que representa um agrupamento de dados separados e não mais utilizado durante o treinamento. Logo, a avaliação para o classificador adequado passa pela etapa de treinamento, para a aproximação do modelo, a validação, para a seleção do modelo e o teste, para o erro de generalização, reforçado pela citação abaixo (FIGUEIREDO e JAIN, 2002).

Hastie (HASTIE *et al.*, 2003), afirmou que:

“(..) caso exista número suficiente de padrões, o erro de generalização pode ser estimado diretamente do conjunto de teste, sem a necessidade do uso de técnicas analíticas ou reutilização dos dados.”

A avaliação do classificador também é feita por meio da matriz de confusão (COSTA e CESAR JÚNIOR, 2001). Trata-se de uma matriz quadrada, onde as linhas e colunas representam as classes. As linhas denotam a classe de um padrão pertencente e as colunas denotam a classe estimada pelo classificador. Os valores associados aos elementos da matriz representam a quantidade de objetos já classificados, onde se verifica a distribuição do erro entre as classes, isto é, a existência de falsos positivos ou falsos negativos, para que seja calculada a matriz diagonal – proporcionalmente equilibrada.

Erro estimado e decisão de Bayes

O erro de Bayes permite a estimação da perda de informações quando os dados são submetidos a uma transformação, em que reduza complexidade de um problema na classificação. Trata-se da diferença entre o erro de Bayes das medidas iniciais e o erro de Bayes das características transformadas (SANTOS, 2006).

Caso o problema não apresente PDF matematicamente em termos de parâmetros, são usados métodos não paramétricos para sua estimação, como a estimação de densidade de Parzen e estimação de densidade do k -ésimo vizinho mais próximo. Ambos medem

localmente a densidade do espaço de características, e a vizinhança considerada é limitada pela distância (*Parzen*) ou limitada pelo número de vizinhos (*k*-ésimo vizinho).

Com isso, é possível calcular o erro de classificação dos dados e então, classificar os padrões segundo algum critério ótimo visto anteriormente, em que as funções de densidade de probabilidade representam a distribuição estatística dos padrões. Logo, o erro de Bayes é calculado para duas classes, dado pela equação 2.33, abaixo:

$$P(\text{erro} | x) = \min [P(\omega_1 | x), P(\omega_2 | x)] \quad (2.33)$$

A probabilidade do erro para a classe ω_i ($i = 1$ ou $i = 2$) é calculado pelo teorema de Bayes que relaciona a probabilidade condicional ou marginal, *a priori* ou *a posteriori*, respectivamente. Logo, o teste da decisão de Bayes retorna a classe mais provável, comparando-se a taxa de máxima verossimilhança $l(x)$.

$$P(\omega_i | x) = \frac{p(x | \omega_i) P(\omega_i)}{p(x)} \quad (2.34)$$

$$l(x) = \frac{p(x | \omega_1)}{p(x | \omega_2)} \rightarrow \frac{P(\omega_1)}{P(\omega_2)} \quad (2.35)$$

Caso a taxa de máxima verossimilhança for maior que a razão das probabilidades acima, entre as classes ω_1 e ω_2 , esta última terá mais chances de ser a correta, do contrário a classe 1 é considerada, e se ambas as classes forem equivalentes, compara-se $l(x)$ com 1.

2.2. Robótica móvel

Robôs móveis (RM) são veículos automatizados que se locomovem adequadamente em vários ambientes, sejam terrestres, aquáticos, aéreos. Quanto ao tipo de estruturas de mobilidade, são classificados pelos atuadores responsáveis pela locomoção, e quanto ao tipo de acionamento são classificados em duas abordagens: *robôs móveis guiados* e *robôs móveis autônomos* (SIEGWART *et al.*, 2011).

Os robôs móveis guiados são capazes de se movimentar com o auxílio ou interação de usuários externos, sendo local ou remotamente, para que a decisão quanto à locomoção e tarefa do veículo seja baseada nas opções de navegação ao condutor. Por outro lado, robôs móveis autônomos se locomovem e realizam tarefas sem a intervenção externa de usuários, mantendo sua autonomia de navegação sem colisão. Há aplicações que exigem tanto autonomia quanto guiagem para as tarefas em determinado robô móvel, caracterizando-o como *veículos guiados autônomos (AGV – autonomous guided vehicle)*.

Para as duas abordagens e a unificação de ambas, são necessários estudos quanto à modelagem matemática e concepção física de robôs móveis empregando teorias científicas para os problemas de locomoção e os mecanismos para tal, modelos algébricos que caracterizam a cinemática e dinâmica, algoritmos e técnicas analíticas, matemáticas, cognitivas e estatísticas para a interpretação de dados na etapa de percepção, auxiliando os processos de localização, mapeamento, navegação e planejamento de trajetórias. Por fim, neste subitem, são relatadas pesquisas sobre as tecnologias existentes no estado da arte, aplicadas na área agrícola.

Basicamente o problema de locomoção do RM consiste em obter um modelo físico ou perceptivo do ambiente, observar e analisar a estrutura do ambiente, determinar sua posição neste, e planejar e executar uma trajetória para uma tarefa específica, realizando sua locomoção. Com isso, é possível ilustrar o problema e sua resolução pelo diagrama de blocos abaixo, na Figura 2.14.

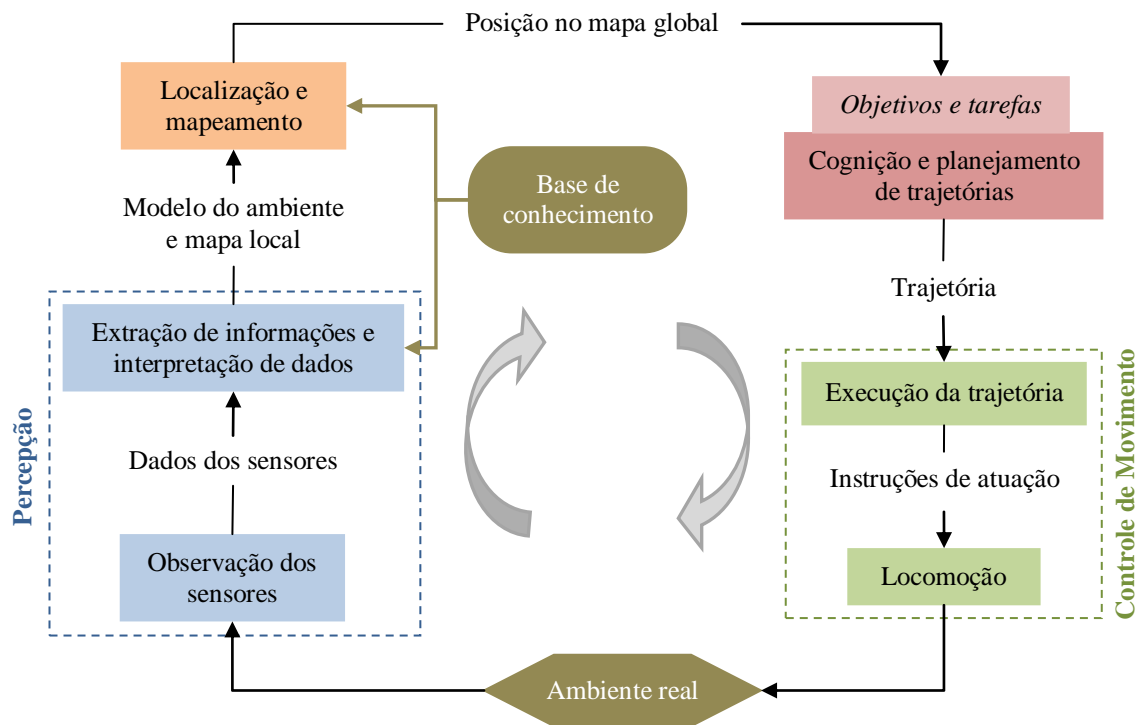


Figura 2.14: Diagrama de blocos para o problema e solução da locomoção de robôs móveis (Adaptado de SIEGWART *et al.*, 2011).

Como as etapas de construção de robôs móveis contemplam aspectos multidisciplinares aos mecanismos de locomoção, acionamentos, modelagens cinemática e dinâmica de sistemas de controle, o contexto estudado neste trabalho será especificamente direcionado à etapa sensorial (percepção), à localização e navegação dos mesmos, detalhando as técnicas matemáticas e algoritmos existentes para esta tarefa, precisamente relacionados à visão computacional.

2.2.1. Percepção e incertezas na estimativa de dados

A etapa de percepção, ou sensorial, lida com a aquisição de dados (medidas) sobre o ambiente que o robô móvel está envolvido. Tais dados são obtidos por sensores dedicados à sua captura, que provêm também a extração de informações úteis na interpretação destes dados. Esta etapa trata diretamente com as características dos sensores, das informações físicas do ambiente e com as incertezas durante a aquisição.

Os sensores usados para tal tarefa têm duas distinções quanto ao tipo de medida, sendo proprioceptivos, que medem valores internos ao robô móvel, e exteroceptivos, que medem valores ou estímulos externos ao robô móvel. Em ambas as abordagens, os sensores são ainda classificados quanto ao modo de leitura (analógico e digital), e quanto à emissão de energia (independente da forma): sensores *passivos* identificam a energia emitida pelo ambiente; sensores *ativos* identificam a energia refletida pelo ambiente, uma vez que a emissão é realizada pelos sensores e estes lidam com a reação aos dados mensurados (THRUN *et al.*, 2005).

As características dos sensores inicialmente são baseadas na quantificação da resposta, seja em frequência ou tempo de atraso, tornando a resolução das medidas dependentes de vários fatores: alcance dinâmico, linearidade, largura de banda, características *in situ* que necessitam da compreensão do ambiente real (sensibilidade, erro sistemático, erro aleatório, exatidão, precisão). Tais fatores originam desvios na leitura, como incertezas de dados, o que torna a validação destes, altamente complexa. A origem das incertezas é devido à inexatidão dos equipamentos/sensores de medida (*bias uncertainty*) e pelas variações aleatórias na medida (*precision uncertainty*).

A incerteza é um problema gradativo e deve ser resolvido localmente, para cada série de medidas, já que sua indefinição provê a propagação de dados incertos, ou seja, a incerteza dos resultados finais depende da incerteza de cada medida individualmente. Logo, o modelo matemático que caracteriza a incerteza máxima (RSS – *root sum squares*) numa relação de medidas à propagação de incertezas é dado pelo condicionamento do somatório empregado às incertezas das variáveis, juntamente com o coeficiente de sensibilidade do sensor (THRUN *et al.*, 2005).

Para estimar, então, incertezas aleatórias, assume-se que a variável medida tem n valores num intervalo de tempo. Calcula-se o desvio padrão das medidas para resolver a curva normal

padrão da média numa distribuição t de *Student* (HELMERT, 1875; HELMERT, 1876a; HELMERT, 1876b). Para incertezas sistemáticas, estas permanecem constantes se a validação é repetida nas mesmas condições, são também independentes do tamanho da amostra, e são especificadas pela calibração dos sensores e sua modelagem matemática, no caso de câmeras CCD. A incerteza total é a combinação das incertezas aleatórias e sistemáticas usando a soma dos quadrados das raízes da solução RSS, para múltiplas amostragens, como para uma única. Neste sentido são adotados modelos estatísticos à resolução das incertezas totais, como o filtro de *Bayes* que denota a predição e correção das incertezas, e distribuições de *Gauss* multivariáveis por elipses de erro, autovalores, matriz de covariância e coeficientes de correlação.

Na robótica móvel, os movimentos dos robôs móveis são inerentemente sujeitos a incertezas das informações adquiridas, sintetizados através do modelo de transição de *Bayes* $p(x | x', \mathbf{u})$, que especifica a probabilidade posterior que a ação u causa no robô móvel de x a x' . A modelagem de $p(x | x', \mathbf{u})$ é baseada nas equações de movimento do RM, cuja configuração é descrita por seis parâmetros (três coordenadas cartesianas e os três ângulos de *Euler*). As equações de movimento são então resolvidas para dois modelos de locomoção: *odometria*, e *velocidade (dead-reckoning)*, os quais são sujeitos à junção de informações dadas pelos sensores proprioceptivos, e a correção das incertezas por sensores exteroceptivos, como no caso de visão computacional.

Com isso, para que o modelo de localização do robô móvel seja ótimo e robusto na solução de incertezas à sua predição, este é definido por abordagens analíticas e estatísticas baseadas na construção de mapas do ambiente. Os detalhes de tais processos são relatados na próxima seção, mas devem seguir a seguinte metodologia: inicialmente o processo de localização depende das observações (dados dos sensores ou características extraídas do ambiente) que são armazenados numa base de conhecimento e atualizam a posição (ou sua

estimativa *a priori*). Como é um sistema a malha fechada, a posição estimada anteriormente é combinada com outras informações provindas da odometria e *dead-reckoning* (sensores proprioceptivos), para que possam ser equiparadas novamente com as leituras atuais de sensores exteroceptivos (sistemas de visão computacional), estimando a posição atual *a posteriori*. Assim, a percepção e o movimento são relacionados às tentativas de resolução das incertezas, como o ruído dos sensores e seu efeito *aliasing*, como os erros determinísticos e não-determinísticos inerentes à odometria.

2.2.2. Localização por mapas, navegação e planejamento de trajetórias

Para a locomoção autônoma de robôs móveis são usados algoritmos estatísticos e técnicas heurísticas de Inteligência Artificial na localização, navegação, mapeamento e planejamento de trajetórias (GOLDBERG *et al.*, 2002; BACZYK *et al.*, 2003; ANDREASSON e DUCKETT, 2004; BONATO, 2007). O uso da teoria probabilística, com tais heurísticas torna os algoritmos para as tarefas de locomoção do robô mais robustos em relação ao grau de incerteza ou leitura de ruídos pelos sensores presentes no ambiente de trabalho (THRUN *et al.*, 2004a).

Navegação e planejamento de trajetórias

No sistema de navegação, deve-se considerar a informação como item fundamental à tomada de decisões elaboradas e eficientes, considerando as seguintes etapas como estrutura ideal ao processo de locomoção veicular (SCATENA, 2003).

a) Direção:

Representado por módulo eletrônico de decisões relacionais, cuja característica é estruturar as atividades de direção em: planejamento, percepção, análise, decisão e controle. O planejamento é atribuído antes da locomoção física do veículo e, em vários casos, necessita-se alto nível de processamento computacional, uma vez que é realizado durante o percurso, em

tempo real. Consideram-se também as prerrogativas de ambientes complexos, acidentados, e suscetíveis a mudanças de composição terrena, tornando os cálculos de percepção e análise mais sofisticados para tomada de decisões e controle mais apurados.

b) Veículos:

Plataformas físicas para a locomoção e consequente navegação, considerando a dependência das aplicações envolvidas (cargas, transportes, tarefas pesadas) e podem-se movimentar por meio de atuadores elétricos, mecânicos, hidráulicos, pneumáticos, supervisionados pela estratégia de direção.

c) Ambiente:

Local externo a ser explorado pelo veículo, caracterizado por vias terrenas urbanas, rurais, acidentadas e de diferentes composições de contato, em locais internos (*indoor*) e externos (*outdoor*).

Assim, considerando as três etapas supracitadas para a locomoção, esta depende de como é feito o posicionamento do veículo, para uma dada tarefa naquele local do percurso e, através dos modelos de localização e mapeamento vistos a seguir, e dos sensores à etapa de percepção à locomoção, o sistema de navegação de robôs móveis necessita de um modelo de aprendizagem de direção para a previsão de itens contidos no ambiente de navegação ou planejamento de ações dentro do ambiente explorado. São métodos que assistem a uma arquitetura de navegação inteligente (SCATENA, 2003).

Reconhecimento de eventos: Dados provindos da etapa de sensoriamento, cuja informação é traduzida em símbolos como eventos.

Aprendizado de padrões: Dados coletados da etapa de sensoriamento, cuja classificação respeita um tipo de padrão associado, que pode ser supervisionado ou não supervisionado.

Predição de dados: Estimativa e avaliação dos dados para situações a serem desenvolvidas.

Estes métodos são fundamentais para uma estratégia de navegação robusta e estabilizada. Aos seus propósitos, são detalhadas a seguir, técnicas e metodologias de localização e mapeamento ao planejamento de trajetórias e navegação desejadas.

Localização

A localização determina a estimação da posição do robô móvel no ambiente, necessária a sua autonomia de trabalho, considerando inicialmente, a base do conhecimento do ambiente por mapas e sua relação com as informações fornecidas pelos sensores. É conhecida também como *position estimation/tracking*, e une informações provindas da fusão sensorial do robô móvel, para evitar ambigüidades nas leituras e níveis de ruído durante a aquisição de dados. A dificuldade encontrada nos cálculos de estimação da localização do robô móvel é variável e depende do tipo de conhecimento disponível *a priori*, e durante o percurso, da localização passiva ou ativa, considerando o ambiente explorado em estático ou dinâmico.

Na localização ativa o algoritmo relaciona seus parâmetros diretamente na locomoção do robô móvel e as dificuldades encontradas pelos obstáculos que, como citado anteriormente, dependem de situações dentro de um ambiente estático, quando todos os objetos são fixos, ou de um ambiente dinâmico, quando há objetos móveis na cena explorada.

Se a informação da posição é obtida inicialmente, a próxima localização é relativamente simples de se denominar, pelas aproximações dos dados anterior, estimado e atual – é chamada também de posição relativa, incremental ou local, e informa a localização pela distância entre o veículo (robô móvel) e algum objeto inerente ao ambiente, cuja leitura é realizada por sensores proprioceptivos, como odometria e sensores exteroceptivos, como opto-receptores.

Se a informação inicial não é conhecida, o cálculo da localização dá-se pela análise dos dados capturados e armazenados, considerando-se mapas do ambiente previamente conhecido ou por marcos de referência (*landmarks*), inerente ao ambiente – é chamada também de

posição absoluta, e informa a localização global através de um sistema de coordenadas absolutas (GPS – *Global Positioning System*) (TANGERINO, 2009).

A diferença entre os dois modelos (relativo e absoluto) consiste na modelagem dos ruídos dos sensores, o que na posição relativa são modelados através de distribuições unimodais (Gaussiana), e na posição absoluta os ruídos são modelados através de distribuições multimodais (filtro de partículas) (BONATO, 2007), que caracteriza a ambigüidade de informações e suas resoluções por múltiplas hipóteses (*multi hypothesis*). Processos de decisão de Markov e Gaussianos são métodos probabilísticos usualmente adotados para a localização (THRUN *et al.*, 2005), assim como Monte-Carlo (KAESS *et al.*, 2009), para os modelos relativo e absoluto.

Mapeamento

Robôs móveis autônomos usam algoritmos de mapeamento para a geração automática de representações do ambiente ao seu sistema de controle de navegação. Segundo DeSouza e Kak (DeSOUZA e KAK, 2002), *apud* Bonato (BONATO, 2007), há três tipos de navegação autônoma:

“(...) a navegação autônoma pode ser classificada em três grupos: a baseada em mapas criados pelo usuário; a baseada em mapas criados pelo próprio robô; e as que não possuem mapa. A navegação com mapa proporciona maior poder de decisão durante a navegação, permitindo, por exemplo, localizar o robô no ambiente e planejar a sua rota de acordo com o mapa. Dentre os três tipos de navegação, os robôs que constroem o seu próprio mapa são considerados mais difíceis de serem desenvolvidos, porém são os que possuem maior nível de autonomia de navegação para lidar com os problemas do mundo real.”

Os algoritmos de mapeamento para os três tipos de navegação acima, dependem diretamente do ambiente explorado (estruturado/desestruturado, interno/externo, estático/dinâmico), e em situações onde este é altamente dinâmico, os objetos móveis são tratados como ruídos do sistema. Em ambientes internos e externos, a preocupação no tratamento das informações adquiridas é desenvolver algoritmos que possam definir de forma

coerente a estrutura do ambiente, cujos dados são relacionados como representações geométricas dos objetos pertencentes à cena, sejam corredores, portas, janelas (*indoor*) e ângulos de guiagem, sinais de trânsito (*outdoor urbano*), bifurcações ou linhas de plantio (*outdoor rural*). Da mesma forma, para a representação topológica do ambiente, devem-se considerar os nós/vértices para objetos significantes da cena e arestas interligando-os, construindo uma estrutura baseada em grafos.

A composição do mapa geométrico dá-se pela informação de sensores proprioceptivos, como a odometria, na transformação bidimensional do ambiente, tornando-o geométrico (SILVA, 2008). Na abordagem topológica, o ambiente é representado por lugares distintos entre si (LIU *et al.*, 2009; SHI *et al.*, 2010), que definem as funções de reconhecimento de lugares dentro do ambiente e o planejamento de trajetórias para a construção destes tipos de mapas. Com isso, a criação de mapas, seja geométrica ou topológica deve interagir com a locomoção precisa do robô no ambiente, o que caracteriza a estimação de sua posição em vários momentos durante o percurso, localizando-o de forma desambigua pelas informações do mapa e coerente dado pela representação relativa/global da sua posição. A essa problemática foi introduzida a técnica SLAM/SLAMB (*Simultaneous Localization and Mapping/Simultaneous Localization and Map Building*) ou CML (*Concurrent Mapping and Localization*), que consiste em construir o mapa do ambiente e estimar a posição do robô móvel neste mesmo mapa simultaneamente, sem informação inicial (SMITH *et al.*, 1988; SMITH *et al.*, 1990; THRUN, 2002; THRUN *et al.*, 2004a; THRUN *et al.*, 2004b; WOLF *et al.*, 2005).

No modelo de mapa geométrico, o ambiente é representado através de marcos de referências ou grades dimensionadas em coordenadas cartesianas, identificadas por células que indicam a presença ou ausência do robô móvel.

Mapas geométricos possuem duas representações: *landmarks* (marcos de referência) (GRITAI *et al.*, 2009; WU *et al.*, 2009) e *occupancy-grid* (grades de ocupação) (KENJI e NAOKI, 2009; RICHTER *et al.*, 2009; WURM *et al.*, 2010).

a) Mapeamento por landmarks:

Landmarks são marcos de referência para a construção do mapa no ambiente explorado. A localização do robô móvel no ambiente é obtida através destas referências, sendo estas naturais (pertencentes ao ambiente) ou artificiais (inseridas no ambiente).

Para o processo de detecção dos marcos de referência é necessário primeiramente a aquisição dos dados do ambiente pela etapa sensorial, e sua análise para a identificação de características dos marcos; logo, tais características são armazenadas na base de dados do sistema de navegação para que sejam referenciais de padrão aos novos *landmarks* adquiridos durante a locomoção do robô móvel.

b) Mapeamento por grades de ocupação:

A representação do ambiente é dada por espaços livres ou espaços ocupados na geração do mapa. Consiste na discretização do ambiente em uma grade de mesma dimensão da área regular, considerando várias divisões geométricas em retângulos ou quadrados, denominados de células (Figura 2.15).

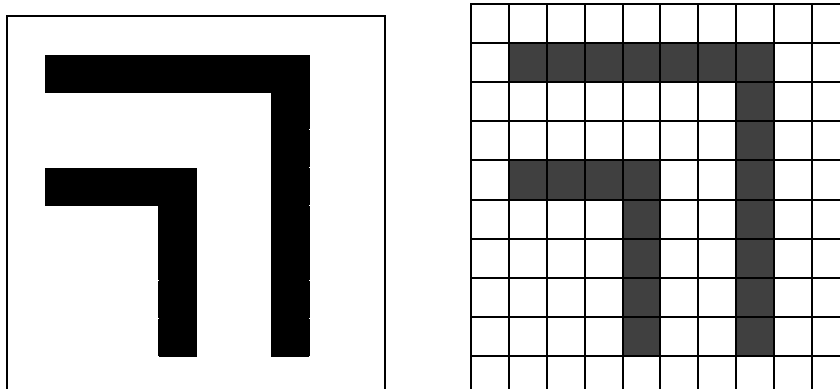


Figura 2.15: Grade de ocupação, com as células representando o espaço ocupado das células.

Assim, cada célula é calculada em sua probabilidade de estar preenchida ou não. Para informações visuais com resolução espacial/cromática muito alta, sua aplicação não é ideal devido ao alto nível de processamento requerido e pela quantidade de *pixels* existentes nas imagens contendo características de cenas naturais.

Para a construção de mapas topológicos, os dados coletados por referência de posição absoluta, possuem considerável nível de informação abstrata, referentes às características do ambiente real explorado, e apresentam assim, grafos para a descrição do mapa (vértices representam os locais identificados e as arestas representam as interligações entre estes).

A construção de um mapa topológico depende inicialmente da leitura dos sensores instalados no robô móvel, dita como inicial a partir do instante de locomoção, o que caracteriza desinformação do ambiente *a priori*. Logo, as interconexões entre os vértices (locais visitados), formalizam a exploração do mapa, durante o percurso. Desta forma, a topologia do mapa a ser construído considera os locais confirmados pela presença do robô móvel e os locais estimados, ainda pendentes de visitaç o do robô móvel (SCATENA, 2003; SILVA, 2008).

Mapas topológicos seguem a estratégia de exploração abaixo:

- a) *Reconhecimento da localização*: Considera a autolocalização do robô móvel dentro do mapa a ser construído, determinando a posição provável ocupada pelo mesmo, e seu deslocamento relativo a determinados locais (nós).
- b) *Detecção de locais inexplorados*: Locais estimados são adicionados no mapa topológico através da orientação do robô móvel.
- c) *Planejamento de trajetórias*: Aplicado pelo algoritmo de *Dijkstra* (DIJKSTRA, 1959), define o trajeto interconectado da posição atual a posição estimada.

- d) *Estimação da posição final*: Locais confirmados no mapa topológico são adicionados mediante o percurso do robô móvel em seu trajeto, considerando limites de distâncias para as posições estimadas evitando obstáculos.
- e) *Estabilidade*: A odometria como etapa sensorial proprioceptiva é suscetível ao acúmulo de erros causados pelo atrito das rodas com o solo, o que invalida a posição atual do robô móvel como a final, durante a construção do mapa.

Segundo Silva (2008), a representação dos mapas topológicos dá-se por duas formas distintas, como ilustrado pela Figura 2.16.

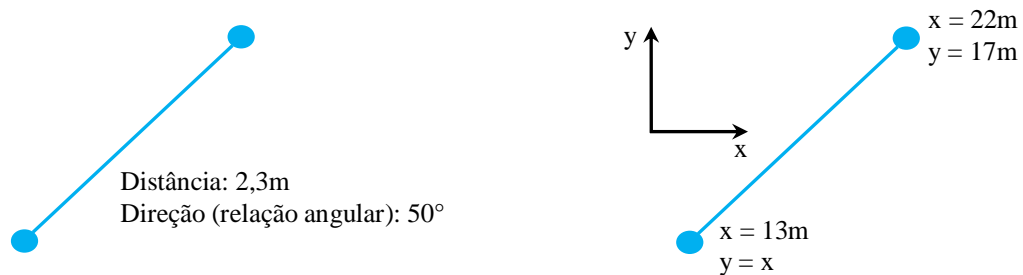


Figura 2.16: Informações métricas adicionadas aos mapas topológicos (Adaptado de SILVA, 2008).

SLAM com informação visual

Neste tópico são apresentadas as aplicações dos algoritmos de SLAM, vistos anteriormente, para a construção de mapas baseados na aquisição da informação visual, por estruturas de sistemas de visão computacional monocular, estéreo e omnidirecional. A relação com as etapas de processamento digital de imagens adquiridas para com os sistemas de visão também é detalhada, identificando como é feita a associação dos dados visuais no gerenciamento e construção do mapa.

A concepção estrutural das etapas de mapeamento e localização simultâneos com o uso da informação sensorial por visão passa, inicialmente, pela aquisição de imagens, análise das componentes de interesse e extração de características, com a detecção de *landmarks* e informações de profundidade (estimação do movimento); logo, são associadas e gerenciadas (adição, remoção e substituição e dados), as informações antes processadas, para a orientação

inicial na construção do mapa; o armazenamento e reconhecimento de informações métricas ou topológicas fornecem a construção final (atual) do mapa.

A aquisição dos dados, sendo a etapa sensorial do sistema de visão/SLAM, é tratada por sensores estereocéptivos em tecnologias como CCD (*Charged Coupled Device*) e CMOS (*Complementary Metal Oxide Semiconductor*) (COMPANY, 2003; DALSA, 2005), sendo a primeira mais flexível na conversão de dados fotométricos (fótons em elétrons) gerando a informação visual - imagem, e fornecendo maior velocidade de processamento e resolução espacial das imagens. As duas tecnologias são associadas com sistemas de visão monocular (uma câmera), estéreo (duas ou mais câmeras), e *omnidirecional* (LULIO *et al.*, 2010a). Nestas duas últimas, a obtenção da informação visual contém dados precisos da profundidade e relações de distâncias entre o robô móvel e objetos pertencentes ao ambiente.

A construção de mapas de ambientes é representada por informações geométricas ou topológicas relevantes para a navegação autônoma de robôs móveis. Tais informações são adquiridas por sensores passivos e ativos para a relação da distância entre os objetos da cena e o robô móvel. O mapa destas distâncias representa a geometria do ambiente e deve atender a três etapas do modelo de construção tridimensional (MANESSIS, 2003 apud DECCÓ, 2004).

- a) *Alinhamento de pontos*: Transformação de duas medidas distintas (provindas de conjuntos bidimensionais ou superfícies tridimensionais) através de suas informações relacionais num sistema de coordenadas;
- b) *Correspondência de pontos*: Associação de pontos entre duas medidas, após o alinhamento. Considerando o sistema mecânico calibrado, usa-se a transformação rígida para o inter-relacionamento dos conjuntos de informações, ou a própria transformação é relacionada para dois conjuntos distintos de dados;
- c) *Integração de pontos*: Junção de todos os pontos correspondidos no sistema de coordenadas criado pra a concepção final do modelo;

Com as etapas necessárias para a construção de mapas de ambientes com ou sem estrutura (interna ou externa), e representados em modelos geométricos ou topológicos, a concepção final do mapa somente será concluída se atenderem às metodologias de organização e associação de pontos do mapa. São elas:

- a) *Interactive Closest Point (ICP)*: Idealizado por Besl e Mackay (BESL e MACKAY, 1992), aborda o critério da mínima distância para a correspondência de pontos entre os mapas existentes e em construção;
- b) *Proximidade angular*: A correspondência de pontos é encontrada pela mínima diferença dos seus ângulos azimutais (YAGI *et al.*, 1995);
- c) *Histograma angular*: A correspondência entre os pontos é feita pela correlação máxima entre duas imagens (NAKAMURA e ISHIGURO, 2002);

Trabalhos no estado da arte de robótica móvel com visão computacional serão analisados com maior abrangência na seção 2.3.3, em aplicações agrícolas.

2.3. Estado da arte em visão computacional aplicada à robótica móvel agrícola

2.3.1. Tecnologia e Agricultura de Precisão

Sistemas dedicados de produção agrícola referenciam os avanços tecnológicos inicialmente desenvolvidos para outras aplicações industriais, favorecendo a alta eficiência e sustentabilidade para tecnologias voltadas a componentes computacionais de aquisição de dados (ZHANG *et al.*, 2002). Estes sistemas compõem a síntese para AP.

Em Menegatti e Molin (MENEGATTI e MOLIN, 2004), a AP é definida como uma metodologia de gestão à produção agrícola, cujas abordagens tecnológicas evidenciam a variabilidade espacial de seus fatores lógicos, para que a lavoura e sistemas de automação

embarcados empregados sejam otimizados e customizados a determinados problemas. Estes sistemas são executados ao processamento de dados (aquisição e tratamento) para a coordenação homogênea das ações de intervenção cultural da lavoura (GODOY, 2007; GOZDOWSK e SAMBORSKI, 2007; TANGERINO, 2009).

Assim, os componentes computacionais instalados em máquinas e implementos agrícolas demandam elevado nível de processamento para a realização de tarefas complexas no âmbito agrário. Estes componentes definem um conjunto de instrumentação de alta tecnologia com o uso de um sistema computacional centralizado denominado Unidade de Controle Eletrônico – UCE (*Electronic Control Unit – ECU*) (GUIMARÃES, 2003; NISSEN e HENNINGER, 2008). Todavia, a aplicação da UCE torna-se complexa à medida que o grau de dificuldade do problema envolvido no processamento e comunicação de dados aumenta e, conseqüentemente, requer a necessidade de se implantar um sistema computacional similar e distribuído, em respeito a normativas técnicas e protocolo específicos, que apresente as características ideais de implementação: robustez, confiabilidade, redução de conexões físicas e expansão de subsistemas modulares (PEREIRA, 2008).

Como sistemas distribuídos requerem interconexões definidas e com relativo grau de transmissões confiáveis de dados, o protocolo usado na automação de sistemas agrícolas denomina-se *Controller Area Network* (CAN), cuja estabilidade de comunicação é validada em camadas de baixo nível por processamento em tempo-real (AUERNHAMMER e SPECKMANN, 2006; PORTILLO *et al.*, 2006; SUVINEN e SAARILAHTI, 2006). Neste tipo de processamento, é necessário o emprego de sistemas eletrônicos de coleta de dados, bem como aplicadores de insumos que variam à taxa de aplicação (*Variable Rate Technology – VRT*) (SOUZA, 2002).

Uma vez definidos os trabalhos das tecnologias existentes relacionadas ao processamento de dados em AP, definem-se a seguir os tópicos científicos voltados ao sensoriamento óptico

de robôs agrícolas móveis para o auxílio à navegação e planejamento de trajetórias, bem como análise e inspeção de culturas agrícolas (cereais e frutas), com o emprego de algoritmos analíticos e estatísticos computacionais.

2.3.2. Análise, inspeção e localização de culturas agrícolas

Trabalhos que relacionam a localização, inspeção e análise de frutos baseados em visão computacional reúnem métodos estereceptivos na aquisição de dados e algoritmos dedicados para o processamento, como citados.

Em um trabalho para a extração de perfil de tomates, Morimoto (MORIMOTO *et al.*, 2000) usou tal abordagem com o propósito de avaliá-lo de forma quantitativa, utilizando o conceito de caos, através de atratores e dimensão fractal, com avaliação de RNA (Redes Neurais Artificiais) recursivas. Modelos caóticos de RNA recursivas são usados para explicar a alternância de perceptrons em objetos com reconhecimento ambíguo, onde cada imagem representa uma transição de dados entre as conexões (FARMER, 2007).

Um sistema de identificação e localização de laranjas foi desenvolvido (PLEBE e GRASSO, 2001; BULANON *et al.*, 2009) para um robô de colheita de frutos. Os autores apontaram que este sistema deve lidar com as seguintes situações: grande amplitude de iluminação; elementos parcialmente ocultos; incertezas relacionadas com a forma da fruta; espaço limitado de operação e restrições de tempo-real. Os autores de ambos os trabalhos inicialmente transformaram os *pixels*, que têm as cores representadas no espaço de cores RGB para um espaço de cores proposto por cada um dos trabalhos. Desta forma, as laranjas foram destacadas dos demais elementos da cena. As laranjas foram encontradas quando os *pixels* foram classificados com auxílio de uma RNA customizada – a estrutura do processo de classificação das componentes RGB por RNA é ilustrada na Figura 2.17. Em trabalho similar, Lopes (LOPEZ *et al.*, 2009) aprimorou a classificação de *pixels* das laranjas após processo de

segmentação por máscara de gradiente de Sobel, por extração de características em relação a média, variância, curtose (achatamento) e classes (QINGZHONG, 2002; DIAZ *et al.*, 2004; BLASCO *et al.*, 2007-a; BLASCO *et al.*, 2007-b).

Na Figura 2.18, observa-se a aproximação poligonal de falhas em superfície do fruto, com a representação por bordas, e a segmentação das áreas para irregularidades encontradas nas laranjas.

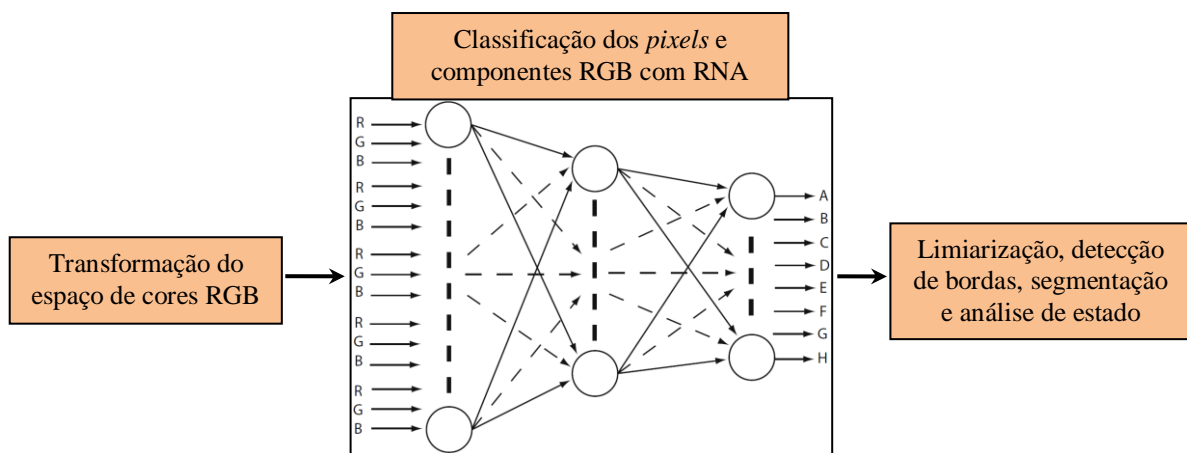


Figura 2.17: Processo de detecção e classificação de frutos.

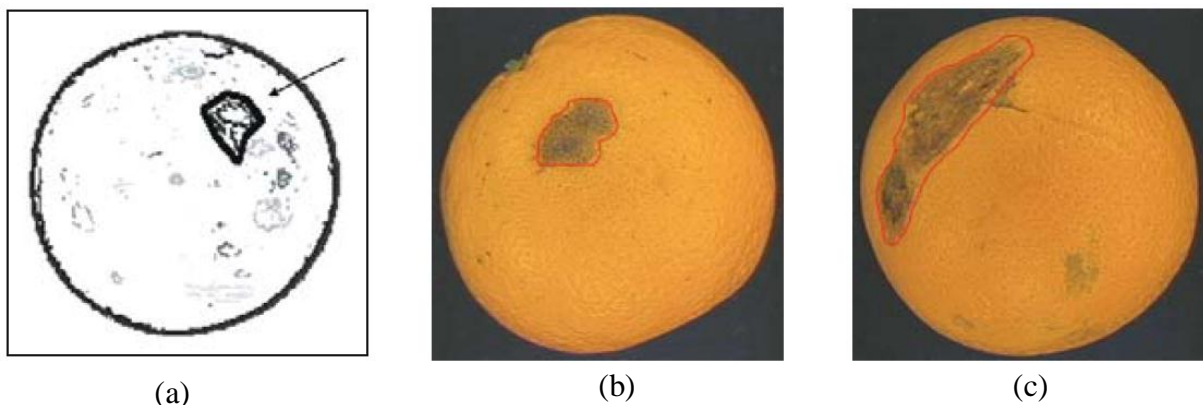


Figura 2.18: (a) Aproximação poligonal de falha em superfície – cada nível de cinza representa uma borda diferente; (b-c) Área de segmentação para dois tipos de irregularidades nos frutos (LOPEZ *et al.*, 2009).

Diversos trabalhos foram desenvolvidos em sistemas de visão computacional similares para localizar e avaliar o tamanho de laranjas (CRUVINEL e MINATEL, 2002; LI *et al.*, 2009; RAO e YING, 2009). As imagens foram digitalizadas com resolução espacial de

256×256 *pixels* e 256 níveis de cinza. Os autores, nos três trabalhos, adotaram protótipos de laranjas com diversos diâmetros, os quais foram usados na operação de correlação e morfologia matemática, a qual foi feita no espaço das frequências devido à restrição temporal. O erro causado pela distância da câmera e da laranjeira foi corrigido utilizando-se um modelo de laranja com tamanho conhecido, para o reconhecimento de classes. Para não envolver este modelo analítico, Blasco (BLASCO *et al.*, 2009) adaptou a inspeção automática de frutos, não apenas quanto ao tamanho, mas quanto à disparidade entre os tipos de doenças presentes em tais frutos (Figura 2.19), aplicando refletância infravermelha e informação espectral por indução ultravioleta.

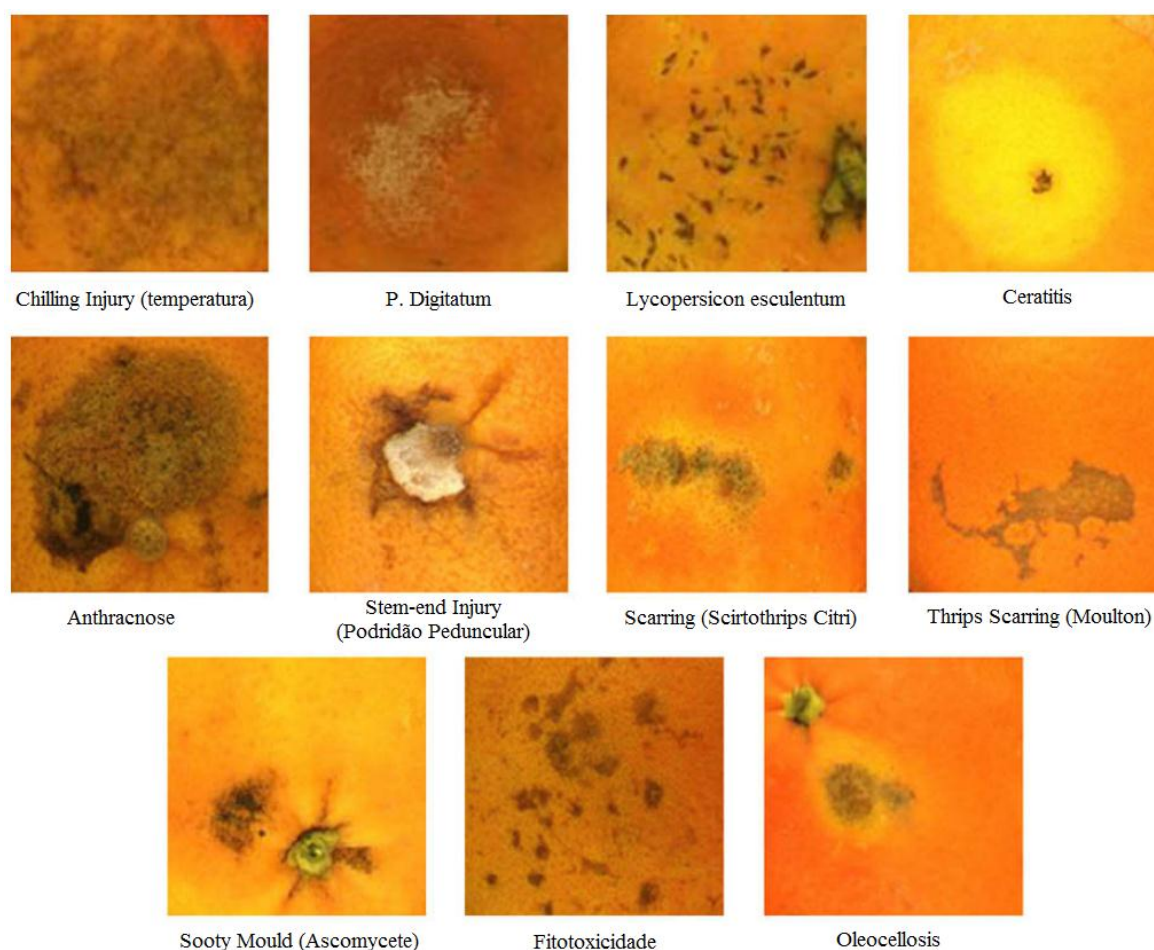


Figura 2.19: Detalhes do tamanho, forma e cor para as doenças estudadas (BLASCO *et al.*, 2009).

Aleixos (ALEIXOS *et al.*, 2002) propôs um sistema de visão computacional para inspeção de frutos (laranjas e limões) em um ambiente industrial. O sistema analisava tamanho, forma,

cor e defeitos externos da fruta em tempo real utilizando uma câmera multiespectral (espectro visível e infravermelho próximo) e um equipamento dotado de dois processadores de sinais digitais (DSP - *digital signal processor*) para processar as imagens. A câmera multiespectral foi construída com duas câmeras de vídeo convencionais, sendo uma responsável pelo espectro visível, para a digitalização de imagens RGB, e a outra pela aquisição de imagens em preto-e-branco. Os defeitos foram detectados utilizando um classificador bayesiano, sendo que as variáveis independentes adotadas foram os valores dos canais do RGB e o valor do infravermelho. Também, foram avaliados classificadores baseados apenas nas componentes do RGB e nas componentes vermelho, verde e infravermelha. A classificação das cores dos frutos foi feita também com um classificador bayesiano utilizando a informação do espectro visível. Seguindo a mesma metodologia, porém com aplicação distinta, Li (LI *et al.*, 2007) desenvolveu um sistema de detecção automática, com a preocupação em classificar laranjas com tonalidades e cores diferentes (alaranjado e amarelo), bem como seu crescimento. A abordagem usada neste trabalho foi por agrupamento de classes baseado na distância *Mahalanobis*, para a detecção dos frutos, após a translação e revolução em escala de nível cinza (PYDIPATI *et al.*, 2005; RAO e YING, 2009).

Bulanon (BULANON *et al.*, 2004) e MacArthur (MACARTHUR *et al.*, 2006) desenvolveram um sistema de localização de maçãs aplicado a colheita automática. As imagens foram adquiridas com resolução espacial de 320×240 *pixels* e resolução cromática de 24-bits, segmentadas em duas partes, sendo a primeira designando as frutas e a outra as demais informações gráficas da cena. Para isto, foram utilizadas análise discriminatória e RNA no primeiro trabalho e RNA no segundo. Nas duas abordagens, os *pixels* das imagens foram classificados utilizando-se as componentes vermelha e verde do espaço de cores NRGB. Após o processo de segmentação, a imagem foi suavizada com um filtro passa baixa, e suas bordas foram extraídas através do operador laplaciano. Na extração, obtiveram os

seguintes dados: áreas, centróides, diâmetros máximo e mínimo. A taxa de acerto para os classificadores foi superior a 80% em todas as condições de iluminação consideradas nos trabalhos. No segundo trabalho ainda, os autores se preocuparam na localização global - realizada por aquisição de imagens aéreas - de cada pomar para a quantidade de laranjas encontradas, salientando a proporção de frutas por *pixels*.

Annamalai (ANNAMALAI *et al.*, 2004) e Lekhal (LEKHAL *et al.*, 2008) desenvolveram um sistema de visão computacional para estimar o número de frutos em pomares de laranjas. A identificação dos frutos foi feita através das seguintes operações: binarização (utilizando a matiz e a saturação), limiarização (considerando-se a área dos elementos), dilatação, erosão, preenchimento e extração das características. As laranjas foram contadas considerando-se o número de elementos não conectados na imagem. No segundo trabalho, a identificação é aprimorada com a classificação de texturas e a implementação de algoritmo genético (AG). Shiigi (SHIIGI *et al.*, 2008) modificou as técnicas acima com a implantação de um sistema de visão com três câmeras coloridas (visão estéreo e central) para, além de localizar os frutos, manipulá-los através de um robô articulado com 3 graus de liberdade, e verificar a inclinação e consistência do pedúnculo de determinado fruto.

Diversos autores reportaram o uso de câmeras de vídeo e sensores de ultra-som para estimar a quantidade e o tamanho de frutos em um pomar de frutas cítricas (REGUNATHAN e LEE, 2005; YIN *et al.*, 2008). O tamanho dos frutos foi estimado com auxílio de um sensor de ultra-som. As cores das imagens eram representadas originalmente no espaço de cores RGB, e posteriormente convertidas para o espaço de cores HSL, para a obtenção de melhores resultados. As imagens foram segmentadas através da classificação dos *pixels* e, posteriormente, estas foram pré-processadas para a remoção de ruídos. Foram construídos três classificadores: bayesiano, baseado em RNA e baseado na discriminação linear de Fisher. Os *pixels* foram classificados considerando-se a matiz (H) e a saturação (S). As funções de

densidade de probabilidade condicionais foram estimadas pelos histogramas dos valores dos *pixels*, que foram quantizados novamente em valores distintos. No segundo trabalho, os autores aprimoraram as técnicas de segmentação de *pixels* para outras quatro derivações: método de segmentação por limiar dinâmico, método de Otsu, método de Otsu combinado com algoritmo genético e segmentação adaptativa baseada na rede LVQ (Quantização Vetorial por Aprendizagem), como podem ser verificadas na Figura 2.20, abaixo.

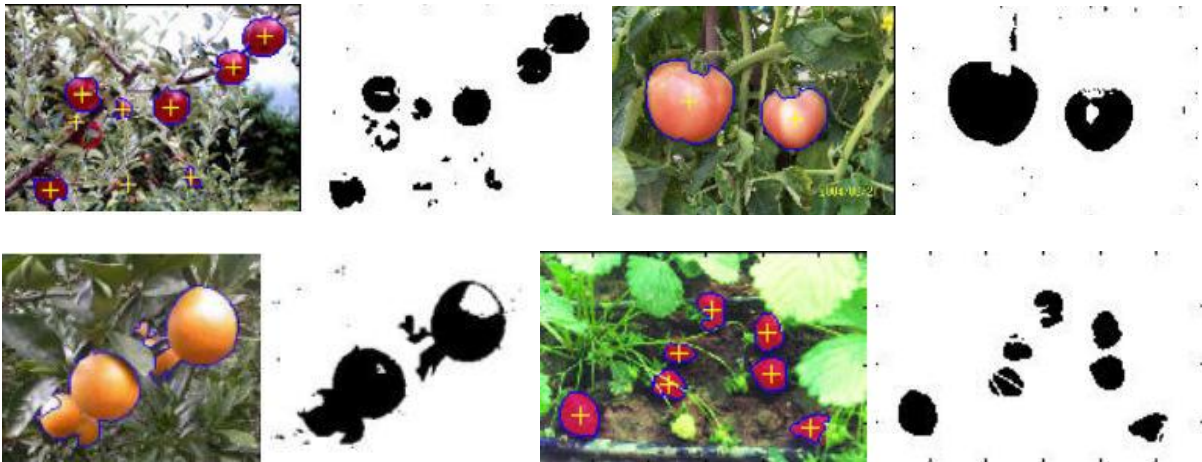


Figura 2.20: Segmentações de maçãs, tomates, laranjas e morangos (YIN *et al.*, 2008).

A classificação de plantas invasoras foi tema de trabalho de Tang (TANG *et al.*, 2003; TANG *et al.*, 2005) e Savakar (SAVAKAR e ANAMI, 2009), que o fizeram por um sistema de visão computacional na classificação em duas classes segundo o tipo de folha. Os autores do primeiro trabalho adotaram esta abordagem, pois a escolha do herbicida depende do tipo da folha e não da espécie da planta invasora, o que simplifica o classificador de padrões. As imagens digitais usadas apresentam resolução espacial de 640×480 *pixels* e resolução cromática de 24-bits, com espaço de cores RGB adotado. As imagens foram feitas em ambiente aberto, assim com o objetivo de minimizar a influência dos *pixels* que representam o solo na análise das folhas, a imagem foi segmentada utilizando a limiarização, ajustada de forma empírica, do índice de excesso de verde modificado. Foram utilizados descritores de texturas baseados na transformada *wavelet* de Gabor nos dois primeiros trabalhos. Os valores obtidos foram normalizados (vetores de características de treinamento) e a norma do número

complexo foi calculada. O classificador adotado foi o perceptron multicamadas, treinado com o algoritmo *backpropagation*. Já no terceiro trabalho, os autores usam o mesmo tipo de classificador, agregando informações das características morfológicas, texturas e cores.

Steward (STEWARD *et al.*, 2004), Orlando (ORLANDO *et al.*, 2004) e Jafari (JAFARI *et al.*, 2004) estudaram e implantaram a distribuição das cores da vegetação em imagens de linhas de plantio para desenvolver algoritmos de agrupamento aplicados à segmentação de imagens. Nos três trabalhos, foram usados tais algoritmos para a identificação de milho e plantas daninhas. Os autores destes trabalhos utilizaram descritores de textura e um classificador bayesiano para as segmentações de imagens. O registro de tais imagens foi feito por duas câmeras de vídeo (imagens coloridas / imagens em tons de cinza). As imagens geradas pela câmera colorida foram digitalizadas com resolução espacial de 1.392×1.039 pixels, 768×494 pixels e 1600×1200 pixels. As imagens geradas pela câmera preto-e-branco foram digitalizadas com resolução espacial de 768×494 pixels. A partir das imagens adquiridas pela câmera colorida foram geradas três imagens: uma somente com o canal verde do RGB, outra com o excesso de verde e a última com o infravermelho próximo, sendo que estas imagens foram quantizadas em 32, 64, 128 e 256 níveis. Foram amostradas regiões de interesse com tamanhos diferentes de cada uma das imagens, porém, posicionadas na mesma região e, desta forma, foram calculadas as matrizes de co-ocorrência, para cada uma destas regiões, com o cálculo dos descritores pela média, momento angular, variância, entropia, correlação, momento do produto e momento inverso da diferença.

Wei (WEI *et al.*, 2009) desenvolveu um sistema de reconhecimento de maçãs baseado em SVM (*Support Vector Machine*). O filtro vetorial mediano é aplicado primeiramente na remoção de ruídos de imagens coloridas das frutas. Os resultados das operações de filtragem e morfológica são vistas na Figura 2.21. Em seguida, é realizada a segmentação de imagens baseada no método de crescimento de regiões, caracterizando as propriedades das cores

segmentadas. Por conseguinte, tais propriedades das cores e também do formato das maçãs são extraídas, para a classificação das imagens usando SVM (ZHAO *et al.*, 2008; WANG *et al.*, 2009).



Figura 2.21: Imagem original, após filtragem e operação morfológica (WEI *et al.*, 2009).

Outro trabalho que aborda métodos analíticos para o reconhecimento de frutos é relatado em Shou (SHOU *et al.*, 2009). Neste trabalho, os autores comentam a dificuldade em se realizar a detecção de bordas, segmentação e recuperação de contornos em frutos para imagens coloridas, e propõem uma nova abordagem na detecção de contornos no espaço de cores RGB numa metodologia rápida e eficaz, baseada no modelo Chan-Vese (JIANYUAN e CHONGYANG, 2009; MEWADA e PATNAIK, 2009; WANG *et al.*, 2010). Neste modelo, os autores customizaram o algoritmo de ordenação de frutos, por apresentar relativa precisão na detecção de bordas das formas geométricas dos frutos, sendo desnecessária qualquer etapa de pré-processamento, mesmo para imagens com ruídos e desfocadas.

Si (SI *et al.*, 2009) e Lee (LEE *et al.*, 2009) propuseram um trabalho de reconhecimento de maçãs baseado no método analítico por *clustering* (algoritmo *k-means*) (MALYSZCO e STEPANIUK, 2009; YE *et al.*, 2009; MALYSZCO e STEPANIUK, 2010) para diferenciar maçãs verdes de folhas com a mesma textura. As imagens possuem blocos definidos com resolução espacial de 8×8 *pixels* para a implementação distinta de cada algoritmo *k-means*. As componentes R e B do espaço de cores RGB foram escolhidas para a caracterização de cores e o valor médio, ao passo que para a caracterização de texturas foram escolhidos o desvio padrão e entropia em imagens com níveis de cinza para a posterior extração de características.

Os experimentos de reconhecimento foram conduzidos à aplicação do algoritmo em 200 amostras de imagens adquiridas sob variações de iluminação distintas, e os autores garantem que, por este método, a percentagem de certeza no reconhecimento é superior a 80%.

Liming (LIMING e TANCHAO, 2010) desenvolveu um sistema de classificação de morangos, considerando três características: forma, tamanho e cor. Neste sistema, os autores comprovaram a obtenção das características da forma geométrica através da manipulação das linhas de contorno e a classificação pelo algoritmo *k-means* e entropia (SIMION *et al.*, 2009; VERMA *et al.*, 2009; ZHANG *et al.*, 2009). A cor do morango foi caracterizada pelo método de cor dominante e o tamanho descrito pelo maior diâmetro encontrado. A automação deste sistema de classificação gera problemas de multicaracterísticas de cores, tamanhos e formas, sendo sua resolução baseada na adaptação de um sistema físico, com uma correia de transmissão, uma câmera, dois sensores fotoelétricos e parafuso de avanço acionado por um motor. Tal sistema é controlado por um microcontrolador e realiza a aquisição da imagem conforme o fruto é conduzido pela correia. Os resultados, segundo os autores, denotaram que o erro na detecção do tamanho dos morangos não passa dos 5%, sendo a precisão na classificação de cores em 89% e a precisão na classificação de formas superior a 90%. A média de tempo na classificação para as três características é de 3s. Abaixo, na Figura 2.22 e Figura 2.23, é possível verificar as etapas de processamento das imagens.



Figura 2.22: Caracterização da imagem (imagem original, imagem em tons de cinza e extração do caule (LIMING e TANCHAO, 2010).

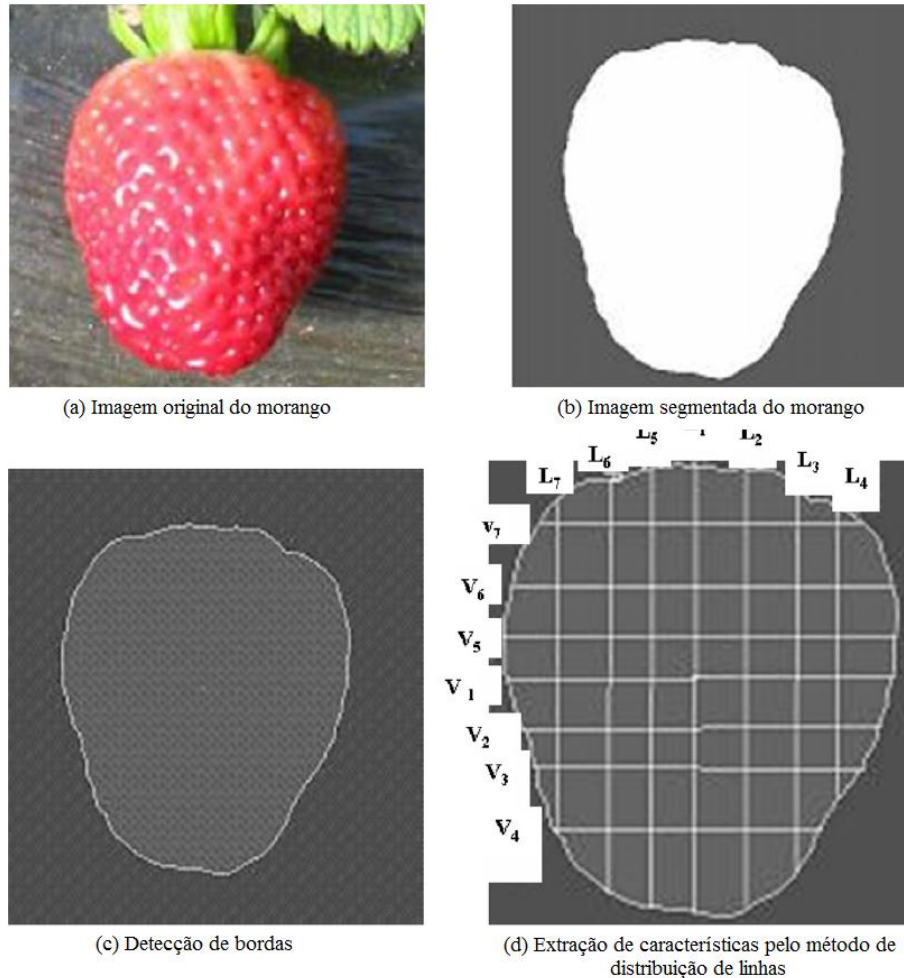


Figura 2.23: Extração de características da imagem (a- imagem original, b- imagem segmentada, c- detecção de bordas, d- extração de características pelo método de distribuição de linhas) (LIMING e TANCHAO, 2010).

A extração de características dos morangos dá-se pela conversão do espaço de cores RGB em níveis de cinza, cujos valores associados ao pedúnculo e receptáculo floral do fruto são maiores que o resto da cena. Para delimitar a região do morango, foram usadas características do tamanho, forma e cor do fruto. A característica da forma foi obtida pelo contorno e a normalização do comprimento das sequências lineares para eliminar a influência de um padrão de tamanho para este tipo de fruto. Assim, a geometria atende a vários tipos de formas de morango, auxiliado também pelo algoritmo *k-means* para a determinação do centróide de área para a classificação da forma do fruto.

Em Lü (LÜ *et al.*, 2010), os autores usaram a diferença G-B de cores em imagens em tons de cinza, para a aplicação do algoritmo Otsu (GAO *et al.*, 2010; XIA e FENG, 2010) na

segmentação de imagens de frutos cítricos maduros e a textura no fundo das imagens. As imagens que apresentam frutas sobrepostas foram separadas pelo método *Watershed* (MEURIE *et al.*, 2010; ZHANG *et al.*, 2010) com a transformação da distância, e foram retificadas com o algoritmo de fecho convexo (*Convex Hull*) (CHEN e CHAN, 2009; SIRAKOV e USHKALA, 2009; WAN, 2009). Após a extração das linhas de contorno das imagens, foi realizada a filtragem da média dos mínimos quadrados com a função de peso e profundidade *Tukey* (SERFLING, 2004), e assim foram obtidas as coordenadas de centróide e do raio da imagem. A base de imagens para a realização dos experimentos apresenta 87 imagens, para o reconhecimento de 592 frutos, com taxa de precisão em 87%, e a acima de 80% para as imagens de frutos sobrepostos.

2.3.3. Locomoção guiada e autônoma de robôs móveis agrícolas

Trabalhos no estado da arte de SLAM, com os sistemas de visão citados nas seções anteriores, são descritos para ambientes *indoor* e *outdoor*, estruturados e desestruturados para ambientes agrícolas.

1) Navegação baseada em visão monocular

As aplicações que envolvem sistemas de visão monocular são limitadas pela aquisição de informações, devido apenas a uma única projeção direcional. São majoritariamente empregados em estratégias de navegação com fusão sensorial, para obtenção de informações mais precisas na localização do robô móvel e mapeamento do ambiente.

Um modelo de localização em faixa virtual para o problema de SLAM foi proposto por Zhang (ZHANG *et al.*, 2008). Os autores destacaram que a restrição no desenvolvimento do modelo do eixo óptico paralelizado é substituída com uma única câmera que manipula os erros de extração da faixa de visão real. Também provê uma medida complementar para a inicialização de *landmarks* como referência para a aquisição de dados à localização e

mapeamento simultâneos. Como a transformação de dados (de *pixels* para valores métricos) capturados pelo sensor óptico é um processo não-linear, o modelo de observação para a caracterização da incerteza ao Filtro de Kalman Estendido não deve ser modelado através da distribuição gaussiana, a qual torna mais complexa a associação de dados para o SLAM. Considerando tal problema, os autores propõem uma nova técnica computacional de associação de dados baseada em transformação homográfica por sequência de imagens, que integra esta associação ao processo de atualização de dados calculados pelo EKF. Os resultados experimentais dos dados reais validaram o desempenho do modelo de localização de faixa virtual, e a nova abordagem de associação de dados.

Outro trabalho que lida com o problema de SLAM com visão monocular é apresentado por Wu (WU *et al.*, 2009). Neste trabalho os autores abordam a proposta computacional direcionada a ambientes externos de grande escala. Descartando as informações sensoriais odométricas, e mantendo apenas a informação visual, são usados referenciais ao modelo de Estrutura por Movimento (*Structure from Motion - SFM*), para a estimação da locomoção do robô móvel passo-a-passo. São usados como *landmarks* naturais, as características invariáveis a transformação em escala, ou SIFT (*Scale Invariant Feature Transformation*), e sua localização tridimensional é construída diretamente pela metodologia de triangulação, após a modelagem do movimento translacional de escala do robô móvel. Logo, a SFM é aplicada matematicamente pelo algoritmo de filtro de partículas *Rao-Blackwellised*, validada pela função de densidade EKF, cuja associação do estado atual de cada *landmark* é atualizada por um filtro EKF independente, que realiza a correspondência no mapa de marcos de referência da partícula analisada. Na Figura 2.24, o gráfico obtido ilustra a construção do mapa para uma partícula de máxima verossimilhança (*maximum likelihood*) ao final da exploração do ambiente. Desconsiderando a remoção de ruídos no mapa, a trajetória vermelha indica o percurso da melhor amostragem, e os pontos azuis são *landmarks* esparsos no mapa.

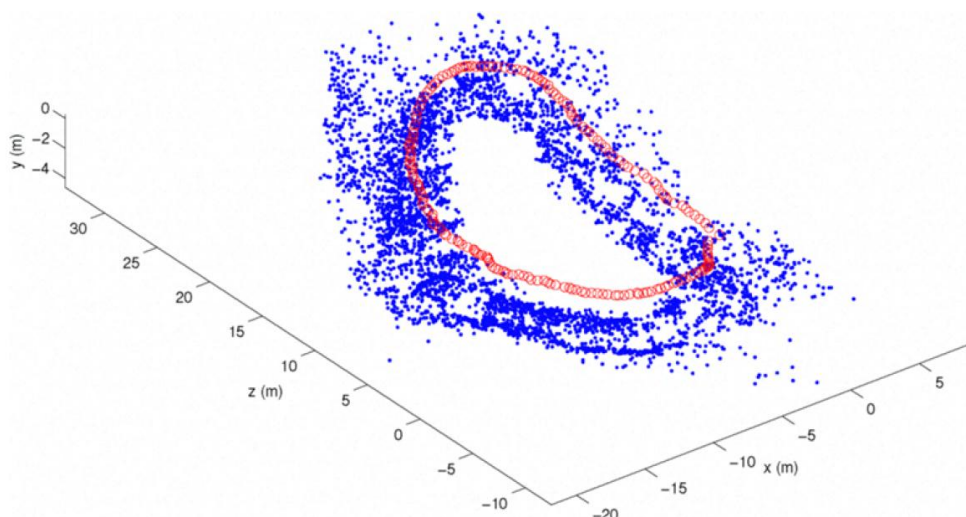


Figura 2.24: Trajetória da câmera e posições dos *landmarks* (WU *et al.*, 2009).

Para a navegação em ambientes externos com deformações no terreno (desestruturados), como o caso de áreas e linhas de plantio, o processo de construção do mapa torna-se mais complexo, à medida que novas entidades dinâmicas se associam na trajetória percorrida pelo robô móvel, apresentando irregularidades no ângulo de guiagem. Tendo estas prerrogativas em base, em Xue (XUE e JU, 2010) foi proposto um novo método de guiagem por visão, através da detecção de linhas de plantio, a qual usa morfologia matemática para processar as imagens e segmentá-las na extração de características de interesse. As linhas de guiagem são associadas com o método dos mínimos quadrados em relação aos centróides de cada característica extraída numa região contínua de percurso da imagem. Tal técnica é empregada para linhas de plantio individuais com espaçamento de regiões, linhas de plantio com regiões contínuas e linhas de plantio com ambas as características.

Os experimentos foram realizados com o auxílio de um robô móvel em uma área de cultivo de vegetais, cujos resultados mostraram que o algoritmo aplicado à navegação do robô obteve a precisão de desvio em $\pm 29,5\text{mm}$, e tempo de processamento das imagens capturadas por um sistema de visão monocular entre 0,7s e 1,3s, que satisfaz as necessidades de produção agrícola. O algoritmo atende às seguintes etapas de implementação, da Figura 2.25:

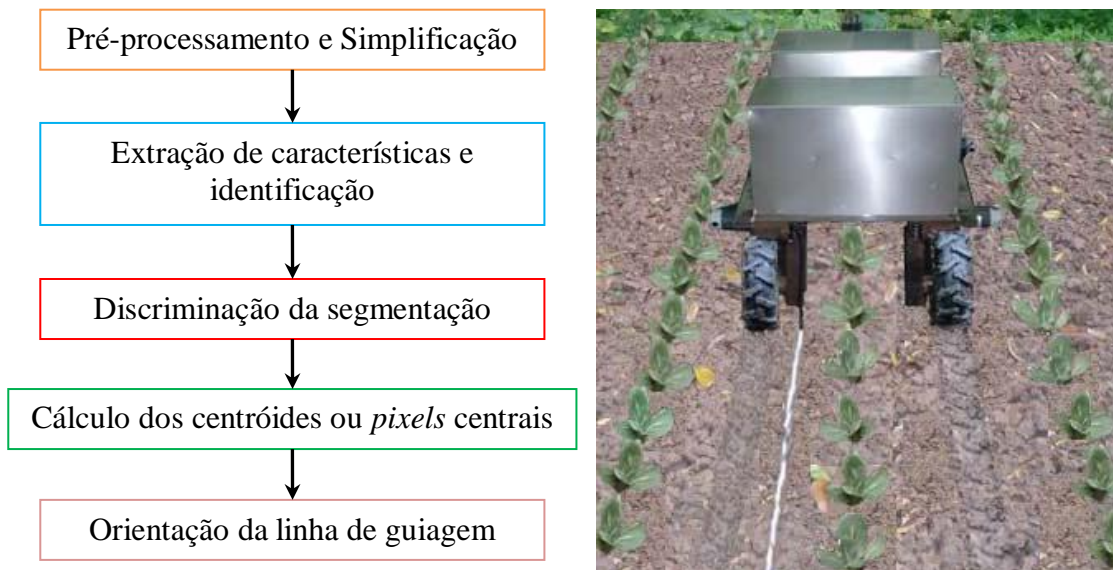


Figura 2.25: Algoritmo de detecção de linhas de guiagem, e teste real com o robô móvel na linha de plantio de vegetais (Adaptado de XUE e JU, 2010).

Na primeira etapa do algoritmo é realizada a remoção de ruídos e pequenas regiões da imagem para transformá-la em tons de cinza com uniformidade de pontos de *pixels*, tornando a textura mais esparsa facilitando a segmentação de regiões de interesse. A operação de reconstrução da morfologia matemática é usada para simplificar as imagens, uma vez que numa imagem que contém informações de cultivo e outras características naturais, a textura e estrutura das plantas cultivadas podem ser identificadas próximas ao solo ou a ervas daninhas, devido a diferentes condições de luminosidade e crescimento natural. A morfologia matemática reconstrói as imagens, eliminando buracos e bordas descontínuas para fixação dos pontos de *pixels* nestes intervalos da cultura, solo e plantas invasoras. Na segunda etapa é realizada a extração de características de acordo a um critério uniforme em tons de cinza e textura, que parametriza a imagem em diferentes regiões para a identificação dos segmentos. Na terceira etapa é realizada a discriminação da segmentação: as bordas das linhas de plantio mantêm regularidade quase uniforme, devido a informações das linhas de cultivo quanto ao tamanho e espaçamentos entre as linhas de cultivo. Os discriminantes são então, empregados na detecção das culturas de forma confiável e precisa. A detecção de bordas não será implementada se o *pixel* detectado não satisfaz a característica do discriminante. Os cálculos

dos centróides para encontrar os *pixels* centrais são obtidos para a orientação da linha de guiagem na quarta e quinta etapas.

A Figura 2.26, abaixo, mostra a aplicação do algoritmo na segmentação e identificação das características nas culturas em plantio contínuo, espaçado e em ambos.

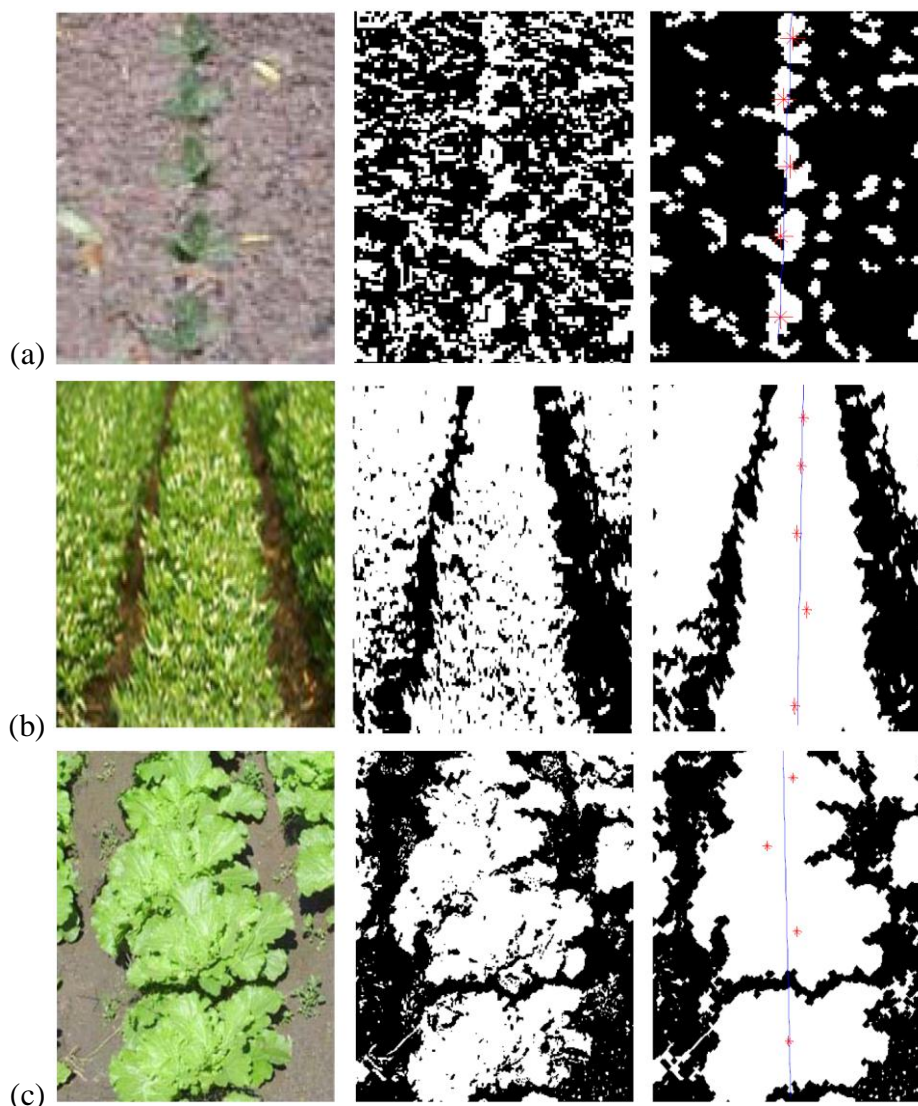


Figura 2.26: Aplicação do algoritmo proposto com as imagens originais, binarizadas e discriminadas após segmentação. (a) linha de plantio espaçada; (b) linha de plantio contínua; (c) ambas (XUE e JU, 2010).

Outro trabalho na área de robótica agrícola para navegação autônoma baseada em visão computacional monocular é apresentado por Bakker (BAKKER *et al.*, 2008). Nele, os autores propuseram uma nova abordagem de reconhecimento de linhas de plantio, baseada na transformada *Hough* em tons de cinza, para imagens combinadas, resultando em desempenho

e processamento otimizados. Uma câmera monocolor foi usada como sensor óptico para a obtenção de imagens em uma plantação experimental de beterrabas em estufa. As imagens coloridas são transformadas em tons de cinza, cujo contraste permite a diferenciação entre as plantas e o solo. Logo, três métodos distintos da transformada *Hough* foram aplicados: as imagens em tons de cinza são divididas em três seções para serem combinadas em uma só imagem, fornecendo menos dados para processamento, mesmo com a informação das três linhas. O algoritmo implementado discrimina as linhas de plantio em vários estágios de crescimento, não diferenciando qual dos três métodos de transformação de cores é usado. O erro médio entre as linhas de plantio estimadas e reais por séries de medidas varia entre 5mm a 198mm. O erro mediano da detecção das linhas de plantio foi de 22mm. As medidas de erros maiores são relacionadas a fatores que não acontecem na prática e devem ser evitadas, como limitações de número e tamanho das plantas cultivadas, superexposição da câmera, e presença de algas verdes pelo cultivo em estufa. Imprecisões criadas pelas pegadas indicam que estruturas lineares na superfície do solo em um campo real, são consideradas como problemas durante a transformação de cores. Em duas séries de medidas que não compartilham estes erros, o algoritmo foi capaz de discriminar as linhas de plantio com erro médio entre 5mm e 11mm, com desvio padrão entre 6mm e 11mm. O tempo de processamento das imagens varia de 0,5s à 1,3s, por imagem.

A Figura 2.27, a seguir, demonstra as etapas de processamento nas imagens adquiridas.

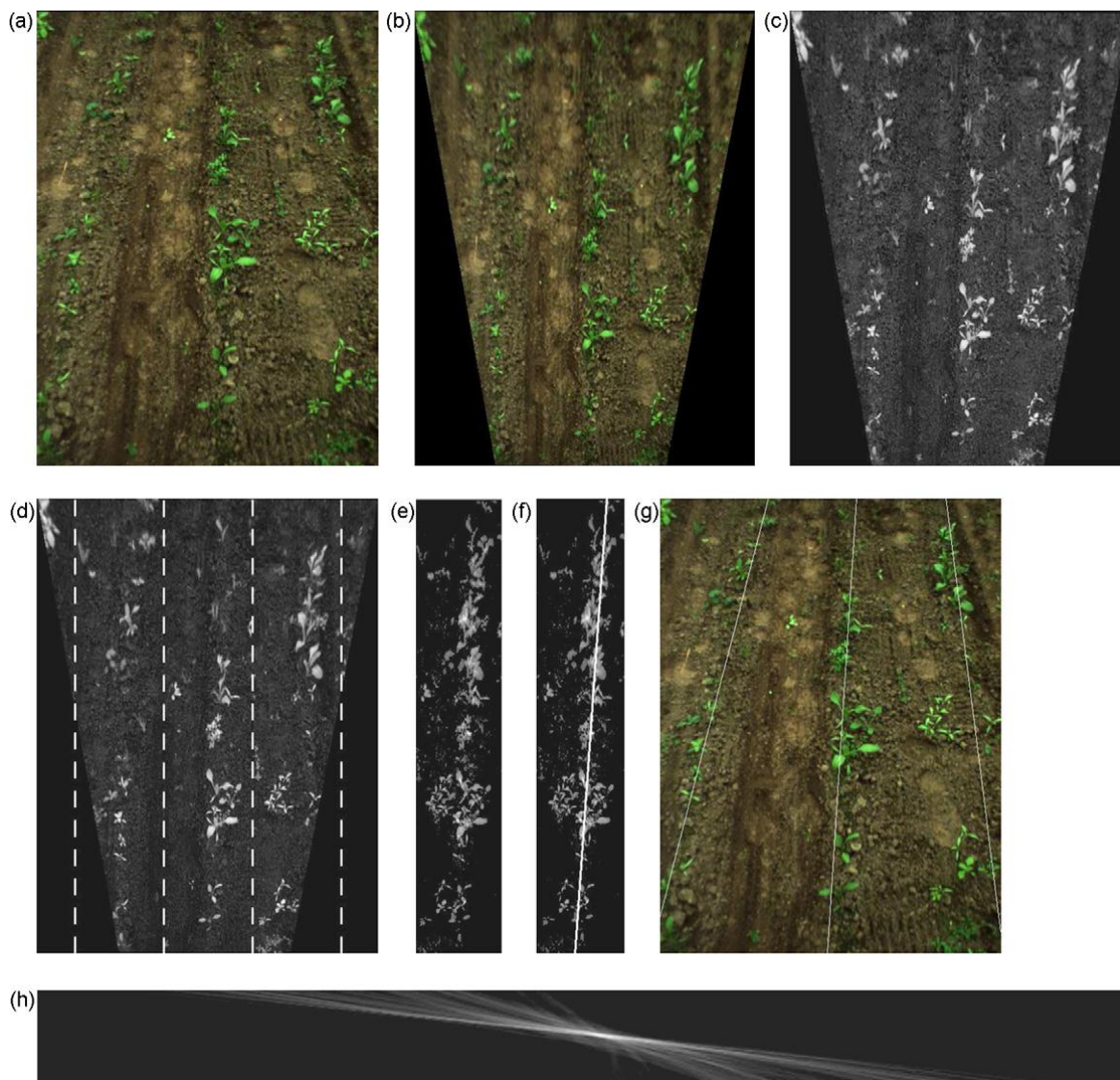


Figura 2.27: Processamento de imagens: (a) imagem original; (b) imagem retificada; (c) imagem em tons de cinza pós cálculo da representação das componentes RGB; (d) seleção dos segmentos; (e) imagens combinadas limiarizadas; (f) imagens combinadas sobrepostas; (g) imagem original sobreposta; (h) acumulador *Hough* (BAKKER *et al.*, 2008).

2) Navegação baseada em visão estéreo

A visão estéreo consiste em associar duas ou mais câmeras para a aquisição de imagens, ou associar aquisições com uma câmera apenas em diferentes intervalos, cuja finalidade principal é obter relações de distância entre as medidas da primeira imagem capturada e a atual, auxiliando nas informações base para algoritmos estatísticos e analíticos ao problema de SLAM em robótica móvel.

Em Rossetot (ROSSELOT *et al.*, 2010), um algoritmo preditivo de detecção de objetos foi desenvolvido a fim de investigar a aplicabilidade do uso de filtros probabilísticos em detecção de objetos por visão estéreo, como idealizado pelos autores de mapa XH. A detecção por este modelo de sistema de visão é inerentemente associada a ruídos e apresenta não linearidade. O algoritmo proposto (mapa XH) adota como modelo matemático de localização o UKF (*Unscented Kalman Filter*) e prediz a associação de *pixels* pela correspondência estérea das disparidades entre os mapas formulados.

Outra proposta de aplicação de filtros estatísticos não-lineares é apresentada por Sibley (SIBLEY *et al.*, 2008), cuja técnica desenvolvida é uma variação do UKF com o filtro de janela deslizante para SLAM incremental. Trata-se de uma técnica dedicada à estimação ótima de ambientes esparsos imediatos, usando uma janela de tempo variável com as medidas atuais dos sensores ópticos. Os autores idealizaram um algoritmo de processamento com tempo constante que mais se aproxima da estimativa de máxima verossimilhança, bem como na variância mínima de limite inferior *Cramer-Rao* (CRLB) – um estimador que engloba noções de otimização estatística (convergência rápida), eficiência (redução de incertezas) e consistência (evasão de confiabilidade). Para tal finalidade, os autores abordam a derivação do problema SLAM pela perspectiva de otimização ao método dos mínimos quadrados para funções gaussianas não-lineares. Os resultados foram aplicados em outros trabalhos de fotogrametria, como filtros de dimensão de variáveis de estado, odometria visual, ajuste de feixes ópticos, e EKF, em um sistema de visão estéreo com medidas inerciais. Os experimentos mostraram que o método aproximado ideal é compatível com o estimador ótimo, ao alcançar a complexidade de tempo constante, empiricamente, que torna indistinguível a diferença entre os desempenhos.

Ao problema do SLAM, são submetidos diversos modelos e técnicas customizadas por filtros probabilísticos, analíticos e cognitivos, de trabalhos que envolvem o mapeamento e

localização simultâneos para a robótica móvel. Neste contexto, Herath (HERATH *et al.*, 2008) desenvolveu uma arquitetura para o problema SLAM em visão estéreo, com atenção aos sensores ópticos usados. Primeiramente, o dispositivo de sensoriamento apresenta modelos de observação altamente não-linear, resultando em estimações de estado inconsistentes quando aplicados filtros derivativos não-lineares do KF, como EKF e UKF. Ademais, as abordagens baseadas em visão tendem a relacionar características de densidade granular, estreito campo de visão e capacidade elevada de armazenamento de dados para as aplicações das técnicas de processamento. A solução proposta para a criação de multi mapas SLAM resolve a questão de inconsistência do filtro aplicado, através da formulação do problema SLAM como uma otimização não linear em lotes (*non-linear batch optimization*). Os resultados podem ser vistos na Figura 2.28, abaixo, cujas estimativas pelo algoritmo proposto, com visão em EKF e covariância na construção do mapa são denotadas pelas elipses verdes e vermelhas, respectivamente.

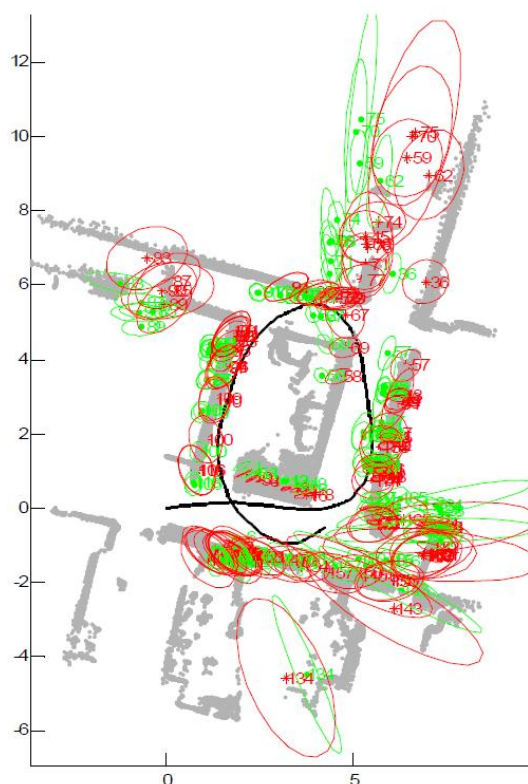


Figura 2.28: Resultados do SLAM usando o algoritmo de otimização em lotes: os pontos vermelhos indicam a estimativa do mapa pelo algoritmo proposto; os pontos verdes indicam as estimativas do mapa pela visão em

EKF; as elipses verdes e vermelhas representam as estimativas de covariância e o emprego do algoritmo, respectivamente (HERATH *et al.*, 2008).

O gerenciamento das características é resolvido pela representação de dois níveis do mapa. Os dois mapas têm atributos únicos associados a eles, representados pelo mapa global (GM), do ambiente percorrido pelo robô móvel, e o mapa local (LM), usado exclusivamente para a navegação em baixo nível entre os pontos locais no horizonte de navegação do robô móvel.

Bais (BAIS *et al.*, 2008) apresentou um trabalho sobre auto-localização de pequenos robôs móveis baseado em visão estéreo, em um ambiente conhecido estruturado, mas altamente dinâmico. O problema abordado consiste no posicionamento do robô móvel com uma estimativa inicial da auto-localização global. O algoritmo permite ao robô encontrar sua posição inicial e verificar sua localização durante toda a locomoção. A posição global do robô móvel é estimada usando-se trilateração baseada em técnicas de extração de características distintas a denominações dos *landmarks*. Medidas de distâncias com poucos *landmarks* são usadas em comparação às medidas dos ângulos, mas as características mínimas associadas a cada *landmark* para a estimação da posição global não são disponibilizadas por todo espaço de estados calculado. Assim, a posição do robô móvel é controlada, uma vez que a estimativa de posição global estiver disponível pelo filtro EKF, usado em fusão sensorial com outros sensores heterogêneos. Os resultados da simulação mostram que o novo método combina a estimação da posição global com o posicionamento de auto-localização dos robôs móveis, apresentando ganho significativo de desempenho e robustez.

A Figura 2.29, a seguir, mostra um experimento no qual um robô se move ao longo de uma trajetória circular, com a pretensão de fazer adequadamente o percurso. Contudo, devido a imperfeições das leituras sensoriais, o trajeto feito no círculo é irregular. O robô começa numa dada posição inicial com relativa incerteza e observa o primeiro *landmark* quando

percorrido um quarto do trajeto. As representações das elipses denotam as incertezas do posicionamento, que crescem ilimitadamente, e assim a posição real do robô acumula os erros de tais incertezas. Quando o robô observa um dos dois *landmarks* no lado esquerdo, a incerteza é reduzida e o posicionamento reajustado. Entretanto, à medida que os *landmarks* são afastados do robô móvel, novamente, as incertezas quanto ao posicionamento real crescem, e voltam a decrescer quando novos *landmarks* são observados.

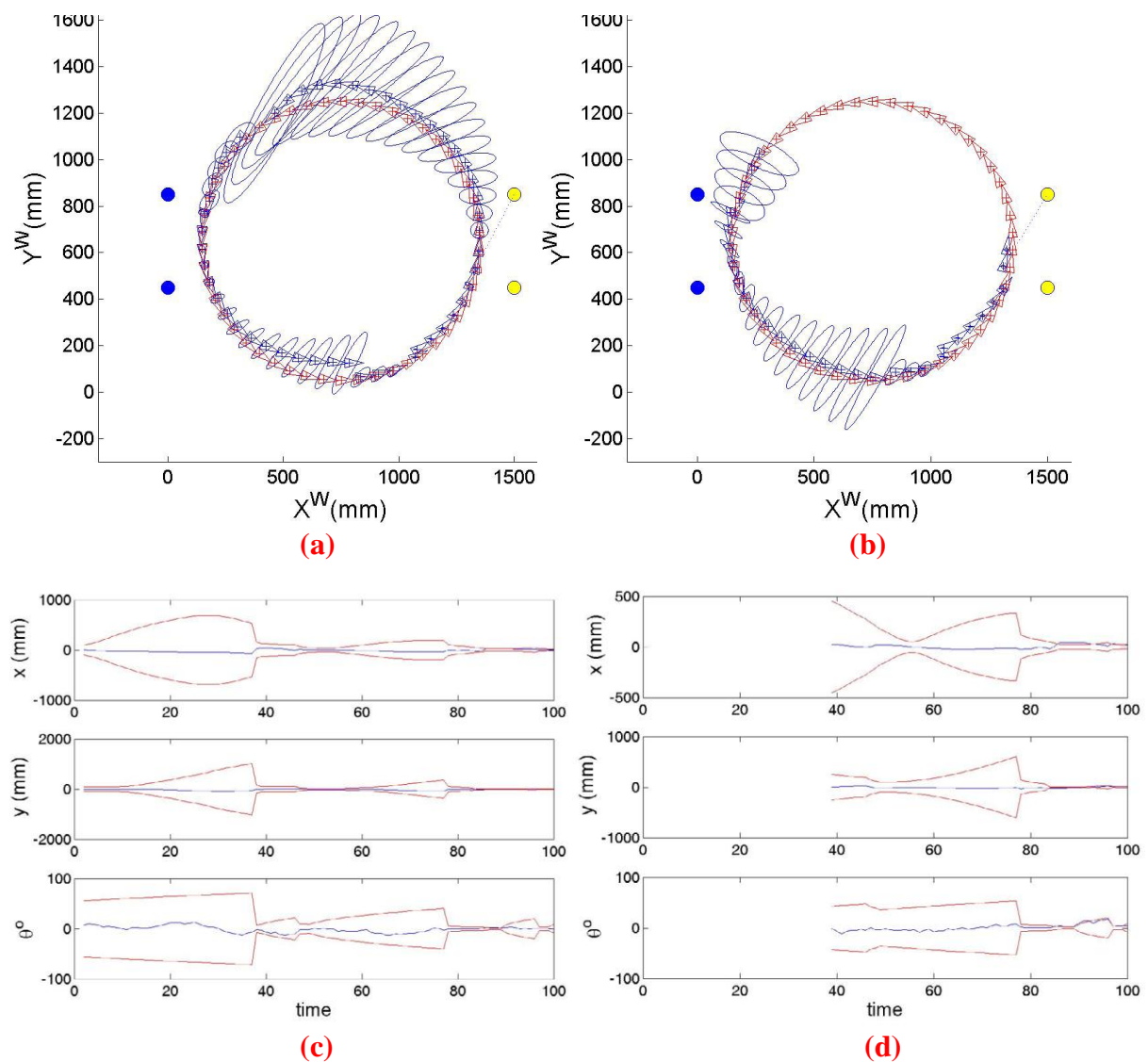


Figura 2.29: Trajetória desejada do robô móvel por um círculo de raio 600mm: (a) posição inicial conhecida; (b) o robô mantém o localizador inativo e procura por *landmarks* a fim de estimar a posição global; (c) limite 3σ no erro em x , y e θ para a posição inicial conhecida; (d) melhoria do posicionamento, quando o algoritmo é aplicado (Adaptado de BAIS *et al.*, 2008).

A preocupação principal em linhas de plantio também é delinear as estruturas de plantio contínuas, formulando projeções geométricas para que a navegação do robô móvel seja possível. Neste contexto, um algoritmo foi proposto para a detecção de linhas baseadas em um sistema de guiagem agrícola por visão estéreo (KISE *et al.*, 2005), que consiste em processamento multifocal, criação de mapas de elevação e determinação de pontos de navegação. O método desenvolvido primeiramente reconstrói um mapa de elevação tridimensional das culturas cultivadas, por meio de uma imagem estéreo e procura por pontos de navegação ótimos neste mapa. O sistema de detecção de linhas de plantio desenvolvido foi testado em uma área de cultivo de plantação de soja para seguir as linhas de plantio retas e curvas em velocidades de operação típicas. Os testes de validação do campo indicaram que o sistema proposto localiza as linhas de cultivo com precisão e segurança em um campo de cultivo com seções descontínuas de plantio de soja. Baseado na informação de localização das linhas de plantio, os autores reforçam que o sistema proposto sobressai na resolução do problema SLAM para linhas retas e curvas, contínuas e descontínuas em diferentes níveis de processamento. A estrutura geométrica da instalação da câmera é ilustrada na Figura 2.30.

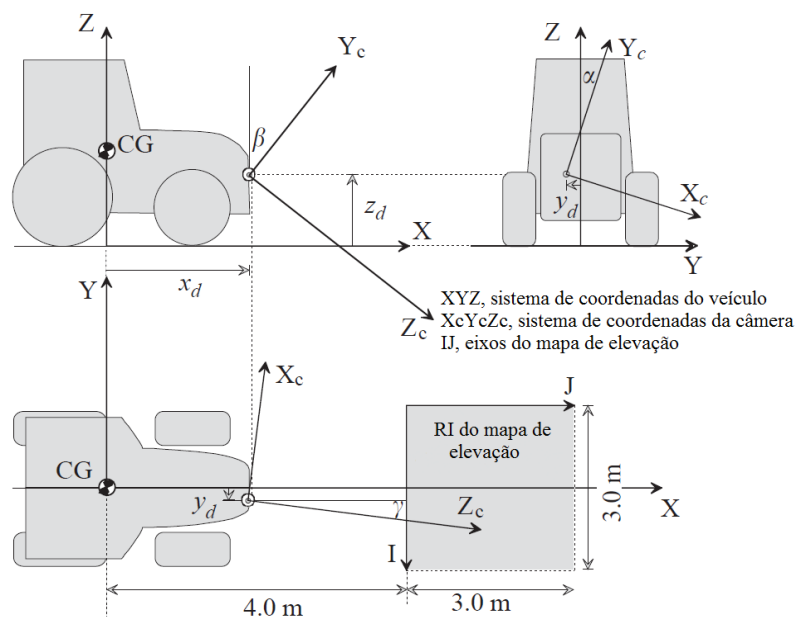


Figura 2.30: Instalação da câmera e sistema de coordenadas; CG – centro de gravidade; RI – região de interesse do mapa de elevação; x_d (longitude), y_d (*offset* horizontal) e z_d (*offset* vertical) representam a localização da câmera no sistema de coordenadas; α , β e γ representam os ângulos de ação da câmera em *roll*, *tilt* e *pan* (KISE *et al.*, 2005).

Abaixo, na Figura 2.31, são ilustradas as etapas da aquisição da imagem, o processamento por disparidade de *pixels*, o mapa de elevação das linhas de plantio detectadas, e sua posterior suavização.

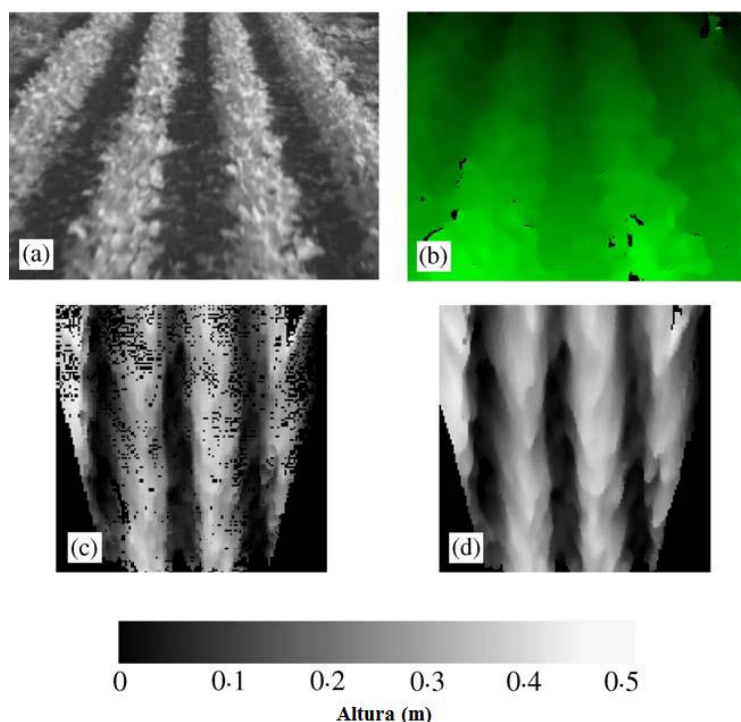


Figura 2.31: (a) imagem original; (b) imagem com disparidade de *pixels*; (c) mapa de elevação das linhas de plantio; (d) mapa de elevação suavizado - estéreo (Adaptado de KISE *et al.*, 2005).

Em trabalho mais recente (KISE e ZHANG, 2008), os mesmos autores aprimoraram o algoritmo de criação do mapa de elevação, fornecendo detalhes de altura e volume da área de cultivo, através de modificações da modelagem matemática na adaptação da câmera estéreo no veículo autônomo.

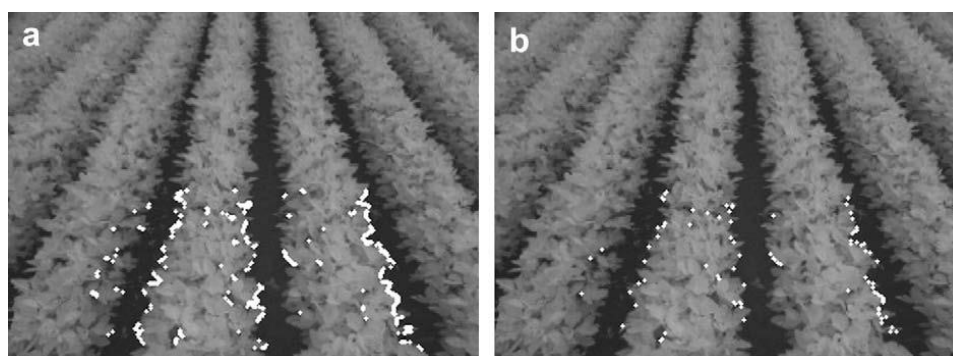


Figura 2.32: Localização dos pontos de características após remodelagem do sensor óptico estéreo: (a) os 400 maiores gradientes de todos os *pixels* foram selecionados como pontos de características; (b) como resultado da localização de tais características, 89 pontos correspondentes foram encontrados na imagem adjacente, fornecendo maior exatidão na descrição das linhas de plantio (KISE e ZHANG, 2008).

Os resultados apresentados na Figura 2.32, acima, denotam a localização dos pontos de características após a remodelagem do sensor óptico estéreo.

Seguindo o mesmo paradigma na concepção das imagens em estéreo para a detecção de linhas de plantio, um algoritmo de validação de disparidade baseada em multiplicidade de imagens foi proposto em Zhang (ZHANG *et al.*, 2005). A aplicação deste algoritmo é direcionada ao reconhecimento de campos de plantação de arroz. As intensidades em tons de cinza e a relação de correlação foram integradas a fim de produzir as disparidades nas imagens em estéreo. A superfície do terreno e o plantio de arroz foram conjugados como dois planos acidentados, e suas disparidades oscilaram numa baixa faixa de variação. As bordas contínuas e descontínuas do plantio de arroz foram primeiramente detectadas e posicionadas nas imagens. Logo, os pontos fora das bordas foram combinados respectivamente aos detectados, resultando nos valores de disparidade estimados, através de um método de correlação. As coordenadas tridimensionais da câmera foram encontradas mediante estas disparidades e as coordenadas do veículo obtidas pela multiplicação das primeiras coordenadas com uma fórmula de transformação.

Para a navegação por visão computacional, Jiang & Zhao (JIANG e ZHAO, 2010) consideraram que em sistemas de navegação autônoma baseada em visão computacional para robôs móveis agrícolas, a obtenção de informações para a guiagem do RM, pela estrutura das linhas de plantio, é o modo mais complexo, porém mais preciso de se trabalhar em técnicas que auxiliem na navegação destes veículos. O projeto proposto apresenta um método de reconhecimento robusto de plantações baseado na transformada de *Hough*. Uma câmera de visão monocular é usada para encontrar o caminho navegável nas áreas de plantio, para que o robô ao mesmo tempo se locomova e siga as linhas de plantio. Primeiramente, a calibração da câmera foi realizada para obter a relação entre as coordenadas da imagem e as coordenadas globais. Logo, os algoritmos de processamento de imagens e reconhecimento de padrões

foram aplicados para se obter uma semi-base de navegação. Por fim, a linha de navegação real foi extraída desta base pela transformada de *Hough*. Os resultados experimentais indicaram que o método propõe um algoritmo simples, mas robusto, que exige baixo nível computacional e atende a navegação dentre as linhas de plantio (Figura 2.33).

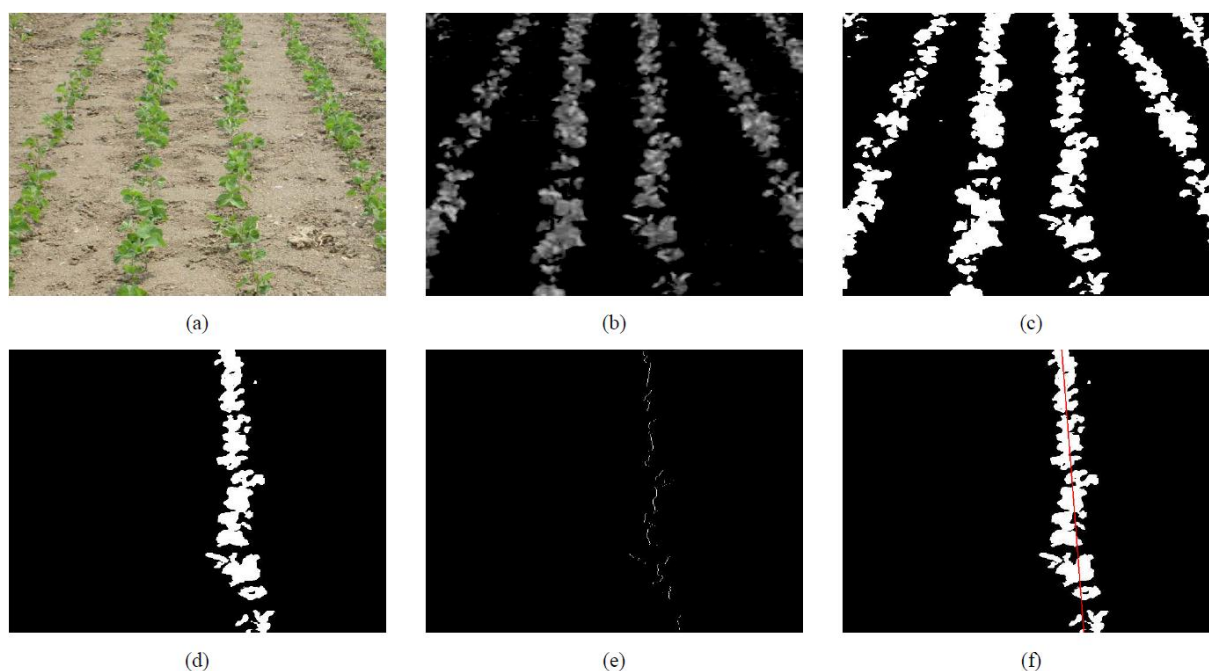


Figura 2.33: Resultado do processamento de imagens: (a) imagem original; (b) transformação da imagem em níveis de cinza; (c) binarização de OTSU; (d) projeção vertical para obter uma linha de plantio apenas; (e) detecção de bordas da linha de plantio; (f) transformada de *Hough* para delimitar a área navegável (JIANG e ZHAO, 2010).

2.4. Considerações finais

Neste capítulo foram fundamentadas as teorias da visão computacional, as técnicas a ela empregadas e a aplicação na área agrícola. A robótica móvel também foi estudada e abordada em relação à percepção e incertezas na etapa sensorial, onde a visão computacional é aplicada. Por fim, vários trabalhos no estado-da-arte foram revisados aplicados à robótica móvel agrícola e a técnicas ou sistemas de visão computacional, sendo referências a elaboração e discussão das propostas no presente trabalho.

3. Metodologia proposta

Neste capítulo são detalhados os materiais e métodos usados para a elaboração de uma arquitetura modular adequada para o problema de localização e inspeção de culturas agrícolas, e para o problema de locomoção do robô móvel entre as linhas de plantio. Nestas duas abordagens são relatados os estágios desenvolvidos:

- a) Caracterização e pré-processamento das imagens adquiridas para a inspeção de laranja, e parâmetros de análise de milho e cana. Caracterização e pré-processamento das imagens adquiridas para a delimitação das áreas de plantio de pomares, milharais e canaviais para a etapa de locomoção;
- b) Processo de segmentação, extração e normalização de características dos segmentos, classificação e reconhecimento de padrões para a quantificação e inspeção visual das culturas, validação dos classificadores aplicados;
- c) Obtenção do ângulo de guiagem, e caracterização de alinhamento, correspondência e integração de *pixels* nos mapas de classe de regiões de interesse para a navegação do robô móvel entre as linhas de plantio de pomares, milharais e canaviais;
- d) Pós-processamento das imagens resultantes com as regiões classificadas dos mapas de classe, para a detecção dos padrões alvos, análise de elementos geométricos e quantificação de objetos.

A arquitetura modular para as duas abordagens foi testada com câmeras CCD na obtenção de imagens estáticas e dinâmicas em campo, sem a adaptação destas no robô móvel agrícola. Os algoritmos usados para todos os estágios acima foram implementados sob a plataforma computacional *Matlab*, com recorrência à técnica de segmentação escrita em C++ e customizada pela biblioteca *OpenCV*, dentro deste ambiente. Ademais, a configuração de *hardware* dos computadores pessoais (PC) usados para os testes com tais algoritmos atende a

requisitos de processador de núcleo duplo e quádruplo (*dualcore/quadcore*) com frequência de operação entre 2,00GHz e 3,50GHz para sistemas operacionais de 32 *bits*, memória volátil RAM de 8GB, memória não-volátil HDD de 500GB e suporte a interfaces *GigaBit Ethernet*, *USB* e *Firewire* para conexões diretas com as câmeras CCD.

Ambas as aplicações são divididas em módulos específicos de desenvolvimento, como mostrado na Figura 3.1, abaixo.

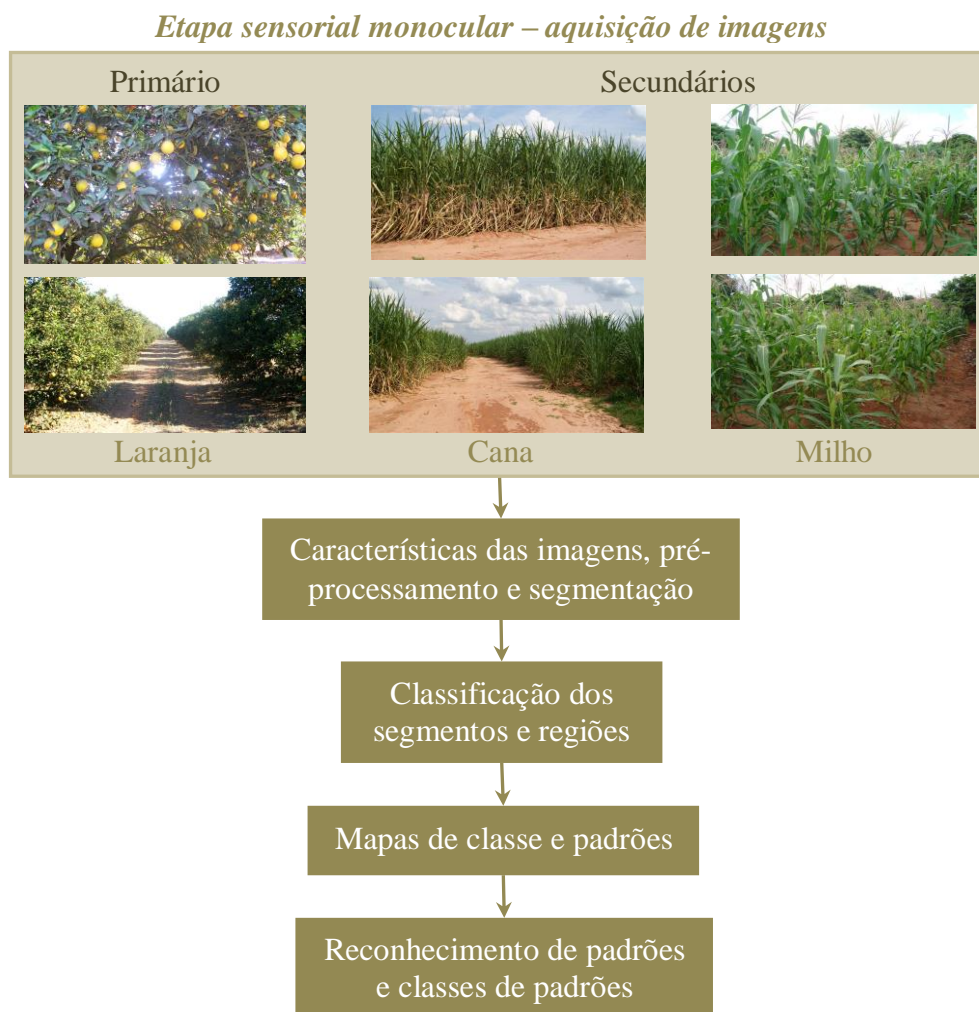


Figura 3.1: Estratégia modular para o processamento de imagens em cenas naturais.

Para a análise e processamento de imagens relativas às três culturas são estabelecidos e estudados critérios intrínsecos a cada uma, correspondendo suas características físicas ora à questão da inspeção *in locu* da qualidade e quantificação da cultura, ora na determinação das linhas de plantio para a navegação do veículo autônomo.

As cenas naturais correspondentes à laranja (prioridade primária para as duas abordagens) são estudadas quanto à sua distribuição não linear em cada árvore do pomar analisado, tendo como parâmetros de identificação o tamanho da área de cada fruto, e conseqüentemente, a estimativa de seu volume, a variação de cor apresentada desde o início da maturação até seu pico, e deformidades de textura ocasionadas por aberrações cromáticas e intensidade de luz incidente. Para a delimitação da área de plantio dos pomares, são analisadas as cenas com visão em perspectiva frontal, as quais identificam possíveis regiões de interesse à segmentação por cor e textura, diferenciando áreas livres para a locomoção do robô móvel, e áreas denominadas intransitáveis (plantações).

As culturas milho e cana são analisadas com prioridade secundária, uma vez que a primeira é direcionada à localização das gramíneas e a dispersão de suas linhas de plantio e a segunda direcionada à navegação. Tal abordagem é caracterizada inicialmente em cenas naturais da distribuição de pomares de laranja, denominando um padrão base ao problema da navegação. As imagens de milho são analisadas em relação à sua localização e formação diagonal entre o colmo e as folhas esparsas de cada gramínea, e sendo característicos pela estimativa do volume da espiga e seu pico, denotado como estigma. As cenas que apresentam as linhas de plantio em perspectiva frontal e lateral são direcionadas ao estudo da caracterização do ângulo de guiagem para a navegação do robô móvel.

As cenas correspondentes à plantação de cana são em perspectiva frontal e lateral, uma vez que o propósito para esta cultura é apenas delimitar as regiões passíveis de navegação das regiões alocadas ao plantio.

3.1. Características de aquisição e análise das imagens

As imagens adquiridas das três diferentes culturas (laranja, milho e cana) atendem a fatores climáticos e temporais diversos, com baixa e alta luminosidade durante o dia (manhã e

tarde), incidência de luz solar e sombra em diferentes momentos, e umidade baixa ou alta em influência à texturização de toda a cena. As imagens dos pomares foram adquiridas durante a época de colheita, em pico de maturação; As imagens dos canaviais foram adquiridas durante a época de plantio, em observação a plantação descontínua das linhas de cultivo, e durante a época de colheita, totalmente aglomeradas em função da concentração do caule em colmos e as folhas com bainhas abertas e esparsas; as imagens dos milharais foram adquiridas durante a época de plantio com ciclo entre as safras de 100 a 115 dias, e durante a época de colheita em linhas de plantio contínuas e densas, também pelo fato de apresentar folhas com bainhas abertas e esparsas.

Pomares

A distância das câmeras é determinada de acordo com a aplicação, mas com tendência livre de aquisição. Para a identificação e inspeção de laranjas, as imagens adquiridas foram estabelecidas em distâncias de 5m à 0,50m entre as linhas de plantio e o robô agrícola para as imagens que lidam com o pomar inteiro, e concentração de folhas, frutos e galhos. Para a navegação entre as linhas de plantio, as imagens atendem a localidades com o corredor central de separação destas linhas e onde circulará o robô móvel, parte do corredor e plantação à direita, parte do corredor e plantação à esquerda, e concentração do solo para ambos os lados de plantação.

A resolução espacial destas imagens é de 3664×2062 *pixels* com resolução cromática de *24-bits*, armazenadas em arquivo do tipo JPEG/JPG adquiridas no espaço de cores RGB. A base de dados para as cenas de pomares, tanto para a inspeção dos frutos, quanto para a navegação entre as linhas de plantio atende a um conjunto de 920 imagens dividido em dois conjuntos: pomares completos individuais e em conjunto (linhas de plantio), e áreas locais de incidência de frutos.

O primeiro grupo para as cenas de pomares completos individuais e em conjuntos mostram o plantio de cada árvore à 4m (perspectiva lateral) e das linhas de plantio (perspectiva frontal), com a representação baixa de *pixels* para frutos, sendo estes parcialmente ocultos por folhas e galhos, submetidos a condições de iluminação intensa ou baixa, e dependendo da textura dos frutos, estes podem refletir a luz natural incidente gerando focos brancos e aberrações cromáticas entre os *pixels* da região de contorno das laranjas e a reflexão de luz. Também refratam a luz devido a regiões que denotem sujeiras ou suspeitas de doença encontradas nos frutos. Outra característica a se considerar é o crescimento desordenado dos galhos de cada pomar para delimitar sua área (irregular) de plantio, o que influencia a iluminação da copa para regiões com muita ou pouca sombra, e muita ou pouca incidência de iluminação. O solo nestas imagens é coberto em sua maioria por vegetação rasteira e muitas folhas e frutos caídos, e nas áreas descobertas apresenta uma mistura de cores entre alaranjado, vermelho e marrom. Em ambas as regiões do solo, há incidências diferentes de luz, e ao fundo, o céu é parcialmente coberto por poucas nuvens e varia em tons de cores entre o branco e azul. Na Figura 3.2, a seguir, são mostrados exemplos destas cenas.



Figura 3.2: Cenas das linhas de plantio de pomares.

O segundo grupo para as cenas de incidência local de frutos mostra o plantio de árvores individuais, representando desde a concepção do caule até a copa, pela representação em perspectiva frontal, à distâncias de 4m a 2m. Novamente, têm-se uma baixa representatividade de *pixels* para os frutos, cobertos em minoria pelas folhas e galhos, que compõem as cenas também, de copas esparsas, e crescimento irregular dos galhos, folhas e ramos, entre os frutos. Na Figura 3.3, a seguir, são mostrados exemplos destas cenas.



Figura 3.3: Cenas de pomares individuais para toda a árvore e áreas locais de incidência de frutos.

Milharais

A base de dados para as cenas de milharais atende a 900 imagens com resolução cromática de 24-bits e resolução espacial de 3072×2304 *pixels*, também em compactação de arquivo JPG. A aquisição de imagens foi feita no espaço de cores RGB.

São dois conjuntos característicos que compõem tal cultura, sendo o primeiro direcionado a imagens em perspectiva lateral, que demonstram as gramíneas em conjunto para a localização das espigas, em contraste com as bainhas abertas e esparsas, em dois períodos de colheita (milho verde e silagem); o segundo grupo apresenta imagens em perspectiva frontal, para cenas compostas de área de plantio e as gramíneas em conjunto, com espaços entre cada

milharal – apresentam períodos de safra e colheita. Na Figura 3.4, a seguir, são mostrados exemplos de tais cenas.



Figura 3.4: (a) Milharal verde em período de colheita; (b) Milharal silagem em período de colheita; (c) Linhas de plantio de milharais em período de safra; (d) Linhas de plantio de milharais em período de colheita.

Canaviais

A base de dados para as imagens de canaviais apresenta projeções em perspectiva frontal dos conjuntos de plantações em linha, justamente com o objetivo de navegação do veículo autônomo. São 750 imagens com resolução espacial de 3664×2062 pixels (24-bits), também adquiridas no espaço de cores RGB.

Pode-se observar nas cenas mostradas na Figura 3.5 que a densidade entre os colmos de cada planta do canavial à direita e à esquerda, é alta, facilitando a quantização de cores e a segmentação por textura. O solo, como área considerável navegável atende pela variação de

cor entre tonalidades de marrom e vermelho, e o céu, invariante à tonalidade branca e azul, é descartada do processo de segmentação, uma vez que somente a delimitação entre as áreas de plantio dos canaviais é adequada ao propósito da navegação veicular, otimizando o tempo de processamento. São cenas que denotam canaviais em períodos de colheita e plantio.



Figura 3.5: Linhas de plantio de canaviais em período de colheita (a) e (b) e em período de safra (c).

3.2. Processamento de imagens

A metodologia adotada neste trabalho, devido a aplicações de reconhecimento de padrões não apenas para cenas naturais em plantio de laranja, milho e cana, estende-se a outras culturas de interesse, o que distingue as etapas de segmentação e classificação das imagens para modelos individuais paramétricos. Logo, os métodos são aplicados em duas diferentes abordagens: cenas naturais das linhas de plantio de laranja, milho e cana, relativas ao problema de locomoção do robô agrícola móvel; e, cenas naturais relativas ao reconhecimento de laranjas e ao reconhecimento de milhos.

Nas duas abordagens, a estrutura adotada mantém-se equivalente, apenas diferenciando-se na sequência de cada etapa, provendo recursos adicionais ao interesse estudado. Tais estratégias são mostradas na Figura 3.6 e Figura 3.7, a seguir.

Cenas naturais das linhas de plantio de laranja, cana e milho à locomoção

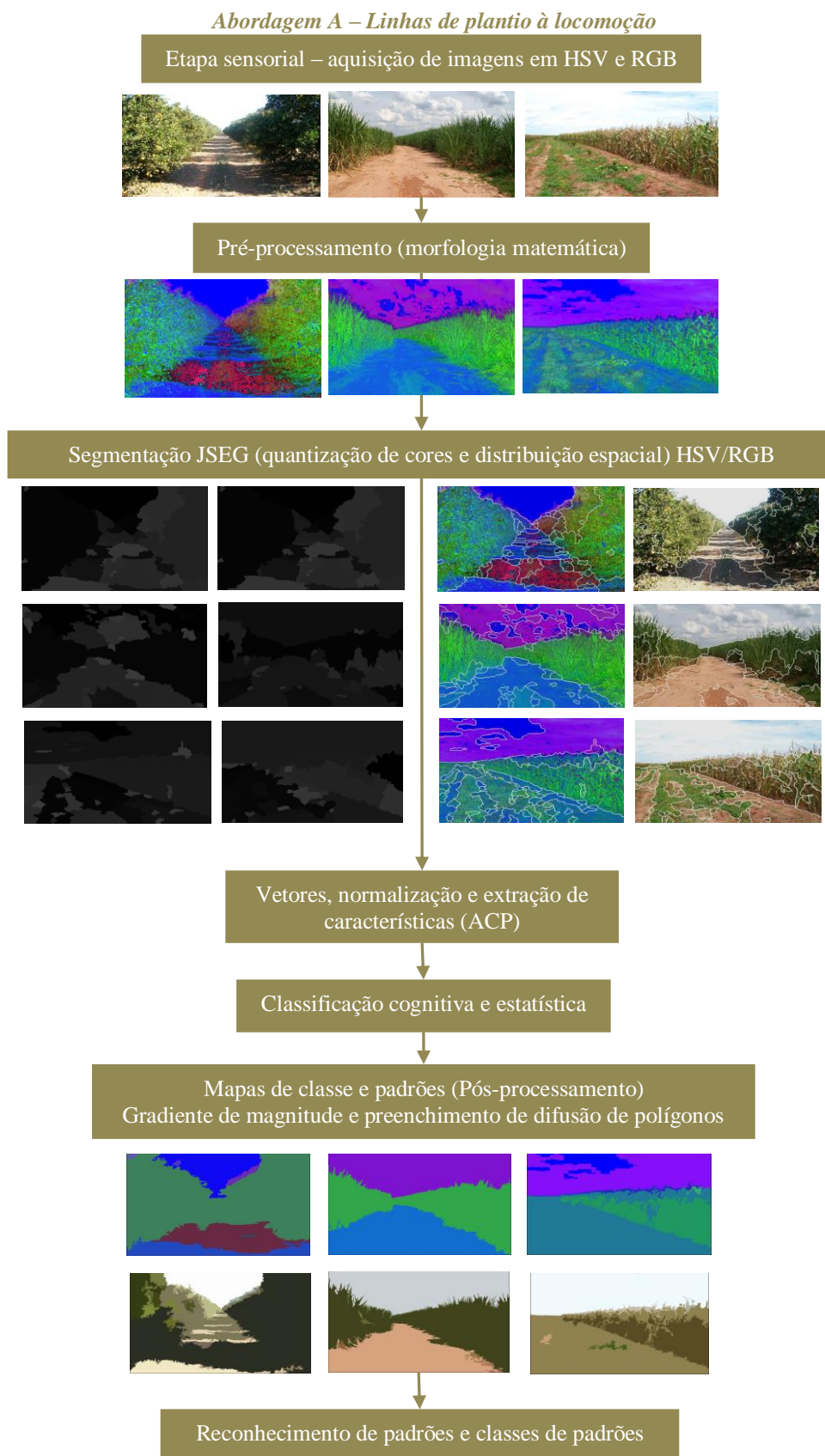


Figura 3.6: Estrutura da primeira abordagem para o processamento e obtenção dos segmentos classificados.

Cenas naturais relacionadas ao reconhecimento de laranjas e milhos

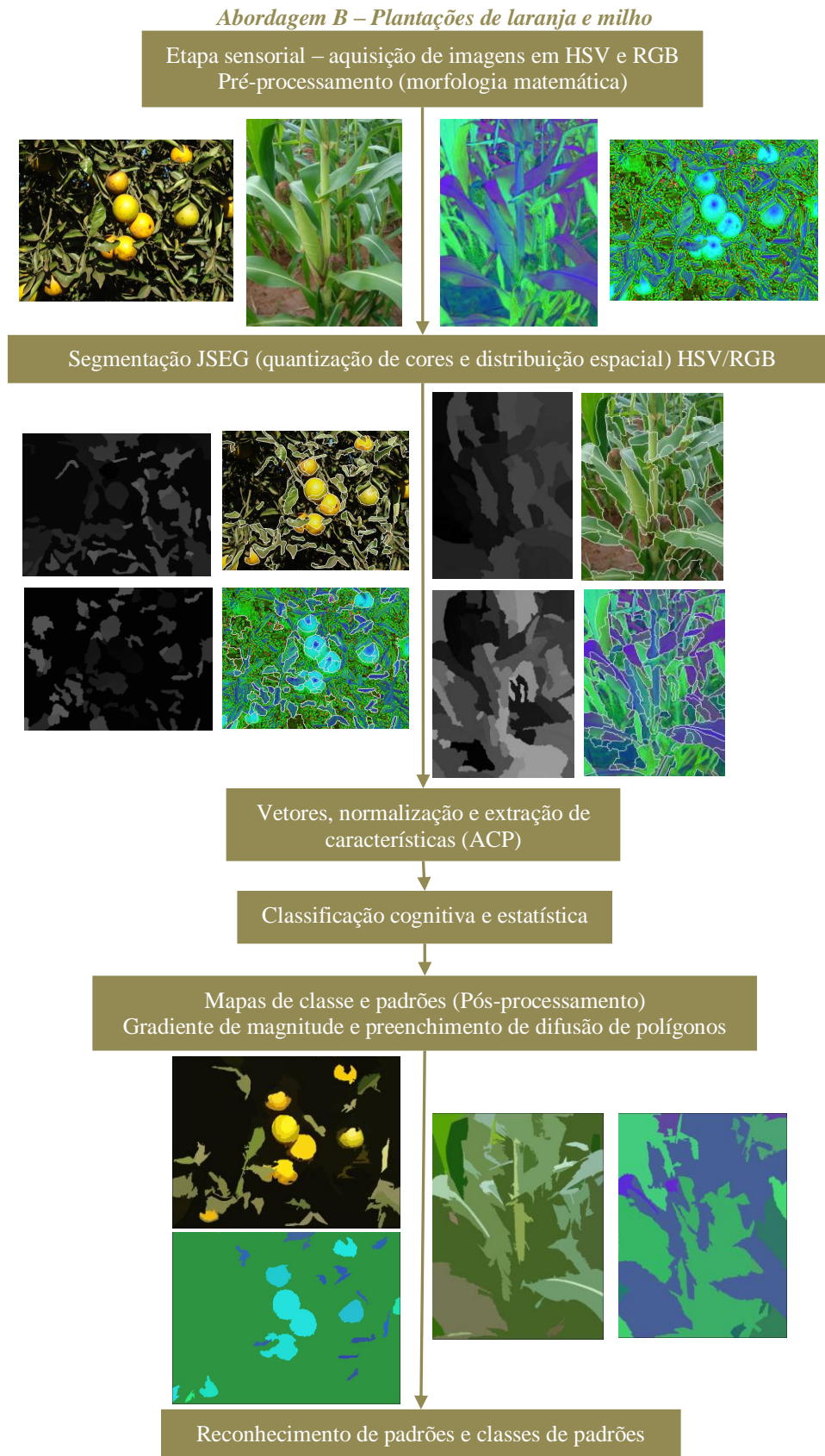


Figura 3.7: Estrutura da segunda abordagem para o processamento e obtenção dos segmentos classificados.

A segmentação é realizada em vários tipos de imagens de forma não supervisionada, não tendo o conhecimento *a priori* das informações do que deve ser segmentado. Assim, o algoritmo dedicado a tal, estende a flexibilidade na aplicação em imagens com características por cores e texturas, ou adaptadas em pós-processamento.

A classificação dos segmentos é feita automaticamente, mas com o prévio conhecimento das características dos segmentos, que une a segmentação com o reconhecimento de padrões, tornando o modelo mais robusto e expansível a outras regiões de interesse na imagem. As características extraídas podem ser classificadas por critério imposto (situações em que os segmentos são membros de determinada classe), por exemplos (situações em que não se conhece a relação dos segmentos com uma classe específica) ou critério aberto (situações em busca de agrupamentos desconhecidos).

Em cenas naturais, para as três culturas citadas, como objetivo deste trabalho, as imagens adquiridas são submetidas às condições destes ambientes como iluminação variável, oclusões, texturas, tonalidades alternáveis entre os frutos/cereais para com as folhas, ramos e galhos, tornando a extração de características mais complexa. Os classificadores adotados então tratam as informações de interesse por exemplos, para que haja supervisão na extração das características da cena em condições ideais de estudo para as culturas laranja, milho e cana.

A metodologia adotada atende à seguinte sequência: aquisição das imagens; pré-processamento para redução de resolução espacial e limiarização na transformação do espaço de cores RGB para HSV, como operações morfológicas incidentes nos valores dos *pixels* destes espaços de cores; segmentação das imagens; cálculo das características dos segmentos para a extração (análise das componentes principais dos vetores de características); classificação dos segmentos; geração dos mapas de classe para os padrões segmentados em ambos os espaços de cores.

3.2.1. Pré-processamento e representações gráficas

O pré-processamento nas cenas de pomares, milharais e canaviais, relativo ao problema de locomoção é aplicado morfologicamente na redução de dimensionalidade para a resolução espacial das imagens, e transformação do espaço de cores.

Com a nova imagem redimensionada, deve-se analisar qual espaço de cores atende às premissas de processamento em cenas naturais, sujeitas à variação de luminosidade, reflexão e sombreamento. O espaço de cores HSV adéqua-se a esta aplicação, uma vez que os canais S e V (saturação/valor) separam as informações cromáticas relevantes à discriminação dos padrões de segmentos desejados destes ruídos. Assim, as imagens em RGB, que serão processadas também, como critério de validação dos classificadores, são transformadas ao espaço de cores HSV, dadas pelas equações 3.1., 3.2 e 3.3:

$$H = \arctan\left(\frac{\sqrt{3}(G-B)}{(R-G)(R-B)}\right) \quad (3.1)$$

$$V = \frac{(R+G+B)}{3} \quad (3.2)$$

$$S = 1 - \frac{\min(R,G,B)}{l} \quad (3.3)$$

3.2.2. Segmentação de imagens em cenas naturais

No processo de segmentação de cenas naturais devem-se considerar a distinção de cores entre as variedades das culturas e os demais entes naturais como folhas, galhos e ramos, além da textura, uma vez que é possível encontrar regiões de cor não homogênea. Imagens coloridas com regiões homogêneas são segmentadas através de um algoritmo de agrupamentos que os separa no espaço de cores pré-processado, isto é, classes de diferentes magnitudes na distribuição espectral, com intensidade específica da radiação eletromagnética visível em vários comprimentos de onda. Este processo é direcionado às três abordagens

citadas anteriormente, uma vez que todas apresentam componentes físicos naturais, mas com relativa desambiguidade de agrupamentos de cores em cada situação. Dessa forma, são desenvolvidas e aplicadas técnicas distintas de segmentação a cada propósito.

A abordagem que lida com a navegação visual do veículo autônomo, atende às três linhas de plantio (laranja, milho e cana), por se tratar de regiões com separação regular de *pixels* em regiões consideradas homogêneas, isto é, nas três plantações são características da cena perspectiva frontal, as linhas de plantio, denotando variação de tonalidades à cor verde e amarela, e os espaços entre as linhas, denotando o a variação de cor e textura do solo.

Segmentação não-supervisionada baseada em regiões homogêneas (JSEG)

A segmentação de imagens texturizadas considera o agrupamento espacial de *pixels* em crescimento de regiões não-homogêneas, tornando-as segmentos homogêneos, através da relação de diferentes escalas no espaço de cores tratado. Em cenas naturais, trata-se de uma técnica não-supervisionada para regiões texturizadas e coloridas, a qual é apropriada na verificação de homogeneidade em padrões de regiões com estas características, e computacionalmente mais viável e eficaz que modelos de estimação de parâmetros.

Neste modelo de segmentação não-supervisionado, são consideradas para a análise da imagem adquirida, as regiões homogêneas de cor e textura, a quantização das informações de cores, e a distinção de cores entre regiões vizinhas. Denominado JSEG (DENG *et al.*, 1999a; DENG e MANJUNATH, 1999c; DENG e MANJUNATH, 2001) (em referência à segmentação da região *J* – agrupamento de classes), tal técnica apresenta duas etapas de processamento: quantização do espaço de cores (*peer-group filtering* - PGF) (DENG *et al.*, 1999b) – processo de redução de números de cores distintas na imagem, e crescimento/agrupamento de regiões com cores semelhantes.

Na primeira etapa, o algoritmo perceptivo de quantização é efetuado com mínima degradação da homogeneidade de cores, definindo regiões em cores reduzidas, cada qual associada a uma classe. Os *pixels* na imagem original são substituídos por tais classes, formando um mapa de classes em composição de texturas. Cada mapa de classe representa uma região J (J -imagem), denotada pela janela de cor quantizada, contendo *pixels* com valores positivos e negativos (valores J) ao contorno e a texturização de suas fronteiras. O crescimento e agrupamento de regiões com cores semelhantes, como segunda etapa, são associados na identificação da similaridade de cores e na distribuição espacial destas cores, delimitando as regiões de contorno da homogeneidade encontrada. A Figura 3.8, a seguir, mostra esquematicamente o funcionamento do algoritmo JSEG.

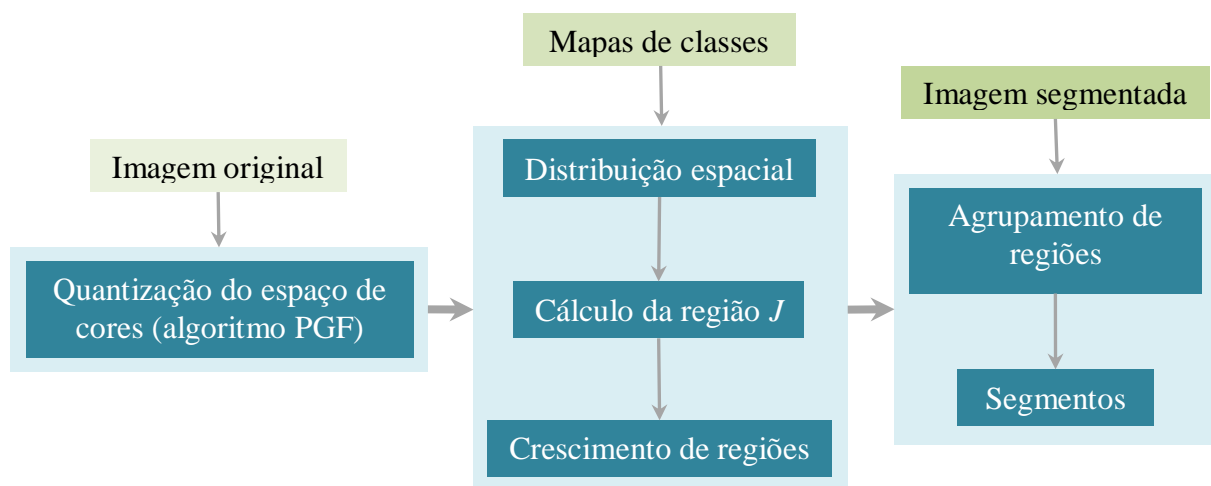


Figura 3.8: Estratégia do algoritmo de segmentação não-supervisionado JSEG.

Os itens abaixo são organizados para a descrição das etapas de desenvolvimento da técnica de segmentação e validação.

a) Validação do algoritmo de segmentação

Como validação do algoritmo de segmentação, as cenas naturais adquiridas apresentam resolução cromática em 24-*bits*, que são densamente quantizadas, preservando a qualidade da imagem. O critério para segmentação ótima na validação é extrair cores representativas da imagem, diferenciando as regiões vizinhas, de forma não-supervisionada. Com isso, a

quantização do espaço de cores por filtragem de pares de regiões é aplicada de forma ponderada em cada *pixel*, para suavizar todas as regiões da imagem e remover ruídos existentes nas mesmas. Logo, os novos valores de *pixels* indicando a suavização das áreas locais são obtidos, e um novo peso é associado a cada *pixel*, priorizando as regiões texturizadas das suavizadas. Tais áreas são identificadas com um vetor de quantização às cores das regiões, baseadas no algoritmo GLA (*general Lloyd algorithm*) (GERSHO e GRAY, 1992), em que o espaço de cores $L*u*v$ é adotado, para preservar a distorção global das regiões, denotada por D , observada na equação 3.4, abaixo, assim como sua forma derivada na equação 3.5.

$$D = \sum_i D_i = \sum_i \sum_n v(n) \|x(n) - c_i\|^2 \rightarrow x(n) \in C_i \quad (3.4)$$

$$c_i = \frac{\sum v(n)x(n)}{\sum v(n)} \rightarrow x(n) \in C_i \quad (3.5)$$

O parâmetro c_i é o centróide do agrupamento (*clustering*) C_i , $x(n)$ e $y(n)$ são os vetores de cor e peso perceptivo para o *pixel* n ; D_i é a distorção local para C_i .

Com o valor do centróide, denotado pela equação 3.5, e após a quantização vetorial de cores e o agrupamento de regiões, os *pixels* com a mesma cor possuem dois ou mais agrupamentos, baseados no cálculo da distorção global GLA. Para a junção de agrupamentos com distância mínima entre os limiares predefinidos para dois centróides (equação 3.5), aplica-se uma derivação do algoritmo GLA, para agrupamentos aglomerados (DUDA e HART, 1970). Após a fusão dos agrupamentos para a quantização de cores, um rótulo é associado para cada região, representando um mapa de classe para os *pixels* de mesma tonalidade de cor.

Dessa forma, a informação necessária para a segmentação, após a quantização de cores, é extraída e realocada em um mapa de classes. Uma região específica contém *pixels* de uma determinada classe de cor, a qual é distribuída em todas as regiões homogêneas, e têm seus

pixels distribuídos em todo o rearranjo espacial, formando um plano bidimensional, para o vetor de posição (x,y) .

Para se calcular então, o valor J , Z é definido como um conjunto de pontos da imagem quantizada, logo, $z = (x,y)$ com $z \in Z$, e considerando m a média para todos os elementos Z . C é o número de classes obtido na quantização, assim, Z é classificada em classes C , e Z_i são os elementos de Z pertencentes à classe i , onde $i = 1, \dots, C$, e m_i são as médias dos elementos em Z_i , dados pelas equações 3.6 e 3.7.

$$m = \frac{1}{N} \sum_{z \in Z} z \quad (3.6)$$

$$m_i = \frac{1}{N_i} \sum_{z \in Z} z \quad (3.7)$$

Portanto, o valor J é dado por:

$$J = \frac{S_B}{S_W} = \frac{(S_T - S_W)}{S_W} \quad (3.8)$$

onde:

$$\begin{aligned} S_T &= \sum_{z \in Z} \|z - m\|^2 \\ S_W &= \sum_{i=1}^C \sum_{z \in Z} \|z - m_i\|^2 \end{aligned} \quad (3.9)$$

O parâmetro S_T representa a soma dos pontos quantizados na imagem, dentro de uma média em todos os elementos Z . Logo, a relação entre S_B e S_W denota as medidas de distância entre os mapas de classe, para distribuições não-lineares arbitrárias. Valores de J altos indicam um aumento na distância entre as classes e os pontos de cada, considerando imagens com regiões de cores homogêneas. Cada região segmentada pode ser recalculada – equação 3.10 abaixo, ao invés de todo o mapa de classe, com novos ajustes de parâmetros para a média \bar{J} . J_K representa o valor J calculado sobre a região k ; M_k é o número de pontos na região k , N

é o número total de pontos no mapa de classes, com todas as regiões na somatória destes mapas.

$$\bar{J} = \frac{1}{N} \sum_k M_k J_k \quad (3.10)$$

b) *Técnica de distribuição espacial*

A minimização global de \bar{J} não é prática, se não aplicada a uma área local do mapa de classes. Com isso, a região J é representada por uma imagem em escala de níveis de cinza, cujos *pixels* são os valores J calculados sobre janelas locais centralizadas nestes *pixels*. Com valores altos aplicados à J , o *pixel* correspondente estará próximo à região de contorno.

A Figura 3.9, a seguir, mostra o fluxograma do algoritmo de distribuição espacial.

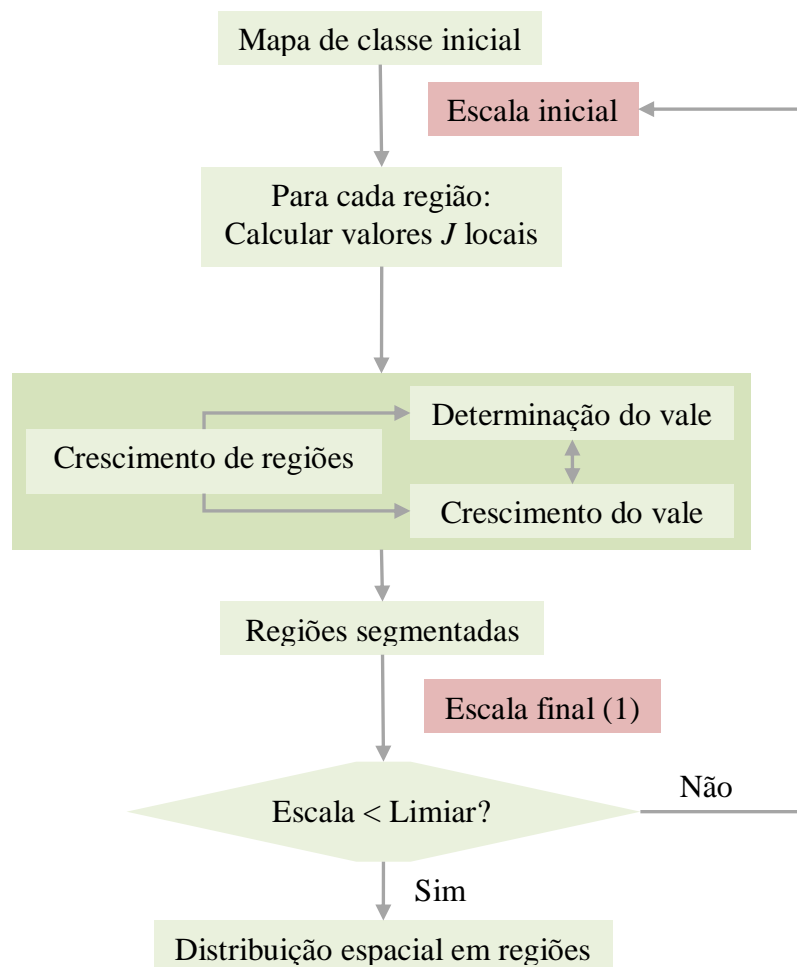


Figura 3.9: Fluxograma do algoritmo de distribuição espacial.

O fluxograma mostrado pela Figura 3.9, inicializa o processo de segmentação pela etapa *Mapa de classe inicial*, correspondente a imagem original de entrada, sem a discriminação de regiões homogêneas. Em seguida, para a etapa *Calcular valores J locais*, em cada região é atribuído um rótulo dimensional com a intensidade de cor no contorno destas regiões, verificando se incidência de *pixels* nas bordas é constante para aquela região. Na etapa *Crescimento de regiões*, ocorrem os processos de determinação e crescimento do vale (item *c*, abaixo), cuja finalidade é determinar se as regiões vizinhas são associadas com os mesmos limiares dos vales e, assim, unir possíveis regiões candidatas a valores homogêneos. Por conseguinte, as regiões homogêneas são separadas através da quantização de cores (etapa *Regiões segmentadas*) e segmentadas até atingir a escala pretendida como limiar de segmentação homogênea (etapa condicional $Escala < Limiar$). Por fim, é realizada a segmentação pela distribuição espacial em regiões homogêneas.

Dessa forma, as dimensões de janelas locais determinam o tamanho das regiões J , para a intensidade das cores em bordas de áreas pequenas, e o oposto acontece para a detecção de regiões de contorno em texturas. A distribuição espacial, então, inicia a segmentação de todas as regiões na imagem, iniciando pelo maior limiar de escala de resolução até a especificação de um dado limiar menor. Este limiar é configurado manualmente para a correspondência com a resolução espacial da imagem em *pixels*.

c) Determinação e crescimento do vale de regiões J

A determinação heurística do vale de regiões J pressupõe que regiões de dimensões pequenas são definidas como os padrões de crescimento iniciais. Estas regiões têm os menores valores J (vales), e seu cálculo segue os métodos abaixo.

- Calcular o desvio padrão e a média de valores J locais na região, denotados por σ_J e μ_J , respectivamente.
- Limiarizar os parâmetros acima.

$$T_j = \mu_j + a\sigma_j \quad (3.11)$$

A condição para se considerar pontos de vale candidatos a *pixels* com valores J locais, é determinada por $T_j > J$. Conectar os pontos de cada região baseados em conectividade de 4 (código de cadeia para a delimitação da região de contorno) e obter os vales.

- Para vales candidatos menores que a relação da distribuição espacial entre a escala e o tamanho da imagem, estes são denotados vales.
- Parâmetros predefinidos empíricos de valores [-0,6; -0,4; -0,2; 0; 0,2; 0,4] são dados pela variável a , o qual fornece a maior estimativa de candidatos a vales.

Após a determinação dos vales, as novas regiões crescem a partir destes, os quais seguem as premissas abaixo:

- *Pixels* não determinados devem ser removidos dos vales, produzindo a média dos valores J locais, nas regiões sobressalentes não segmentadas. Logo, os *pixels* são conectados abaixo da média para formar áreas de crescimento, e se estas forem adjacentes a um e apenas um vale, são associadas a este vale.
- Cálculo dos valores J locais aos *pixels* restantes na escala correspondente para localizar as regiões de contorno.
- Crescimento dos *pixels* restantes no final da escala. *Pixels* não classificados nas fronteiras dos vales são armazenados num *buffer*. Tal *buffer* é atualizado quando todos os *pixels* são classificados.

Uma segmentação inicial da imagem é obtida, após o crescimento de regiões, apresentando áreas super-segmentadas, combinadas em suas similaridades de cor. As cores quantizadas representam os feixes dos histogramas de cor, e suas medidas de distância são calculadas pela extração de regiões (distância euclidiana entre duas regiões vizinhas). Os pares de regiões com distâncias mínimas são combinados juntos. Todas as distâncias são armazenadas em uma base de dados e atualizadas quando se calcular a estimativa dos vetores

de características e a região correspondente. O processo continua até atingir um limiar máximo para a distância que, após a junção de regiões, tem-se a segmentação final com todas as regiões J calculadas.

O algoritmo, escrito originalmente em C++ (DENG e MANJUNATH, 2001), foi recompilado e customizado na aplicação de funções e bibliotecas em *OpenCV* (BRADSKI e KAEHLER, 2008), dedicadas ao processamento de imagens, na tentativa de melhorar a distribuição de regiões e o seu reagrupamento, com a fusão de outras técnicas de segmentação baseadas em detecção de bordas. Ademais, os parâmetros necessários à execução do algoritmo após a compilação são: *limiar de quantização de cores*, que corresponde ao número de cores usado na quantização; *número de escalas*, que refina os detalhes de segmentação na imagem; *limiar de agrupamento de regiões*, que une regiões segmentadas com cores semelhantes.

Seguindo a ordem, os parâmetros ajustados empiricamente a valores consideráveis na obtenção de resultados ótimos são: *limiar de quantização de cores* (270) – este valor varia numa faixa entre 0 à 600, e tendo sua adoção próxima a condições pretendidas ao número de regiões que serão quantizadas; *números de escalas* (3), 3 sendo a escala inicial e a 1 a escala final, consolidando um limiar que agrupa as regiões de interesse, e evitando que outras regiões sejam super-segmentadas, em 90% da aplicação; *limiar de agrupamento de regiões* (0,1) variando entre 0,1 e 0,7 – menores valores significam agrupamento de regiões com equivalência de *pixels*, favorecendo a classificação manual dos segmentos. A resolução espacial das imagens, após as etapas de pré-processamento, é de 852x480 *pixels*.

3.2.3. Vetores, normalização e extração de características

Como os segmentos das imagens são usados como características a cada classificador, a imagem total é desconsiderada, uma vez que a identificação se dá pelos elementos que

compõem a cena, e estes são representados por descritores que formam as componentes de interesse ao vetor de características dos segmentos.

As cores e texturas em cada espaço de cores são usadas como parâmetros na descrição ao vetor de características, através da relação tridimensional dos canais de cores HSV e RGB para a formação de histogramas das imagens segmentadas. A aplicação dos histogramas é feita automaticamente por funções ordenadas no *MATLAB*, formando uma matriz de histogramas. Com os histogramas normalizados sendo os descritores para o vetor de características de cada imagem segmentada é possível confirmar a projeção dos *pixels* da cor de interesse em cada histograma. Logo, são usadas todas as dimensões de cada canal dos espaços de cores HSV e RGB, agregando o valor de 256×3 (768) à quantidade de categoria.

Assim, são particionados conjuntos de treinamento para cada abordagem nos espaços de cores RGB e HSV. O desempenho dos classificadores cognitivos e estatísticos é avaliado para os casos onde os vetores de características estão mais distantes entre os centróides de cada classe. Como abordado por Cavani (CAVANI, 2007), as porcentagens relativas das distâncias euclidianas entre os centróides de cada classe favorecem uma maior estimativa da dispersão dos vetores aos segmentos. Todavia, nas duas abordagens deste projeto, buscou-se parametrizar os vetores na classificação através da redução da dimensionalidade pela Análise das Componentes Principais (ACP), uma vez que as classes de entrada crescem proporcionalmente ao aumento das dimensões (maldição da dimensionalidade). Logo, os vetores mais distantes foram eliminados e os conjuntos de treinamento extraídos dos vetores. Dessa forma, dados redundantes do conjunto de treinamento são eliminados e comprimidos com padrões de interseção entre os conjuntos de teste e validação.

A ACP é aplicada inicialmente na extração dos autovetores da matriz de covariância - a média dos vetores de entrada em um vetor linha. Pela diferença de cada vetor média, encontra-se um padrão de dimensões da matriz de covariância. Em seguida, os autovetores

são recalculados e os autovalores são definidos na diagonal principal da matriz resultante, normalizando enfim as dimensões dos vetores de características das imagens de entrada no conjunto de treinamento.

Os histogramas obtidos nos espaços de cores RGB e HSV pela etapa de quantização de cores na segmentação das imagens são apresentados a seguir, nas Figuras 3.10, 3.11, 3.12, 3.13, 3.14 e 3.15, que denotam a aplicação em um conjunto de treinamento com três imagens em cada espaço de cor – proporção de aproximadamente 20:1 do conjunto treinado em 20% da base de dados total.

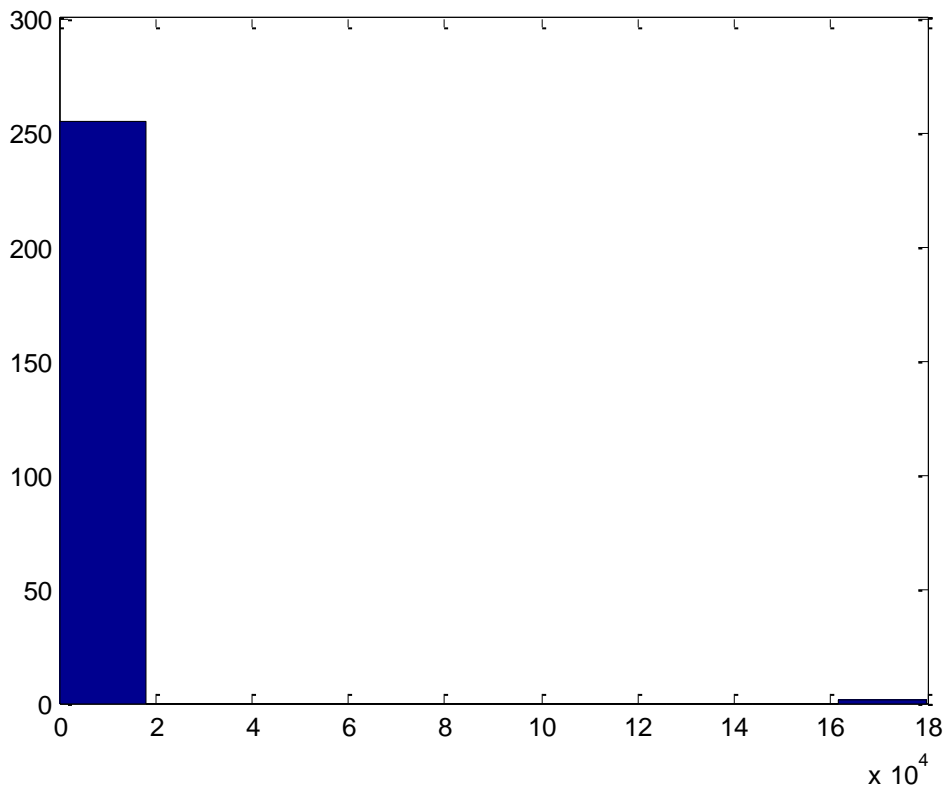


Figura 3.10: Histograma imagem quantizada 01 RGB.

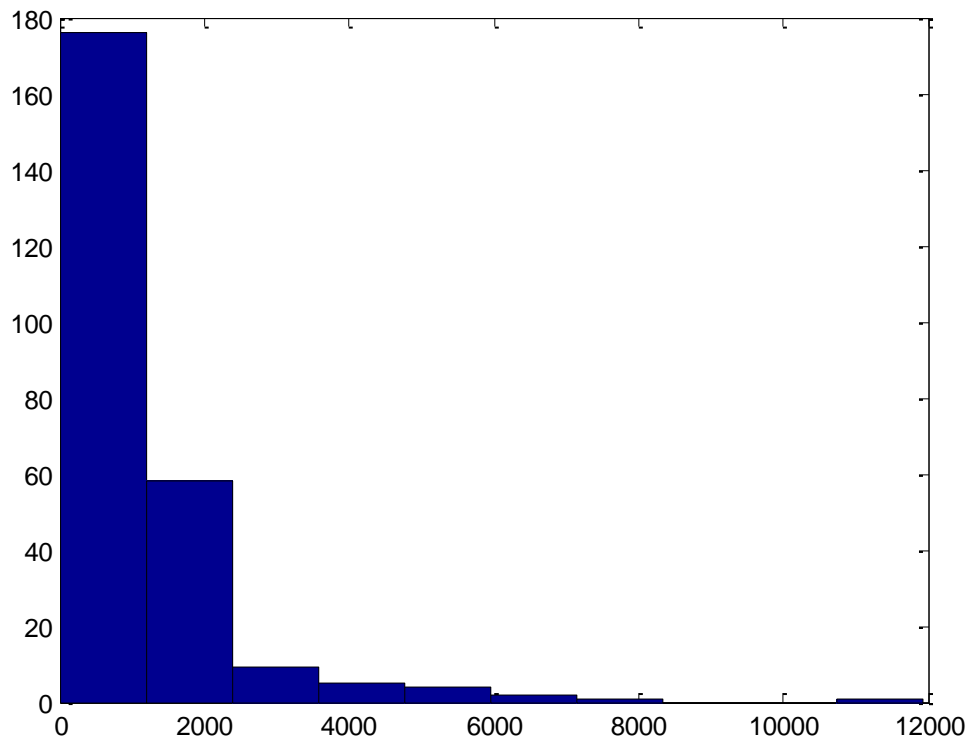
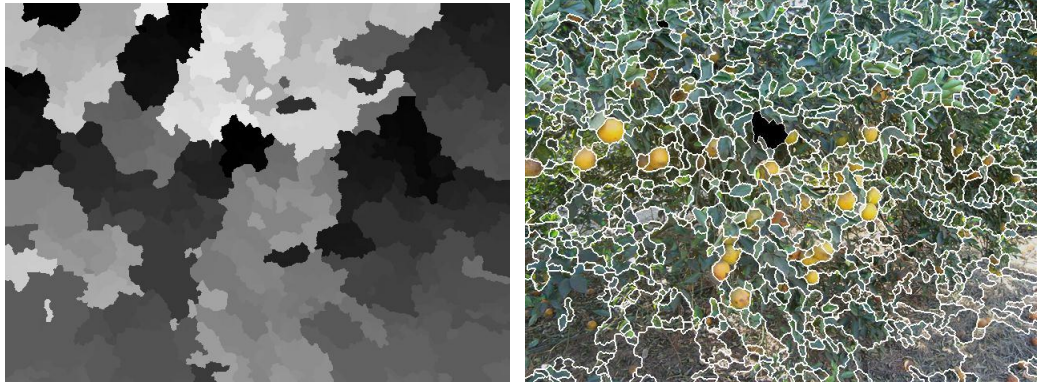
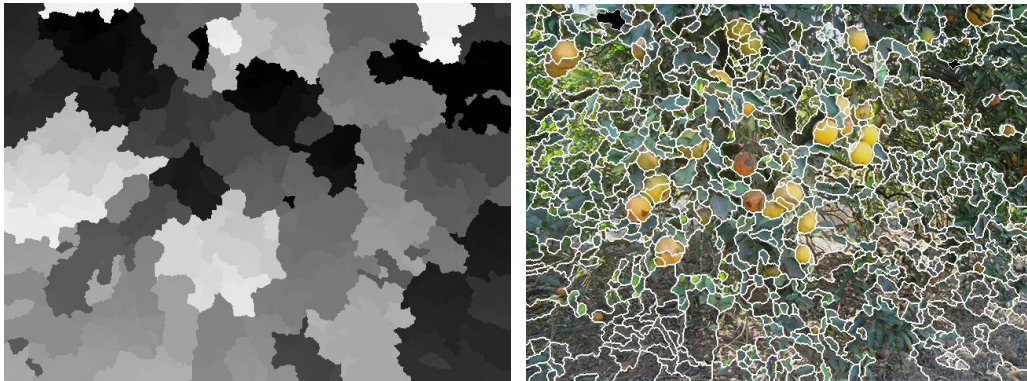


Figura 3.11: Histograma imagem quantizada 02 RGB.



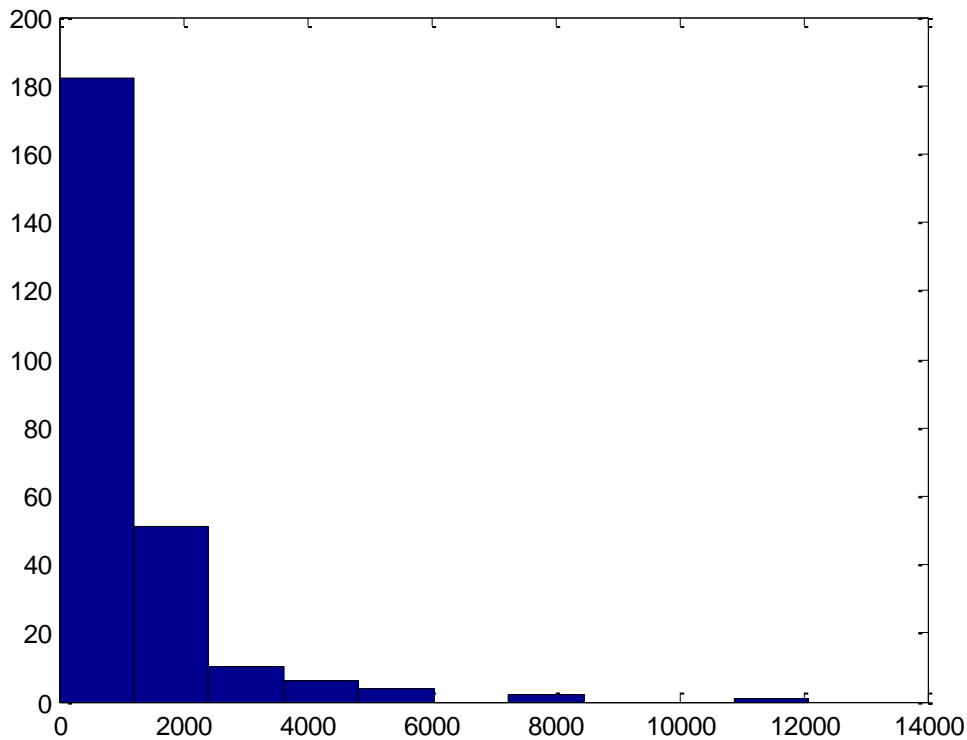


Figura 3.12: Histograma imagem quantizada 03 RGB.

Os histogramas das Figuras 3.10, 3.11 e 3.12 acima, foram obtidos no espaço de cor RGB, quantizado pelas regiões não homogêneas durante a primeira etapa de segmentação do algoritmo JSEG. Estes histogramas ilustram a representação gráfica em barras verticais da intensidade dos segmentos com mesma equivalência de *pixels* nas regiões quantizadas, ou seja, a quantidade de dimensões não normalizadas é dada pelo eixo das abscissas, enquanto a intensidade do nível de cinza em cada região é dada pelo eixo das ordenadas.

Na Figura 3.10, observa-se que a equivalência de *pixels* é maior para a tonalidade entre alaranjado e amarelo (maior intensidade de nível de cinza), que caracterizam as laranjas, e em menor grau, os demais itens da cena, como folhas e galhos, devido a não homogeneidade entre estas regiões. Quanto menor o número de dimensões na representação das regiões, menores são as chances de redundância de dados para que estes sejam normalizados pela ACP.

Para o caso das Figuras 3.11 e 3.12, as imagens adquiridas atendem a maior parte da árvore até o solo, permitindo que durante a etapa de quantização de cores, haja mais regiões não homogêneas. Para as regiões de baixa densidade de *pixels*, na vegetação rasteira e nos picos das folhas, há maior intensidade de nível de cinza, identificando homogeneidade acentuada nestas regiões. Porém, nas regiões onde há maior disparidade de tonalidade entre cores e texturas, conseqüentemente, o número de regiões quantizadas e segmentadas é maior, dada a sua não homogeneidade.

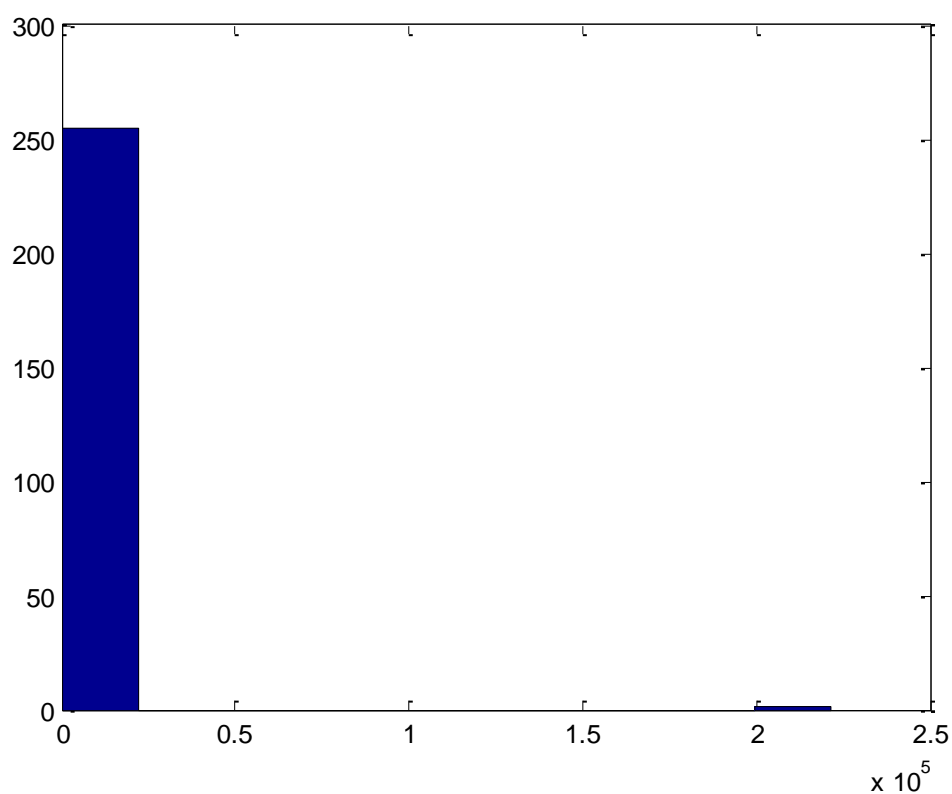
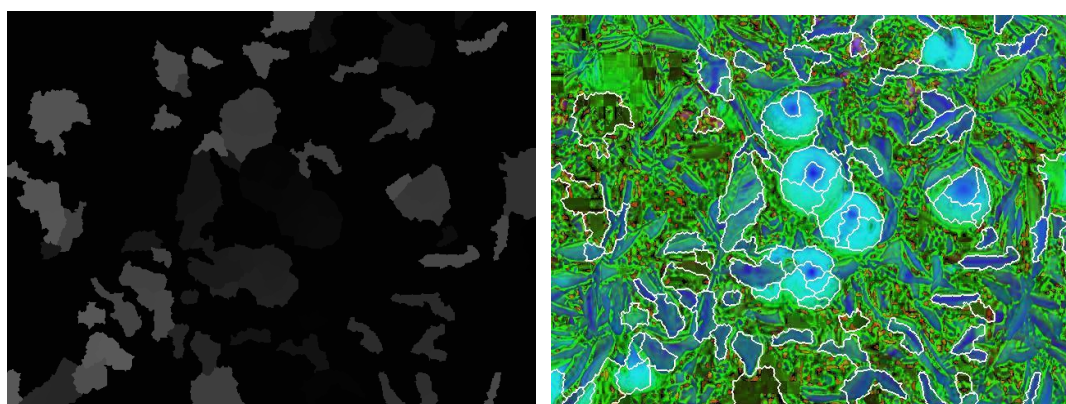


Figura 3.13: Histograma imagem quantizada 01 HSV.

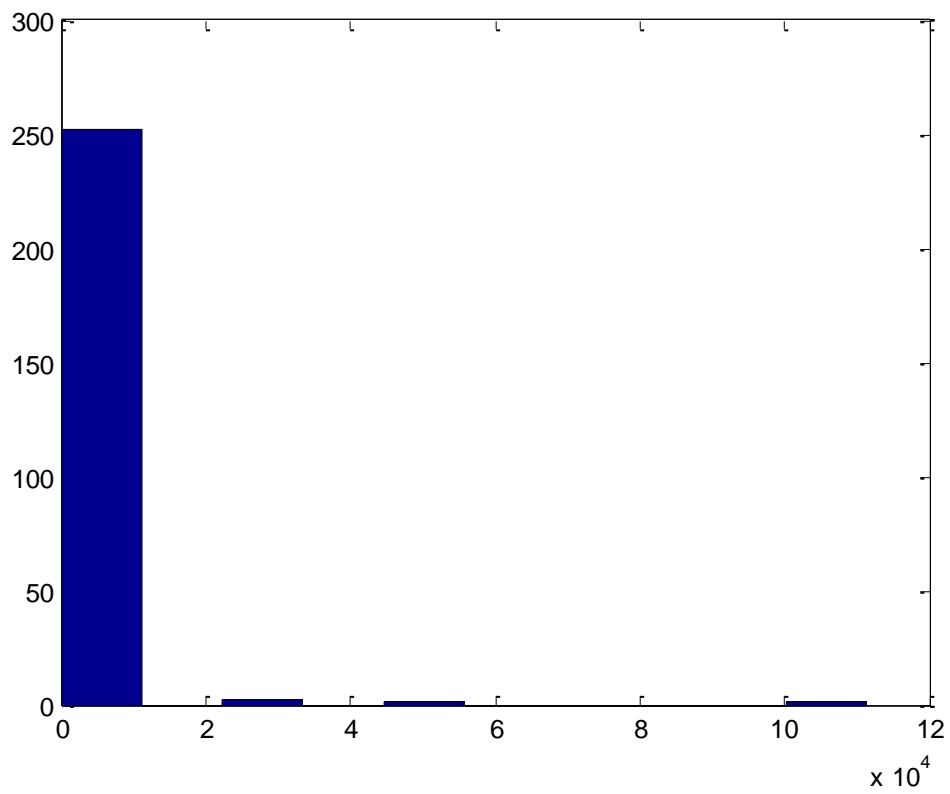
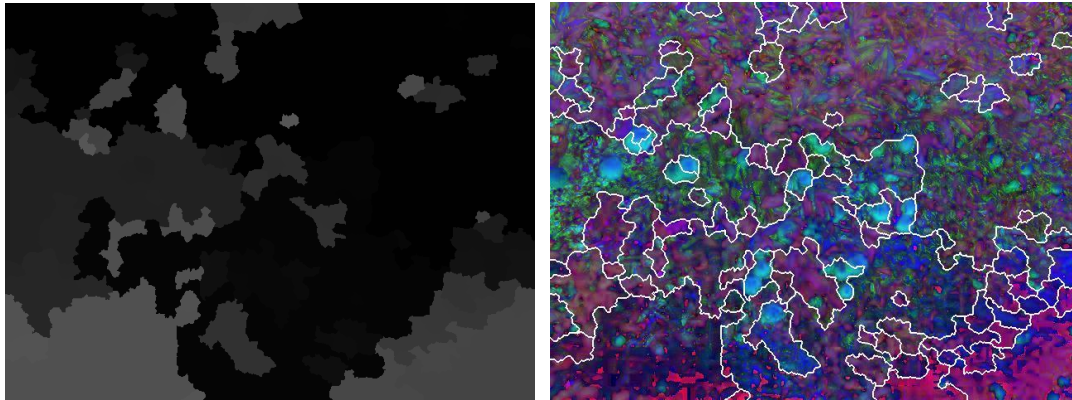
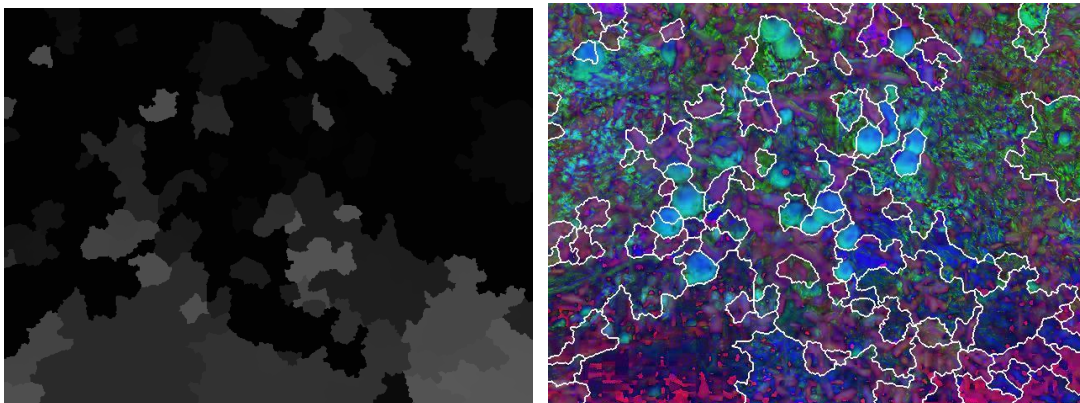


Figura 3.14: Histograma imagem quantizada 02 HSV.



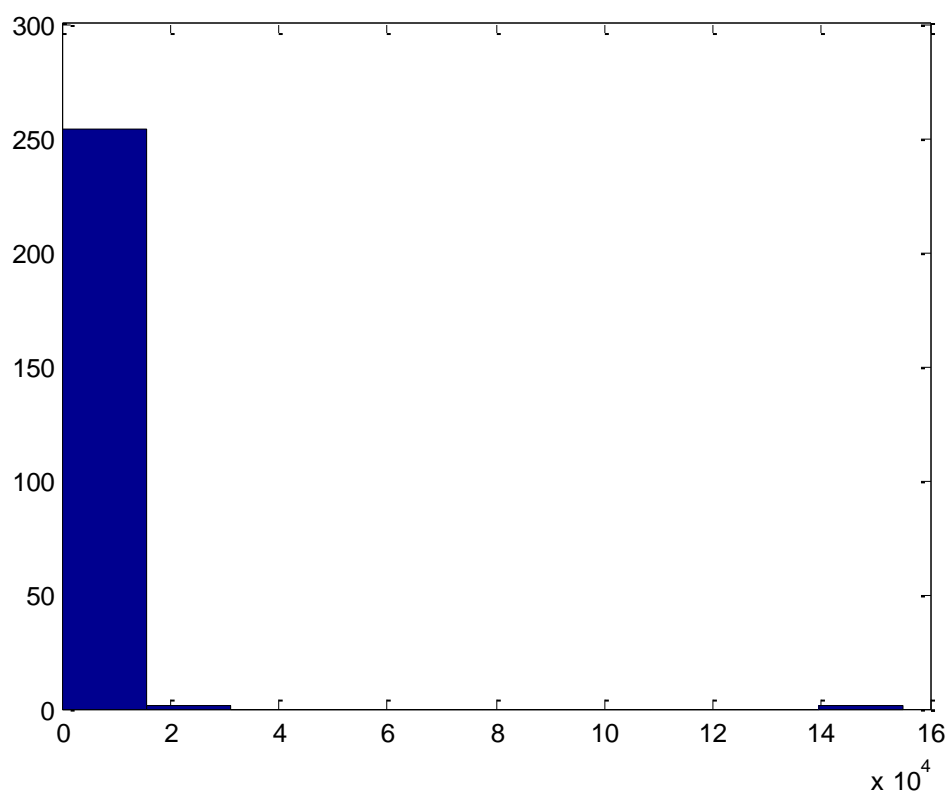


Figura 3.15: Histograma imagem quantizada 03 HSV.

As Figuras 3.13, 3.14 e 3.15 apresentam os histogramas nas cenas reamostradas no espaço de cor HSV.

Na Figura 3.13, da mesma forma que para o caso RGB, há maior incidência de equivalência de *pixels* entre os segmentos correspondentes aos maiores níveis de cinza da quantização (maior barra – menor quantidade de dimensões para maior nível de cinza), logo, as laranjas são agrupadas em regiões adjacentes, tornando-as homogêneas, e distintas das demais regiões adjacentes, que contemplam as folhas, ramos e galhos.

Nas Figuras 3.14 e 3.15, o maior índice de nível de cinza em ambas as imagens denotam regiões homogêneas agrupadas que destacam as possíveis regiões dos frutos, e para quantidades maiores de dimensões, o restante da cena, apresentando regiões não homogêneas.

Para estas regiões nas imagens acima, a relação da quantidade de dimensões e a intensidade do nível de cinza apresentam valores não paramétricos, considerando regiões de

contorno, adjacentes e pigmentos de textura distinta no restante da cena. Como trabalhar com valores altos e redundantes é inviável pelo tempo de processamento e resposta, e no armazenamento de dados expansivo, é necessário aplicar a Análise das Componentes Principais, com a finalidade de comprimir os dados mais significativos e eliminar a repetitividade destes, normalizando os vetores de características que são usados como dados de entrada nos classificadores propostos.

As duas abordagens são detalhadas através das tabelas 3.1, 3.2, 3.3, 3.4, 3.5 e 3.6, 3.7, 3.8, 3.9 e 3.10, a seguir.

A distribuição dos vetores de características às classes de pomares no espaço de cores RGB atende a aproximações entre as componentes das classes: **Área navegável** e **Área não navegável**, para os conjuntos entre 10% a 50%. Os vetores de características são maioria na segunda classe para todas as percentagens, já que as imagens analisadas nesta abordagem apresentam mais regiões não navegáveis, como as linhas de plantio nas laterais e o céu na parte superior (Tabela 3.1).

Tabela 3.1: Distribuição dos vetores de características às classes de Pomares em RGB.

Pomares (RGB)			
%	Área navegável (solo)	Área não navegável (céu e plantação)	Total
10	3768	4128	7896
20	3876	4024	7900
30	3900	4054	7954
40	3923	4067	7990
50	3941	4154	8095
75	4001	4244	8245
100	4056	4401	8457

Da mesma forma para o caso dos pomares em RGB (Tabela 3.2), os vetores de características em HSV denotam maioria na classe **Área não navegável** devido ao mesmo comportamento acima, mas para as percentagens 75% e 100%, a diferença entre a quantidade total de vetores em relação às percentagens menores, aumenta, dada a concentração dos

vetores quando a região de interesse da primeira classe (**Área navegável**) é limiarizada nos canais H, S e V deste espaço de cores.

Tabela 3.2: Distribuição dos vetores de características às classes de Pomares em HSV.

Pomares (HSV)			
%	Área navegável (solo)	Área não navegável (céu e plantação)	Total
10	3770	4132	7902
20	3789	4126	7915
30	3800	4154	7954
40	3832	4197	8029
50	3845	4210	8055
75	3890	4249	8139
100	3916	4414	8330

Tabela 3.3: Distribuição dos vetores de características às classes de Canaviais em RGB.

Canaviais (RGB)			
%	Área navegável (solo)	Área não navegável (céu e plantação)	Total
10	3890	4007	7897
20	3903	4031	7934
30	3954	4000	7954
40	3967	3999	7966
50	3991	4042	8033
75	4099	4024	8123
100	4134	4066	8200

Tabela 3.4: Distribuição dos vetores de características às classes de Canaviais em HSV.

Canaviais (HSV)			
%	Área navegável (solo)	Área não navegável (céu e plantação)	Total
10	3888	3912	7800
20	3903	3922	7825
30	3967	3944	7911
40	4087	3856	7943
50	4099	3920	8019
75	4113	4031	8144
100	4201	4139	8340

Nas tabelas 3.3 e 3.4, acima, em cenas de canaviais, a relação dos vetores de características para a primeira classe apresenta maior quantidade nas percentagens 40%, 50%, 75% e 100%, pois conforme o conjunto de vetores aumenta, seja nos espaços de cores RGB e HSV, a concentração das regiões para a classe **Área não navegável** torna-se homogênea,

dada a densidade das linhas de plantio dos colmos. Dessa forma, a classificação é facilitada, com taxas de acerto próximas de 1,0.

Nas tabelas 3.5 e 3.6 os vetores de características para a classe **Área navegável**, denotam quantidade superior no caso HSV em relação ao RGB, sendo que a homogeneidade das regiões à navegação mostra aumento gradativo neste espaço de cores, conforme as percentagens dos conjuntos também aumenta. Por se tratar de milharais, esta cultura gramínea atrai plantas invasoras rasteiras, principalmente entre as linhas de plantio, dificultando a extração de características normalizadas da primeira classe para o caso RGB, considerando a aplicação da técnica de segmentação para regiões homogêneas.

Tabela 3.5: Distribuição dos vetores de características às classes de Milharais em RGB.

Milharais (RGB)			
%	Área navegável (solo)	Área não navegável (céu e plantação)	Total
10	3222	4123	7345
20	3529	3971	7500
30	3645	3922	7567
40	3699	3911	7610
50	3830	3970	7800
75	4180	3715	7895
100	4234	3769	8003

Tabela 3.6: Distribuição dos vetores de características às classes de Milharais em HSV.

Milharais (HSV)			
%	Área navegável (solo)	Área não navegável (céu e plantação)	Total
10	3845	3455	7300
20	3901	3420	7321
30	3942	3455	7397
40	3994	3413	7407
50	4188	3232	7420
75	4234	3267	7501
100	4329	3449	7778

A distribuição dos vetores de características nos dois casos RGB e HSV para a localização de frutos (tabelas 3.7 e 3.8) mostra quantidades superiores para a classe **Folha, ramos e galhos**, sendo que na análise destas cenas naturais, os objetos associados a esta classe são dispostos em toda a imagem, e a tarefa de associar frutos entre todos os objetos da cena torna-

se complexa, à medida que a quantidade de elementos dos conjuntos aumenta. No caso RGB, para os conjuntos 50%, 75% e 100% a quantidade de vetores relacionada com a classe **Frutos** é superior que os correspondentes no caso HSV, dada a não homogeneidade de elementos de cores compostas neste espaço de cores.

Observa-se que para o caso HSV, a quantidade de dimensões para a representação dos frutos é menor que a mesma relação para o caso RGB, e a quantidade de dimensões para folhas, ramos e galhos em HSV é maior que no caso RGB, o que proporciona ao classificador, maior robustez durante a associação das classes de padrões com os segmentos alvos.

Tabela 3.7: Distribuição dos vetores de características às classes na localização de frutos em RGB.

Localização de frutos (RGB)				
%	Frutos	Folhas, ramos e galhos	Céu	Total
10	1873	4057	309	6239
20	1903	4383	357	6643
30	1933	4383	377	6693
40	1988	4393	402	6783
50	2023	4414	408	6845
75	2094	4554	465	7113
100	2100	4737	489	7326

Tabela 3.8: Distribuição dos vetores de características às classes na localização de frutos em HSV.

Localização de frutos (HSV)				
%	Frutos	Folhas, ramos e galhos	Céu	Total
10	1359	5551	290	7200
20	1400	5528	306	7234
30	1422	5564	355	7341
40	1459	5545	384	7388
50	1468	5579	403	7450
75	1493	5586	452	7531
100	1577	5976	491	8044

Da mesma forma que na localização de frutos, o mesmo caso se aplica para o caso de reconhecimento de gramíneas (tabelas 3.9 e 3.10), cuja distribuição de vetores na classe **Folhas e ramos** é majoritária das demais classes **Milhos** e **Solo ou Céu**. Para o caso HSV, as formas dos colmos são equivalentes para as cores e texturas das espigas, diminuindo assim, a normalização das características da classe **Milhos**, ou seja, as espigas de milho não são própria e devidamente segmentadas. Como a classificação de gramíneas não é objetivo

específico deste trabalho, a extração de características desta cultura para a localização atende apenas como requisito de avaliação e propostas de outros algoritmos de segmentação, classificação e pós-processamento.

Para ambos os espaços de cores, a amostragem dos vetores de características normalizados é feita de forma generalizada para os períodos de milho verde e silagem. Neste último período, a definição do colmo da gramínea e a textura da espiga com os demais objetos da cena torna a associação das classes de padrões com os segmentos alvos menos robusta e suscetível a erros de delimitação.

Tabela 3.9: Distribuição dos vetores de características às classes na localização de milhos em RGB.

Localização de milhos (RGB)				
%	Milhos	Folhas e ramos	Solo ou Céu	Total
10	1234	5688	301	7223
20	1244	5735	344	7323
30	1290	5704	398	7392
40	1324	5656	450	7430
50	1388	5594	476	7458
75	2450	4601	514	7565
100	2654	4691	532	7877

Tabela 3.10: Distribuição dos vetores de características às classes na localização de milhos em HSV.

Localização de milhos (HSV)				
%	Milhos	Folhas e ramos	Solo ou Céu	Total
10	1309	5098	501	6908
20	1345	5144	532	7021
30	1388	5073	599	7060
40	1402	5100	627	7129
50	1434	5094	648	7176
75	1589	4978	692	7259
100	1673	5294	730	7697

Após a normalização pela ACP nos conjuntos de treinamento, e sua devida classificação, os conjuntos de teste também foram normalizados pela mediana das dimensões para que a validação apresentasse dados paramétricos de mesma unidade, através da média e do desvio padrão das distâncias.

Em todas as tabelas mostradas anteriormente, é notável que o número de vetores de características cresça com baixa razão entre 10%, 20%, 30%, 40% e 50%, e entre 75% e 100% as taxas de crescimento são maiores. Logo, as taxas de acerto serão semelhantes para os dois últimos conjuntos e também semelhantes entre os cinco primeiros conjuntos.

3.2.4. Classificadores cognitivos e estatísticos

Para todos os conjuntos de treinamento, os vetores de características normalizados são submetidos à classificação com, no máximo, 20 dimensões, possibilitando a exatidão do número de segmentos classificados às regiões de interesse.

São estudados e avaliados os classificadores cognitivos, baseados em redes neurais artificiais e lógica nebulosa preditiva; e, classificadores estatísticos, baseados no teorema de Bayes e Bayes ingênuo.

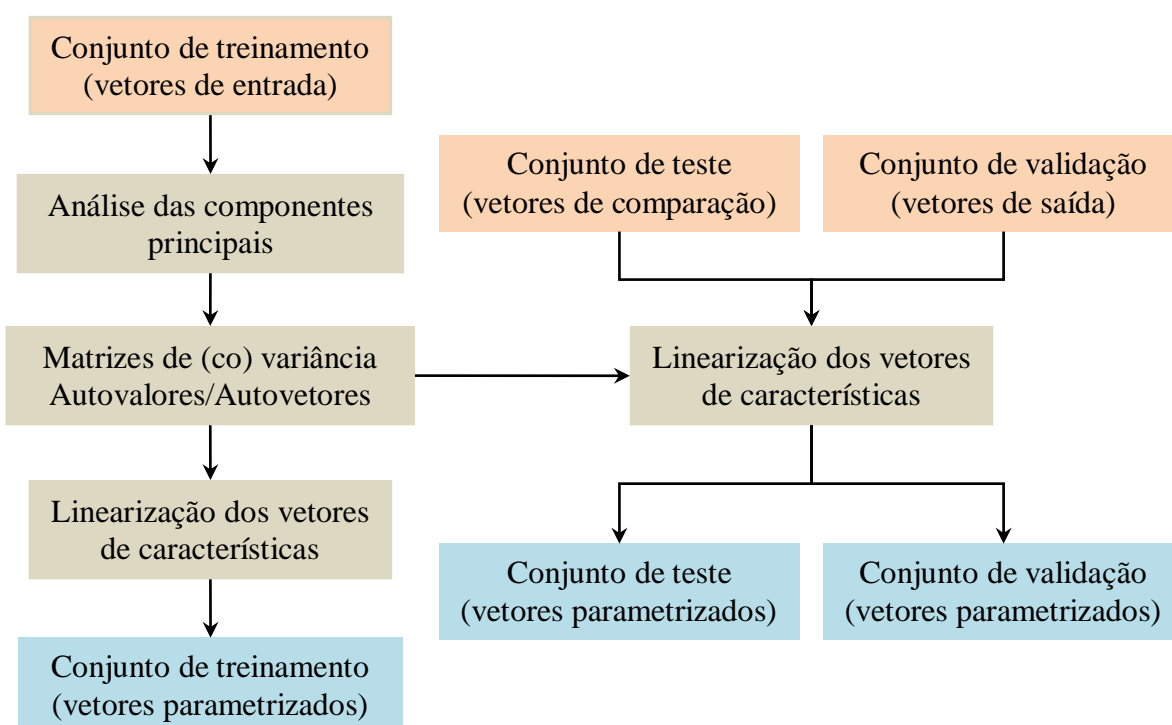


Figura 3.16: Etapas de extração e normalização de características pela ACP nos conjuntos de treinamento, teste e validação.

A Figura 3.16, acima, o processo de linearização dos vetores de características para o treinamento, teste e validação. Inicialmente, com os dados do conjunto de treinamento, são

extraídas as características das cores nos espaços de cores RGB e HSV, através dos histogramas das imagens. Logo, para eliminar redundâncias e comprimir os dados foi aplicada a Análise das Componentes Principais (ACP), gerando assim uma linearização dos vetores de características e, conseqüentemente, a equivalência de matrizes de dados com os conjuntos de teste e validação.

a) Redes neurais artificiais:

Os vetores de características para todas as percentagens de distância aos centróides das classes, e em cada espaço de cor são classificados em topologia adequadas de redes neurais MLP customizadas com o *iRPROP*.

Cada topologia da primeira abordagem (locomoção) apresenta uma camada de entrada contendo n dimensões relativa aos vetores de características, camada escondida de 20 neurônios, e camada de saída com dois neurônios (área navegável e área não navegável), direcionadas a três linhas de plantio diferentes, no espaço de cores RGB e HSV, totalizando 840 tipos de RNA. Já na segunda abordagem (localização de frutos), a mesma estrutura de topologia é empregada, porém com três neurônios na camada de saída, totalizando 210 tipos de RNA.

As topologias foram treinadas cinco vezes, para o número máximo de iterações (épocas) em 100.00 ou o erro quadrático médio for inferior a 0,0001. A ponderação sináptica foi iniciada com valores aleatórios, e os padrões de treinamento foram aplicados arbitrariamente. A função de ativação é tangente hiperbólica variando entre 0,7 e 0,9. A programação das topologias foi feita através da *toolbox* de redes neurais artificiais (ANN) do MATLAB.

A *toolbox* de RNA do MATLAB possui uma biblioteca dedicada ao reconhecimento de padrões. Nela, os dados de entrada e testes são relacionados como matrizes de características. Cada linha desta matriz corresponde aos elementos ou características e cada coluna corresponde às amostras ou dimensões, definida pelos cálculos da ACP, também projetada em

MATLAB. A percentagem entre os conjuntos de amostras para o treinamento, testes e validação é dada respectivamente por: 70%, 15% e 15%.

A topologia da RNA para a abordagem de locomoção é ilustrada na Figura 3.17, abaixo.

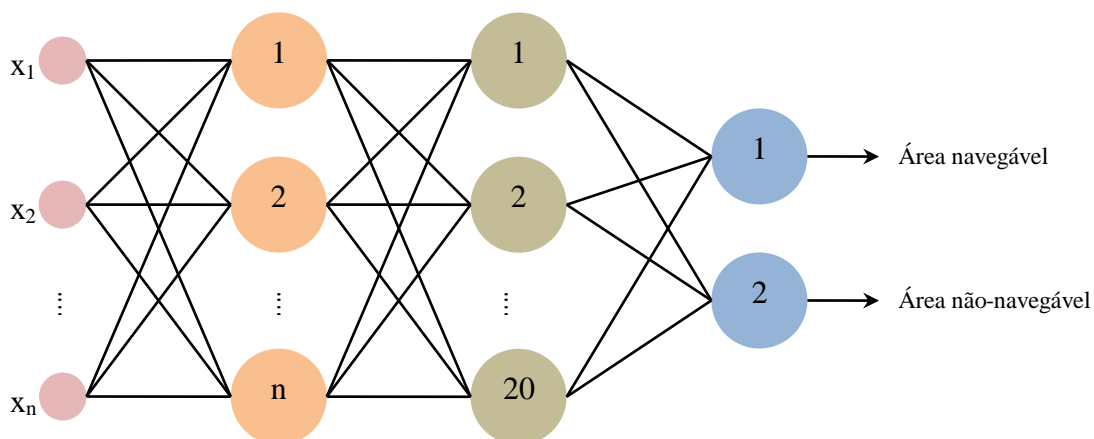


Figura 3.17: Topologia RNA à primeira abordagem.

A topologia da RNA para a abordagem de localização de laranjas e milhos é ilustrada na Figura 3.18, abaixo.

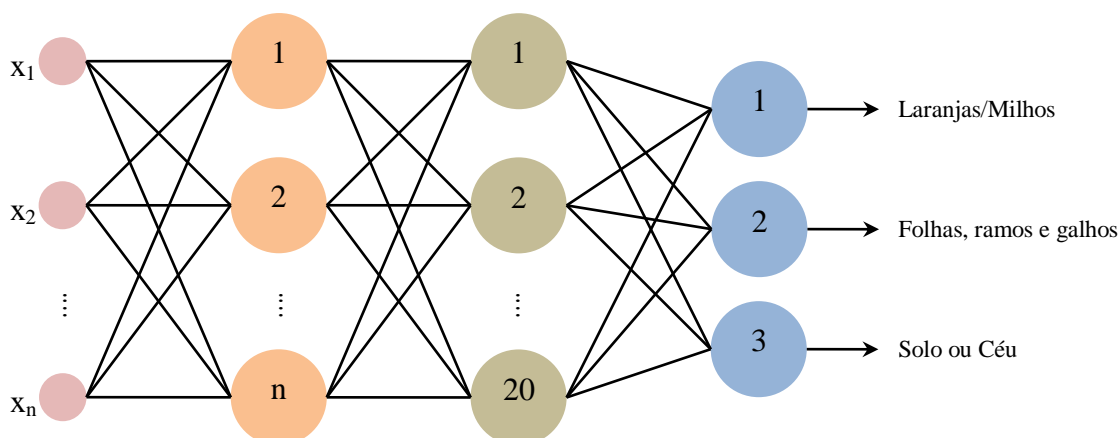


Figura 3.18: Topologia RNA à segunda abordagem.

As entradas x_n correspondem ao número de dimensões dos vetores de características após a normalização, que cada segmentação apresenta a quantização de cores em RGB e HSV.

b) Lógica nebulosa preditiva:

O controlador *fuzzy* preditivo manipula as variáveis de entrada (*fuzzyficação*), para o escalonamento da normalização dos vetores de características, resultando em variáveis

linguísticas dentro de um intervalo que correspondem a padrões desejados nas imagens analisadas. Com isso, tais variáveis obedecem a uma base de regras para a adequação dos padrões em um processo de inferência preditivo, que após analisados, são classificados de acordo com as abordagens trabalhadas.

Na primeira abordagem, os padrões podem ser classificados em área navegável ou solo, área de plantio ou céu. Na segunda abordagem, os padrões são classificados em frutos, gramíneas, ramos, galhos e folhas, conjuntos de frutos, conjuntos de milhos, céu, solo, vegetação rasteira. Estes padrões são então, *defuzzyficados*, tornando as variáveis linguísticas em números representativos dos *pixels* analisados na imagem. Nas duas abordagens o método de *defuzzyficação* aplicado é o centróide (CDA) que retorna o centro de área dos conjuntos difusos de saída.

A aplicação do controle *fuzzy* foi feita através da *toolbox* de lógica difusa no MATLAB (*Anfis*). As relações de entrada, conjunto de regras, funções-membro e processos de inferências são descritas a seguir. Os gráficos que relacionam as funções membro e seus respectivos estados de inferência ao grau de pertinência são vistos no **APÊNDICE A**.

Como as imagens analisadas são adquiridas e pré-processadas nos espaços de cores RGB e HSV, os vetores de características destas imagens correspondem aos valores normalizados dos histogramas aplicados após a quantização de espaço de cor durante a segmentação. Logo, as entradas e saídas de ambas as abordagens são analisadas mediante tais condições.

Primeira abordagem (locomoção do veículo):

As entradas definidas são:

- Reta diagonal direita lateral (*RDDL*) – imagem lateral direita
- Reta diagonal esquerda lateral (*RDEL*) – imagem lateral esquerda
- Reta diagonal direita frontal (*RDDF*) – imagem frontal direita
- Reta diagonal esquerda frontal (*RDEF*) – imagem frontal esquerda

As saídas de interesse são:

- Área navegável (AN)
- Área não navegável (AnN)

A Figura 3.19, a seguir, ilustra a estrutura das entradas e saídas desejadas, ao modelo *fuzzy* preditivo, sendo o sistema de inferência configurado como *mamdani*, uma vez que a expectativa da classificação nas saídas desejadas seja atribuída em conjuntos difusos de *pixels* durante a varredura da imagem, e não apenas a característica de determinado *pixel*.

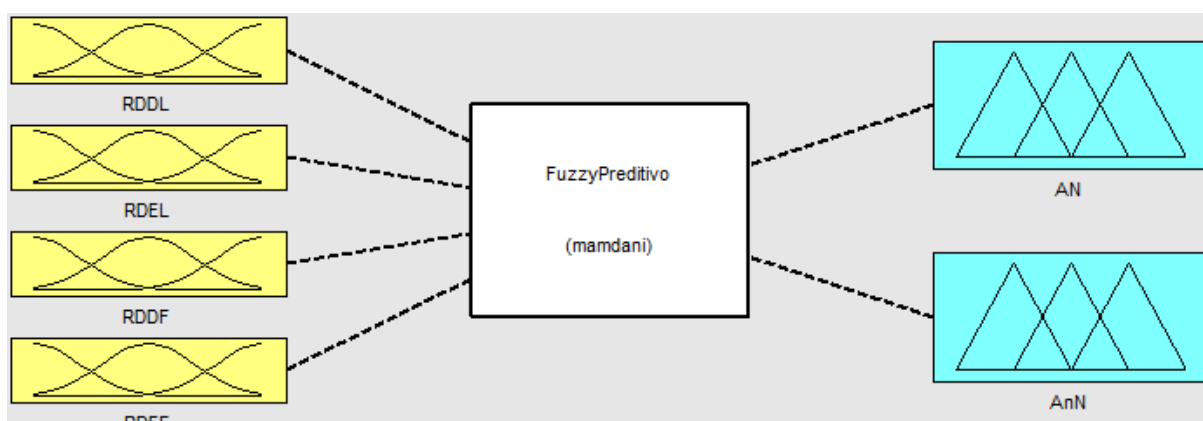


Figura 3.19: Sistema de inferência *fuzzy* para a primeira abordagem.

As funções-membro para cada valor de entrada são definidas pelo modo de aquisição das imagens, sendo *RDDL* e *RDEL* para imagens laterais (direita e esquerda), respectivamente, e *RDDF* e *RDEF* para imagens frontais com os ângulos de guiagem à direita e esquerda.

Logo, nas duas primeiras entradas (*RDDL* e *RDEL*) são atribuídas as funções-membro:

- *acutângulo*: varia em estado de inferência 0 à 0,3 em padrão linear constante, para grau de pertinência máximo (trapezoidal), no caso do ângulo de guiagem ser predominantemente agudo para a determinação da área navegável (AN), e de 0,3 à 0,5 em padrão linear decrescente;

- *retângulo*: grau de pertinência máximo (triangular) em inferência 0,5, com padrão linear crescente de 0,1 até 0,5, e padrão linear decrescente de 0,5 até 0,9. Este estado denota baixa incidência de imagens que possuem ângulos de guiagem retos.

- *obtusângulo*: varia em estado de inferência 0,5 à 0,9 em padrão linear crescente, e 0,9 à 1 em padrão linear constante, ao máximo grau de pertinência (trapezoidal). Aquisições de imagens laterais direitas e esquerdas apresentam mínimos ângulos de guiagem obtusos, porém existentes em desníveis de plantios maciços.

Em cenas de aquisição frontal, as duas entradas *RDDF* e *RDEF* são inferidas em funções-membro com estados de inferência geométricos dados os lados de um triângulo, que denota a área navegável. Como os ângulos de guiagem e desvio lateral são irregulares, três são as funções-membro para este caso:

- *escaleno*: varia em estado de inferência 0 à 0,4 em padrão linear decrescente, para máximo grau de pertinência (triangular), uma vez que este estado é minoritário geometricamente.

- *equilátero*: varia em estado de inferência 0,6 à 1 em padrão linear crescente, para máximo grau de pertinência (triangular), uma vez que este estado é minoritário geometricamente.

- *isósceles*: varia de 0,2 à 0,3 em padrão linear crescente e 0,7 à 0,8 em padrão linear decrescente. O estado de inferência 0,3 à 0,7 em máximo grau de pertinência (trapezoidal) denota estado majoritário na forma geométrica das linhas de plantio com mesmo lado, delimitando a área navegável (*AN*) e a área excluída da inferência, como a área não navegável (*AnN*).

As saídas *AN* e *AnN* são aplicáveis a todas as entradas, para aquisições de cenas frontais e laterais. Na primeira, apenas duas funções-membro foram definidas, considerando a textura porosa e a tonalidade das cores em RGB amarelo e marrom, denotando a variedade de cores do solo e vegetação rasteira entre estas duas cores. Logo, a função-membro *tex_p/amarelo_C* possui estado de inferência 0 à 0,4 em padrão linear constante ao máximo grau de pertinência (trapezoidal), e inferência 0,4 à 0,6 em padrão linear decrescente. Da mesma forma, a função-

membro *tex_p/marrom_C* possui estado de inferência 0,4 à 0,6 em padrão linear crescente, e 0,6 à 1 em padrão linear constante ao máximo grau de pertinência (trapezoidal).

A saída *AnN* apresenta três funções-membro apenas para representar a localização das plantações, em esquerda, direita e horizonte, restringindo o local de área não navegável. O grau de pertinência máximo (trapezoidal) às funções-membro *plantação_E*, *plantação_H* e *plantação_D* são proporcionais à representação das imagens e seus estados de inferência variam em taxas similares com padrões lineares crescente, decrescente e constante.

Segunda abordagem (localização de frutos e gramíneas):

As entradas definidas são:

- Círculo área 100 (*CA100*) – círculos com área de 100 *pixels*
- Círculo área 200 (*CA200*) – círculos com área de 200 *pixels*
- Arcos adjacentes 300 (*AA300*) – conjuntos de círculos iguais ou maiores que 300 *pixels*
- Área irregular média (*AIM*) – regiões irregulares que denotam folhas, galhos e ramos
- Área irregular baixa (*AIB*) – regiões irregulares que denotam o solo e vegetação rasteira
- Losango área 300 (*LA300*) – áreas em losango que denotam a espiga de milho
- Losango área 100 (*LA100*) – áreas em losango que denotam partes das espigas de milho

As saídas de interesse são:

- Fruto (laranja) (*FL*)
- Folhas, ramos e galhos (*FRG*)
- Solo e vegetação rasteira (*SVR*)
- Conjunto de frutos (*CF*)

- Gramínea (milho) (*GM*)
- Folhas e caule (*FC*)
- Conjunto de gramíneas (*CG*)

A estrutura das entradas e saídas na segunda abordagem é mostrada na Figura 3.20, a seguir.

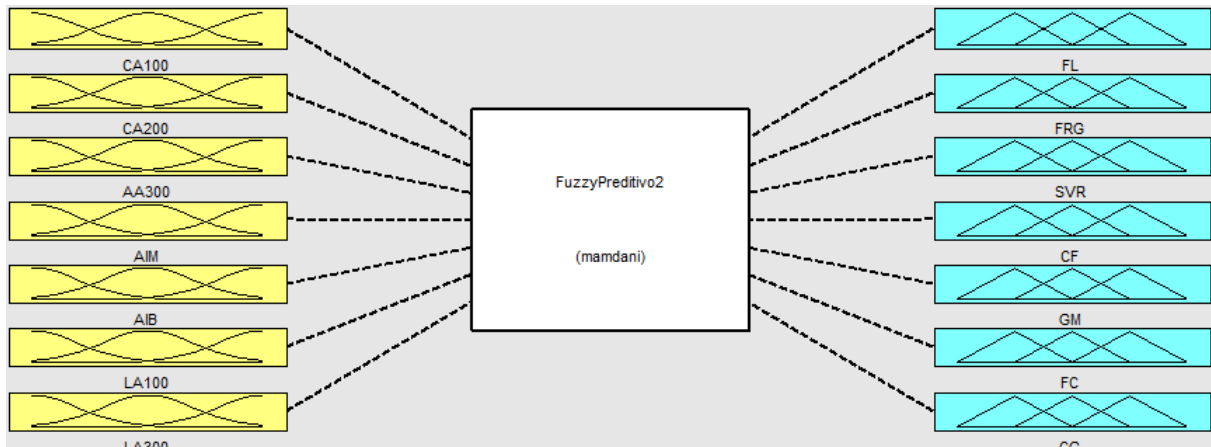


Figura 3.20: Sistema de inferência *fuzzy* para a segunda abordagem.

As entradas *CA100*, *CA300*, *AA300* são aplicadas à localização de frutos. As entradas *LA100* e *LA300* são aplicadas à localização de gramíneas, e as entradas *AIM* e *AIB* são aplicadas a ambas.

No problema de localização de frutos, as funções-membro para cada valor de entrada são definidas em seus respectivos limiares de ação (Figura 3.20). As entradas *CA100*, *CA200* e *AA300* variam o estado de inferência entre o padrão linear constante para a detecção percentual do fruto em semi-arco de 0 à 0,1 ao grau de pertinência máximo (trapezoidal), e 0,1 à 0,5 em padrão linear decrescente. A função-membro *arco100* que referencia toda a circunferência do fruto apresenta estado de inferência 0,5 em grau de pertinência máximo (triangular). A função-membro *arco75* apresenta estado de inferência 0,5 à 0,6 em padrão linear crescente, e 0,6 à 1 em padrão linear constante (trapezoidal), que simboliza a maior parte do fruto, o que majoritariamente ocorre nas cenas de pomares localizados. Com isso, arcos de circunferência detectáveis nas imagens segmentadas com áreas correspondentes até

100 e 200 *pixels* são relacionadas como possíveis candidatos a frutos – dada a distância focal das cenas adquiridas, bem como conjuntos de frutos próximos na mesma região segmentada.

No problema de localização de gramíneas, a entrada *LA100* infere de 0 à 0,4 em padrão linear decrescente à função-membro *triângulo_S* (triângulo superior), de 0,6 à 1 em padrão linear crescente à função-membro *triângulo_I* (triângulo inferior), e de 0,4 à 0,6 em padrão linear constante ao grau de pertinência máximo (trapezoidal) para a função-membro *triângulo_{ED}* (adjacências entre triângulos direito e esquerdo). Esta entrada é direcionada à formações iniciais de espigas de milho, sem a projeção ápice do estigma. Por sua vez, a entrada *LA300* infere aos mesmos índices crescente e decrescente, respectivamente às funções-membro *losango_A* (área da espiga e tonalidades em verde) e *losango_V* (área da espiga e tonalidades em amarelo). Como o cume do estigma é variável não apenas pela textura, como pela cor e pelo tamanho, a função-membro *estigma* relaciona o estado de inferência não linear de 0,2 à 0,8, sendo 0,5 o grau de pertinência máximo, em função gaussiana.

A entrada *AIM* possui as seguintes funções-membro:

- *polígono_M*: regiões que denotam galhos e ramos em tonalidades da cor marrom. Apresenta estado de inferência 0 à 0,4 em padrão linear decrescente (grau de pertinência triangular);

- *polígono_V*: regiões que denotam folhas em tonalidades da cor verde. Apresenta estado de inferência 0,6 à 1 em padrão linear crescente (grau de pertinência triangular);

- *regiões_P*: regiões segmentadas tanto para folhas, galhos e ramos entre os pomares. Apresenta estado de inferência 0,3 à 0,7 em padrão linear constante para o grau de pertinência máximo (trapezoidal);

A entrada *AIB* possui as seguintes funções-membro:

- *marrom100* e *marrom200*: solo e vegetação rasteira em níveis de 100 e 200 *pixels* para as tonalidades da cor marrom. Apresenta estados de inferência 0 à 0,1 e 0,9 à 1 em padrão linear constante ao grau de pertinência máximo (trapezoidal);

- *regiões_M*: todas as regiões adjacentes e conjuntas com áreas em 100 e 200 *pixels* para as tonalidades da cor marrom. Apresenta estado de inferência 0,3 à 0,7 em padrão linear constante ao grau de pertinência máximo (trapezoidal). Estas regiões representam o solo e vegetação rasteira nos pés de pomares e milharais.

As saídas de interesse à localização de frutos são: *FL*, *CF*, *FRG*. Para a localização de gramíneas, as saídas são: *GM*, *CG*, *FC*. E para ambas as situações, a saída de interesse é *SVR*.

Na localização de frutos, as saídas *FL* e *CF*, que respectivamente representam a detecção de unidades de laranja e conjuntos de laranja, possuem as funções-membro *tex_F/alaranjado_E*, *tex_F/alaranjado_C* e *tex_F/alaranjado*. As três possuem grau de pertinência máximo trapezoidal para $\frac{1}{5}$ do estado de inferência total. As funções-membro são aplicadas para a classificação em textura fosca e tonalidades da cor alaranjado.

A saída *FRG* apresenta variedades na aquisição de cenas com folhas, ramos e galhos, destacando-os pela sua tonalidade nas cores verde e marrom para regiões esparsas ou em conjuntos. As funções-membro *esparsas_V* e *esparsas_M* possuem grau de pertinência máximo trapezoidal aos limites inferior e superior dos estados de inferência. A função-membro *conjuntos_{VM}* apresenta estado de inferência 0,4 à 0,6 em grau de pertinência máximo trapezoidal. Desta forma, as folhas, ramos e galhos espalhados pela cena são classificados em relação aos frutos.

Na localização de gramíneas, as saídas *GM* e *CG* possuem mesmo grau de pertinência máximo (trapezoidal) para os limites inferior e superior dos estados de inferência, respectivamente às funções-membro *tex_F/verde_C* (textura fosca e tonalidade verde claro) e *tex_F/verde_E* (textura fosca e tonalidade verde escuro), e função-membro *tex_F/amarelo_C* (textura

fosca e tonalidade amarelo claro) com estado de inferência máximo em 0,5 para grau de pertinência máximo triangular. Assim, as espigas de milho são classificadas pela textura fosca e tonalidades das cores verde e amarela.

A saída *FC* apresenta graus de pertinência máximos triangulares nos estados de inferência 0,1, 0,5 e 0,9 para as funções-membro $tex_F/cinza_C$, $tex_F/amarelo_C$ e $tex_F/marrom_C$, respectivamente, uma vez que as folhas e caule são classificados pela textura fosca e tonalidades das cores cinza, amarelo e marrom.

A saída *SVR* à ambos frutos e gramíneas, atende à duas funções-membro $texP/marromE$ (marrom escuro) e $texP/marromC$ (marrom claro), com estados de inferência 0 à 0,3 ao grau de pertinência máximo trapezoidal, e 0,4 á 0,9 ao grau de pertinência máximo trapezoidal, respectivamente. Logo, o solo e a vegetação rasteira são classificados mediante a textura porosa e as tonalidades de marrom escuro e claro.

Através do editor *Anfis* é realizada a adição das variáveis de entrada e saída, funções-membro, definidas como triangulares, trapezoidais e gaussianas, e a *defuzzyficação* pelo método do Centróide (centro de gravidade dos conjuntos *Fuzzy*) em ambas as abordagens.

Abaixo, a Tabela 3.11 ilustra para a primeira abordagem, a base de regras entre a condição lógica ‘E’ das variáveis de entrada (*RDDL*, *RDEL*, *RDDF* e *RDEF*), e suas respectivas funções-membro para cada associação, na extração das saídas desejadas (*AN* e *AnN*) para a classificação das classes **Área navegável** e **Área não navegável**.

Tabela 3.11: Base de regras para as entradas e saídas à primeira abordagem.

		RDDL				
		<i>acutângulo</i>	<i>retângulo</i>	<i>obtusângulo</i>		
RDEL	<i>obtusângulo</i>	<i>AN</i> $tex_P/amarelo_C$	<i>AnN</i> $plantação_E$	<i>AnN</i> $plantação_H$	<i>isósceles</i>	
	<i>retângulo</i>	<i>AN</i> $tex_P/amarelo_C$	<i>AN</i> $tex_P/marrom_C$	<i>AnN</i> $plantação_D$	<i>equilátero</i>	
	<i>acutângulo</i>	<i>AnN</i> $plantação_H$	<i>AN</i> $tex_P/marrom_C$	<i>AN</i> $tex_P/marrom_C$	<i>escaleno</i>	
		<i>escaleno</i>	<i>equilátero</i>	<i>isósceles</i>	RDEF	

As tabelas Tabela 3.12 e Tabela 3.13, a seguir, apresentam as bases de regras para a segunda abordagem.

Tabela 3.12: Base de regras para as entradas e saídas à segunda abordagem (A).

		CA100/CA300				
		arco25	arco75	arco100		
LA100	triângulo _{ED}	FL <i>(tex_F/alaranjado_C)</i>	FL <i>(tex_F/alaranjado_C)</i>	FL <i>(tex_F/alaranjado_C)</i>	arco100	AA300
		CF <i>(tex_F/alaranjado_C)</i>	CF <i>(tex_F/alaranjado_C)</i>	CF <i>(tex_F/alaranjado_C)</i>		
		GM <i>tex_F/verde_C</i>	GM <i>tex_F/verde_C</i>	GM <i>tex_F/verde_C</i>		
		CG <i>tex_F/verde_C</i>	CG <i>tex_F/verde_C</i>	CG <i>tex_F/verde_C</i>		
	triângulo _I	FL <i>(tex_F/alaranjado_C)</i>	FL <i>(tex_F/alaranjado_C)</i>	FL <i>(tex_F/alaranjado_C)</i>	arco75	
		CF <i>(tex_F/alaranjado_C)</i>	CF <i>(tex_F/alaranjado_C)</i>	CF <i>(tex_F/alaranjado_C)</i>		
		GM <i>tex_F/verde_C</i>	GM <i>tex_F/verde_C</i>	GM <i>tex_F/verde_C</i>		
		CG <i>tex_F/verde_C</i>	CG <i>tex_F/verde_C</i>	CG <i>tex_F/verde_C</i>		
	triângulo _S	FL <i>(tex_F/alaranjado_C)</i>	FL <i>(tex_F/alaranjado_C)</i>	FL <i>(tex_F/alaranjado_C)</i>	arco25	
		CF <i>(tex_F/alaranjado_C)</i>	CF <i>(tex_F/alaranjado_C)</i>	CF <i>(tex_F/alaranjado_C)</i>		
		GM <i>tex_F/verde_C</i>	GM <i>tex_F/verde_C</i>	GM <i>tex_F/verde_C</i>		
		CG <i>tex_F/verde_C</i>	CG <i>tex_F/verde_C</i>	CG <i>tex_F/verde_C</i>		
		losango_V	losango_A	estigma	LA300	

Tabela 3.13: Base de regras para as entradas e saídas à segunda abordagem (B).

		AIM						
		polígono _M	polígono _V	regiões _P				
AIB	regiões _M	FRG <i>esparsas_M</i>	FRG <i>esparsas_M</i>	FRG <i>conjuntos_{VM}</i>	SVR <i>tex_P/marrom_E</i> <i>tex_P/marrom_C</i>			
		FC <i>tex_F/marrom_C</i>	FC <i>tex_F/amarelo_C</i>	FC <i>tex_F/marrom_C</i>				
		SVR <i>tex_P/marrom_E</i>	SVR <i>tex_P/marrom_E</i>	SVR <i>tex_P/marrom_E</i>				
	Marrom200	FRG <i>esparsas_M</i>	FRG <i>esparsas_V</i>	FRG <i>conjuntos_{VM}</i>				
		FC <i>tex_F/amarelo_C</i>	FC <i>tex_F/amarelo_C</i>	FC <i>tex_F/marrom_C</i>				
		SVR <i>tex_P/marrom_E</i>	SVR <i>tex_P/marrom_E</i>	SVR <i>tex_P/marrom_E</i>				
	Marrom100	FRG <i>esparsas_M</i>	FRG <i>esparsas_V</i>	FRG <i>conjuntos_{VM}</i>				
		FC <i>tex_F/cinza_C</i>	FC <i>tex_F/amarelo_C</i>	FC <i>tex_F/cinza_C</i>				
		SVR <i>tex_P/marrom_C</i>	SVR <i>tex_P/marrom_C</i>	SVR <i>tex_P/marrom_C</i>				
			FRG (<i>esparsas_M</i> ; <i>esparsas_V</i> ; <i>conjuntos_{VM}</i>)					
			FC (<i>tex_F/cinza_C</i> ; <i>tex_F/amarelo_C</i> ; <i>tex_F/marrom_C</i>)					

Na Tabela 3.12, são associados os estados de entrada e as funções-membro para a localização de frutos (*CA100*, *CA300* e *AA300*) e gramíneas (*LA100* e *LA300*). Observa-se que para cada função-membro dos estados de entrada existem 4 saídas com suas respectivas funções-membro, aos dois casos. Na Tabela 3.13, são associados os estados de entrada para as áreas irregulares de vegetação rasteira, folhas, galhos e ramos (*AIM* e *AIB*). As funções-membro das saídas (*FRG*, *FC* e *SVR*) nestes casos correspondem à formação, cor e texturas destes objetos.

A estrutura do MCP (*Model Predictive Control*) (NIKOLAU, 2001) nos conjuntos *Fuzzy* atende os intervalos otimizadores das funções-membro às saídas pretendidas nas duas abordagens, como as funções objetivo. Não é elaborada uma lei de controle que estabeleça tal relação, apenas a correção do modelo linear permite corresponder os estados de inferência como as saídas previstas, sendo estas corrigidas pela referência com o grau de pertinência máximo e as variáveis de interesse na saída.

c) *Teorema de Bayes:*

Considerando que a probabilidade condicional é indisponível nos casos em que a probabilidade *a priori* seja usada para a estimação de conjuntos de treinamento, o teorema de *Bayes* é dado pela função densidade de probabilidade (PDF), na equação 3.14.

$$P(C_i | y) = \frac{p(y | C_i)P(C_i)}{\sum_{j=1}^K p(y | C_j)P(C_j)} \quad (3.14)$$

Acima, $p(y/C_i)$ é a PDF condicional da classe i , sendo y é um valor escalar ou vetorial, para várias características, cujas variáveis aleatórias são denominadas de multi-variadas com PDF conjunta (PDFc). Logo, a resolução da PDFc depende do número de componentes da mistura (k) e seus parâmetros (q), que são encontrados através de *maximum-likelihood* (ML) ou através da estimativa Bayesiana (PAALANEN *et al.*, 2006).

Abaixo, observa-se o algoritmo customizado à classificação pelo teorema de *Bayes*.

Entradas: $k_{min}, k_{max}, \epsilon$, parâmetros iniciais $\hat{\theta}(0) = \{\hat{\theta}_1, \dots, \hat{\theta}_{k_{max}}, \hat{\alpha}_1, \dots, \hat{\alpha}_{k_{max}}\}$

Saídas: Modelo de mistura em $\hat{\theta}_{melhor}$

$T \leftarrow 0, k_{nz} \leftarrow k_{max}, \zeta_{min} \leftarrow +\infty$

$u_m^{(i)} \leftarrow (p(y(i) | \hat{\theta}_m))$, para $m = 1, \dots, k_{max}$ e $i = 1, \dots, n$

enquanto $k_{nz} \geq k_{min}$ **faça**

repita

$t \leftarrow t + 1$

para $m = 1, \dots, k_{max}$ **faça**

$w_m^{(i)} \leftarrow \hat{\alpha}_m u_m^{(i)} \left(\sum_{j=1}^{k_{max}} \hat{\alpha}_j u_j^{(i)} \right)^{-1}$, para $i = 1, \dots, n$

$\hat{\alpha}_m \leftarrow \max \left\{ 0, \left(\sum_{i=1}^n w_m^{(i)} \right) - \frac{N}{2} \right\} \left(\sum_{j=1}^k \max \left\{ 0, \left(\sum_{i=1}^n w_j^{(i)} \right) - \frac{N}{2} \right\} \right)^{-1}$

$\{\hat{\alpha}_1, \dots, \hat{\alpha}_{k_{max}}\} \leftarrow \{\hat{\alpha}_1, \dots, \hat{\alpha}_{k_{max}}\} \left(\sum_{m=1}^{k_{max}} \hat{\alpha}_m \right)^{-1}$

se $\hat{\alpha}_m > 0$ **então**

$\hat{\theta}_m \leftarrow \arg \max \log p(Y, W | \theta)$

$u_m^{(i)} \leftarrow p(y^{(i)} | \hat{\theta}_m)$, para $i = 1, \dots, n$

senão

$k_{nz} \leftarrow k_{nz} - 1$

fim se

fim para

$\hat{\theta}(t) \leftarrow \{\hat{\theta}_1, \dots, \hat{\theta}_{k_{max}}, \hat{\alpha}_1, \dots, \hat{\alpha}_{k_{max}}\}$,

$\zeta[\hat{\theta}(t), Y] \leftarrow \frac{N}{2} \sum_{m: \hat{\alpha}_m > 0} \log \frac{n \hat{\alpha}_m}{12} + \frac{k_{nz}}{2} \log \frac{n}{12} + \frac{k_{nz} N + k_{nz}}{2} - \sum_{i=1}^n \log \sum_{m=1}^k \hat{\alpha}_m u_m^{(i)}$

até $\zeta[\hat{\theta}(t-1), Y] - \zeta[\hat{\theta}(t), Y] < \epsilon$ **ou** $\zeta[\hat{\theta}(t-1), Y]$

se $\zeta[\hat{\theta}(t), Y] \leq \zeta_{min}$ **então**

$\zeta_{min} \leftarrow \zeta[\hat{\theta}(t), Y]$

$\hat{\theta}_{melhor} \leftarrow \hat{\theta}(t)$

fim se

$m^* \leftarrow \arg \min_m \{\hat{\alpha}_m > 0\}, \hat{\alpha}_{m^*} \leftarrow 0, k_{nz} \leftarrow k_{nz} - 1$

fim enquanto

Figura 3.21: Algoritmo customizado à classificação pelo teorema de *Bayes*.

Desta forma, a aplicação do algoritmo de maximização da esperança (*expectation-maximization*) na obtenção da ML, restringe as estimativas pela insensibilidade à inicialização

e a necessidade da escolha do número de componentes da mistura, considerado como sub-adaptação ou super-adaptação. A customização deste algoritmo foi feita por (FIGUEIREDO e JAIN, 2002) que corrige tais problemas.

A equação para a obtenção da PDFc é dada a seguir. Trata-se de um modelo de misturas de funções gaussianas finitas, sendo k o número de misturas, α_m são as probabilidades ponderadas da mistura, θ_m são os parâmetros da função gaussiana (FIGUEIREDO e JAIN, 2002).

$$\hat{p}(y|\theta) = \sum_{m=1}^k \alpha_m p(y|\theta_m) \quad (3.15)$$

O número de componentes é definido arbitrariamente na inicialização do algoritmo, cujas médias destas são distribuídas no espaço formado pelo conjunto de treinamento. Estas componentes são eliminadas quando não suportadas pelos dados, isto é, são saturadas. Com apenas o número mínimo de componentes, o algoritmo atende ao modelo mais eficiente dada a ponderação da função objetivo, que deve ser minimizada em relação a θ .

A estimação da PDFc é direcionada para cada quantidade de dimensões, em todos os conjuntos percentuais para as duas abordagens. A classificação do conjunto de testes é feita com a equação (0.14) e implica no resultado da na maior taxa de acerto. A programação deste algoritmo é feita também como uma biblioteca ao *toolbox* de Estatística do *Matlab*.

d) Teorema de Bayes ingênuo:

Como as características são independentes nesta variação do teorema de *Bayes*, a suposição é dada pelo grau de incerteza da PDF para cada classe ou conjunto de classes.

$$p(y|C_i) = \prod_{j=1}^n p(y_j|C_i) \quad (3.16)$$

Sendo $p(y/C_i)$ a função de densidade marginal condicional (PDFc), o treinamento e a avaliação deste classificador estatístico são feitos para várias características na discriminação da PDF, resultando em várias PDFs marginais discretas, para cada dimensão analisada,

baseado na estimação por *máxima verossimilhança* (*maximum-likelihood estimation*). Esta função critério é dada pelas medidas de média e variância de cada classe dos padrões de treinamento como entrada de dados, dado pela distribuição normal na estimação da PDF de cada classe.

As distâncias de *Bhattacharyya*, *Jeffries-Matusita* e *mínima* (*k-means*) são direcionadas às medidas de separabilidade entre duas ou mais classes, para a estimativa de correspondência entre as regiões classificadas pelo teorema de *Bayes*, *Bayes ingênuo*, RNA e lógica *Fuzzy*.

As distribuições para as distâncias acima, são relacionadas a seguir.

v) *Distância de Bhattacharyya*:

$$B_{ij} = \frac{1}{8} (M_j - M_i)^T \cdot \left(\frac{C_i + C_j}{2} \right)^{-1} \cdot (M_j - M_i) + \frac{1}{2} \ln \left(\frac{\left| \frac{C_i + C_j}{2} \right|}{\sqrt{|C_j| |C_i|}} \right) \quad (3.17)$$

vi) *Distância de Jeffries-Matusita*

$$J_{ij} = 2 \left(1 - e^{-B_{ij}} \right) \quad (3.18)$$

vii) *Distância mínima*

$$e'(Z, M_i)^2 = (Z - M_i)^T \cdot P_i^{-1} \cdot (Z - M_i) \quad (3.19)$$

3.2.5. Modelos de validação dos classificadores

Os métodos de avaliação dos classificadores aplicados são estabelecidos na análise dos resultados para a validação dos conjuntos de testes com os conjuntos de treinamento. A classificação ótima é resultante das provas de treino por cinco vezes de cada classificador.

A avaliação é feita sob inspeção das taxas de acerto calculadas para os conjuntos validados, o número de dimensões (vetores de características), percentagem do conjunto de treinamento, modelo do classificador, espaço de cores, matrizes de confusão às classes alvo, gráficos de dispersão das componentes principais, e a abordagem adotada.

As taxas de acerto são calculadas pelo número de segmentos classificados corretamente em relação ao número total de segmentos, cujo desempenho é verificado a partir de duas análises ao conjunto de treinamento e ao conjunto de testes. Os gráficos de dispersão denotam a quantidade de segmentos para cada representação às duas componentes principais da ACP.

A influência de todos os dados (segmentos) processados com a taxa negativa do total de segmentos é definida pelas matrizes de confusão, aos conjuntos de treinamento, teste e validação. Estes métodos são dedicados à avaliação quantitativa dos segmentos.

A avaliação qualitativa é feita pelos mapas de classe ao reconhecimento gráfico de padrões. Para cada imagem de entrada nas duas abordagens, um mapa é construído pela atribuição de um *pixel* naquele espaço de cores por preenchimento de área (*floodfill* – determinação da área de um dado nó *intra-pixels* em uma matriz multidimensional) (HECKBERT, 1990). A técnica *floodfill* necessita de três parâmetros: nó de início, nó alvo e cor de reposição/preenchimento. Inicia-se pela procura de todos os nós na matriz conectada ao nó alvo, e assim realiza o preenchimento de cor na área delimitada desta matriz. Tal algoritmo é construído por estruturas de dados baseadas em pilha ou fila.

Assim, o processo exige que, para cada classificador aplicado nos dois espaços de cores estudados, sejam avaliados dois critérios: limiar de tolerância para evitar a presença de segmentos desconhecidos (classificadores estatísticos em 0,977 e classificadores cognitivos em 0,907); tamanho de agrupamento máximo que verifica qual classe é maior dentre as demais classes alvo.

3.3. Considerações finais

Os materiais e métodos propostos foram apresentados neste capítulo, detalhando as etapas e aspectos que compõem a metodologia aplicada nas técnicas de visão computacional, tais como o pré-processamento, segmentação de imagens, modelos de extração e normalização de

características, tipos de classificadores e reconhecimento de padrões, para as duas abordagens trabalhadas.

4. Resultados

Os resultados aqui apresentados são organizados na seguinte sequência: pré-processamento em realce de imagens; segmentação das cenas de culturas agrícolas e linhas de plantio; normalização e extração de características dos vetores paramétricos; classificação cognitiva e estatística dos segmentos; avaliação dos classificadores implementados e validação dos conjuntos de teste e treinamento; pós-processamento em morfologia matemática para extração de objetos de interesse na quantificação, dimensionamento e comparação de culturas; reconhecimento de padrões através dos mapas de classes e mapas de padrões.

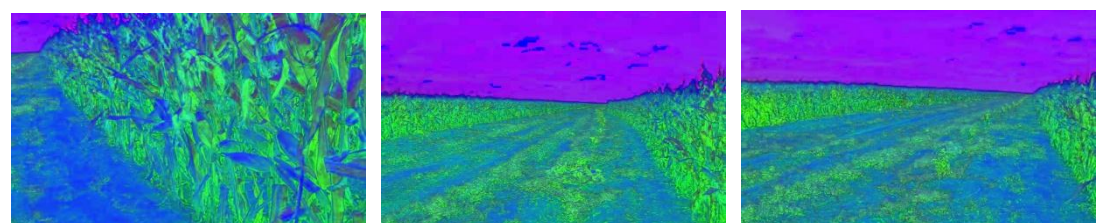
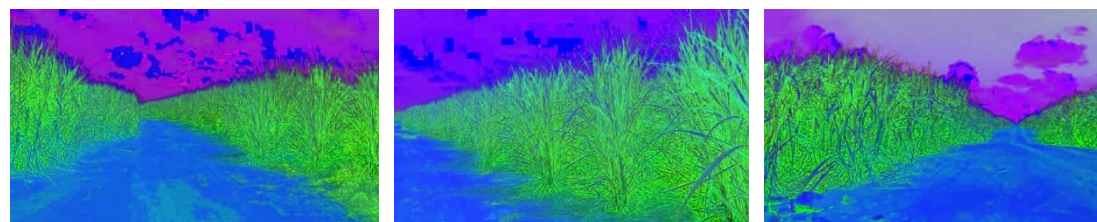
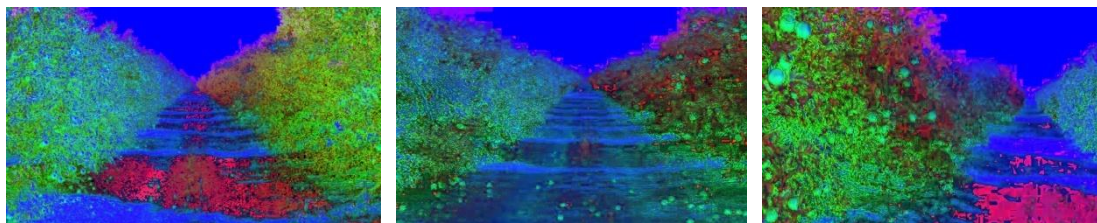
4.1. Pré-processamento em realce de imagens

Quanto maior a resolução espacial das imagens, maior é o tempo de processamento e a precisão na etapa de segmentação. Como a análise do tempo computacional nesta proposta é relativamente alta, cada imagem teve sua resolução espacial reduzida em uma proporção de 4x1 para linhas e 6x1 para colunas (852x480 *pixels*). Nesta nova dimensão, os *pixels* que representam as culturas foram mantidos em maior densidade de regiões, favorecendo estimativas mais convergentes de sua localização.

As imagens com as novas dimensões foram pré-processadas no espaço de cores RGB, em fator de escala para valores individuais na intersecção de cada *pixel* em todas as linhas e colunas. Para tal, foi utilizado o método de interpolação bicúbica (convolução) (BITTENCOURT de, 1998), o qual assinala ao valor do *pixel* da imagem de saída, uma média ponderada dos *pixels* em vizinhança 4x4 mais próxima.

Os resultados das imagens pré-processadas com a redução da resolução espacial e no espaço de cores HSV são mostrados na Figura 4.1, a seguir. Os conjuntos de imagens (a), (b),

(c), (d) e (e) representam cenas de linhas de plantio de pomares, canaviais, milhoais, gramíneas e frutos, respectivamente, e abaixo das mesmas cenas são relacionados os resultados das transformações do espaço de cores RGB para o espaço de cores HSV.



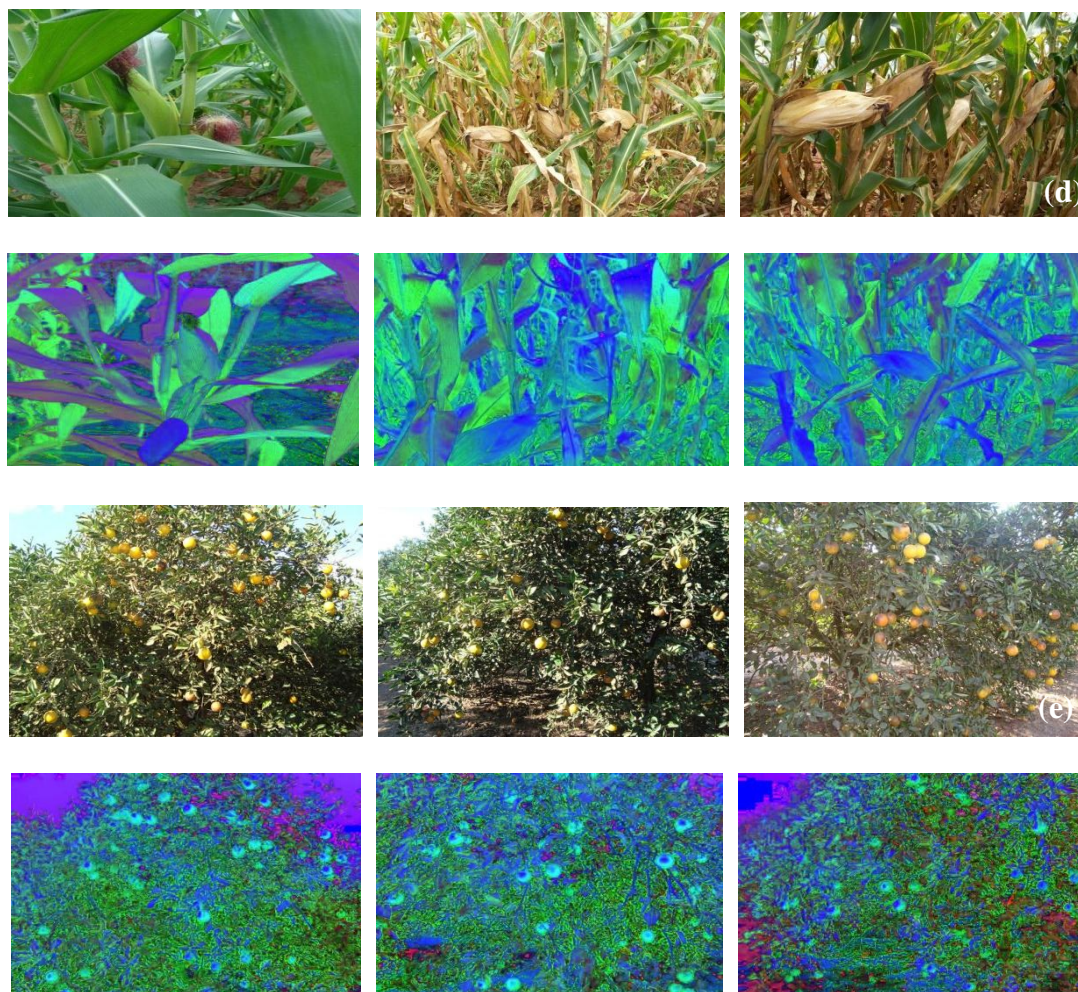


Figura 4.1: Imagens adquiridas e processadas nos espaços de cores RGB e HSV: (a) linha de plantio de pomares; (b) linha de plantio de cana; (c) linha de plantio de milho; (d) plantação de milho; (e) plantação de laranja.

Os processos morfológicos no pré-processamento das imagens em ambos os espaços de cores foram realizados pela plataforma matemática computacional *Matlab* (MATHWORKS, 2010), com auxílio dos *toolboxes* de Processamento de Imagens e Processamento de Sinais.

4.2. Segmentação das cenas de culturas agrícolas

Dados os parâmetros ajustados após a compilação com algumas diretivas próprias do *OpenCV*, nas cenas de pomares, canaviais e milharais, o algoritmo de segmentação é processado em C++, de modo recursivo como biblioteca do *Matlab*, separando os *pixels* relativos à classe *área navegável* (solo) e à classe *área não navegável* (guias limites do plantio das árvores e céu), sendo a primeira submetida a condições de variação de

sombreamento, no caso da posição da origem da luz. A transformação inicial em escala de cores HSV provê a tentativa de agrupar os segmentos com maior tonalidade (sombra) com os segmentos de maior incidência de luz, constituindo uma classe de padrões (área navegável); os demais segmentos, alocados geometricamente na imagem, por desvios laterais e superiores são relacionados à segunda classe de padrões (área não-navegável).

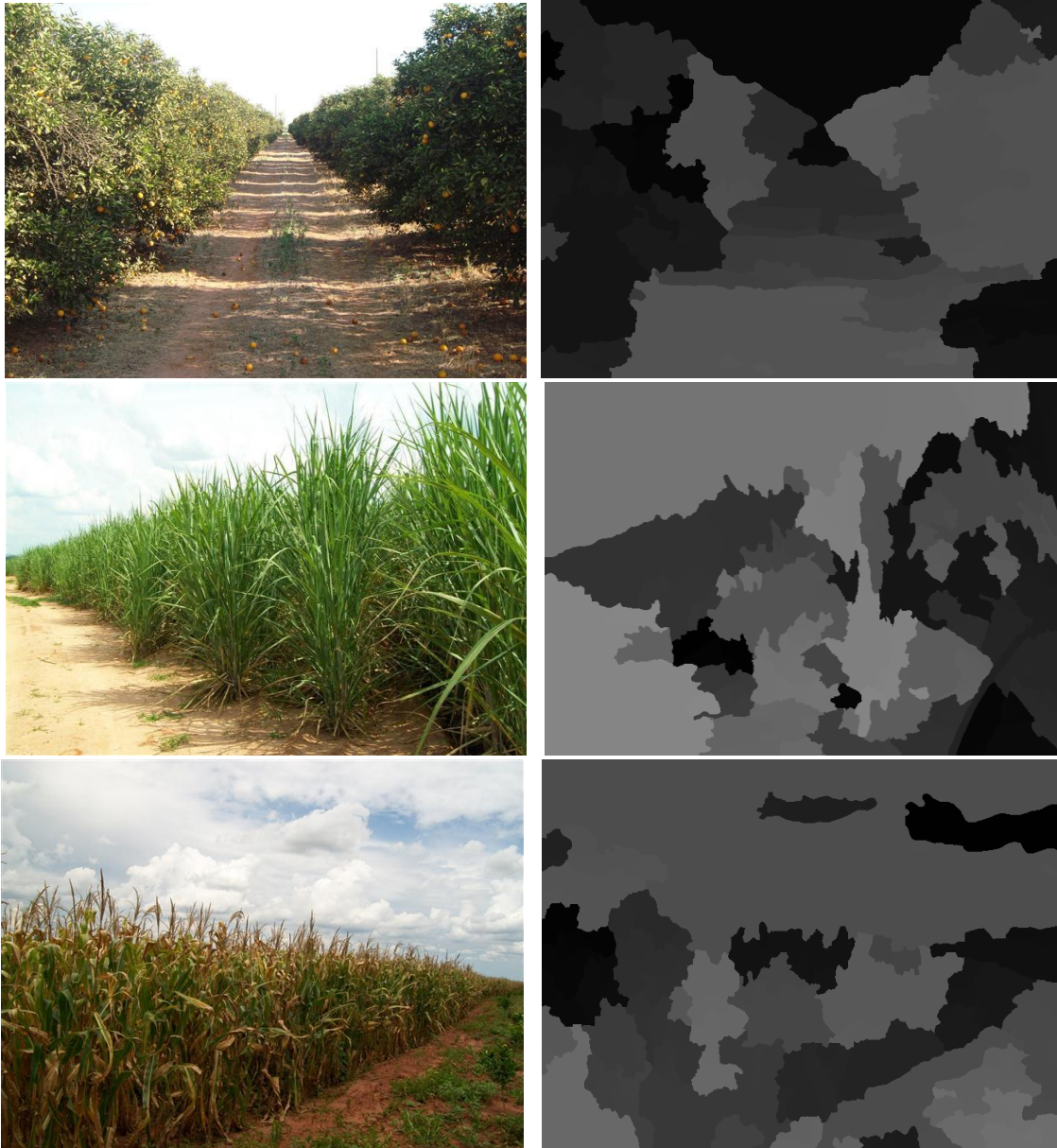


Figura 4.2: Imagem original à esquerda e quantizada em regiões de cores por níveis de cinza (direita). Observa-se que existem regiões na área navegável não desejados, posteriormente agrupados numa classe de padrões.

A Figura 4.2, acima, ilustra as regiões quantizadas para as cenas de pomares, canaviais e milharais. Nota-se que a quantização de cores é feita de forma não homogênea, com realce em regiões de mesma tonalidade de cores, e que devem ser atribuídos à mesma classe, mas tal compensação é relativamente empírica com o reajuste do *limiar de agrupamento de regiões* (0,5) e o *número de escalas* (2). Esta adoção é feita para se obter um equilíbrio de regiões de interesse em imagens com formato geométrico de polígonos à esquerda, à direita e centro. As fronteiras destas regiões são redimensionadas formando uma nova classe de padrões e com relativa redução de ruídos.

Pode-se observar na Figura 4.3, a seguir, as primeiras imagens com máxima distribuição espacial são passíveis de super-segmentação. Logo, a aquisição de imagens é readequada quanto à distância focal da câmera, para se obter distribuições dos padrões pretendidos em toda a cena, integradas em maior escala. Isso resulta em cenas menos esparsas, com menos padrões para as duas classes, viabilizando menor de tempo de processamento durante a segmentação.

Dessa forma, conclui-se que a partir da quantização de cores dada pela filtragem de grupo de pares de *pixels*, a primeira etapa do algoritmo não supervisionado (JSEG) realizou a divisão adequada de todas as regiões não homogêneas nas imagens. Experimentalmente, a partir destes valores, foi possível reajustar os parâmetros de *limiar de agrupamento de regiões* e o *número de escalas* na fusão de regiões, até atingir baixa equivalência de *pixels*, resultando numa distribuição espacial menos densa, simples e com os padrões apropriados, na segmentação final ao problema de locomoção do robô agrícola móvel.



Figura 4.3: Imagens à esquerda mostram a segmentação (distribuição espacial) final, com os parâmetros inicialmente ajustados para encontrar a maior parte de irregularidades de *pixels*; à direita, as mesmas cenas com reajuste dos parâmetros para agregar a maior parte das regiões em concentrações simplificadas aos classificadores.

4.3. Segmentação das cenas de linhas de plantio

Os ajustes iniciais dos parâmetros necessários ao algoritmo de segmentação para as culturas laranja e milho foram os mesmos adotados na primeira abordagem. Com isso, observou-se que as espigas e laranjas foram representadas pela maior quantidade de *pixels* da imagem, dadas suas características de alto contraste com os demais objetos da cena.

A Figura 4.4, a seguir, mostra três categorias de cenas em pomares. A primeira identifica a maior parte da árvore, evidenciando também o solo. Nesta categoria, o *limiar de quantização* foi reajustado a maiores valores para que a fusão de regiões com mesma tonalidade de cores entre colmos, ramos, galhos e solo fosse evitada. A segunda denota detalhes em determinados conjuntos nas cenas de pomares, excluindo regiões mais escuras do restante da cena. Não apenas as irregularidades de cada folha são segmentadas, como também anormalidades de tons no próprio fruto, possibilitando posterior análise de doenças características de laranjas. A terceira categoria identifica, a exemplo da primeira, a maior parte das árvores, porém com maior incidência da copa e regiões do céu.

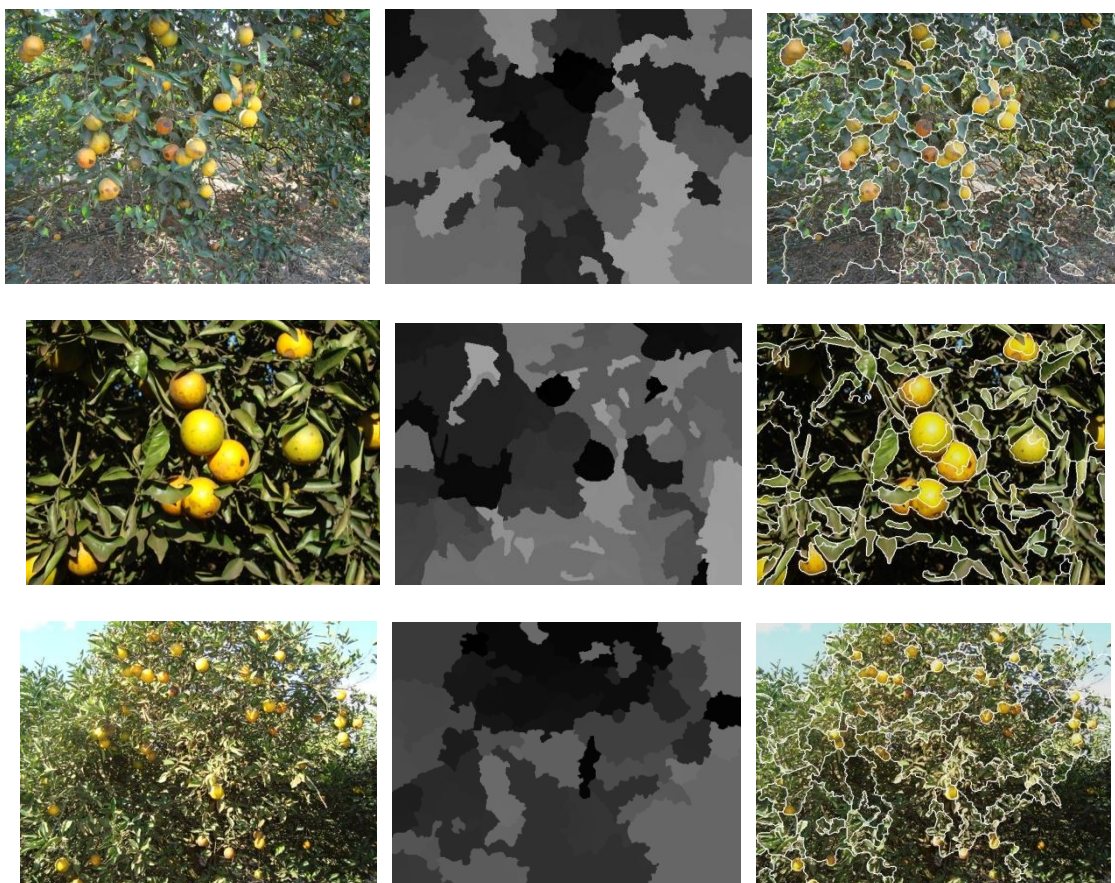


Figura 4.4: Cenas de pomares originais (coluna 1); imagens com cores quantizadas (coluna 2); imagens segmentadas (coluna 3).

As cenas mostradas na Figura 4.5, a seguir, ilustram as plantações de milho em duas ocasiões de colheita (milho verde e silagem). Observa-se que no período de silagem (duas primeiras cenas da coluna 1), o algoritmo de segmentação proposto não é eficaz na etapa de

quantização de cores (coluna 2), tampouco eficiente na distribuição espacial (coluna 3), uma vez que as características de cor e textura destas cenas se confundem com os demais objetos contidos nas imagens – as espigas de milho neste período apresentam características físicas que requerem outras técnicas pré-processamento.

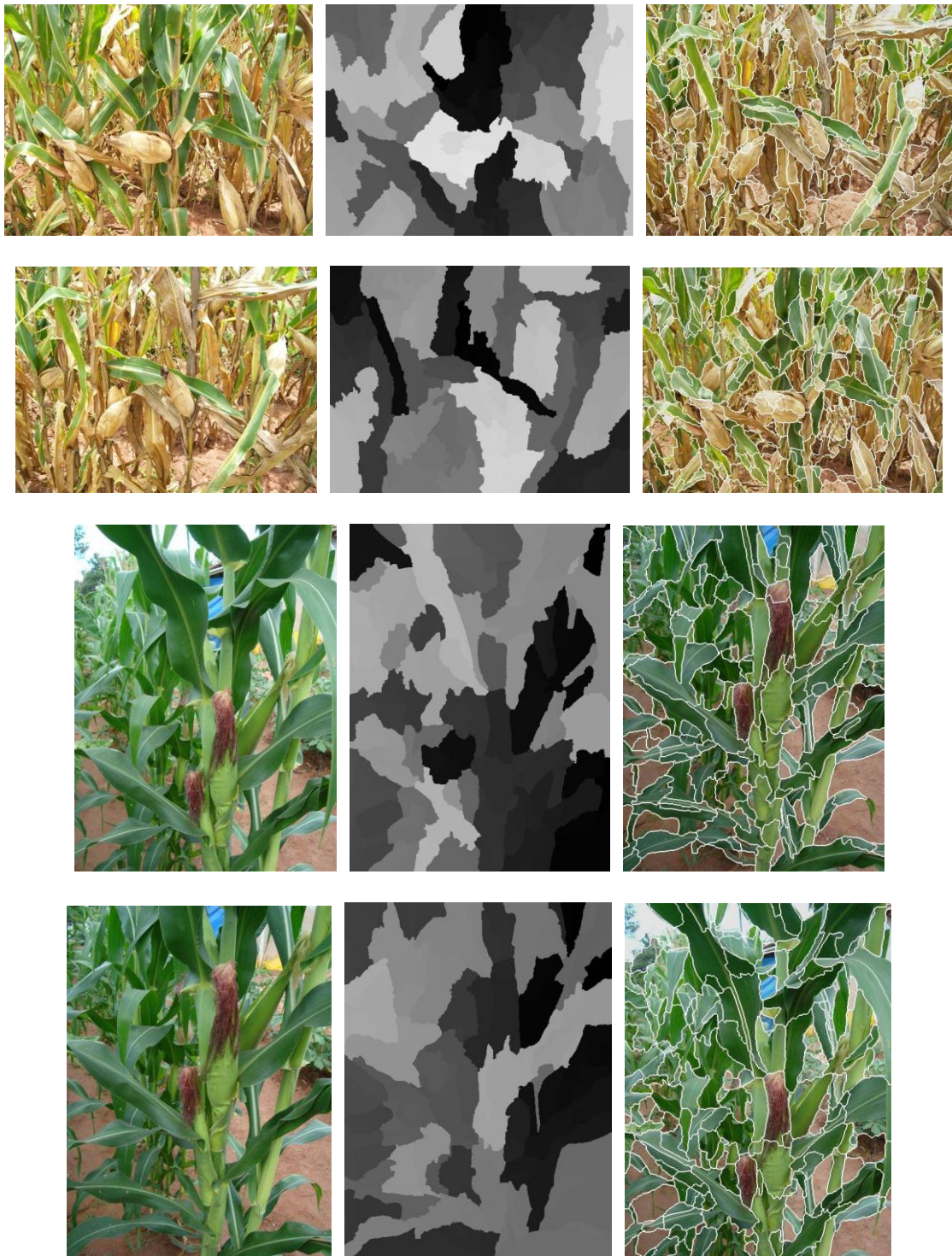


Figura 4.5: Plantação de milho em período de colheita: as duas primeiras linhas de imagens denotam a colheita em silagem e as duas últimas linhas denotam a colheita em milho verde.

No período de milho verde (duas últimas cenas da coluna 1), a segmentação não supervisionada aplicada necessita ser customizada com outras técnicas em morfologia matemática e segmentação por área, considerando os *pixels* do estigma da espiga, como região inicial de contorno, verificadas pelos resultados da quantização de cores (coluna 2) e distribuição espacial (coluna 3).

As cenas naturais de plantação de milho para todas as épocas, como mencionado no capítulo 3 (Metodologia proposta) são referenciadas como padrões distintos da quantização de cor e distribuição espacial durante a segmentação pelo algoritmo JSEG, apresentando características indesejadas de tonalidades na cor e textura, que os pomares de laranjeiras não apresentam. Desta forma, são propostas técnicas de fusão em segmentação por área e textura que podem auxiliar o algoritmo JSEG na composição final da distribuição espacial de toda a cena, como complemento e continuação deste projeto, uma vez que são os objetivos secundários de processamento.

Logo, as cenas apresentadas aqui apenas comprovam que a segmentação não-supervisionada é majoritariamente ineficiente quando somente aplicada a esta cultura, não necessitando *a priori* sua classificação.

As imagens no período de colheita em silagem são compostas por padrões equivalentes quase em toda a cena, uma vez que as folhas mais próximas das espigas têm a mesma tonalidade de cor e são confundidas durante a quantização de cor. Por outro lado, as imagens no período de colheita em milho verde são compostas por características de texturas e na observação da forma geométrica das espigas. A segmentação baseada em grafos em conjunto com JSEG tende a apresentar resultados satisfatórios para a delimitação da região de contorno de cada gramínea e o restante dos objetos da cena, pois com o cálculo da área da espiga,

iniciado pela região superior do colmo até a região do seu estigma, fornece uma estimativa do volume e disposição espacial.

4.4. Normalização e extração de características

Os gráficos a seguir denotam a razão percentual da aplicação da ACP, entre a soma dos autovalores até a dimensão indicada, como também a soma dos autovetores para os conjuntos de treinamento das duas abordagens propostas nos espaços de cores RGB e HSV.

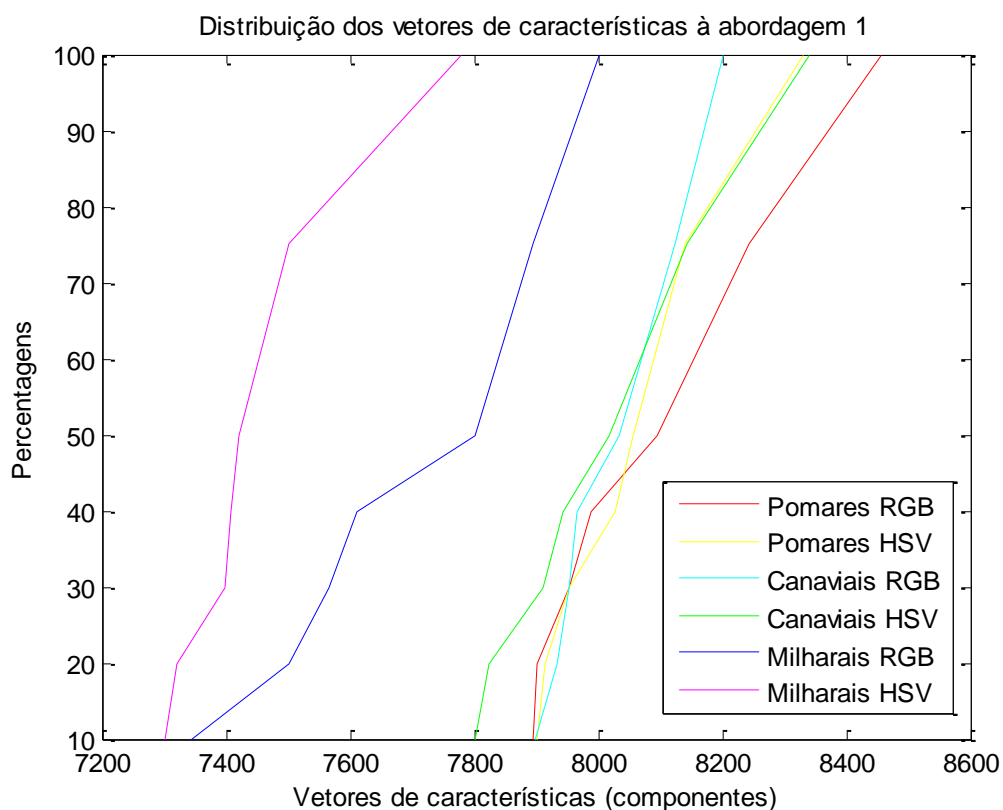


Figura 4.6: Distribuição dos vetores de características à abordagem 1.

O gráfico ilustrado na Figura 4.6, acima, apresenta a distribuição dos vetores de características da primeira abordagem, para as três situações de locomoção nas linhas de plantio de pomares, canaviais e milharais. Observa-se que para todas as situações nos casos RGB e HSV o aumento da quantidade de vetores é proporcional ao aumento da porcentagem dos conjuntos de treinamento, tornando-os maiores em amostragens e em representação vetorial.

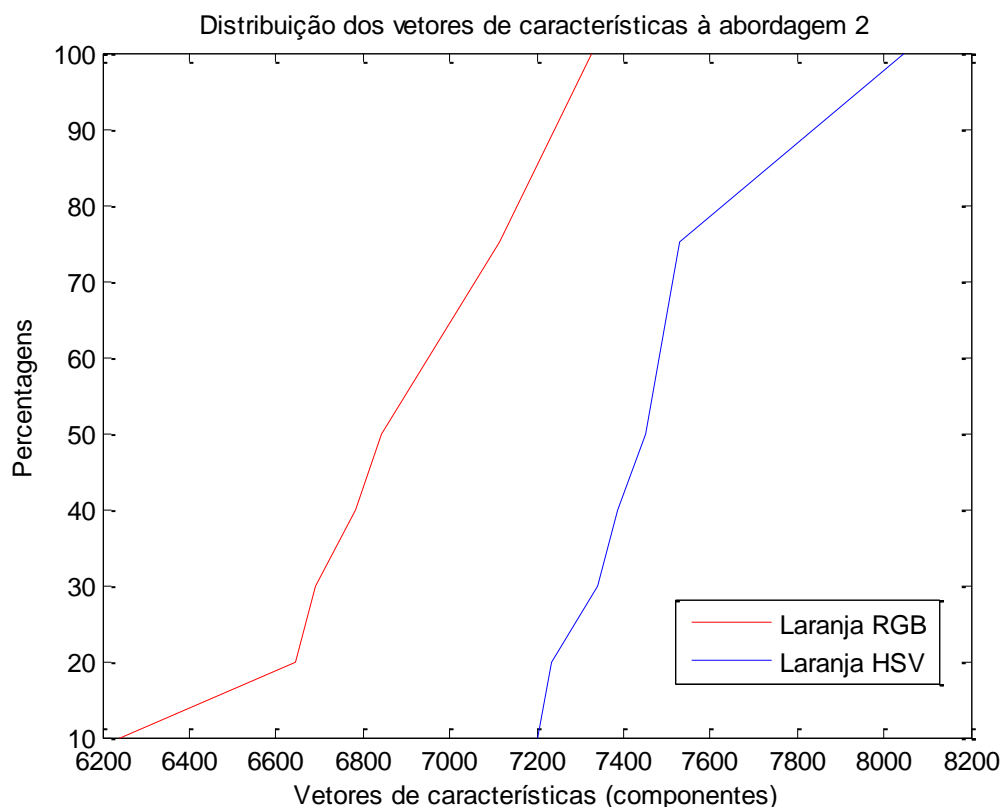


Figura 4.7: Distribuição dos vetores de características à abordagem 2.

Na Figura 4.7, a distribuição vetorial da segunda abordagem apresenta superioridade na quantidade de vetores para a representação dos objetos contidos nas cenas naturais no caso HSV, em relação ao caso RGB, devido à concentração e homogeneidade dos histogramas neste espaço de cores.

Os gráficos abaixo mostram a quantidade de dimensões para cada conjunto em 10%, 20%, 30%, 40%, 50%, 75% e 100% às duas abordagens nos espaços de cores RGB e HSV.

A Figura 4.8, a seguir, ilustra a quantidade de dimensões (vetores de características) no caso RGB, para cada classe, indicando as máximas quantidades relativas a cada percentagem. As duas classes são próximas quanto às suas quantidades, com menos dimensões para a primeira classe. Na Figura 4.9, logo abaixo, a mesma relação é válida para o caso HSV, porém a disparidade entre as duas classes é maior.

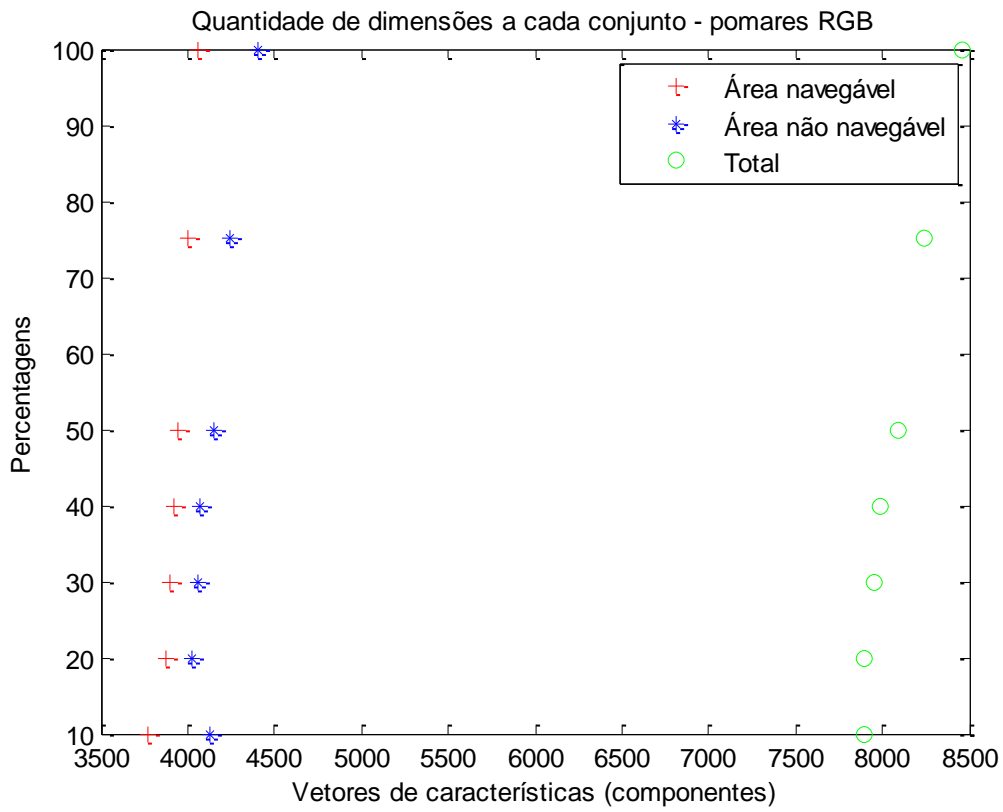


Figura 4.8: Quantidade de dimensões a cada conjunto – pomares RGB.

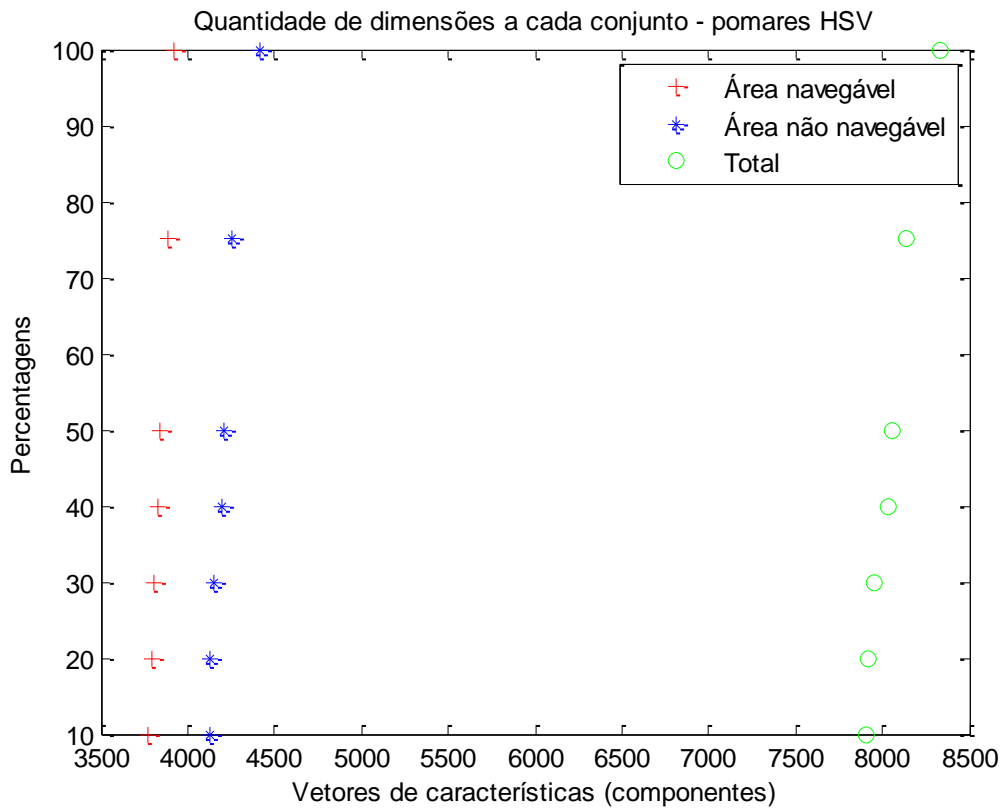


Figura 4.9: Quantidade de dimensões a cada conjunto – pomares HSV.

Nas linhas de plantio de canaviais, para o caso RGB, a Figura 4.10 mostra a relação equiparável entre as quantidades de dimensões para cada percentagem das duas classes. A alternância entre as quantidades dá-se pela característica da composição das imagens deste tipo de cultura, identificando com maior precisão as linhas de plantio como os ângulos de guiagem, à medida que os conjuntos aumentam.

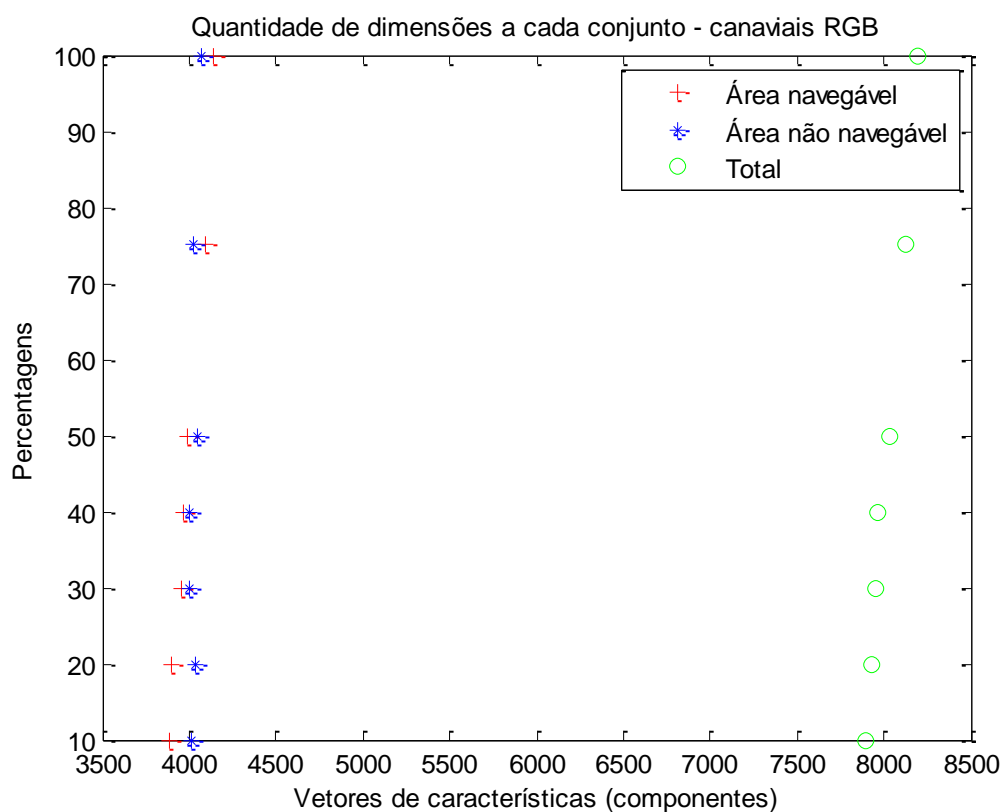


Figura 4.10: Quantidade de dimensões a cada conjunto – canaviais RGB.

No caso HSV (Figura 4.11), em quatro conjuntos de vetores (40%, 50%, 75% e 100%), a quantidade de dimensões para a primeira classe é maior que a segunda, pois o critério de homogeneidade tende a agrupar mais vetores para a classe *Área navegável*, cuja área que representa o solo, na sua identificação é delimitado como uma única região.

Na Figura 4.12, acima, para o caso RGB, a primeira classe tende a aumentar a quantidade de dimensões, conforme a percentagem de cada conjunto aumenta.

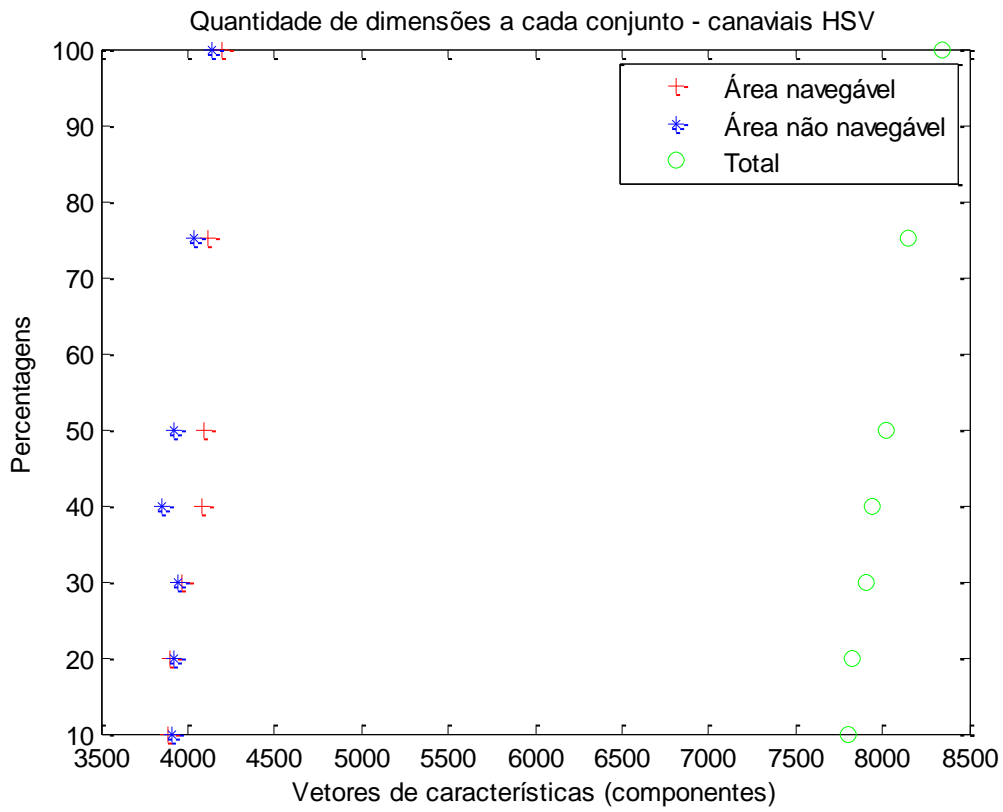


Figura 4.11: Quantidade de dimensões a cada conjunto – canais HSV.

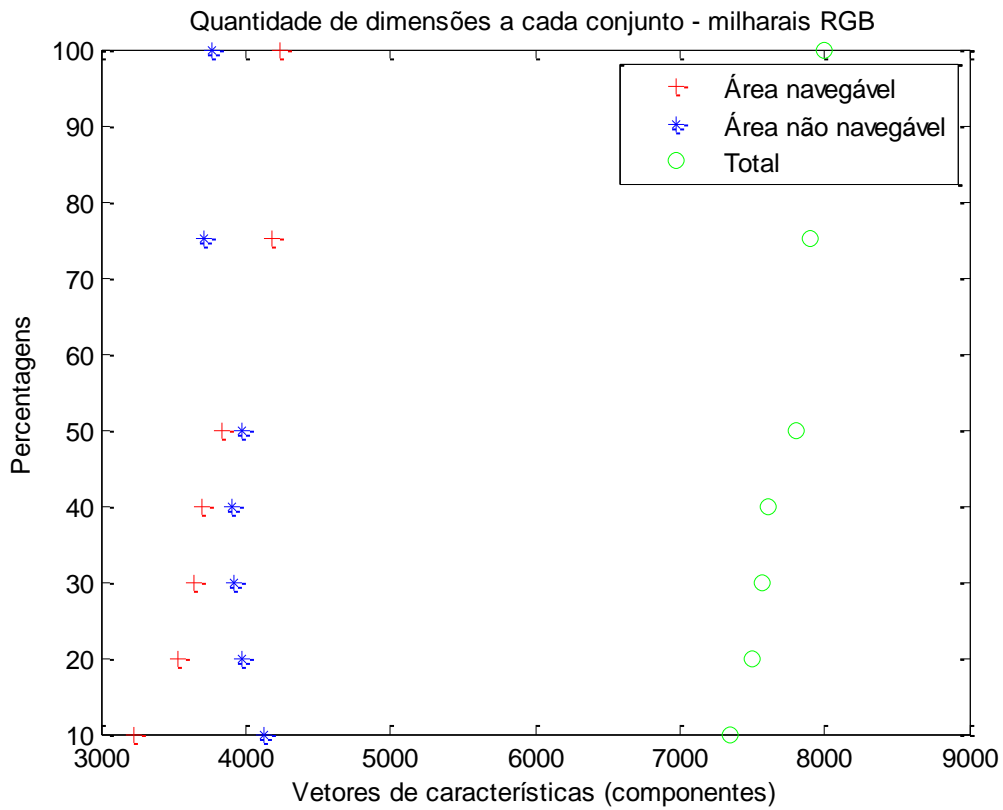


Figura 4.12: Quantidade de dimensões a cada conjunto – milhares RGB.

Novamente, quanto maior a relação de vetores de características para a representação de cada classe, maior é a tendência da primeira classe na obtenção de componentes que identifiquem as áreas das regiões correspondem à área passível de navegação. Observa-se que nos primeiros cinco conjuntos, a aproximação entre a quantidade de dimensões das duas classes é intensificada.

Na Figura 4.13, a seguir, para o caso HSV de milharais, a primeira classe já atende a quantidades de dimensões maiores que a segunda classe em percentagens baixas, e mais componentes são adicionados, proporcionalmente ao aumento de conjuntos de treinamento.

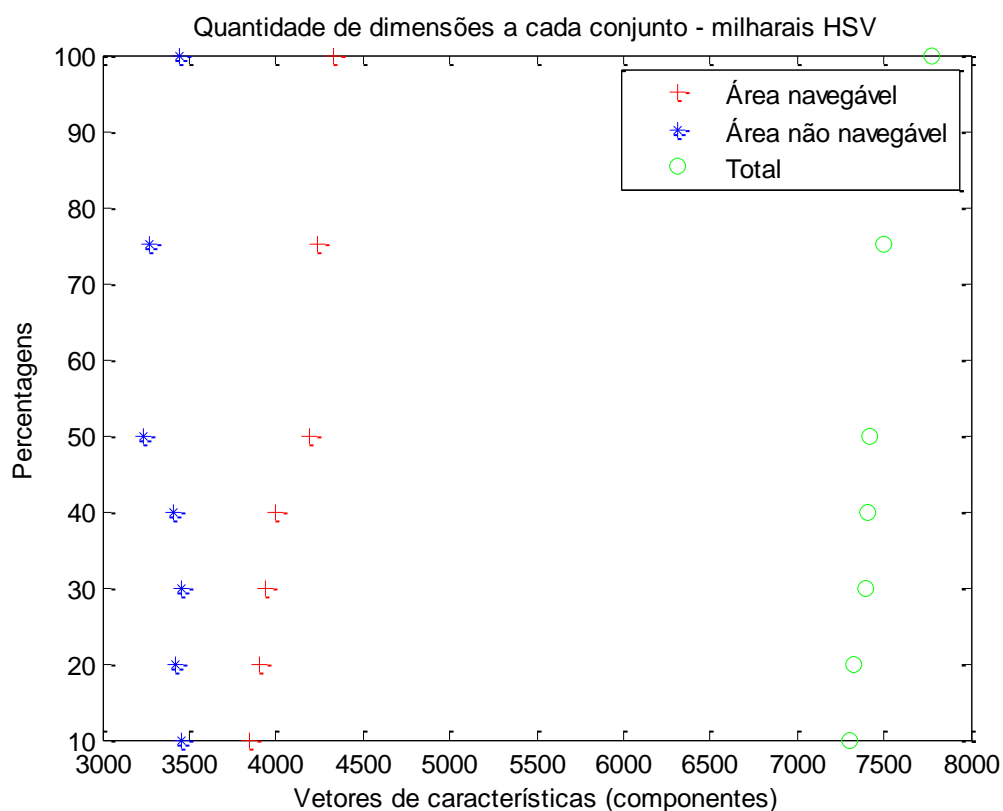


Figura 4.13: Quantidade de dimensões a cada conjunto – milharais HSV.

Na Figura 4.14 e Figura 4.15, para a segunda abordagem, no caso RGB, a discriminação das classes *Frutos*, *Céu* e *Folhas, ramos e galhos*, atende a quantidades constantes, proporcionais ao aumento dos conjuntos de treinamento. Tal quantidade, no caso HSV, é diminuída para a classe *Frutos*, pois a dispersão dos *pixels* é maior neste espaço de cores.

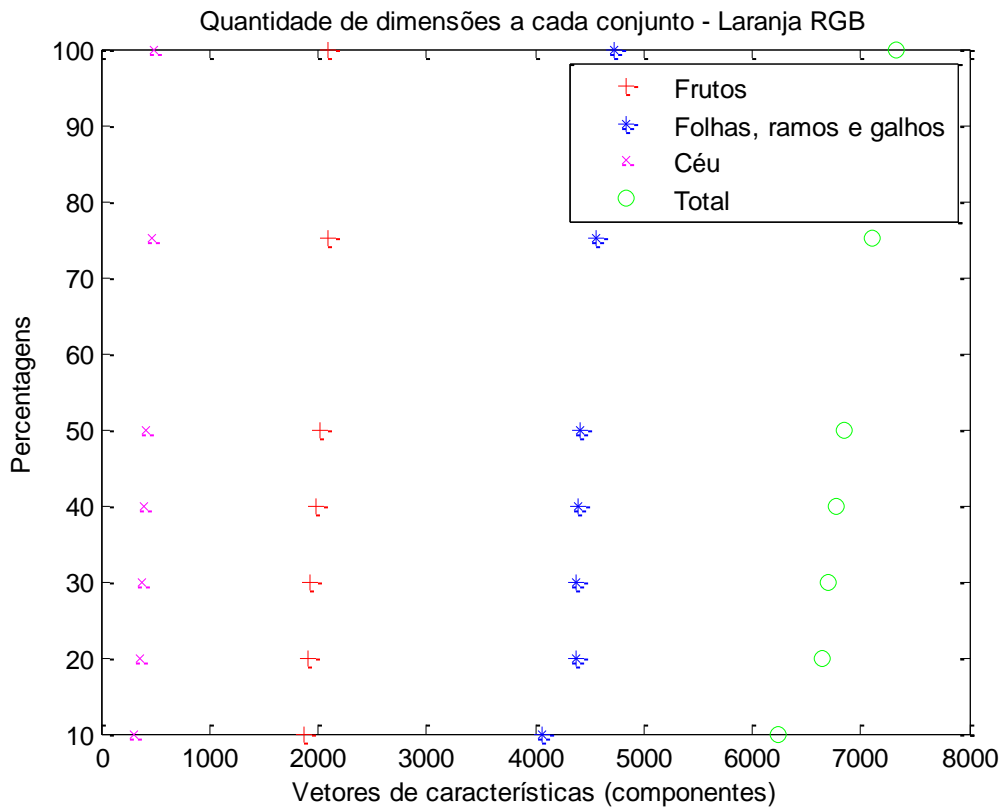


Figura 4.14: Quantidade de dimensões a cada conjunto – laranja RGB.

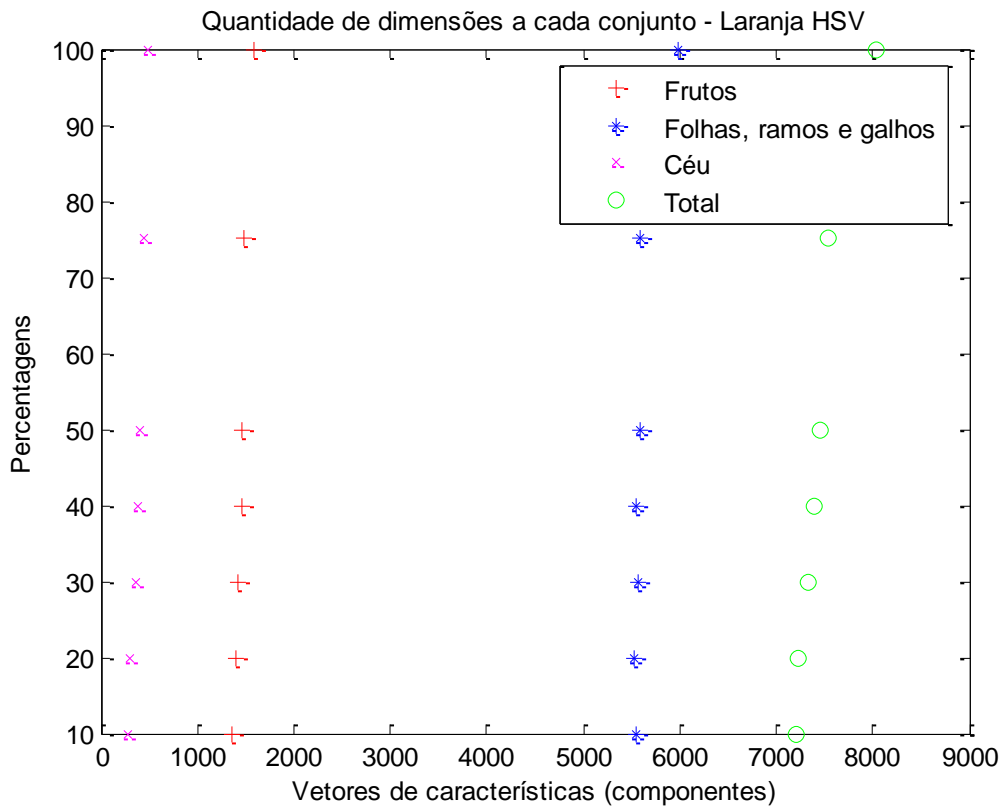


Figura 4.15: Quantidade de dimensões a cada conjunto – laranja HSV.

4.5. Classificação dos segmentos para as duas abordagens

4.5.1. Classificador baseado em redes neurais artificiais

Na primeira abordagem, os melhores resultados das RNA são obtidos com o espaço de cores HSV, pois este delimita satisfatoriamente as regiões de contorno onde é caracterizada a navegação do veículo autônomo, em perspectiva frontal. Porém, em casos de perspectiva lateral, 35 casos em HSV e 40 casos em RGB, o algoritmo não foi capaz de aprender a classe *área navegável*, alocando-a nas regiões adjacentes (plantações), formando apenas uma classe (área não navegável). Essa verificação dá-se pelos gráficos de regressão das saídas alvo, baseadas nas matrizes de confusão dos testes nas duas abordagens (**APÊNDICE B**).

Os gráficos de regressão denotam a quantização dos vetores ou saídas de interesse em relação às dimensões treinadas, testadas e validadas, nas duas abordagens. A proximidade das classes quanto sua dispersão é notável nos gráficos, uma vez que os conjuntos de treinamento e teste podem ser inter-relacionados com similaridades de dimensões, tornando o classificador mais robusto, porém com maior tempo de resposta.

Na segunda abordagem, para imagens em perspectiva que retratam detalhes da incidência de frutos, o espaço de cores HSV foi mais bem avaliado, devido às condições de luminosidade e dispersão com outras classes (folhas, ramos e galhos). Algumas cenas não foram corretamente classificadas em percentual aceitável ao reconhecimento de padrões, para o espaço de cores HSV.

Os segmentos classificados corretamente para ambas as abordagens resultam em valores próximos da quantidade total para HSV e RGB (locomoção), e em valores acima da média total para RGB e HSV (localização), com média baixa de falsos positivos no total para todas as iterações. Logo, as médias das taxas de acerto para os classificadores de menores dimensões de entrada apresentam os melhores resultados, uma vez que a dispersão dos

vetores de características foi facilmente evitada nas percentagens de aproximação entre 10%, 20%, 30%, 40% e 50%. Para a locomoção, o conjunto com 75% e 20 dimensões, obteve a melhor classificação. Para a localização, o conjunto com 100% e 15 dimensões, obteve a melhor classificação.

Para as avaliações acima foram usadas as taxas de acerto nas duas abordagens, para as 20 dimensões normalizadas, e o número de iterações do algoritmo de treinamento, sendo responsável pela melhoria do aprendizado, ajustando os pesos sinápticos de acordo com a necessidade das abordagens, minimizando o erro de classificação.

No **APÊNDICE C** é possível analisar os erros quadráticos médios (EQM) dos conjuntos de treinamento. Os EQM denotam a relação das médias de treinamento, teste e validação para as duas abordagens, em que ambas apresentam média baixa no caso HSV e alta em RGB. Com o aumento de neurônios na camada escondida, o classificador RNA torna-se mais eficiente, ao invés de expandir a relação de dimensões e amostras no conjunto de treinamento, uma vez que o tempo de resposta na denominação dos vetores que distam dos centróides das classes desejadas pode aumentar.

4.5.2. Classificador baseado em lógica nebulosa preditiva

Os modelos de variação dos estados de inferência para cada estado de entrada, correspondentes aos valores de saída desejados, são apresentados na Figura 4.16, para o problema da primeira abordagem.

Observa-se o comportamento dos estados de entrada e respectivos valores desejados nas saídas para a determinação das áreas de navegação, irregulares ou não. Como a variação de inferência está relacionada com a média de probabilidades na detecção geométrica dos triângulos, várias situações podem ser simuladas para valores diferentes de fator inicial, e correspondência com os processos de saída esperados.

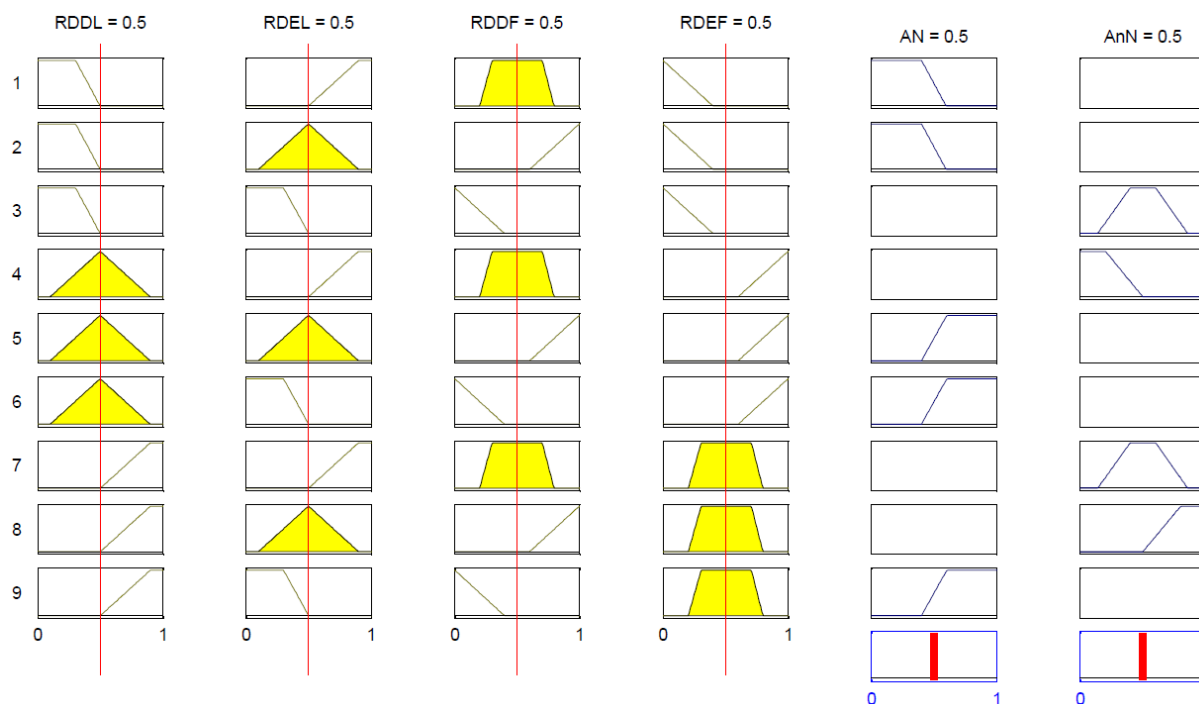


Figura 4.16: Variação dos estados de inferência das entradas e saídas para as regras formuladas à primeira abordagem.

Ainda na Figura 4.16, considerando as regras 4, 5 e 6 para a entrada RDDDL, as regras 2, 5 e 8 para a entrada RDEL, as regras 1, 4 e 7 para a entrada RDDF e as regras 7, 8 e 9 para a entrada RDEF, observa-se o máximo grau de pertinência para as funções membro de cada situação, ou seja, a saída desejada AN (Área navegável) é detectável com inferência máxima em quatro regras (3, 4, 7 e 8), para pelo menos dois graus de pertinência máximos numa mesma regra. Já nas regras 1, 2, 5, 6 e 9, em que o estado de inferência é ascendente ou descendente, o grau de pertinência não é máximo, em que o classificador não associa corretamente as funções membro de saída para a *defuzzificação*. Nestas funções membro, são associadas funções objetivo do modelo preditivo para a correção dos estados de inferência, que variam de 0,1 à 1,0, até a determinação do máximo grau de pertinência.

Abaixo, para a segunda abordagem, a variação dos estados de inferência para os estados de entrada e respectivos valores de saída é observada na Figura 4.17.



Figura 4.17: Variação dos estados de inferência das entradas e saídas para as regras formuladas à segunda abordagem.

Todas as entradas e saídas foram consideradas como um conjunto único de regras, tanto para a detecção de frutos quanto para a detecção de gramíneas, e os objetos inerentes aos dois casos, como folhas, ramos, galhos, caules, solo e vegetação rasteira. Como no capítulo 3, seção 3.2.4, foram apresentadas duas tabelas de regras distintas para a segunda abordagem, que denotam tal distinção, a aplicação conjunta da base de regras é realizada de forma autônoma quando o classificador for submetido à uma das duas situações.

Assim, são 27 regras ao todo que contemplam a segunda abordagem. Nove regras para a detecção de frutos, nove para a detecção de gramíneas e nove para a detecção dos objetos comuns aos dois casos. Analisando o primeiro caso, cinco regras apresentam grau de pertinência máximo (1, 2, 3, 6 e 9) para as saídas *FL* e *CF*, isto é, os frutos individuais e possíveis conjuntos com dois ou mais frutos são possíveis candidatos à detecção, e apenas na regra 3 o grau de pertinência denota a detecção.

Nas regras 4, 5, 7 e 8, que são os estados de inferência críticos, as saídas apresentam grau de pertinência não ótimo e são reajustados por suas próprias funções objetivo.

No segundo caso, as regras 10, 11, 12, 15 e 18, são aplicáveis aos candidatos à gramíneas, com grau de pertinência máximo apenas na regra 12; nas regras 13, 14, 16 e 17, não há candidatos a gramíneas devido aos estados de inferência não apresentarem intersecção entre as funções membro das entradas.

No terceiro caso, para a detecção de objetos pertencentes às duas cenas, o mesmo ocorre com as regras aplicáveis aos possíveis candidatos (5 regras) e não candidatos (4 regras). Apenas na regra 21 as saídas *FRG*, *SVR* e *FC* são classificadas corretamente.

No **APÊNDICE D**, as funções de superfícies são apresentadas para cada regra nas duas abordagens. Estas denotam a relação de condição das regras formuladas quanto aos possíveis valores de saída: funções membro geométricas por triângulos formados e suas condições de textura e cor, quanto à área denominada navegável ou área não navegável, da primeira

abordagem; funções membro geométricas das entradas *CA100/CA300/AA300* para os arcos incidentes, e das entradas *LA100/LA300* para os triângulos formados e suas condições de textura e cor, quanto aos frutos, gramíneas e seus conjuntos, bem como para as entradas *AIM* e *AIB* e sua relação com folhas, ramos e galhos esparsos e em conjuntos, folhas, caules, solo e vegetação rasteira por textura e cor, da segunda abordagem.

4.5.3. Classificador baseado no teorema Bayes e Bayes ingênuo

No teorema *Bayesiano* a classificação das probabilidades *a priori* estimadas à primeira abordagem apresenta maior incidência na classe *área navegável* nas aquisições lateral direita e esquerda, uma vez que os índices de cor e textura nestas são pouco variáveis, destacando a divisão de regiões de interesse à navegação. Nas imagens por aquisição frontal, a textura e variação de cor no solo são super-segmentadas, ocasionando crescimento de regiões não homogêneas e aumentando a incerteza da classificação.

Para o processamento no espaço de cores HSV, o grau de incerteza na classificação da área navegável é menor, uma vez que as regiões para a navegação apresentam baixa taxa de falsos negativos, ao contrário do que ocorre com o espaço de cores RGB, em que a classe *área não navegável* tende a apresentar maiores taxas de acerto. A quantidade de dimensões nesse caso provê melhores taxas de acerto, com classificação coerente das duas classes em HSV e, em menor grau, em RGB, mas como o conjunto de treinamento é maior, a predição requer maior tempo de processamento.

Já na segunda abordagem a probabilidade *a priori* dos conjuntos de treinamento denota maior estimativa dos frutos e gramíneas (avaliação), uma vez que irregularidades de textura e cor nestes objetos não são super-segmentadas, apresentando regiões homogêneas e distintas. Folhas, caules, ramos e galhos são segmentados com maior quantidade de regiões e, conseqüentemente, a probabilidade de sua classificação é incerta. Novamente, no espaço de

cores HSV, a taxa de acerto é maior, para a classificação de frutos, folhas, ramos e galhos. Em RGB, folhas, ramos e galhos são classificados como frutos, apresentando maiores taxas de falsos negativos.

Os resultados ainda desta abordagem, para o espaço de cores RGB mostram que a taxa de acerto é diretamente proporcional ao aumento do conjunto de treinamento, observado em 50% e 75%, já que os classificadores provêm conhecimento dos vetores mais distantes dos centróides das três classes. Por outro lado, no espaço de cores HSV, os conjuntos em 30%, 40% e 50% apresentam as melhores taxas de acerto, considerando que nestes casos, os classificadores usam a quantidade crescente de vetores e dimensões para a determinação da PDFc.

As tabelas abaixo mostram as probabilidades *a priori* estimadas dos conjuntos de treinamento para as duas abordagens.

Tabela 4.1: Probabilidade *a priori* estimadas dos conjuntos de treinamento de pomares.

Locomoção (pomares)								
Espaço de cores	Classes	10%	20%	30%	40%	50%	75%	100%
RGB	Área navegável	0,4129	0,4273	0,3128	0,4112	0,4223	0,4122	0,4187
	Área não navegável	0,5871	0,5727	0,6872	0,5888	0,5777	0,5878	0,5813
HSV	Área navegável	0,3593	0,4544	0,4384	0,5209	0,5232	0,7733	0,8274
	Área não navegável	0,6407	0,5456	0,5616	0,4791	0,4768	0,2267	0,1726

Na Tabela 4.1, acima, a probabilidade da primeira classe *Área navegável* para o caso RGB é estimada em menor índice que a segunda classe, para todos os conjuntos de treinamento. Para o caso HSV, a mesma primeira classe apresentou índice de estimação superior que a segunda classe para os conjuntos de treinamento em 40%, 50%, 75% e 100%.

Tabela 4.2: Probabilidade *a priori* estimadas dos conjuntos de treinamento de canaviais.

Locomoção (canaviais)								
Espaço de cores	Classes	10%	20%	30%	40%	50%	75%	100%
RGB	Área navegável	0,1242	0,6545	0,6753	0,6788	0,3423	0,4588	0,2398
	Área não navegável	0,8758	0,3455	0,3247	0,3212	0,6577	0,5412	0,7602
HSV	Área navegável	0,5565	0,5466	0,5648	0,5654	0,4456	0,6654	0,7543
	Área não navegável	0,4435	0,4534	0,4352	0,4346	0,5544	0,3346	0,2457

Na Tabela 4.2, a estimação das duas classes para as linhas de plantio de canaviais mostrou-se eficiente para a primeira classe no caso RGB, apenas para os conjuntos de treinamento em 20%, 30% e 40%, ao passo que, em 10%, 50%, 75% e 100%, a probabilidade da segunda classe é superior. No caso HSV, apenas no conjunto em 50%, a primeira classe é inferior, confirmando que para este espaço de cores, a probabilidade se confirma em maior índice para as áreas de regiões homogêneas, que são passíveis de navegação.

Tabela 4.3: Probabilidade *a priori* estimadas dos conjuntos de treinamento de milharais.

Locomoção (milharais)								
Espaço de cores	Classes	10%	20%	30%	40%	50%	75%	100%
RGB	Área navegável	0,2677	0,2556	0,4564	0,4666	0,4605	0,6221	0,6353
	Área não navegável	0,7323	0,7444	0,5436	0,5334	0,5395	0,3779	0,3647
HSV	Área navegável	0,2344	0,3673	0,3458	0,6567	0,6645	0,8733	0,8744
	Área não navegável	0,7656	0,6327	0,6542	0,3433	0,3355	0,1267	0,1256

Ainda na primeira abordagem, para a locomoção entre as linhas de plantio de milharais, a Tabela 4.3 denota aumento da taxa de acerto no caso RGB, para a primeira classe, sendo confirmada nos conjuntos de treinamento em 75% e 100%. No caso HSV, apenas nos conjuntos em 10%, 20% e 30% a segunda classe é superior, e nos demais conjuntos, dado o crescimento do número de dimensões, a estimação é mais eficiente na primeira classe.

Na Tabela 4.4, a probabilidade para a segunda abordagem (localização de frutos) apresenta maior estimativa da taxa de acerto nos conjuntos de treinamento em 50% e 75% para a primeira classe (*Frutos*), no caso RGB. Já os conjuntos de treinamento em 30%, 40% e 50% denotam melhores estimativas de acerto para a primeira classe, no caso HSV, em comparação com a segunda classe (*Folhas, ramos e galhos*). A terceira classe (*Céu*) apresenta baixa estimativa nos dois espaços de cores.

Tabela 4.4: Probabilidade *a priori* estimadas dos conjuntos de treinamento de laranjas.

Localização (laranjas)								
Espaço de cores	Classes	10%	20%	30%	40%	50%	75%	100%
RGB	Frutos	0,1324	0,1745	0,1234	0,2213	0,3943	0,3743	0,2485
	Folhas, ramos e galhos	0,8637	0,8211	0,8716	0,7733	0,5983	0,6233	0,7476
	Céu	0,0039	0,0044	0,0050	0,0054	0,0074	0,0024	0,0039
HSV	Frutos	0,1327	0,1745	0,3535	0,3613	0,3643	0,1452	0,1346
	Folhas, ramos e galhos	0,8640	0,8191	0,6397	0,6331	0,6259	0,8460	0,8620
	Céu	0,0033	0,0064	0,0068	0,0056	0,0098	0,0088	0,0034

Como a classificação de gramíneas/milhos por modelo estatístico é parcial e depende de outras técnicas de processamento de imagens, que não as usadas em frutos, essas não foram estimadas.

Com estas tabelas foi possível parametrizar as esperanças de cada classe para a abordagem pretendida, em que nos dois espaços de cores apresentam similaridade no número de segmentos da classe folhas e frutos, com menos dimensões. Já na primeira abordagem, as classes área navegável e não navegável demonstraram alta regularidade e disparidade, favorecendo a classificação estatística.

A classificação pelo teorema *Bayes Ingênuo* apresenta majoritariamente os mesmos resultados para ambas as abordagens. Na primeira, o comportamento quanto à taxa de acerto e sua relação com a quantidade de dimensões é idêntico para o caso RGB do classificador *Bayesiano*, uma vez que a dispersão das classes aumenta com a percentagem dos conjuntos de

treinamento. Com considerável diferença, nos conjuntos 75% e 100%, tanto para RGB, como para HSV, as taxas de acerto foram melhores, pois como as características são independentes, a suposição da classificação de objetos a determinada classe é dada pelo grau de incerteza da PDF e sua distribuição normal, relacionando as distâncias de *Bhattacharyya* e *Jeffries-Matusita*.

Na segunda abordagem, o mesmo problema acontece com o aumento de dimensões e a taxa de acerto. O tempo de processamento também é comprometido, mas conforme as percentagens dos conjuntos aumentam para RGB e HSV, as classes são classificadas de forma correta.

Em porcentagens menores para o caso HSV, a classificação não foi linear para as últimas dimensões dos conjuntos de treinamento, porém sua média na taxa de acerto é maior no caso RGB, uma vez que a classificação destes segmentos é prejudicada pela quantidade de segmentos e suas correspondências com as classes erradas.

Os gráficos abaixo mostrados nas figuras Figura 4.18, Figura 4.19, Figura 4.20, Figura 4.21, Figura 4.22, Figura 4.23, Figura 4.24 e Figura 4.25, apresentam os parâmetros e vetores na distribuição da mistura dos dados e na estimativa dos conjuntos de treinamento para cada percentual, espaço de cor e abordagens. Os parâmetros independentes (vermelho) simbolizam as taxas de acerto da razão da estimativa das classes alvo para com o crescimento do número de dimensões das classes de entrada, ao classificador *Bayes ingênuo*. Os vetores de características (azul) também associados à taxa de acerto com o número de dimensões e as classes estimadas, demonstram os resultados do classificador *Bayes*.

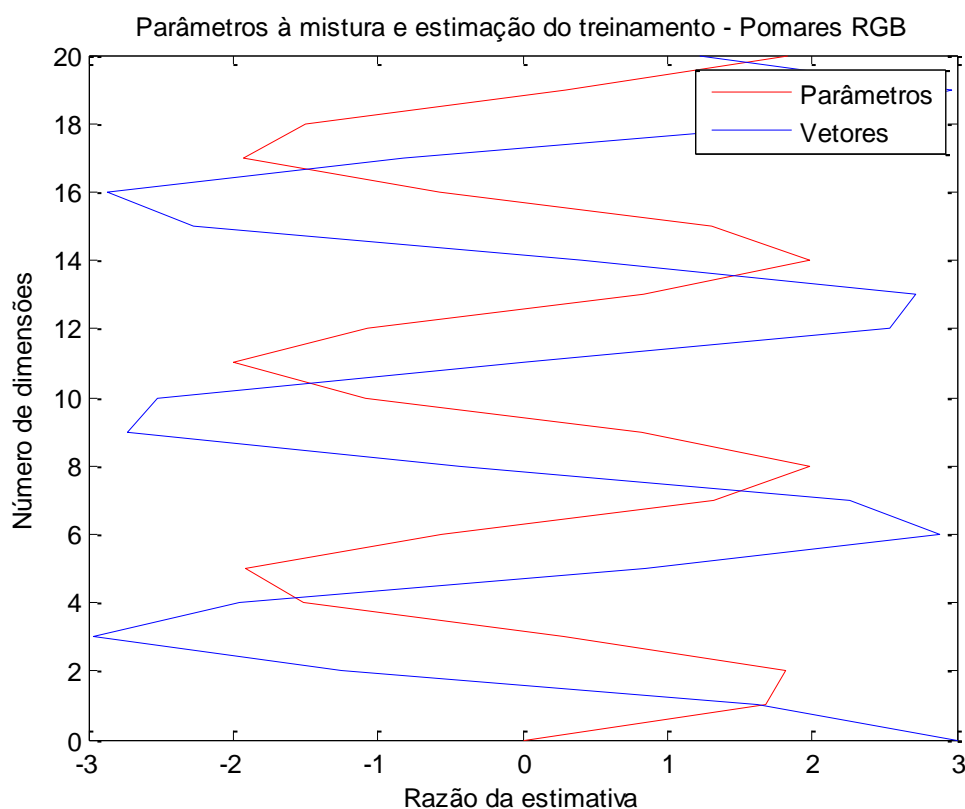


Figura 4.18: Parâmetros à mistura do treinamento – Pomares RGB.

Na Figura 4.18, acima, a estimativa dos classificadores *Bayes* e *Bayes ingênuo* para a primeira abordagem, em pomares e no caso RGB, dá-se pela baixa variação da razão probabilística dos parâmetros independentes ao modelo ingênuo, em comparação com os vetores de características até o número máximo (20) de dimensões. Tal razão simboliza que, quanto menor a variação tanto dos vetores de características, como dos parâmetros independentes, maior a taxa de acerto até o número desejado de dimensões.

Na Figura 4.19, a seguir, a razão da estimativa em relação ao número de dimensões é menor para o classificador *Bayes ingênuo*, no espaço de cores HSV, ainda para as linhas de plantio de pomares. Em comparação com o caso RGB, os índices de probabilidade apresentam melhores taxas de acerto neste espaço de cores, uma vez que a PDF e a distribuição normal associam os objetos em suas respectivas classes com menor grau de incerteza.

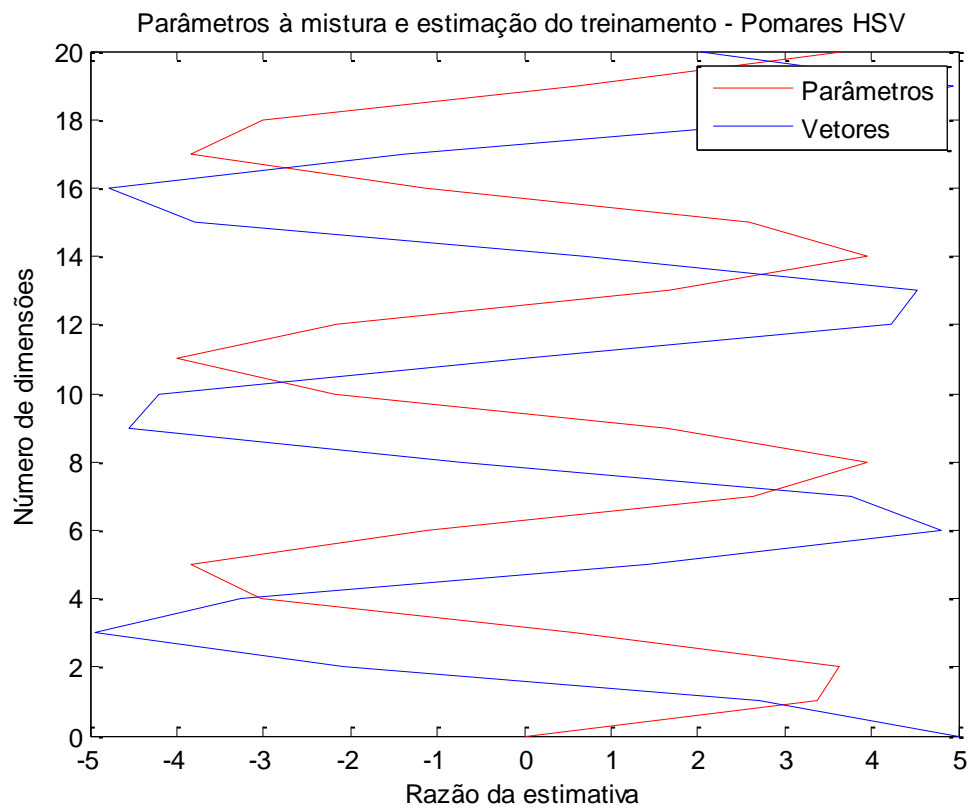


Figura 4.19: Parâmetros à mistura do treinamento – Pomares HSV.

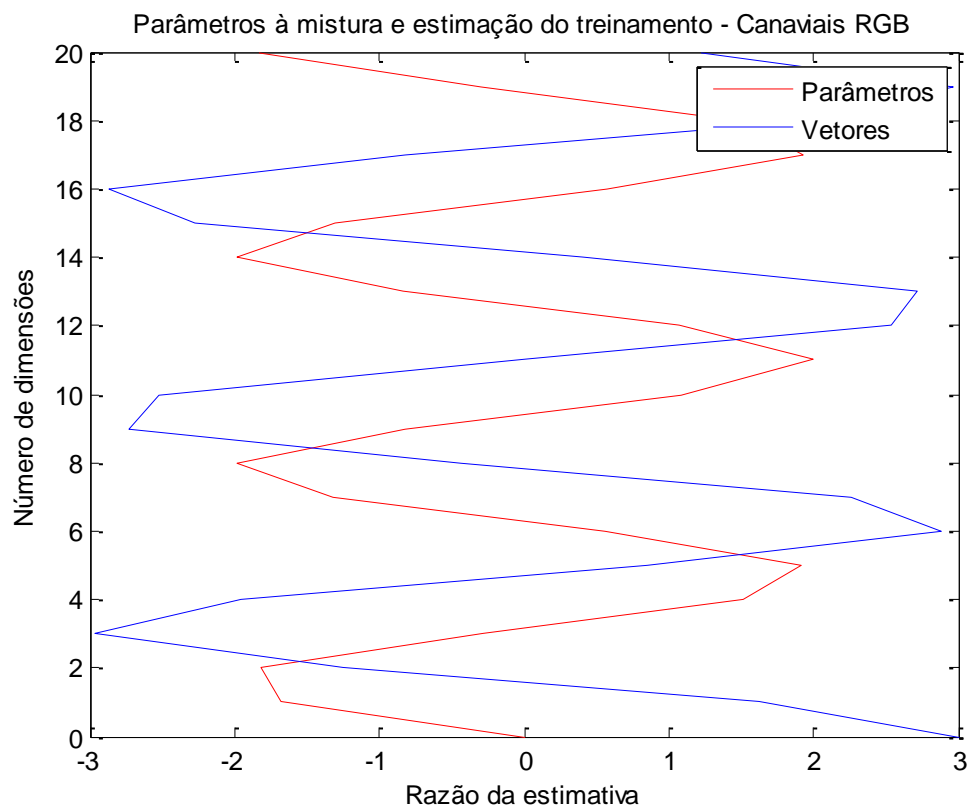


Figura 4.20: Parâmetros à mistura do treinamento – Canaviais RGB.

Para a locomoção entre as linhas de plantio de canaviais, na Figura 4.20, no caso RGB, e na Figura 4.21, no caso HSV, a menor razão da estimativa de classificação em relação com o número de dimensões é dada pelo classificador *Bayes ingênuo*. Novamente, o classificador *Bayes* foi ineficaz para todas as dimensões, pois a PDF calculada para cada dimensão não atende corretamente a distribuições de confiança para as duas classes, ou seja, a discriminação de áreas para a navegação (classe *Área navegável*) é associada às áreas de plantio (classe *Área não navegável*).

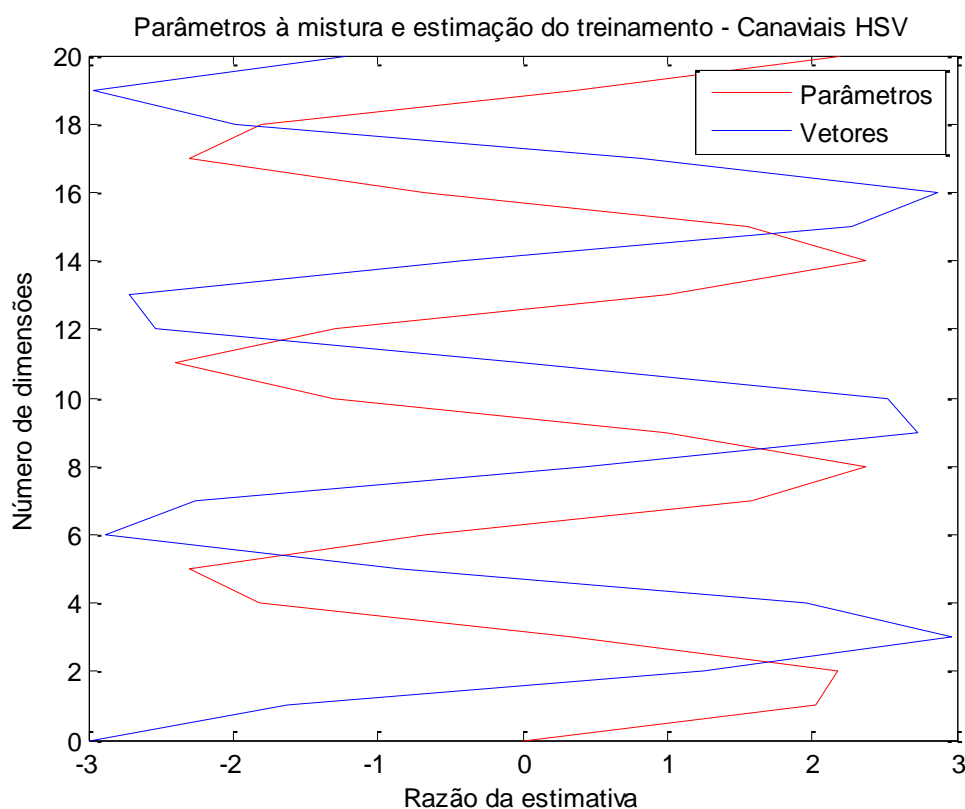


Figura 4.21: Parâmetros à mistura do treinamento – Canaviais HSV.

Em milhares, para o caso RGB (Figura 4.22), o classificador *Bayes* obteve melhor desempenho da razão da estimativa, associando mais componentes das distribuições normais da PDF de cada dimensão às áreas de navegação e plantio, com suas respectivas classes. Para o caso HSV (Figura 4.23), a menor razão de estimativa para a obtenção das melhores taxas de acerto é dada pelo classificador *Bayes ingênuo*.

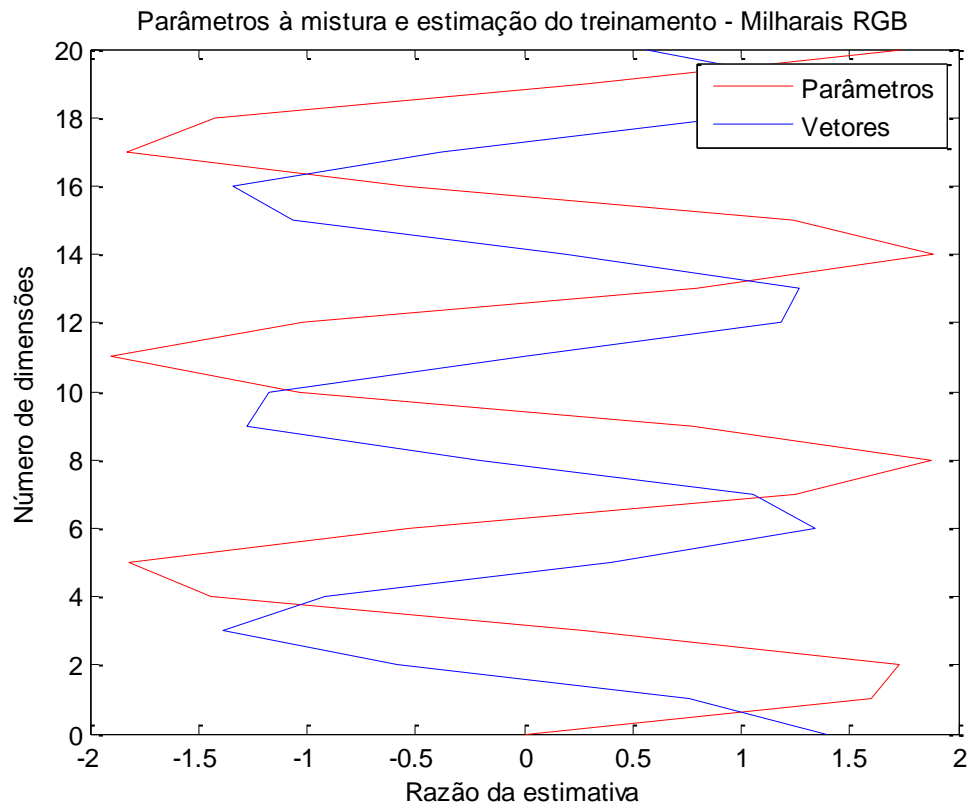


Figura 4.22: Parâmetros à mistura do treinamento – Milharais RGB.

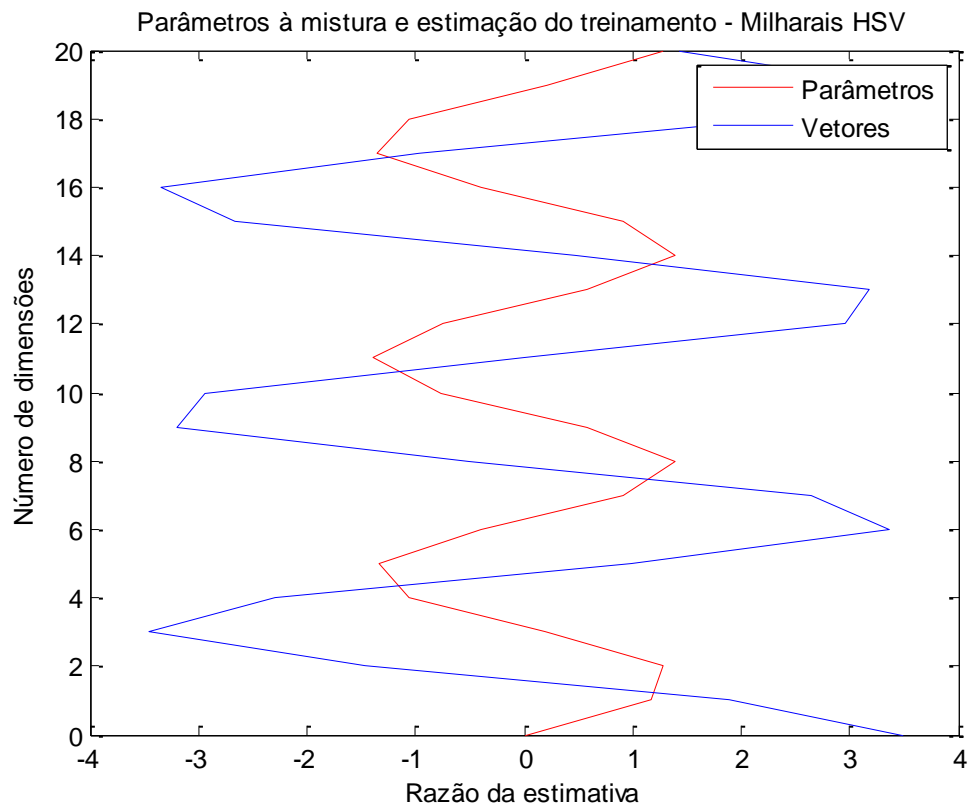


Figura 4.23: Parâmetros à mistura do treinamento – Milharais HSV.

Na Figura 4.24, para a segunda abordagem, no caso RGB, os melhores resultados foram obtidos com o classificador *Bayes*, apresentando menor razão de estimativa em relação ao número de componentes analisadas. Neste espaço de cores, a estimação no reconhecimento dos objetos relacionados aos frutos é dada pela PDF de cada dimensão, corrigindo os valores atuais pela esperança de cada área não correspondida à respectiva classe.

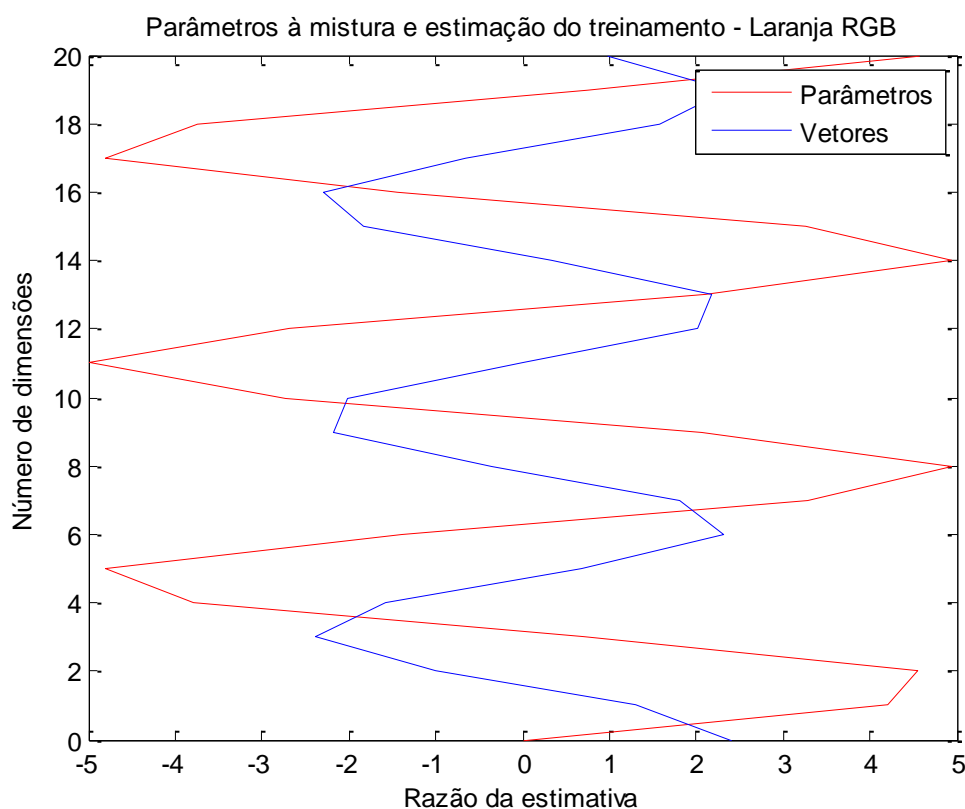


Figura 4.24: Parâmetros à mistura do treinamento – Laranja RGB.

Na Figura 4.25, o reconhecimento de frutos para o caso HSV apresenta equilíbrio nos resultados dos dois classificadores, mas com compensação da taxa de acerto, para margens menores da razão de estimativa ao classificador *Bayes*, sendo que este permite a correção dos próximos resultados pela aproximação das estimativas *a priori*, na PDF de cada dimensão.

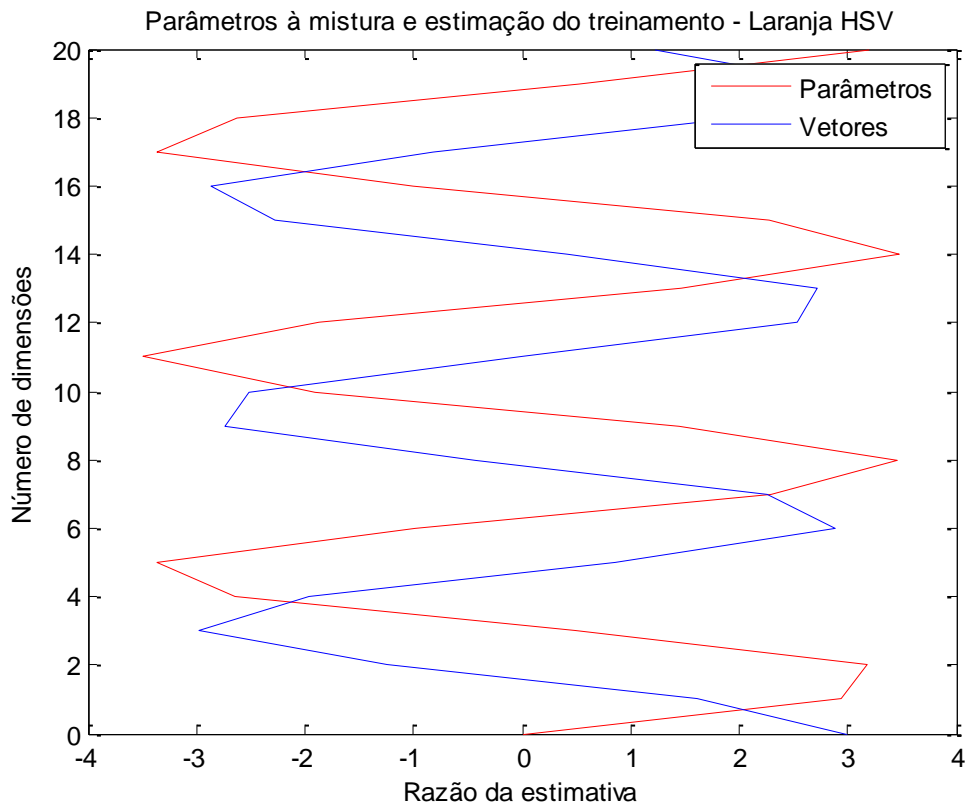


Figura 4.25: Parâmetros à mistura do treinamento – Laranja HSV.

É possível observar que, a razão da estimativa deve ser menor para o aumento do número de dimensões e sua posterior classificação, para todos os casos.

4.5.4. Avaliação dos classificadores

A análise das taxas de acerto e dos gráficos dos conjuntos de treinamento indica que os classificadores *Bayes ingênuo* e RNA para o espaço de cores HSV obtiveram melhor desempenho na abordagem de localização de frutos, ao passo que os classificadores *Bayes* para o caso HSV em canaviais e milharais, e para caso RGB em pomares, e *Bayes ingênuo* para o espaço de cores RGB obtiveram melhores resultados na abordagem da locomoção autônoma. A taxa de acerto para todos os classificadores se torna relativamente equiparável à medida que os segmentos para a classe folhas, galhos e ramos é separada da classe frutos, e a classe área não navegável para os segmentos que denotam o céu, também é distinta da região de contorno onde é limitada a locomoção do robô móvel.

Quanto ao espaço de cores, os quatro tipos de classificadores para a primeira abordagem mostraram comportamentos semelhantes, no caso RGB, que a taxa de acerto em relação ao número de dimensões cresceu menos que no caso HSV, entre 10% e 50%. No caso HSV, a taxa de acerto cresceu menos que a RGB, entre 50% e 100%. Já na segunda abordagem, o espaço de cores HSV foi consideravelmente superior ao RGB, pois a luminosidade e os efeitos de sombreamento foram eliminados durante a segmentação e, posteriormente pela ACP na normalização dos vetores de características.

Os classificadores estatísticos que usam a PDFc (*Funções de densidade conjunta de probabilidade*) apresentaram melhores resultados que os cognitivos RNA e *Fuzzy*, na primeira abordagem, e na segunda abordagem este mesmo classificador é equiparável em número de iterações das topologias RNA, pois conseguiram distinguir com precisão as classes alvo.

Abaixo, nas figuras 4.26 a 4.29, são mostrados os gráficos para as duas abordagens. A avaliação dos classificadores dá-se pela relação entre a quantidade de dimensões dos vetores e a taxa de acerto. Quanto mais próximo de 20 vetores de características o classificador atingir a correspondência à maior taxa de acerto, melhor este será considerado, uma vez que antes da normalização pela ACP, o critério discriminado para a classificação cognitiva e estatística é a redução e padronização dos vetores de características. Se a quantidade de dimensões for tão menor quanto a pretendida e alcançar a máxima taxa de acerto, conclui-se que a classificação não é exata, considerando que para o problema de locomoção a área não navegável é considerada como navegável, e folhas, ramos e galhos são considerados como frutos, na segunda abordagem.

Caso a quantidade de dimensões seja maior que a pretendida, os classificadores neste intervalo necessitam de mais dados aos conjuntos de treinamento e teste para classificar de forma correta.

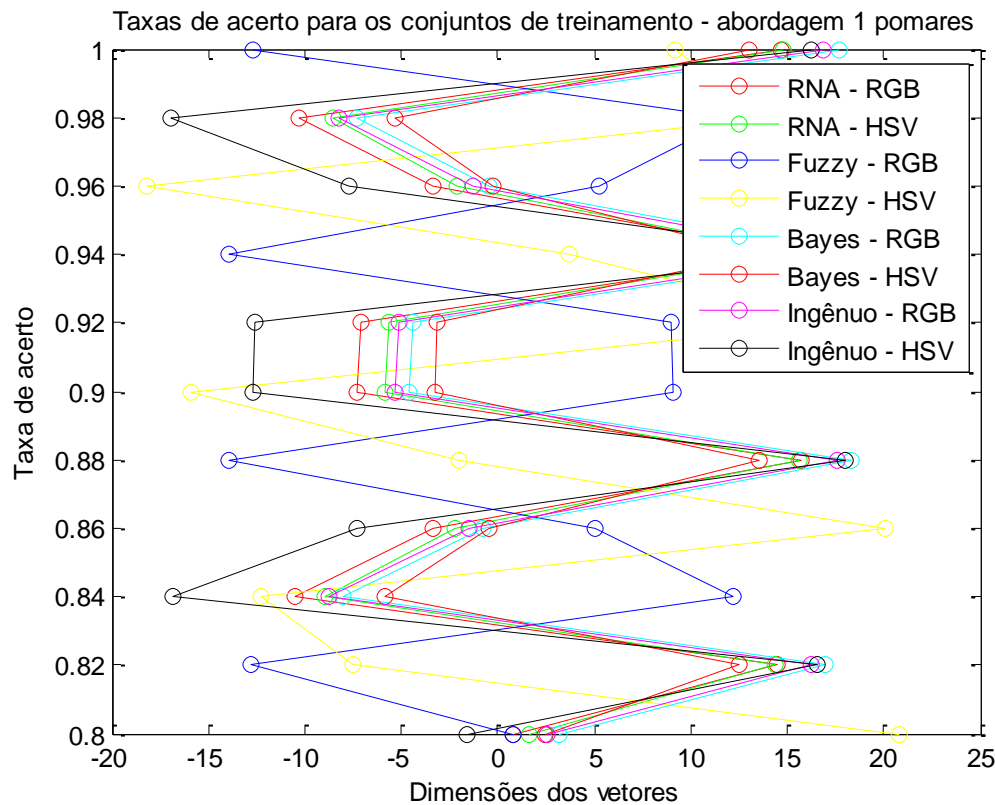


Figura 4.26: Taxa de acerto para os conjuntos de treinamento – abordagem 1 (pomares).

A Figura 4.26, acima, apresenta a taxa de acerto para a primeira abordagem, em locomoção através dos pomares. Observa-se que o classificador *Bayes* para RGB obtém a melhor condição de classificação, sendo equiparável à sua variação (*Ingênuo*) também ao mesmo espaço de cores. Todavia, para variação da taxa de acerto entre 0,8 e 1,0, a média de classificação, considerando o número máximo de dimensões, os melhores resultados são fornecidos pelo classificador *Bayes ingênuo*.

Para as linhas de plantio de canaviais (Figura 4.27), o classificador *Bayes* para HSV obteve melhor desempenho que os demais, com compatibilidade ao classificador *Bayes Ingênuo* para RGB, em relação ao número máximo de dimensões. A mesma análise vale para as linhas de plantio em milharais, na Figura 4.28. A média de acerto entre os classificadores *Bayes* (HSV) e *Bayes ingênuo* (RGB) denota que, tanto para milharais, quanto para canaviais, a classificação estatística é mais precisa que a cognitiva.

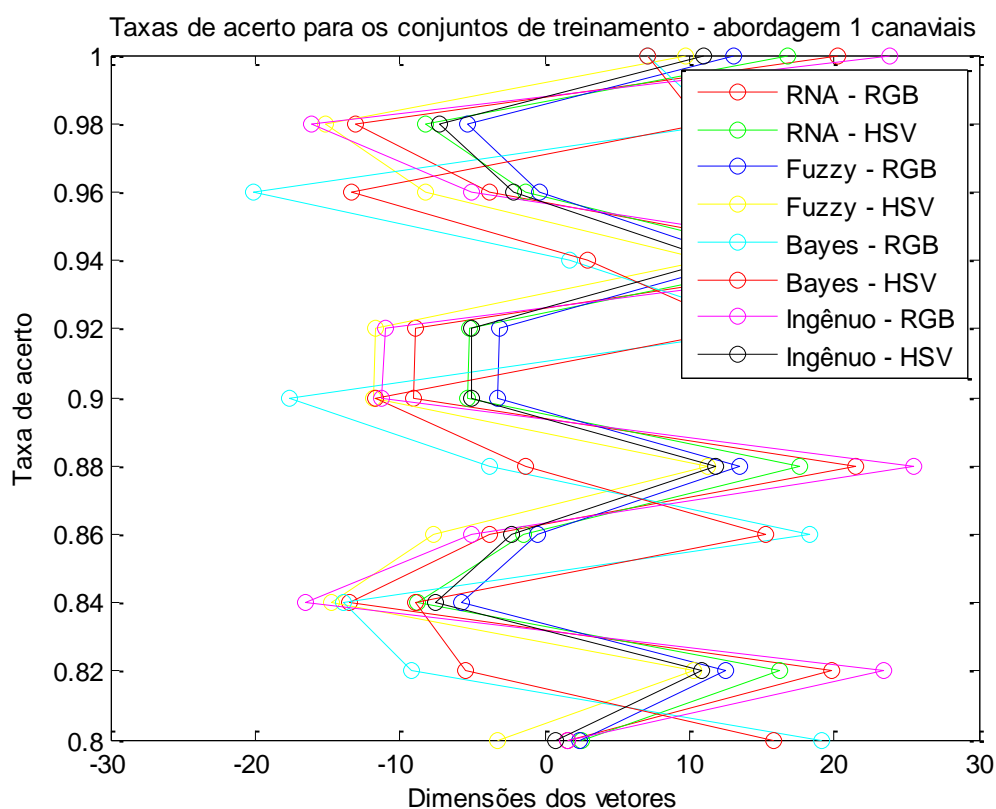


Figura 4.27: Taxa de acerto para os conjuntos de treinamento – abordagem 1 (canaviais).

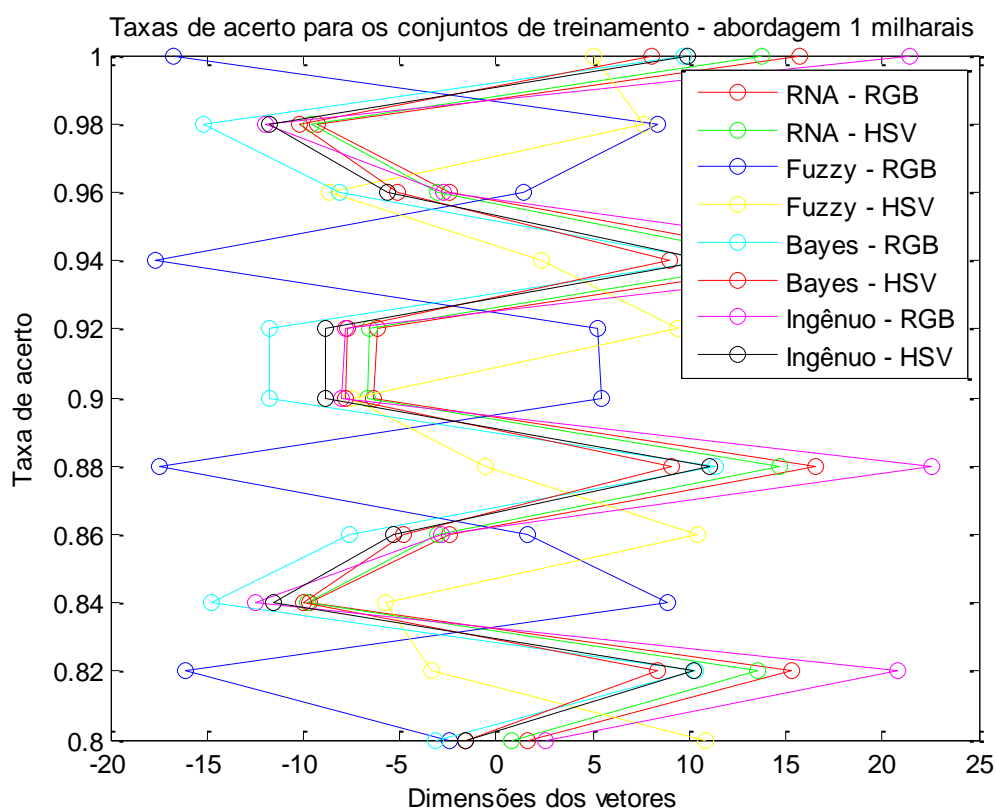


Figura 4.28: Taxa de acerto para os conjuntos de treinamento – abordagem 1 (milharais).

Na segunda abordagem, como apresentado pela Figura 4.29, a seguir, as melhores taxas de acerto para com a quantidade de dimensões pretendida são dos classificadores RNA e *Bayes Ingênuo* para o caso HSV. A RNA, neste caso, por ter número de iterações equiparáveis ao modelo probabilístico de *Bayes*, apresenta as melhores condições na discriminação das classes reconhecidas e regiões segmentadas.

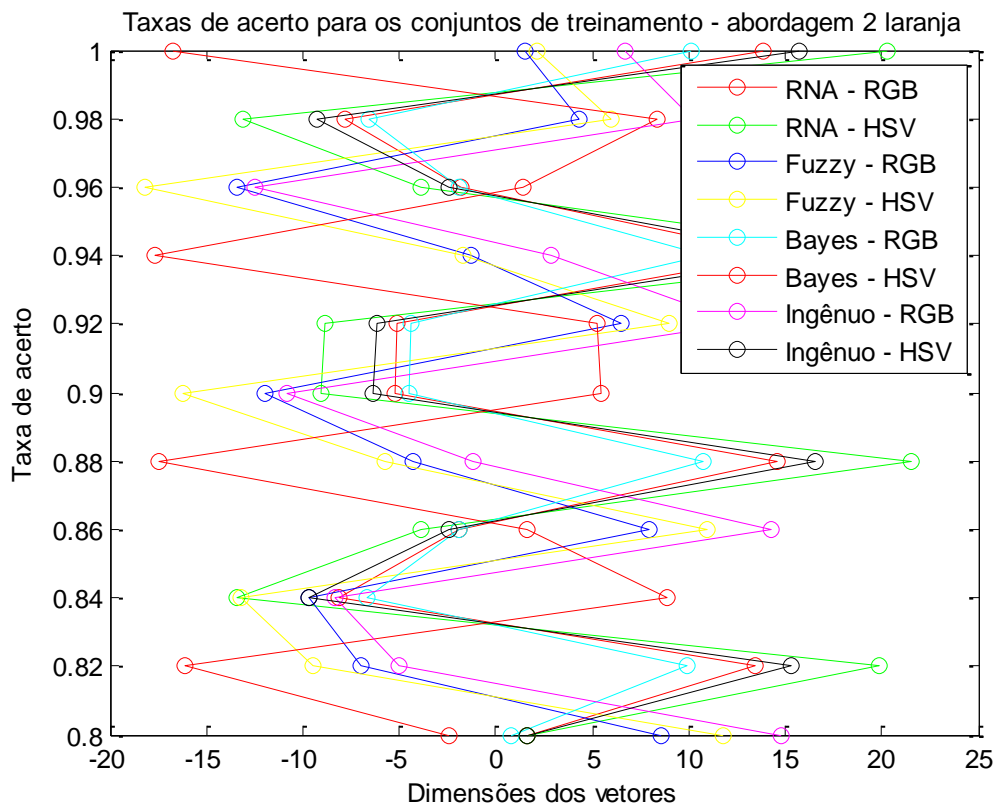


Figura 4.29: Taxa de acerto para os conjuntos de treinamento – abordagem 2 (laranja).

4.6. Mapas de classes para reconhecimento de padrões

Os mapas de classe (**APÊNDICE E**) representam a formação de regiões classificadas em sua maioria pelas classes pretendidas de cada abordagem. São apresentadas imagens, através das figuras 4.30 a 4.49, para cenas de laranjas, cenas de milhos, e as linhas de plantio (pomares, canaviais e milharais), e suas respectivas classificações em Redes Neurais

Artificiais, Fuzzy preditivo, filtro de Bayes e filtro de Bayes ingênuo. Os dois espaços de cores RGB e HSV também são amostrados.

Laranja – imagem 01

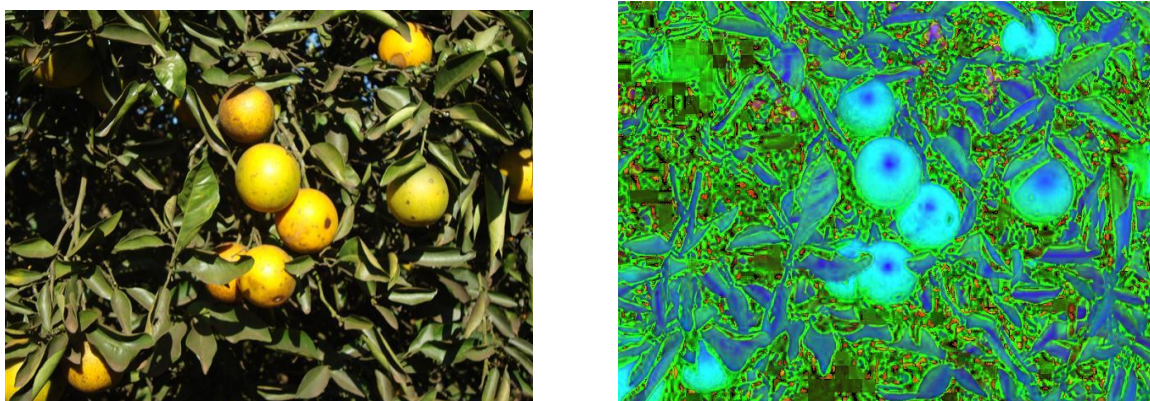


Figura 4.30: Imagem original RGB e HSV (laranja 01).

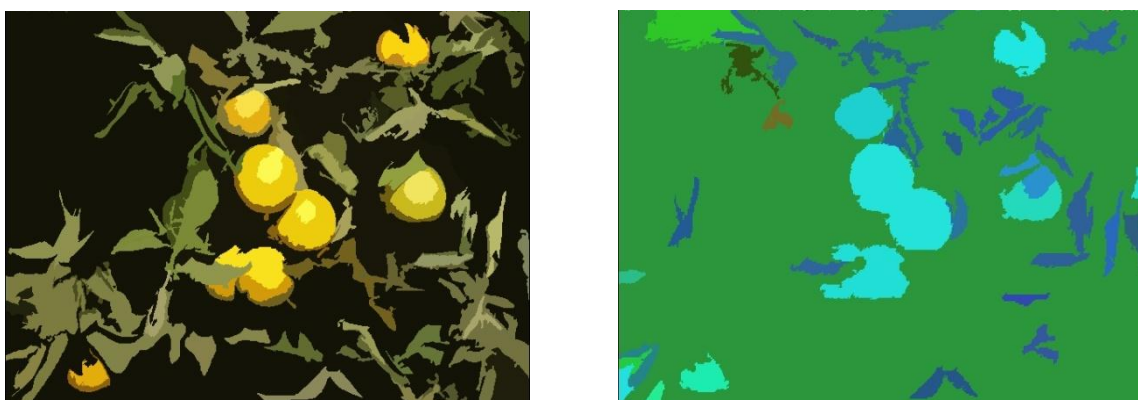


Figura 4.31: Mapa de classe 1 (Bayes RGB e HSV).



Figura 4.32: Mapa de classe 1 (Bayes ingênuo RGB e HSV).

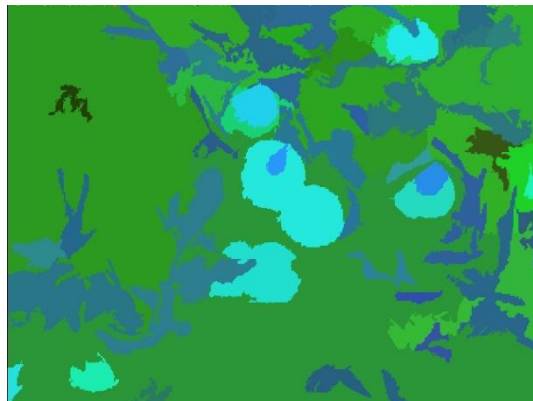


Figura 4.33: Mapa de classe 1 (Fuzzy preditivo RGB e HSV).

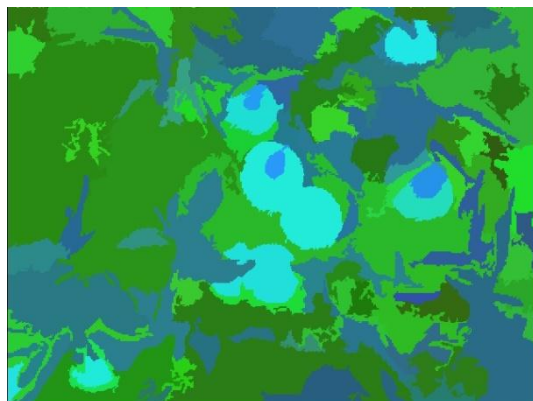


Figura 4.34: Mapa de classe 1 (RNA RGB e HSV).

Linhas de plantio – Pomares 01

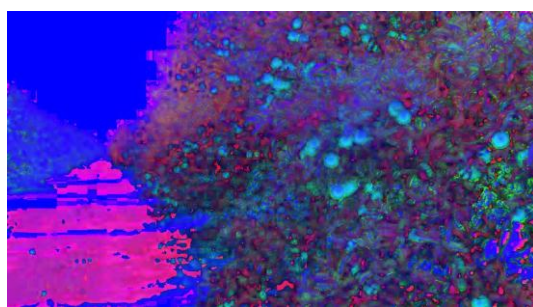


Figura 4.35: Imagem original RGB e HSV (pomares 01).

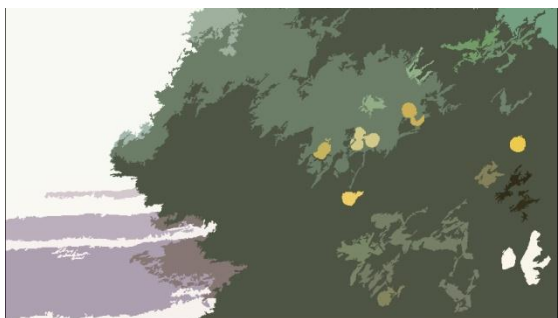


Figura 4.36: Mapa de classe 1 (Bayes RGB e HSV).



Figura 4.37: Mapa de classe 1 (Bayes ingênuo RGB e HSV).



Figura 4.38: Mapa de classe 1 (Fuzzy preditivo RGB e HSV).



Figura 4.39: Mapa de classe 1 (RNA RGB e HSV).

Linhas de plantio – Milharais 01

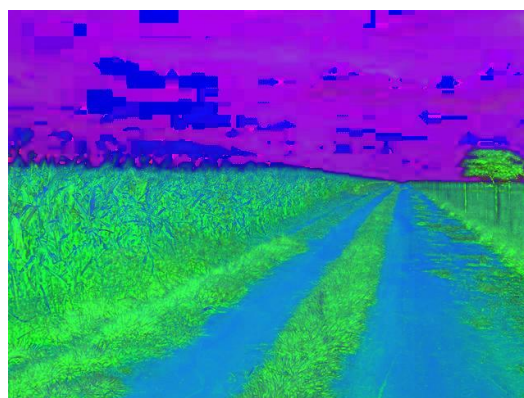


Figura 4.40: Imagem original RGB e HSV (milharais 01).

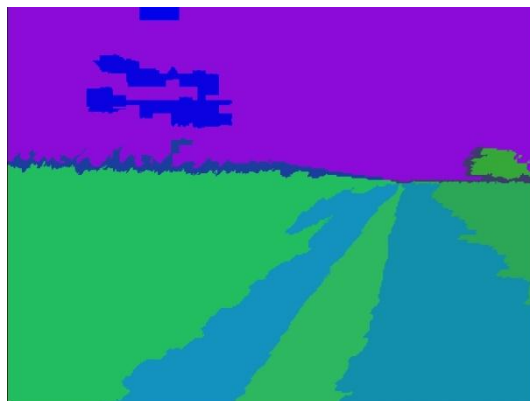
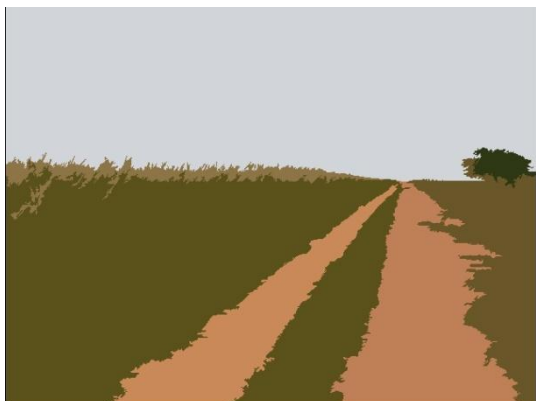


Figura 4.41: Mapa de classe 1 (Bayes RGB e HSV).

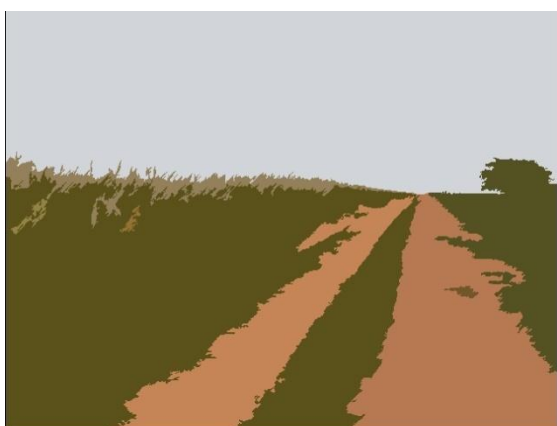


Figura 4.42: Mapa de classe 1 (Bayes ingênuo RGB e HSV).

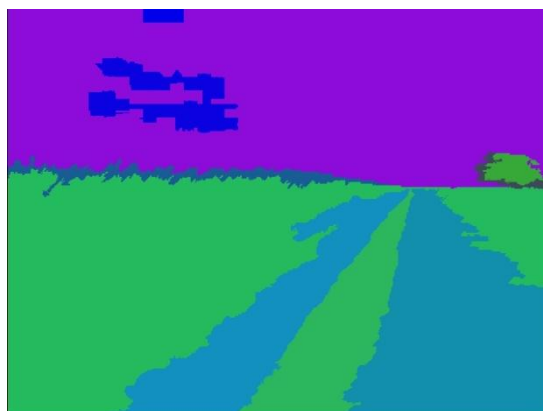


Figura 4.43: Mapa de classe 1 (Fuzzy preditivo RGB e HSV).

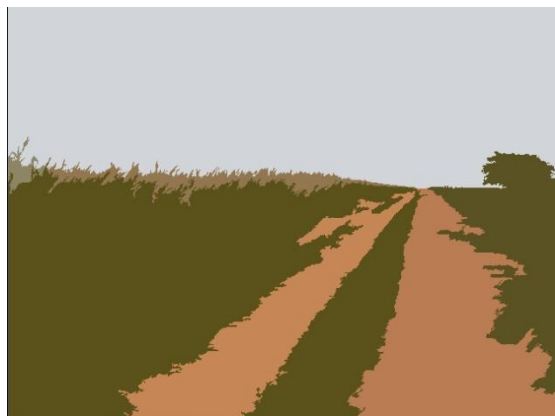


Figura 4.44: Mapa de classe 1 (RNA RGB e HSV).



Linhas de plantio – Canaviais 01



Figura 4.45: Imagem original RGB e HSV (canaviais 01).

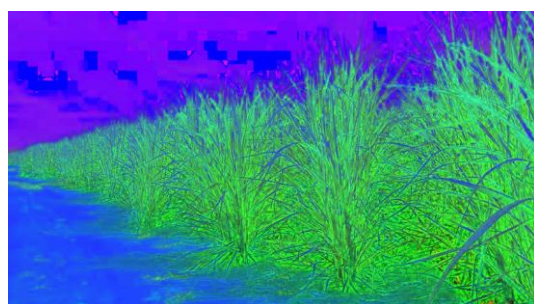


Figura 4.46: Mapa de classe 1 (Bayes RGB e HSV).

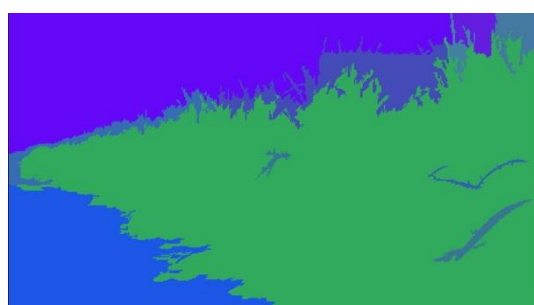
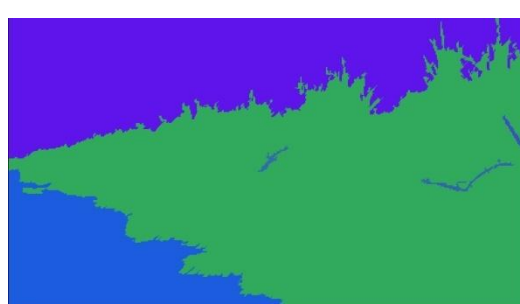


Figura 4.47: Mapa de classe 1 (Bayes ingênuo RGB e HSV).



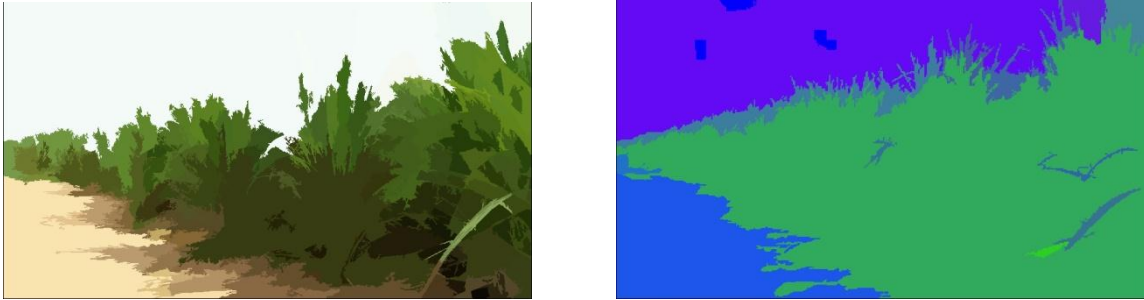


Figura 4.48: Mapa de classe 1 (Fuzzy preditivo RGB e HSV)

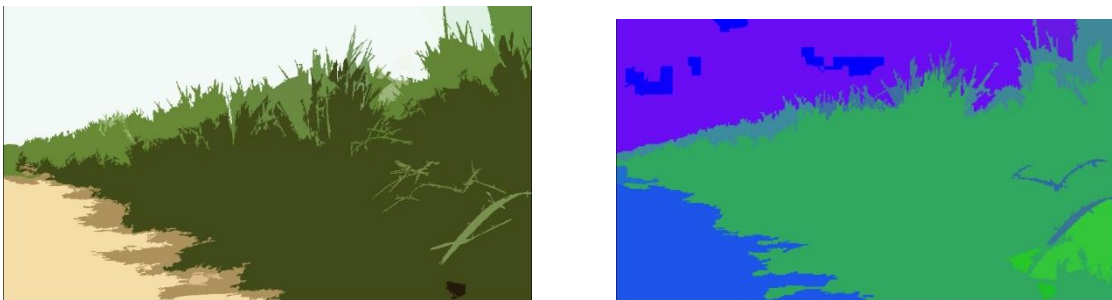


Figura 4.49: Mapa de classe 1 (RNA RGB e HSV).

4.7. Pós-processamento na quantificação de objetos

Esta etapa é direcionada apenas à segunda abordagem, para a localização e quantificação de frutos. As técnicas usadas foram customizadas do *Image Processing Toolbox (MATLAB)*.

Os mapas de classe são processados, uma vez que a representação pelo preenchimento de área (*floodfill*) agrega apenas as regiões sólidas que são quantificadas.

Inicialmente, é realizada uma conversão em níveis de cinza da imagem, para limiarizar as regiões que serão contornadas. Em seguida, para a determinação dos rótulos dos elementos conectados, devem-se excluir objetos com área maior a 200 ou 300 *pixels*, dependendo da distância focal. Logo, é necessário identificar cada elemento menor que este limiar, e calcular as propriedades destes objetos, tais como área, centróide, e a região de contorno. Com isso, os objetos que apresentarem áreas próximas à geometria circular serão rotulados e quantificados como frutos.

Para a determinação das métricas e delimitação dos objetos passíveis a cultura de laranja, a segmentação baseada em representação de grafos (GONZALEZ e WOODS, 2007) foi

aplicada. Esta provê a relação de adjacência A entre os valores binários dos *pixels*, e suas respectivas posições, destacando as propriedades locais geométricas da imagem.

O conjunto de *pixels* $A(p)$ representa ao agrupamento de *pixels* adjacentes ao *pixel* p em relação a A , ou seja:

$$q \in A(p) = (p, q) \in A \quad (3.12)$$

Assim, a relação de adjacência provê a definição de um grafo $G = (D, A)$ para a imagem, onde D é o agrupamento de *pixels*, e π uma sequência de *pixels* adjacentes (p_1, p_2, \dots, p_n) , para $(p_i, p_{i+1}) \in A, i = 1, 2, \dots, n-1$. Logo, a formação do grafo atende ao seguinte sistema.

$$\begin{cases} (p, q) \in A, d(p, q) \leq p \\ (p, q) \in A, (q - p) \in [(-1, 1), (1, -1)] \\ (p, q) \in A, |x_p - x_q| + |y_p - y_q| \leq 1 \rightarrow |f(p) - f(q)| \leq l \end{cases} \quad (3.13)$$

Com isso, para a relação simétrica da delimitação das regiões, sem variações de translação durante o agrupamento dos *pixels*, são usadas vizinhanças de *pixel* para três situações:

Vizinhança-4: $p = 1$; Vizinhança-8: $p = \sqrt{2}$; Dispersão de *pixels* vizinhos: $p = \sqrt{5}$.

No primeiro caso, áreas correspondentes às regiões pequenas, como frutos parcialmente ocultos (laranjas) de textura e cor equivalentes às folhas são excluídas. No segundo caso, os elementos de interesse já são agrupados em sua totalidade, quando ocupam segmentos representando as formas passíveis de laranja. No terceiro caso, o agrupamento é feito para regiões que detectam dois ou mais segmentos de áreas passíveis a laranja.

Com esta formulação, é possível realizar a busca por regiões de diversos tamanhos, bem como, no caso de laranjas, se os segmentos correspondem à geometria circular. A quantificação é realizada também para o total de regiões de interesse agrupadas na imagem.

A segmentação baseada em grafos é tratada como etapa de pós-processamento posterior à criação dos mapas de classe pelos classificadores, uma vez que com as regiões devidamente

classificadas, torna-se viável a aplicação do algoritmo de conexão das componentes principais de saída.

Como os melhores resultados de classificação, relacionados à segunda abordagem foram através da RNA no espaço de cores HSV, apenas os mapas de classe resultantes destes classificadores serão apresentados para a localização e quantificação de objetos, com comparação ao caso RGB.

A seguir, para os casos RGB e HSV são apresentadas, através das figuras 4.50 a 4.61, as imagens nos seus respectivos mapas de classe, o pré-processamento para a limiarização com áreas menores que 100 e maiores que 300, as métricas de aproximação geométrica para a detecção de objetos circulares, as regiões de contorno com os centróides de cada objeto, e finalmente o rótulo associado ao fruto.

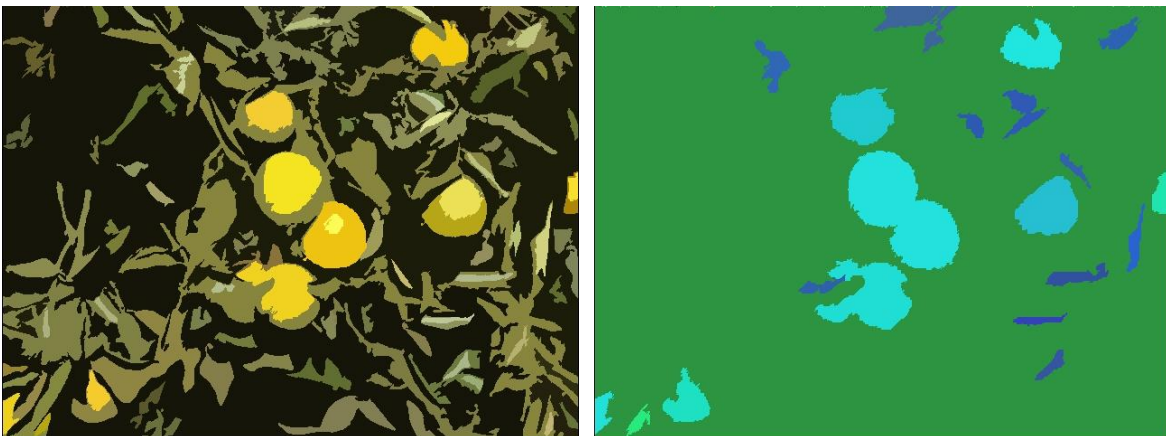


Figura 4.50: Mapas de classe RGB (esquerda) e HSV (direita) – cena 1.

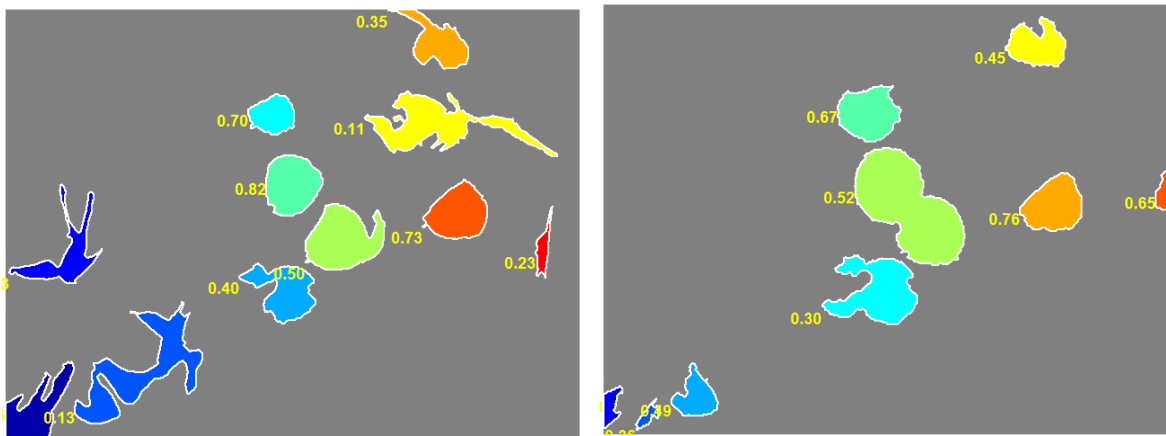


Figura 4.51: Métricas próximas do limiar geométrico circular 1,0 para RGB (esquerda) e HSV (direita) – cena 1.



Figura 4.52: Representação da área e centróide dos objetos para RGB (esquerda) e HSV (direita) – cena 1.

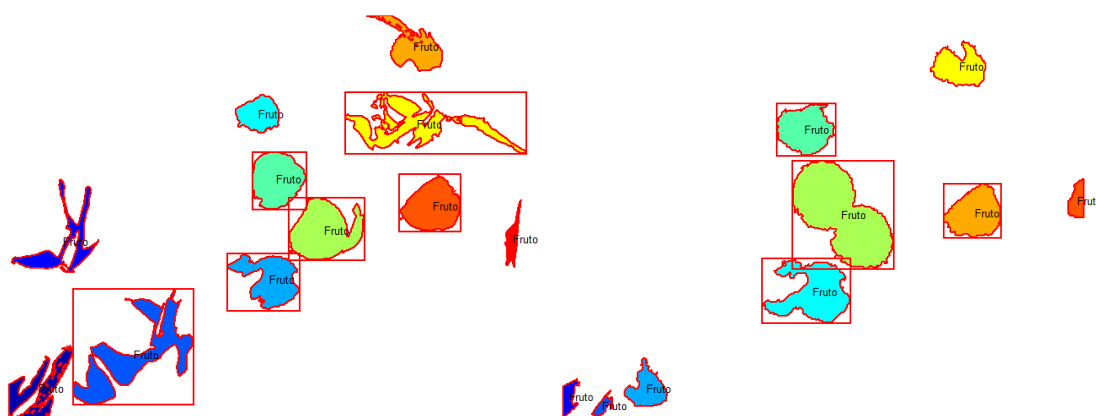


Figura 4.53: Rótulo da associação do elemento fruto nos objetos encontrados aos dois casos – cena 1.



Figura 4.54: Mapas de classe RGB (esquerda) e HSV (direita) – cena 2.

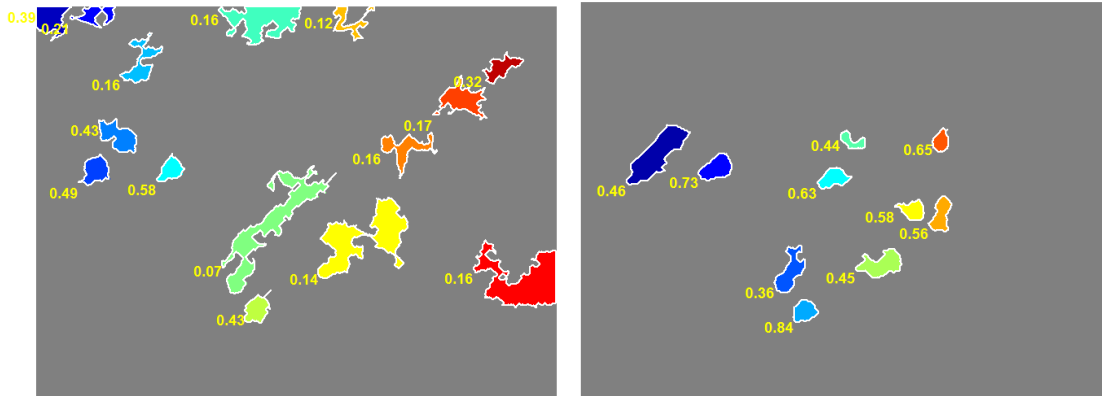


Figura 4.55: Métricas próximas do limiar geométrico circular 1,0 para RGB (esquerda) e HSV (direita) – cena 2.

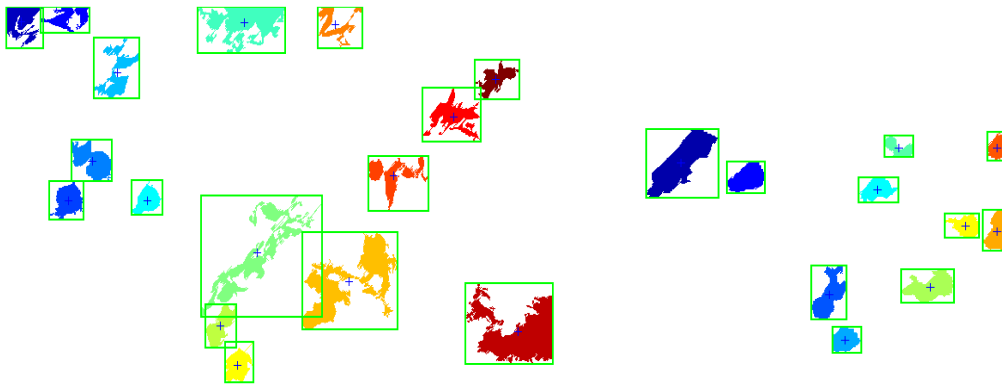


Figura 4.56: Representação da área e centróide dos objetos para RGB (esquerda) e HSV (direita) – cena 2.

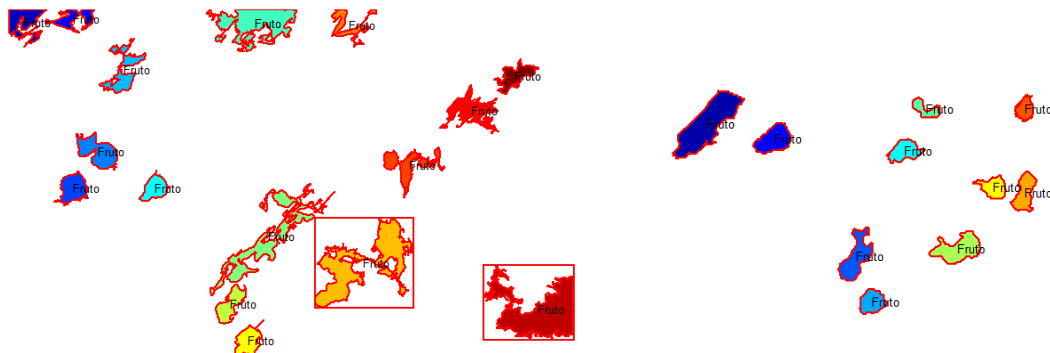


Figura 4.57: Rótulo da associação do elemento fruto nos objetos encontrados aos dois casos – cena 2.

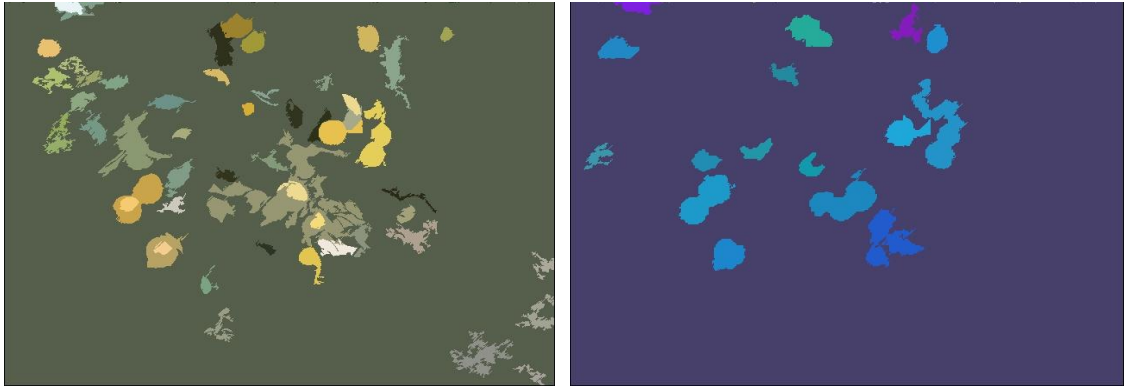


Figura 4.58: Mapas de classe RGB (esquerda) e HSV (direita) – cena 3.

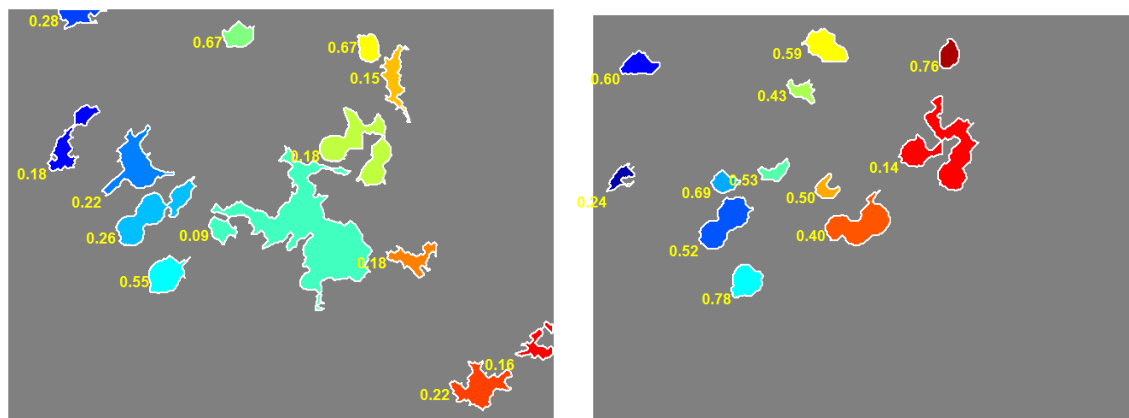


Figura 4.59: Métricas próximas do limiar geométrico circular 1,0 para RGB (esquerda) e HSV (direita) – cena 3.

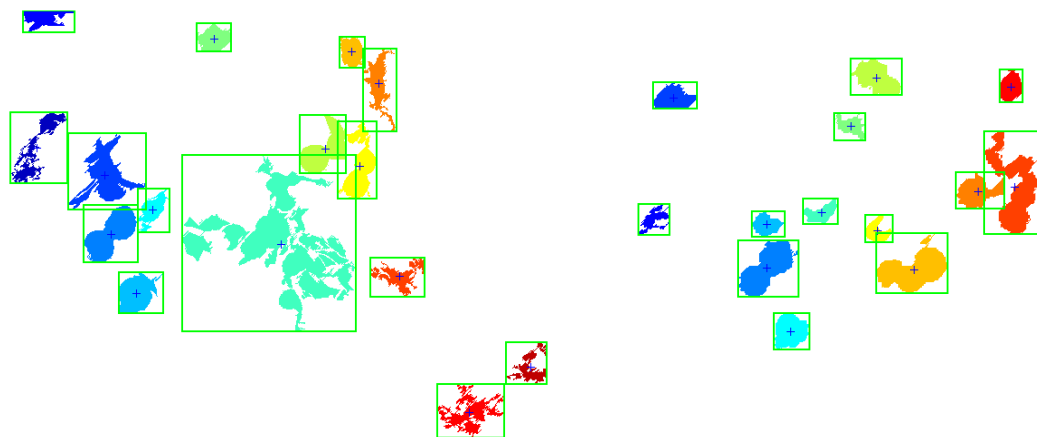


Figura 4.60: Representação da área e centróide dos objetos para RGB (esquerda) e HSV (direita) – cena 3.

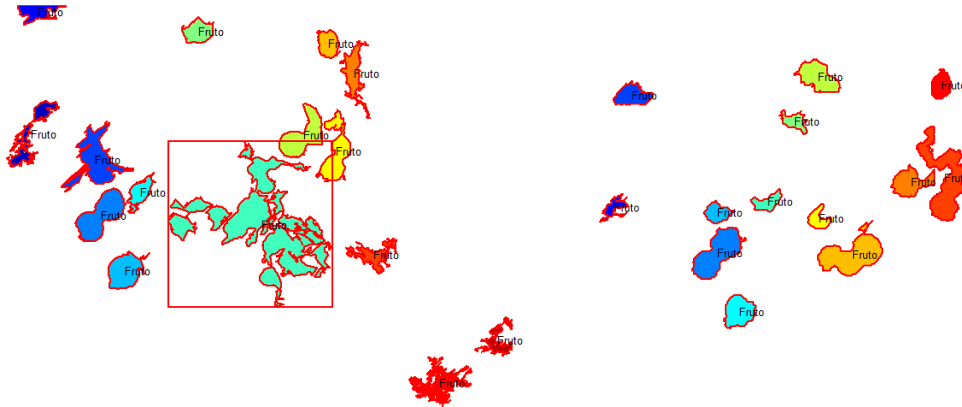


Figura 4.61: Rótulo da associação do elemento fruto nos objetos encontrados aos dois casos – cena 3.

As funções do *MATLAB* usadas para a rotulação customizadas com o algoritmo desenvolvido para a segmentação baseada em representação de grafos são: *bwlabel* (rótulo dos elementos conectados – identificador e quantidade), *regionprops* (cálculo das propriedades de cada objeto – área, centróide e as regiões de contorno), *bwareaopen* (eliminar áreas menores que determinado valor de *pixel* – conectividade de 4, 8, 6, 18 e 26), *bwboundaries* (impressão do perímetro de cada objeto).

No **APÊNDICE F** são apresentados os resultados de localização e quantificação para 20 imagens com os mapas de classe em RGB e HSV ao classificador *Bayes Ingênuo*.

4.8. Considerações finais

Neste capítulo os resultados para a aquisição e tratamento inicial das imagens foram apresentados, bem como as demais etapas da arquitetura proposta de visão computacional, tanto na abordagem de locomoção do robô móvel agrícola, quanto da localização e inspeção de frutos. Os mapas de classes para tais abordagens e a etapa de pós-processamento na quantificação de objetos, aplicada na localização de frutos também foram apresentados.

5. Conclusões

5.1. Considerações finais e discussões

Técnicas de pré-processamento de imagens, segmentação, extração e classificação de características, reconhecimento de padrões, e pós-processamento agregam os procedimentos aplicados de visão computacional neste trabalho, com os objetivos focados na locomoção autônoma de robôs agrícolas móveis, e na localização e inspeção de frutos.

No pré-processamento, foram realizadas transformações na resolução cromática e nos espaços de cores das imagens adquiridas, possibilitando alternativas distintas de análise e segmentação em cenas naturais.

O algoritmo de segmentação JSEG atendeu às necessidades da primeira abordagem, identificando as regiões homogêneas como de uma mesma classe, e as associando com os objetos de interesse na delimitação das áreas passíveis de navegação. Na segunda abordagem, as cenas que apresentaram pomares distantes foram devidamente segmentadas, dadas as condições de cor e textura, em oposição aos demais objetos da cena (folhas, ramos e galhos). Em cenas com partes das árvores, os frutos foram super-segmentados, com regiões externas aos mesmos que identificam manchas ou supostas incidências de doenças. Neste caso, os parâmetros do algoritmo foram customizados na biblioteca *OpenCV* para caracterizar apenas as regiões de contorno dos frutos.

Na extração dos vetores de características, a redução da dimensionalidade para valores empiricamente aceitáveis como parâmetros de entrada aos conjuntos de treinamento, foi realizada de forma adequada pela ACP, fornecendo informações necessárias aos classificadores, através da relação entre a taxa de acerto e a quantidade de dimensões.

Os resultados obtidos com os classificadores cognitivos (RNA e *Fuzzy*) foram inferiores aos obtidos com os classificadores estatísticos (*Bayes* e *Bayes ingênuo*), pois os primeiros não atendem a seleção de modelos equivalentes entre os conjuntos de treinamento, teste e validação. O teorema de *Bayes* e *Bayes ingênuo* foram capazes de estimar os componentes das misturas entre estes conjuntos, de forma que a relação entre as regiões classificadas e as classes pretendidas apresentasse melhores resultados para o reconhecimento de padrões.

Na etapa de pós-processamento, para a localização e quantificação de frutos, os algoritmos elaborados no ambiente *Matlab* foram capazes de processar as imagens dos mapas de classe nos espaços de cores RGB e HSV, sendo que neste último, os objetos identificados atenderam a maiores porcentagens de acerto em relação aos frutos.

Por fim, a proposta do presente trabalho, para o reconhecimento de cenas naturais na locomoção de Robôs Móveis Agrícolas e localização de frutos mostrou-se adequada e aplicável de acordo com a metodologia empregada. As regiões processadas para a navegação do robô agrícola móvel, bem como as relacionadas à localização de frutos foram devidamente segmentadas e reconhecidas, identificando as vantagens e desvantagens dos classificadores cognitivos e estatísticos aplicados.

5.2. Propostas de trabalhos futuros

Algumas observações em relação aos resultados obtidos e a metodologia aplicada devem ser ponderadas e analisadas, como propostas de continuidade ao presente trabalho, bem como alternativas de reestruturação das técnicas de pré-processamento e segmentação de imagens, possibilitando a implantação de novos descritores ao reconhecimento de padrões.

Na primeira abordagem, para a locomoção do robô agrícola móvel, deve-se reestruturar o algoritmo de segmentação aplicado, e elaborar alternativas em conjunto com outras técnicas

estatísticas que realizam a extração de características mais precisas do ângulo de guiagem (linhas de plantio) e desvio lateral (delimitação da área de navegação).

A redução da dimensionalidade pela ACP em fusão com outras técnicas é outro fator que deve ser realizado na caracterização dos conjuntos de treinamento para os classificadores cognitivos e estatísticos, sendo que sua diminuição provê menor tempo de processamento, visando melhores médias de acerto.

Os descritores e vetores de características com texturas e histogramas tridimensionais em outros espaços de cores, que não os trabalhados, são propostos também como alternativas à melhoria dos classificadores, precisamente os estatísticos com relacionamentos intra-*pixels* de distâncias gaussianas.

Na segunda abordagem, para a localização de gramíneas, devem-se considerar outras técnicas de segmentação de imagens, baseadas em morfologia matemática e em formação de agrupamentos no espaço de características deste tipo de cultura. Assim, é possível extrair informações qualitativas dos descritores que representam os milhos, a formação do estigma, região da espiga e distinção entre os colmos. Para a localização e análise de cana, pretende-se com auxílio de um sistema de visão omnidirecional catadióptrico, estimar períodos de maturação e colheita, necessitando-se observar o canavial em visão superior, uma vez que há áreas de plantio de cana que exigem maior elevação de pico de crescimento em toda plantação, ou seja, denota a diferença de crescimento e densidade entre regiões de plantio à função de colheita, cuja característica de colheita dá-se pela linearidade da densidade do plantio, o que reforça a viabilidade do uso deste tipo de sistema de visão artificial.

APÊNDICE A. Gráficos das variáveis de entrada e saída do controlador Fuzzy preditivo às duas abordagens

Os gráficos são ajustados dados os estados de inferência (abscissa) em relação ao grau de pertinência (ordenada) para entradas e saídas à primeira abordagem.

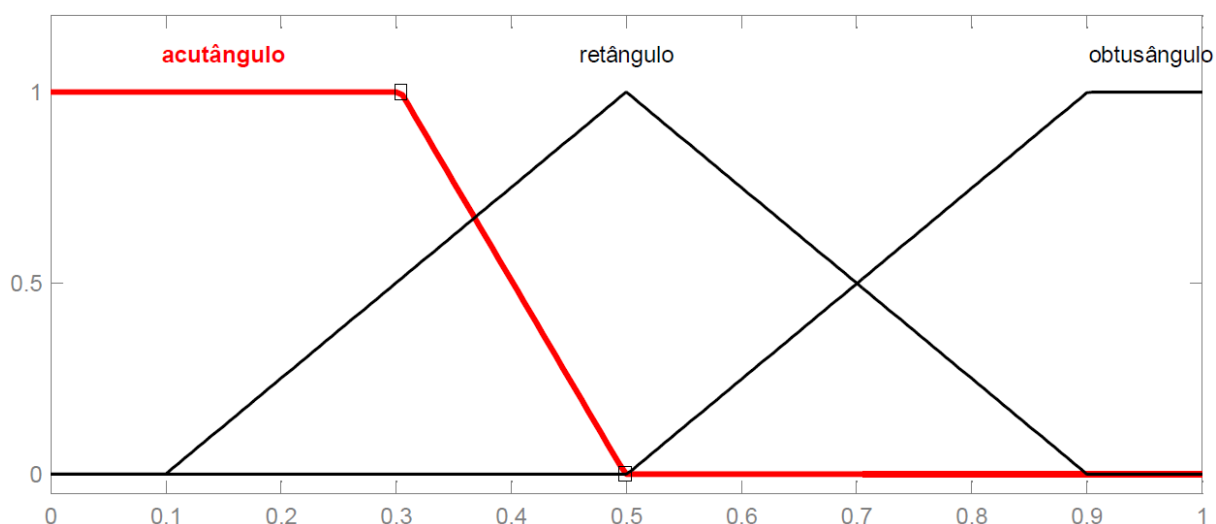


Figura A.1: Gráfico da entrada "RDDL" para os estados de inferência *acutângulo*, *retângulo* e *obtusângulo*, à primeira abordagem.

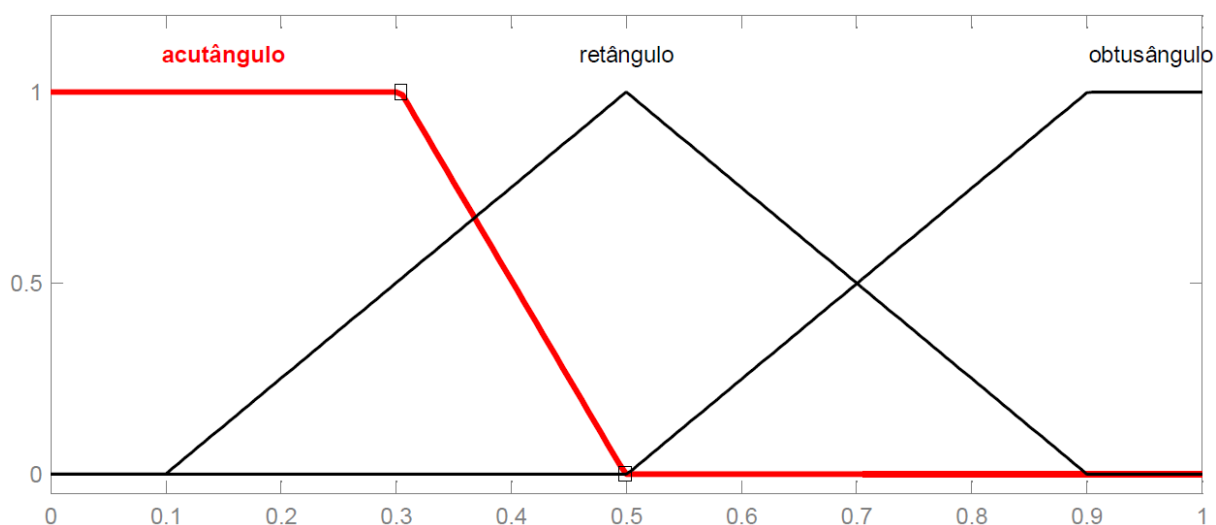


Figura A.2: Gráfico da entrada "RDEL" para os estados de inferência *acutângulo*, *retângulo* e *obtusângulo*, à primeira abordagem.

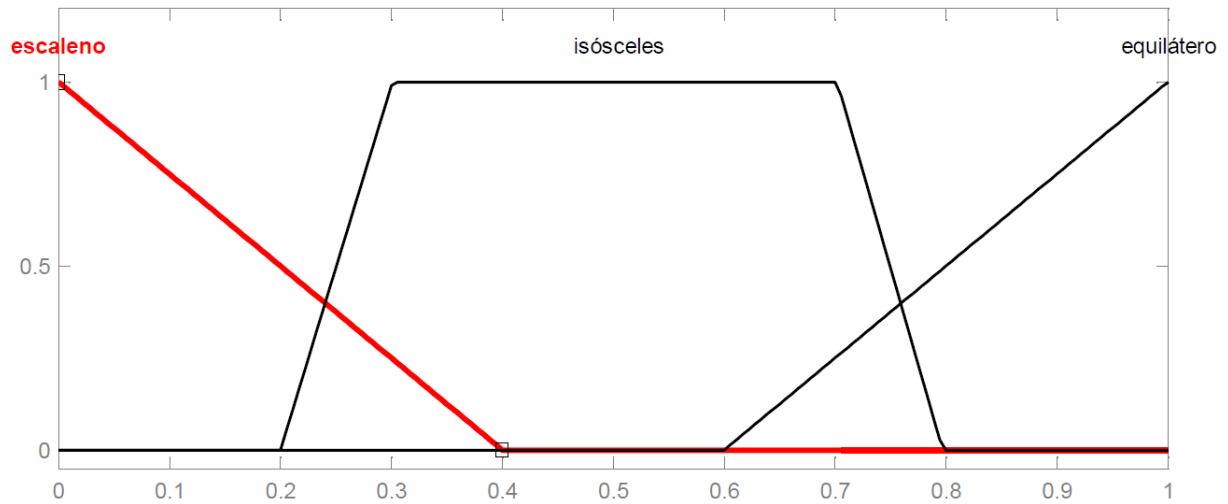


Figura A.3: Gráfico da entrada “RDDF” para os estados de inferência *escaleno*, *isósceles* e *equilátero*, à primeira abordagem.

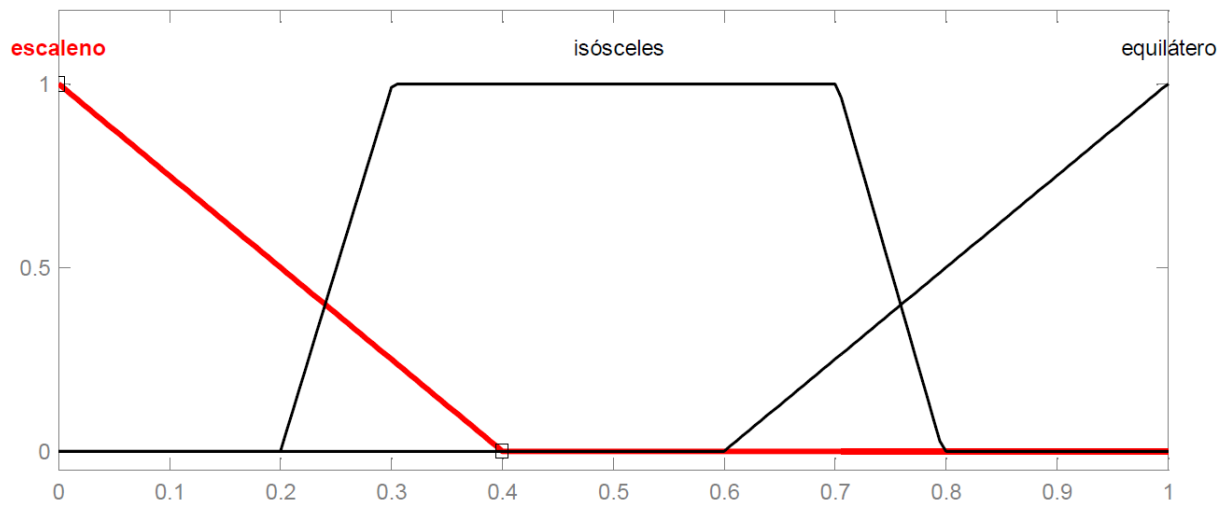


Figura A.4: Gráfico da entrada “RDEF” para os estados de inferência *escaleno*, *isósceles* e *equilátero*, à primeira abordagem.

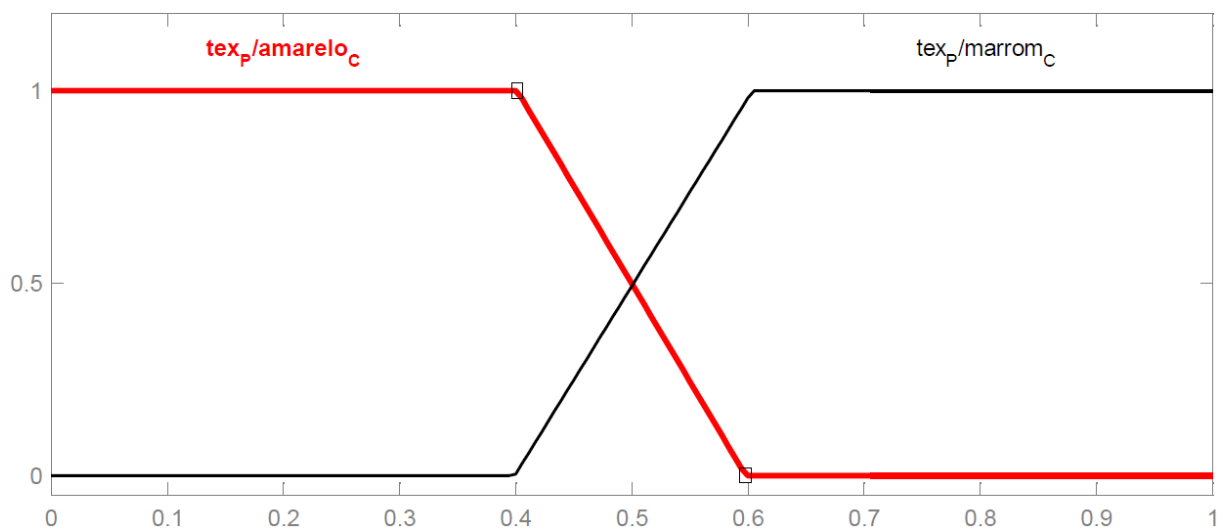


Figura A.5: Gráfico da saída “AN” para os estados de inferência *tex_p/amarelo_C* e *tex_p/marrom_C*, à primeira abordagem.

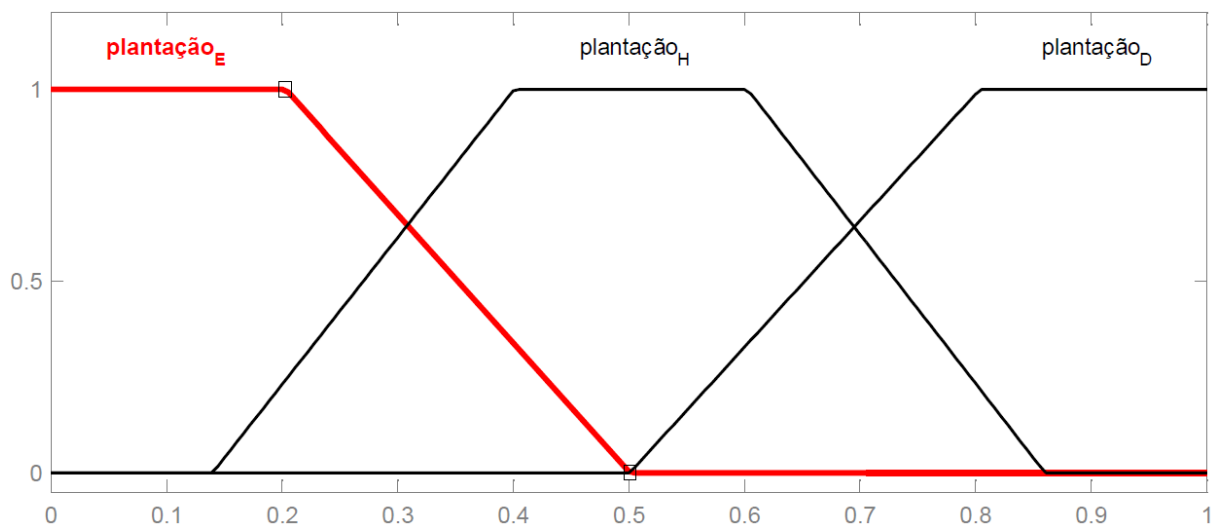


Figura A.6: Gráfico da saída “AnN” para os estados de inferência *plantação_E*, *plantação_H* e *plantação_D*, à primeira abordagem.

Os gráficos são ajustados dados os estados de inferência (abscissa) em relação ao grau de pertinência (ordenada) para entradas e saídas à segunda abordagem.

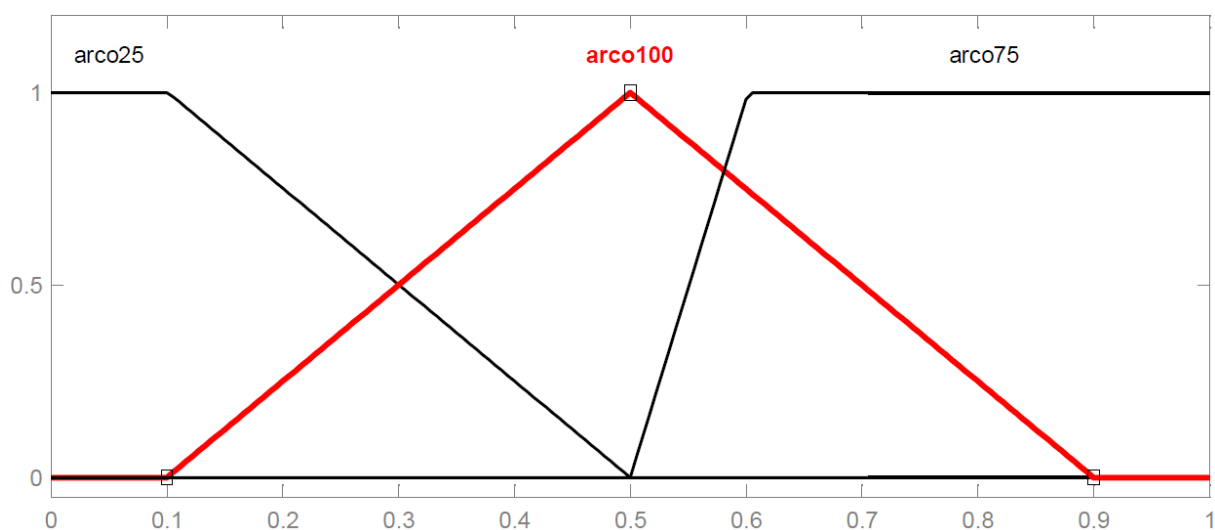


Figura A.7: Gráfico da entrada “CA100” para os estados de inferência *arco25*, *arco75* e *arco100*, à segunda abordagem.

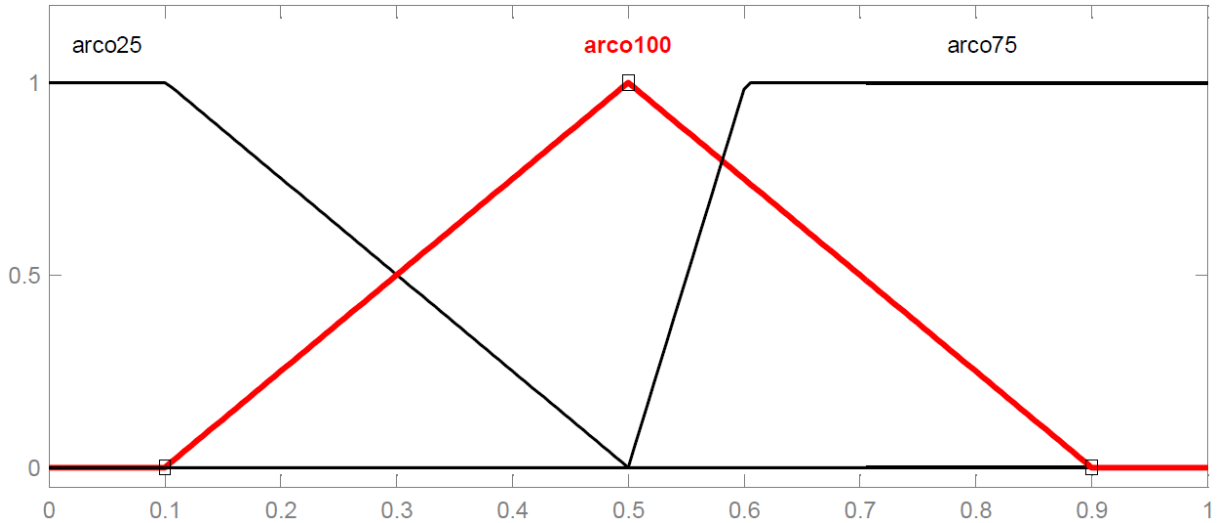


Figura A.8: Gráfico da entrada “CA300” para os estados de inferência *arco25*, *arco75* e *arco100*, à segunda abordagem.

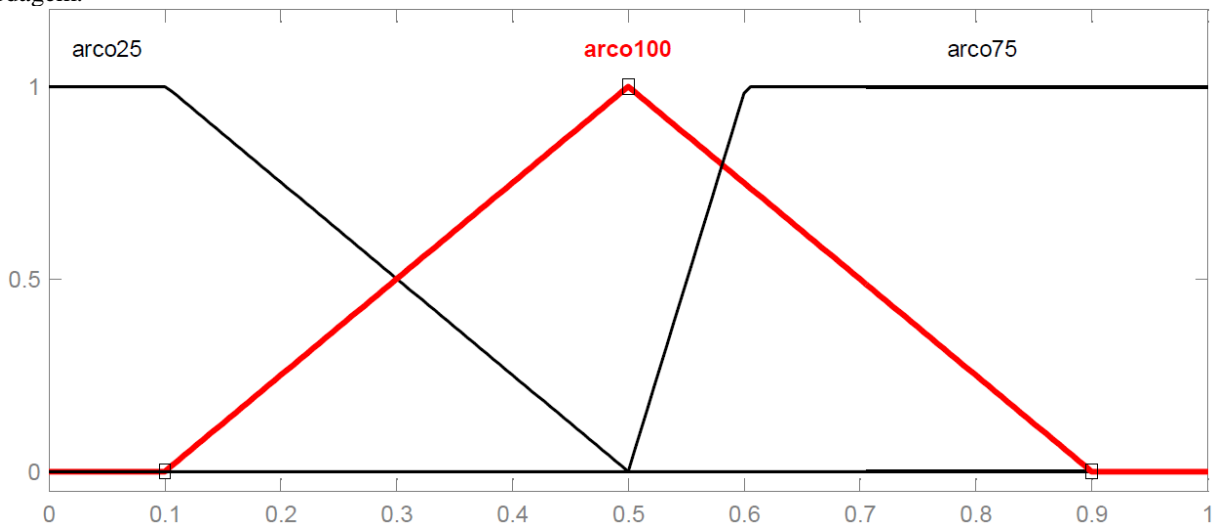


Figura A.9: Gráfico da entrada “AA300” para os estados de inferência *arco25*, *arco75* e *arco100*, à segunda abordagem.

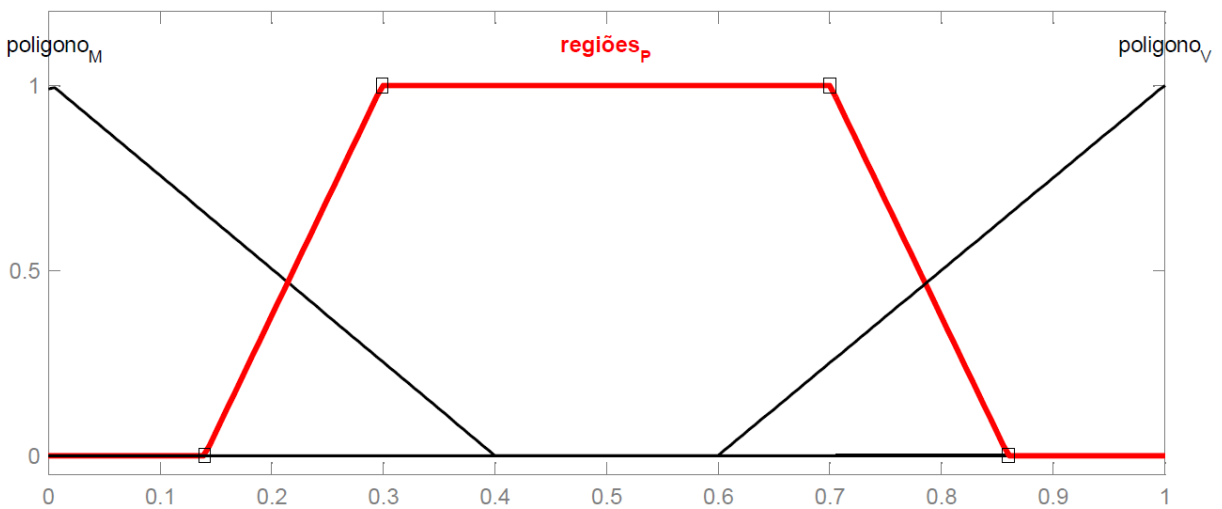


Figura A.10: Gráfico da entrada “AIM” para os estados de inferência *arco25*, *arco75* e *arco100*, à segunda abordagem.

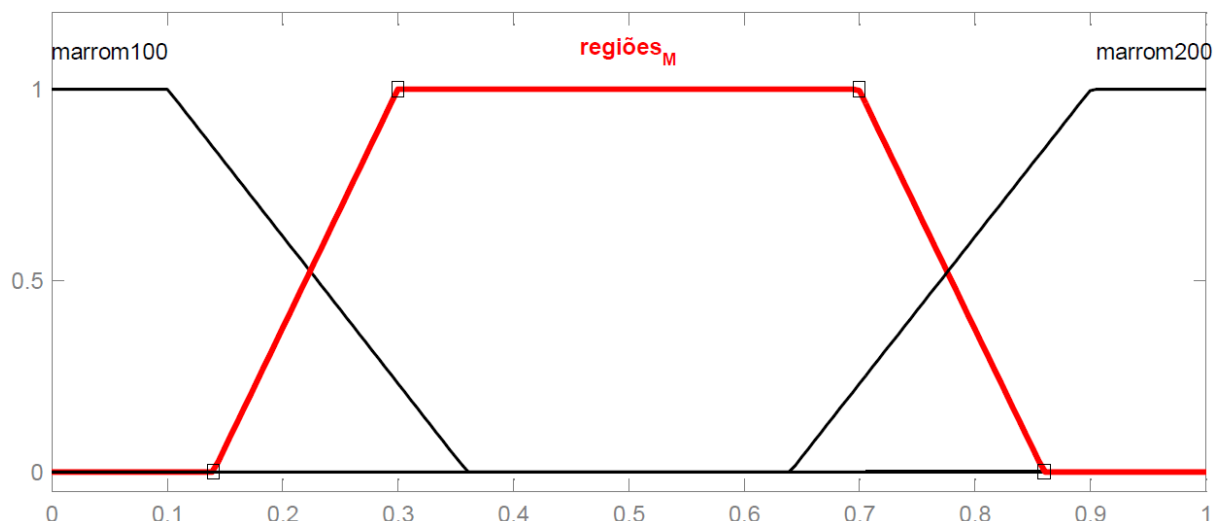


Figura A.11: Gráfico da entrada “AIB” para os estados de inferência arco25 , arco75 e arco100 , à segunda abordagem.

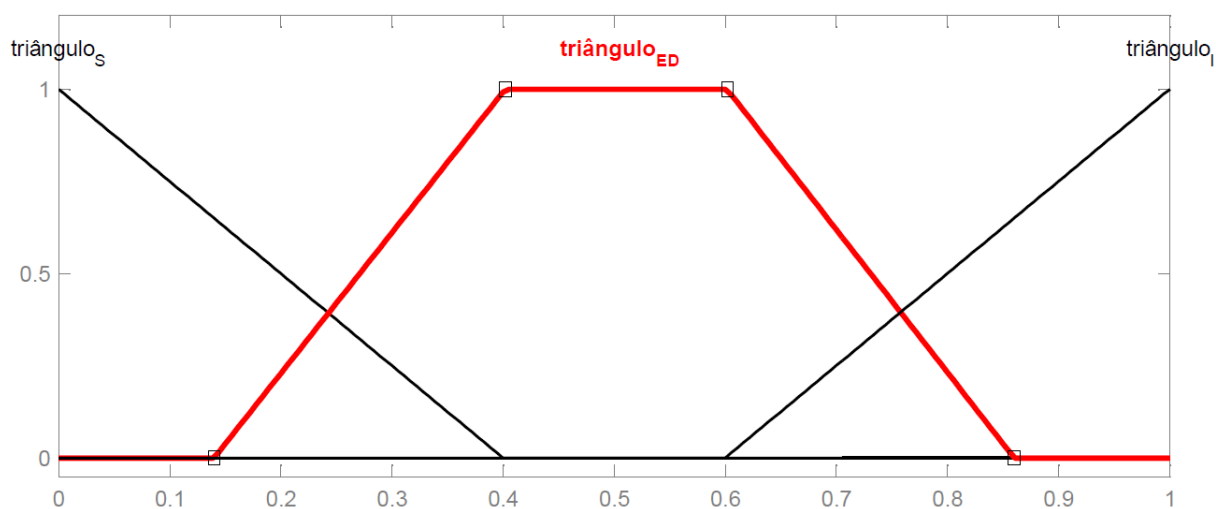


Figura A.12: Gráfico da entrada “LA100” para os estados de inferência arco25 , arco75 e arco100 , à segunda abordagem.

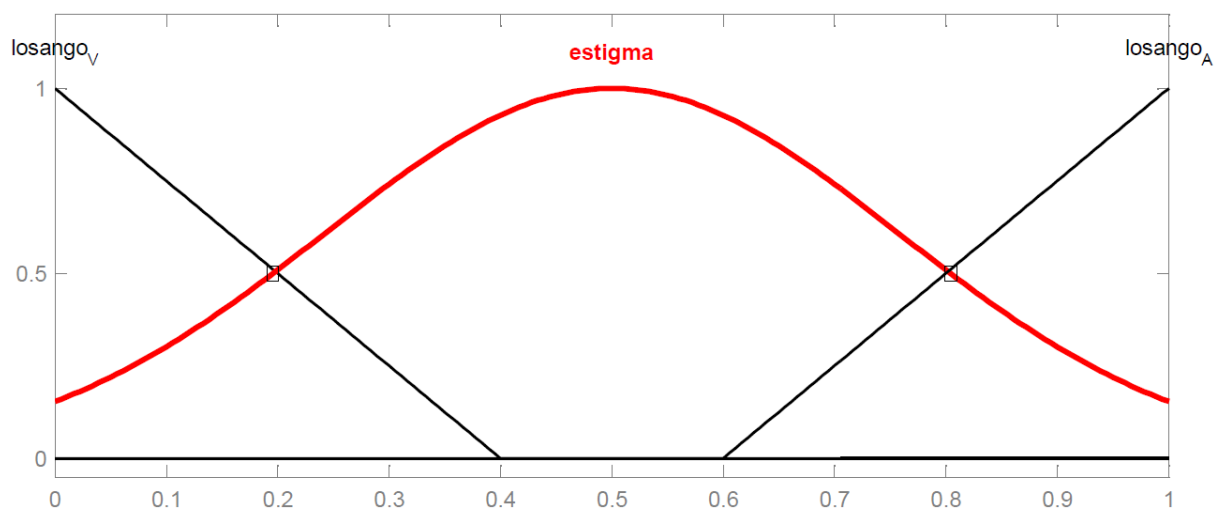


Figura A.13: Gráfico da entrada “LA300” para os estados de inferência arco25 , arco75 e arco100 , à segunda abordagem.

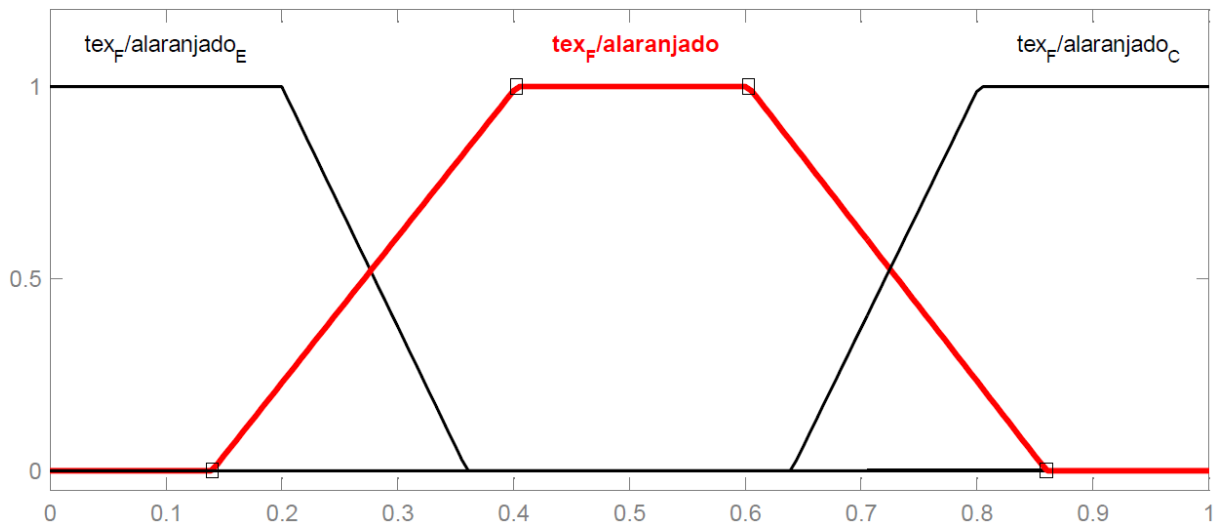


Figura A.14: Gráfico da saída “FL” para os estados de inferência $arco25$, $arco75$ e $arco100$, à segunda abordagem.

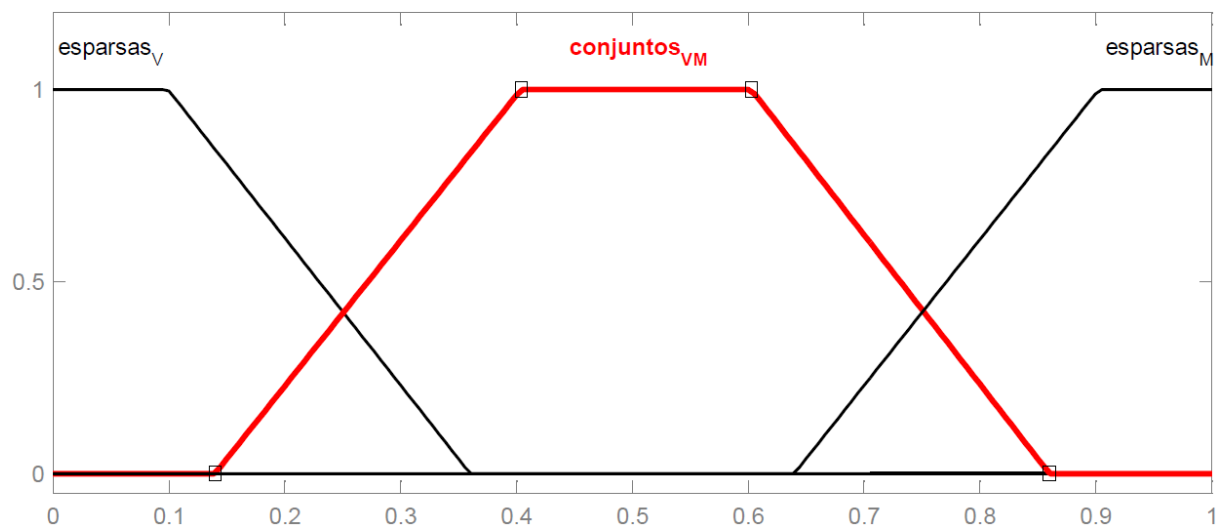


Figura A.15: Gráfico da saída “FRG” para os estados de inferência $arco25$, $arco75$ e $arco100$, à segunda abordagem.

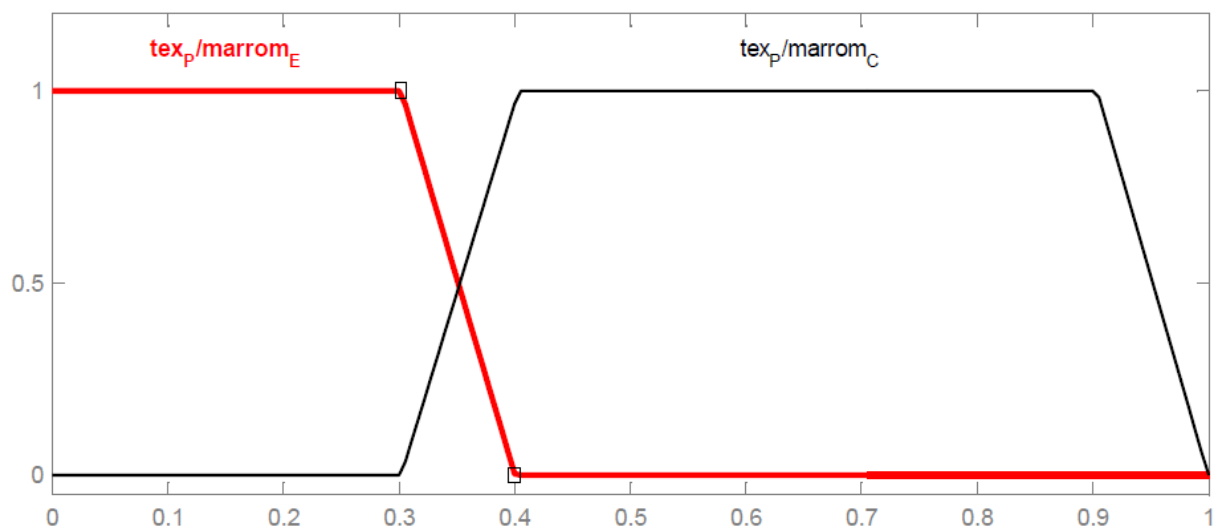


Figura A.16: Gráfico da saída “SVR” para os estados de inferência *arco25*, *arco75* e *arco100*, à segunda abordagem.

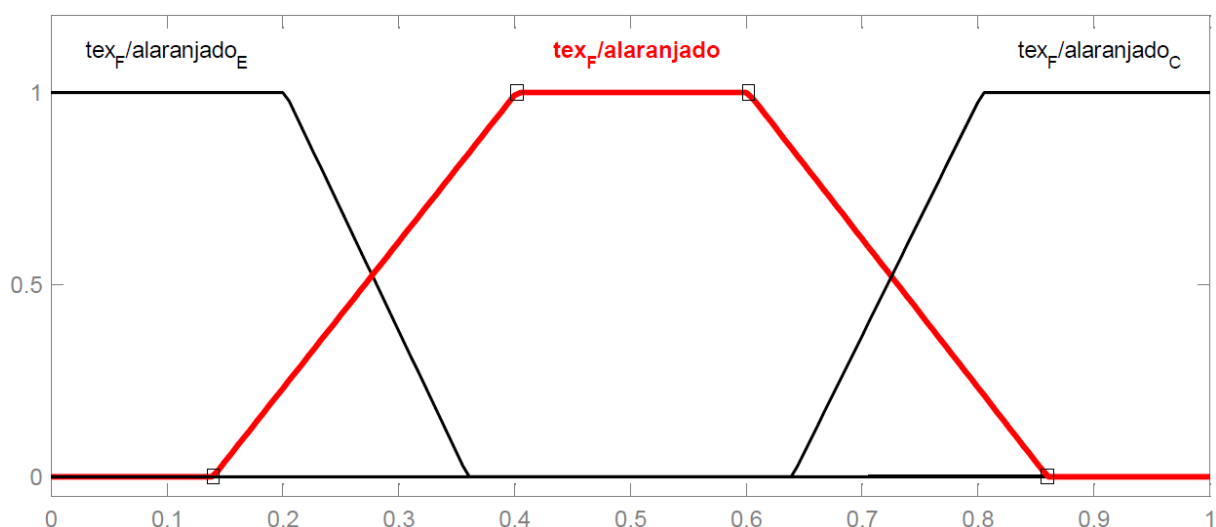


Figura A.17: Gráfico da saída “CF” para os estados de inferência *arco25*, *arco75* e *arco100*, à segunda abordagem.

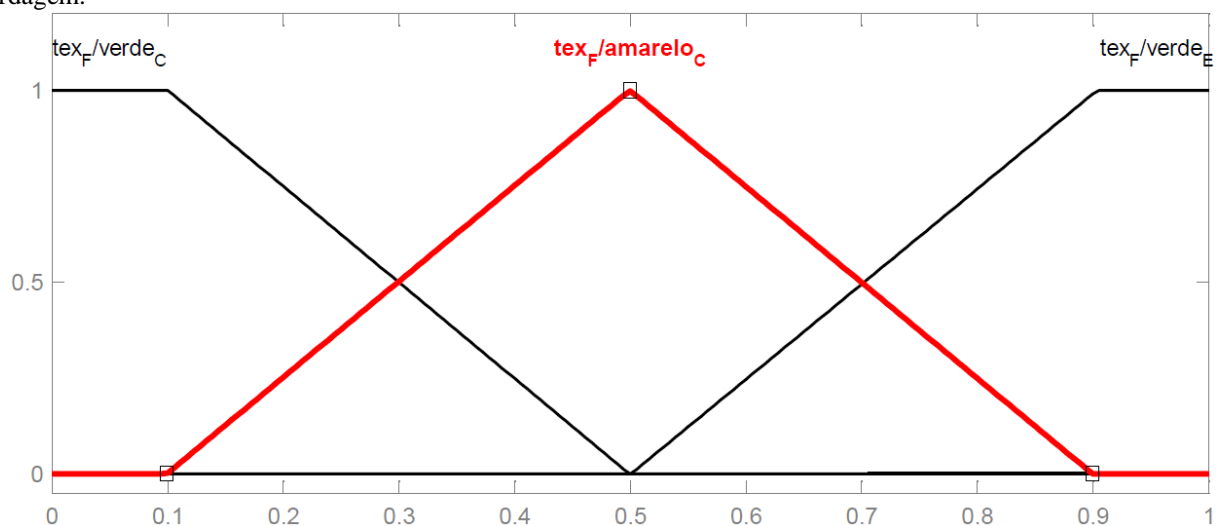


Figura A.18: Gráfico da saída “GM” para os estados de inferência *arco25*, *arco75* e *arco100*, à segunda abordagem.

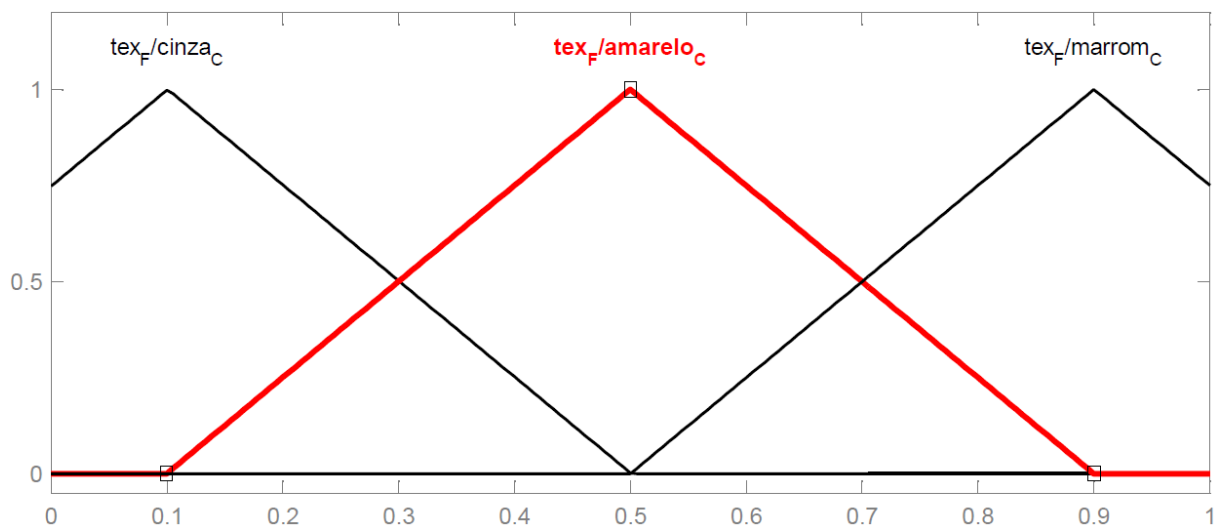


Figura A.19: Gráfico da saída “FC” para os estados de inferência *arco25*, *arco75* e *arco100*, à segunda abordagem.

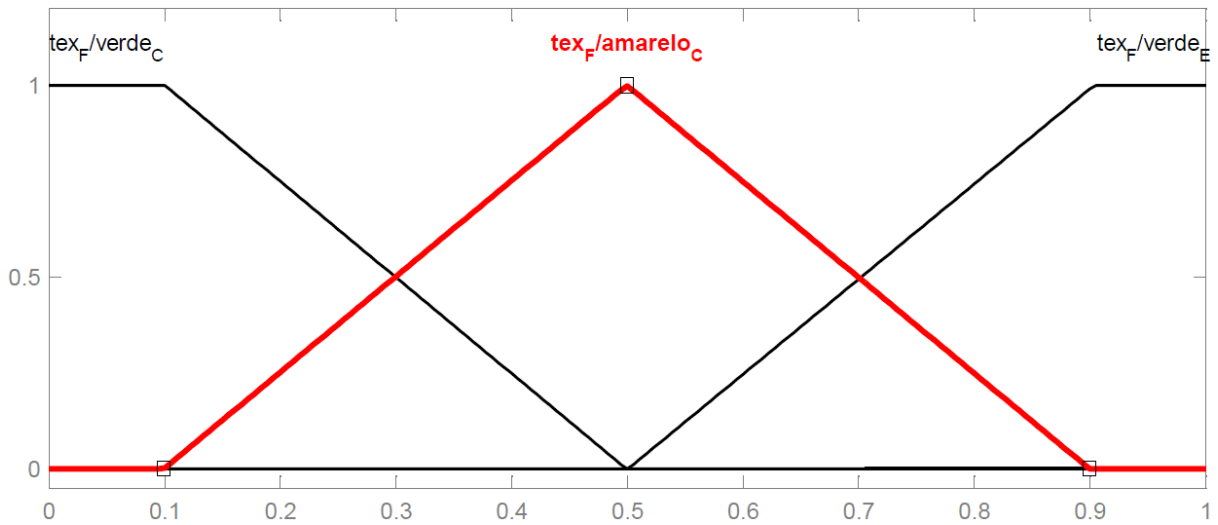


Figura A.20: Gráfico da saída “CG” para os estados de inferência *arco25*, *arco75* e *arco100*, à segunda abordagem.

APÊNDICE B. Gráficos de regressão das saídas alvo para o classificador RNA

Os gráficos abaixo mostram a regressão para as saídas alvo do classificador RNA, nas duas abordagens. Quanto maior a concentração dos dados no cruzamento das perspectivas de tendência e $Y = T$ (saída igual ao período de amostragem), menor a regressão dos dados classificados, baseados nas matrizes de confusão para cada conjunto de dimensões.

Os tempos de resposta são dados para os conjuntos de treinamento, teste, validação e todos juntos.

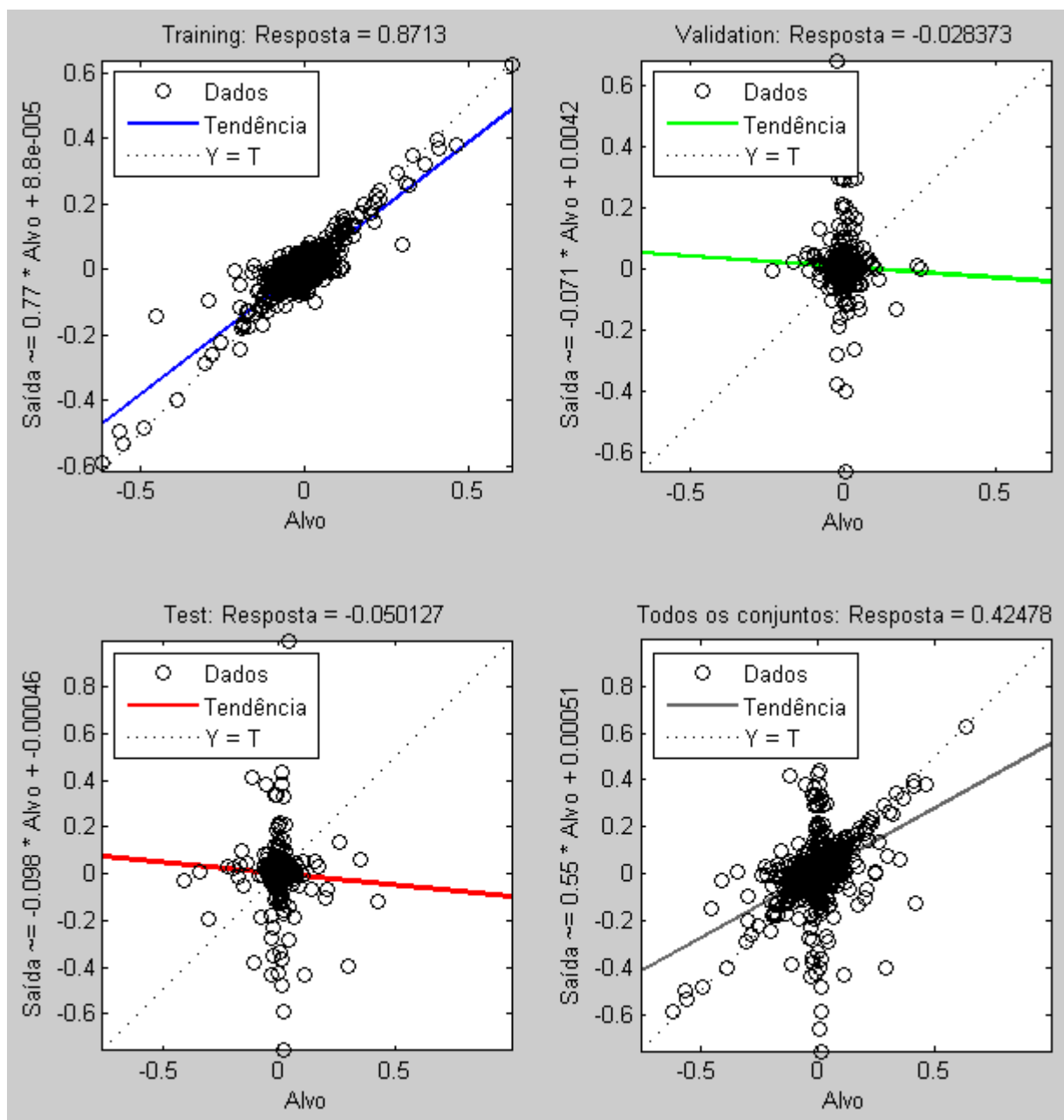


Figura B.1: Regressão das saídas alvo 10% para a segunda abordagem RGB.

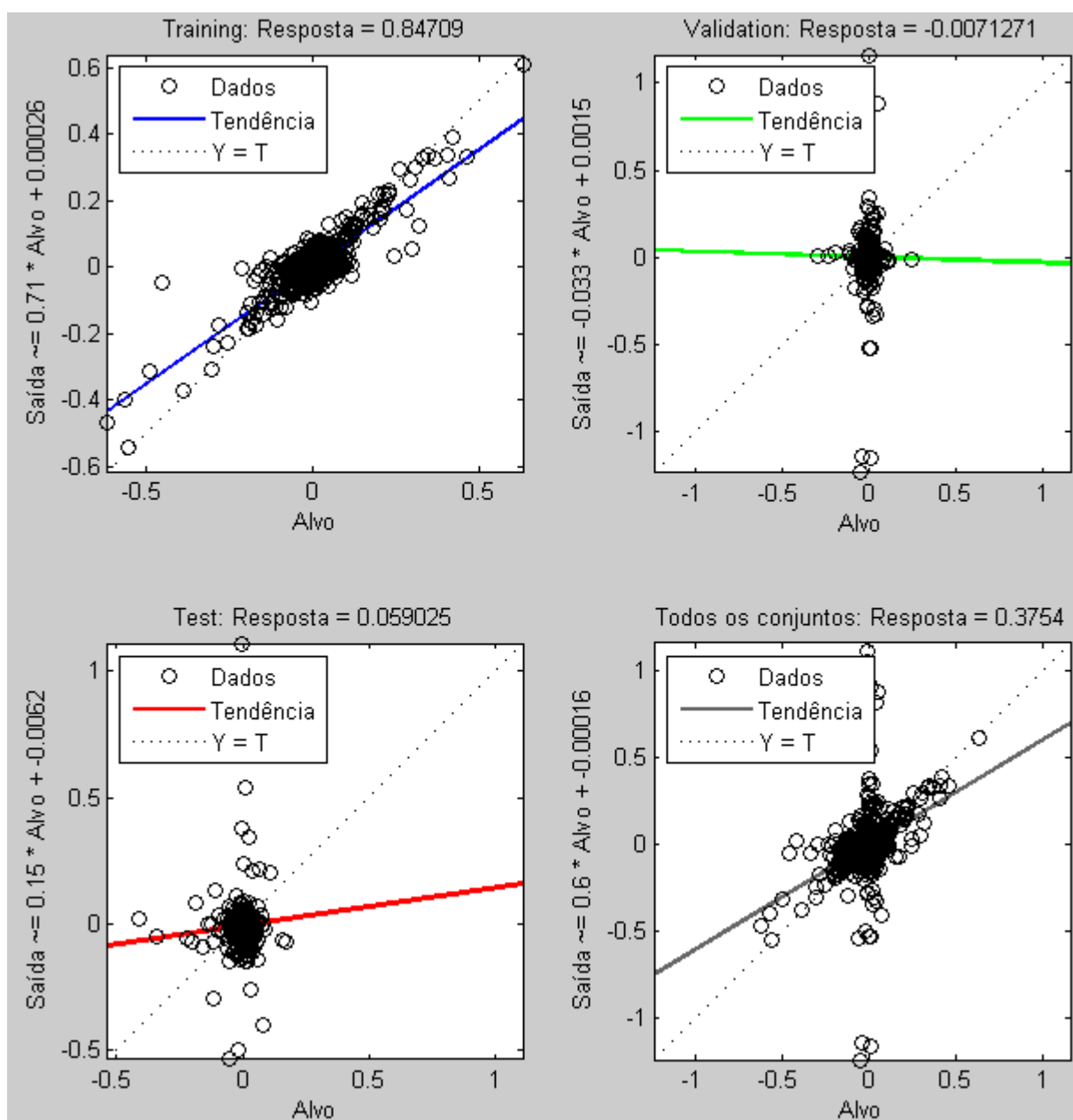


Figura B.2: Regressão das saídas alvo 20% para a segunda abordagem RGB.

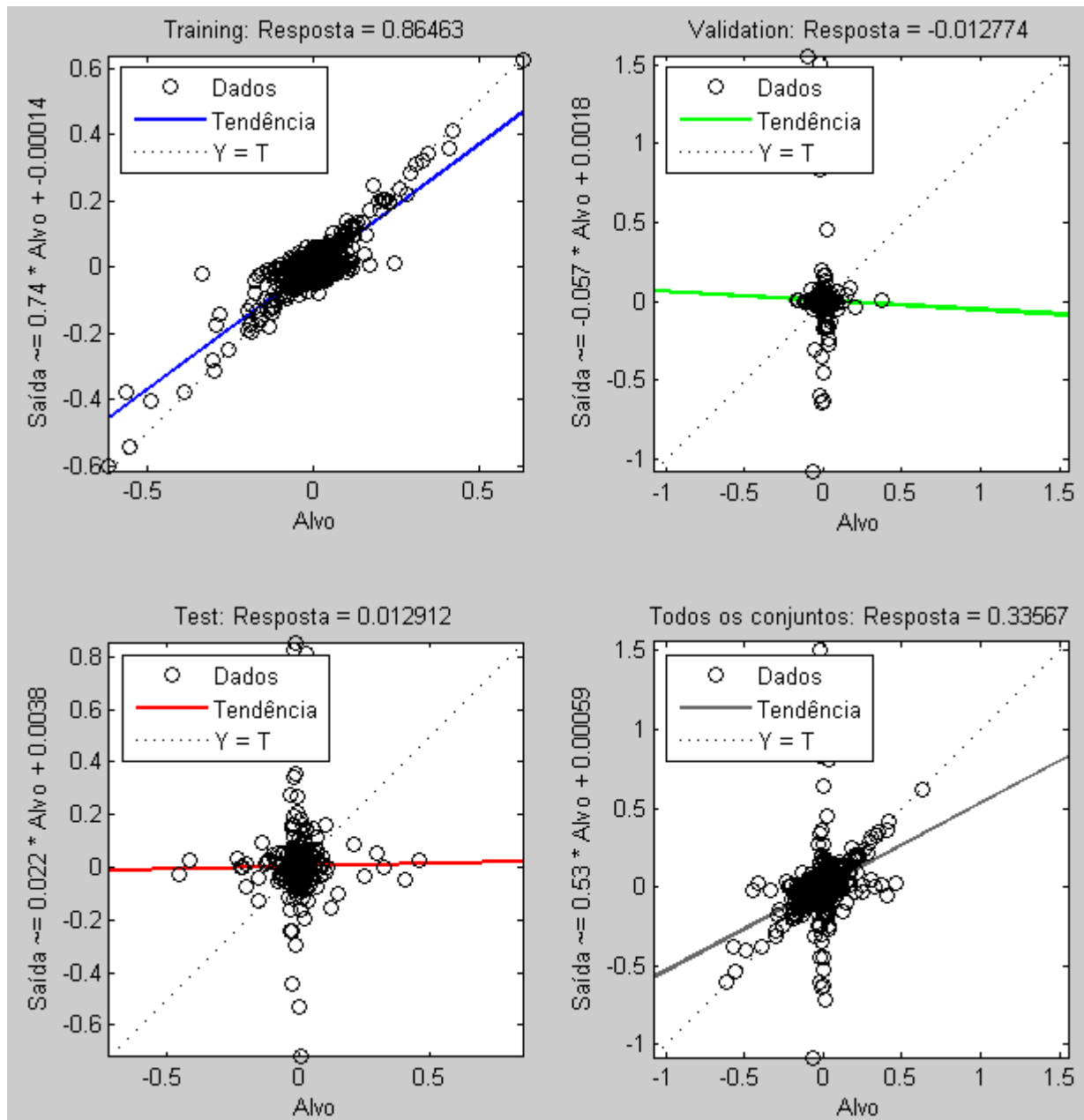


Figura B.3: Regressão das saídas alvo 30% para a segunda abordagem RGB.

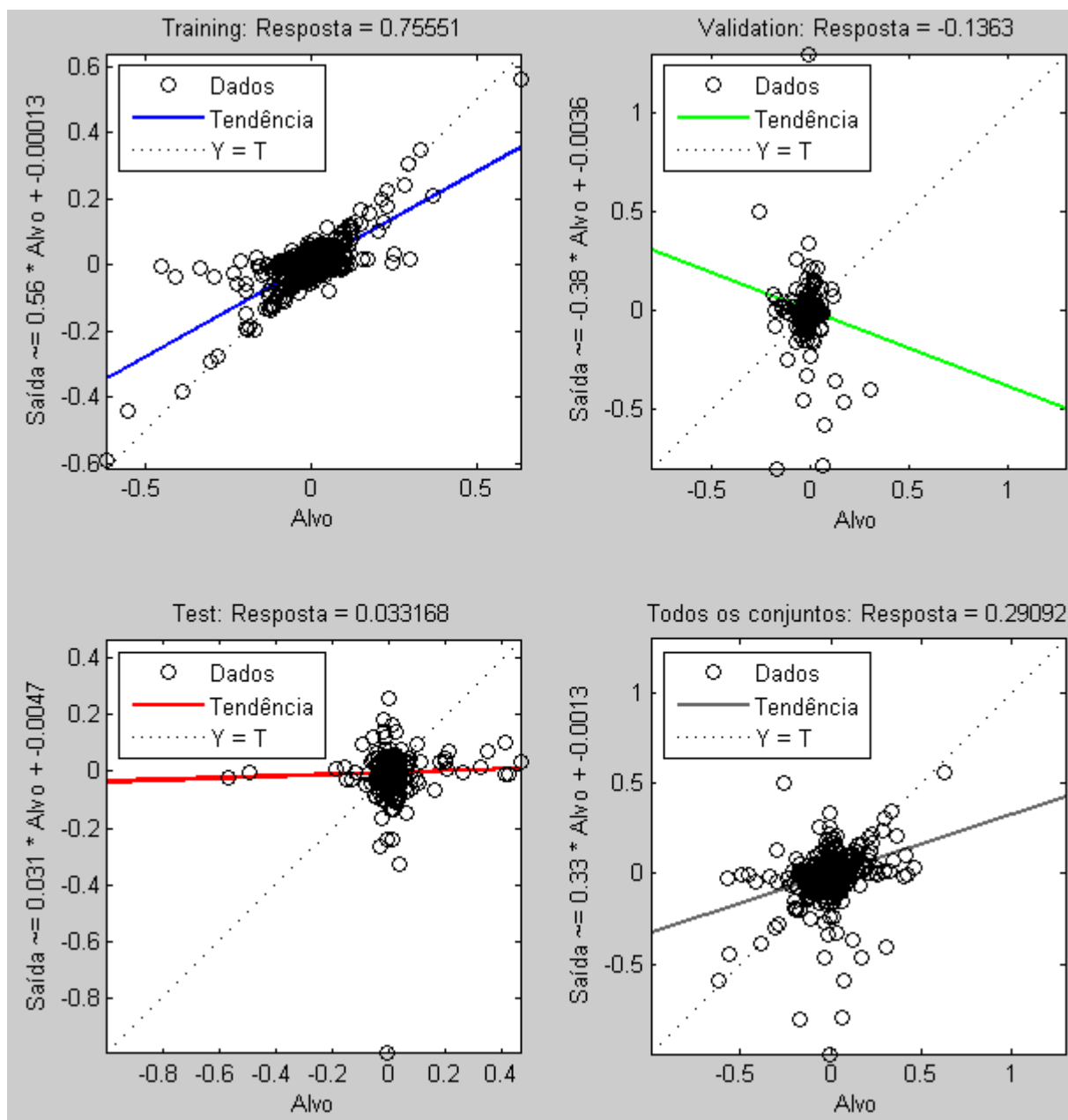


Figura B.4: Regressão das saídas alvo 40% para a segunda abordagem RGB.

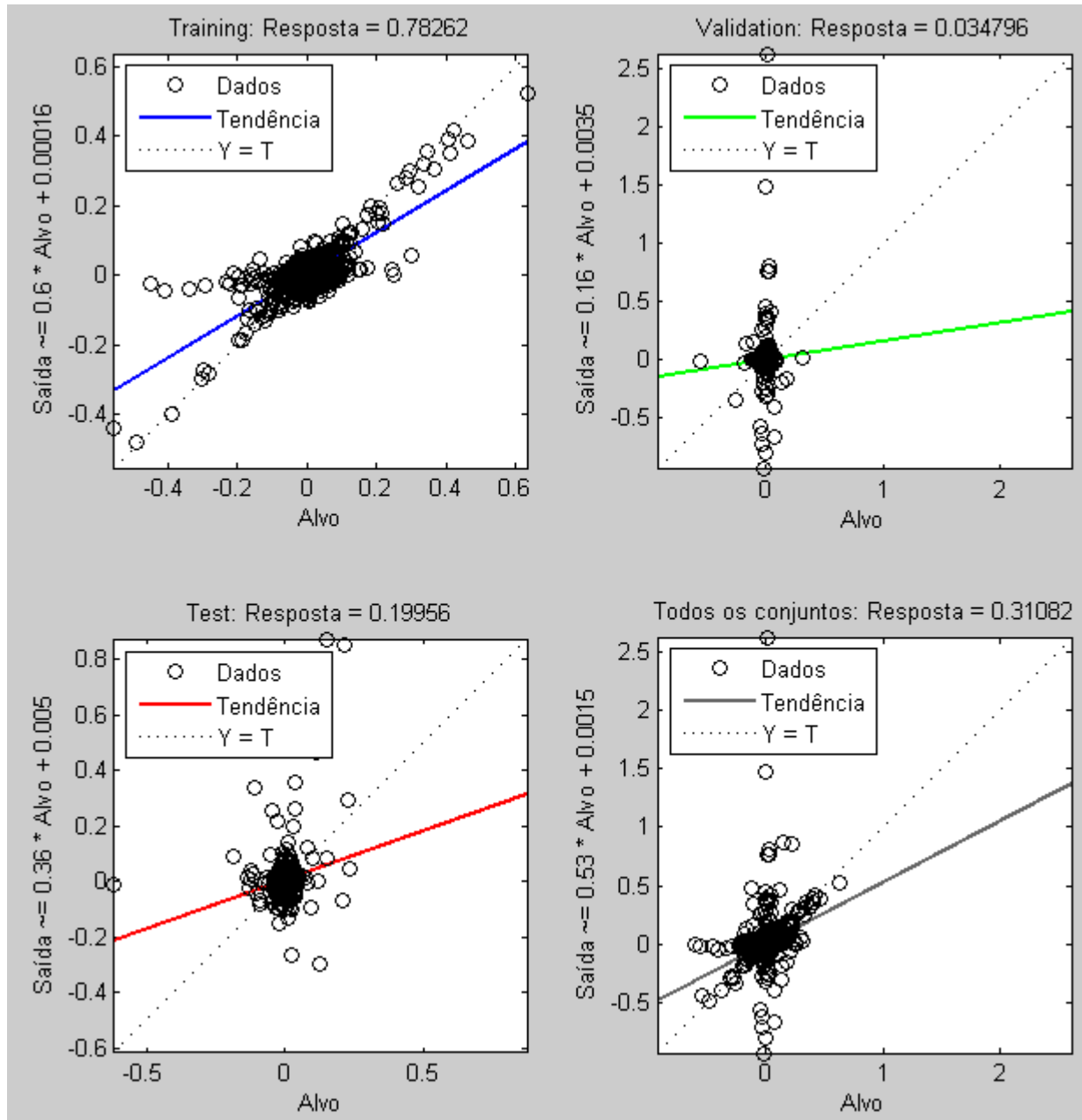


Figura B.5: Regressão das saídas alvo 50% para a segunda abordagem RGB.

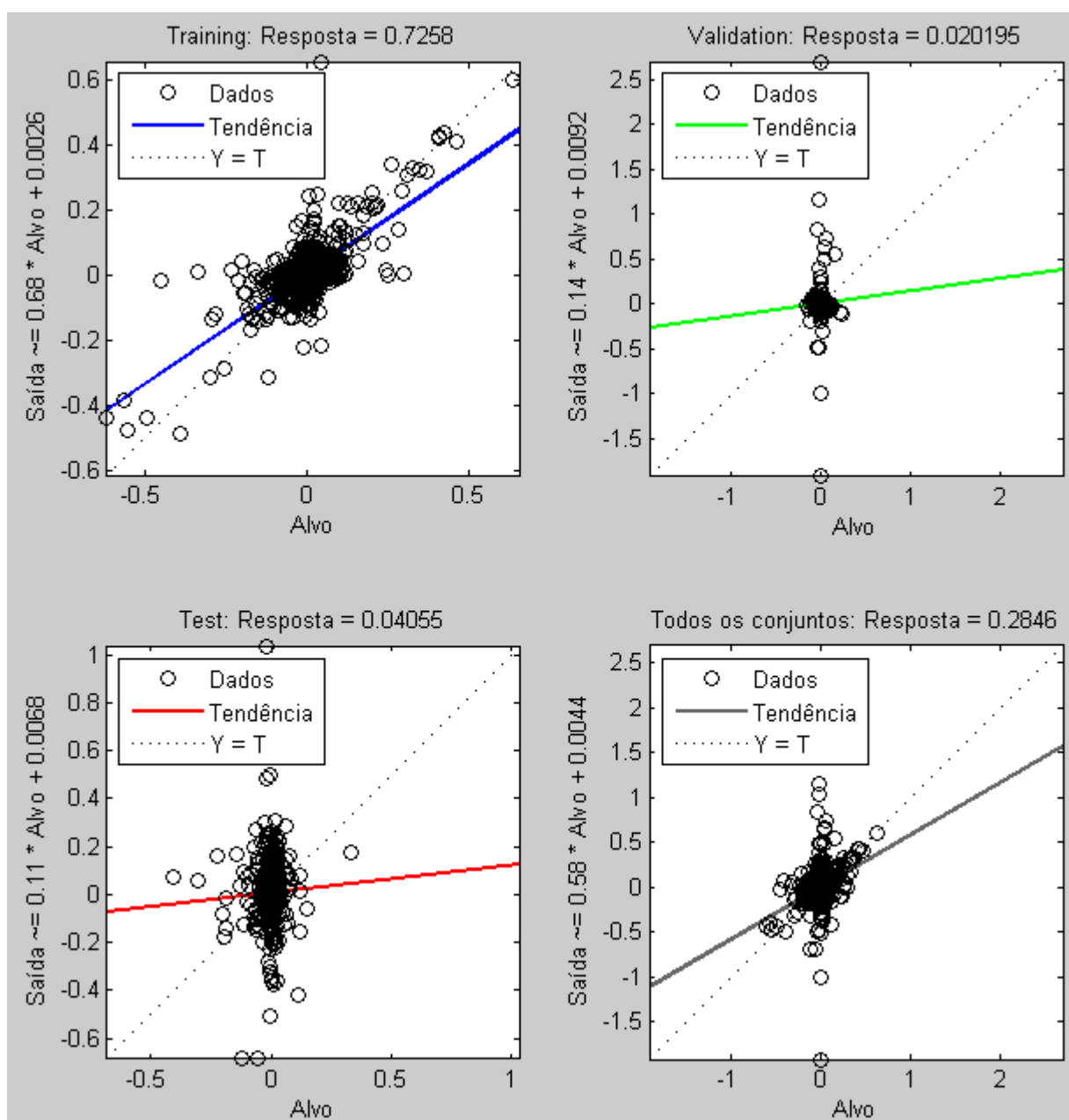


Figura B.6: Regressão das saídas alvo 75% para a segunda abordagem RGB.

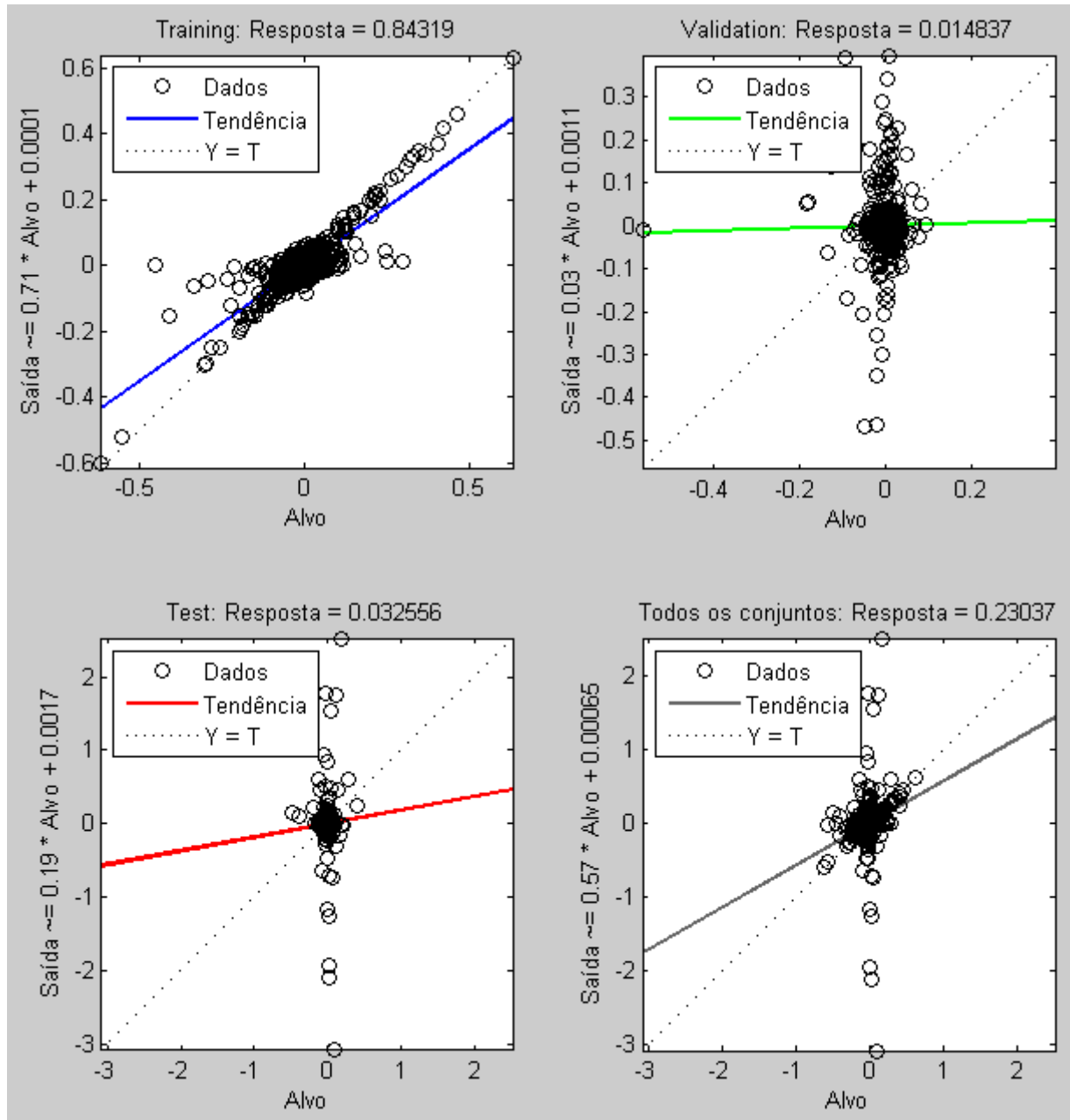


Figura B.7: Regressão das saídas alvo 100% para a segunda abordagem RGB.

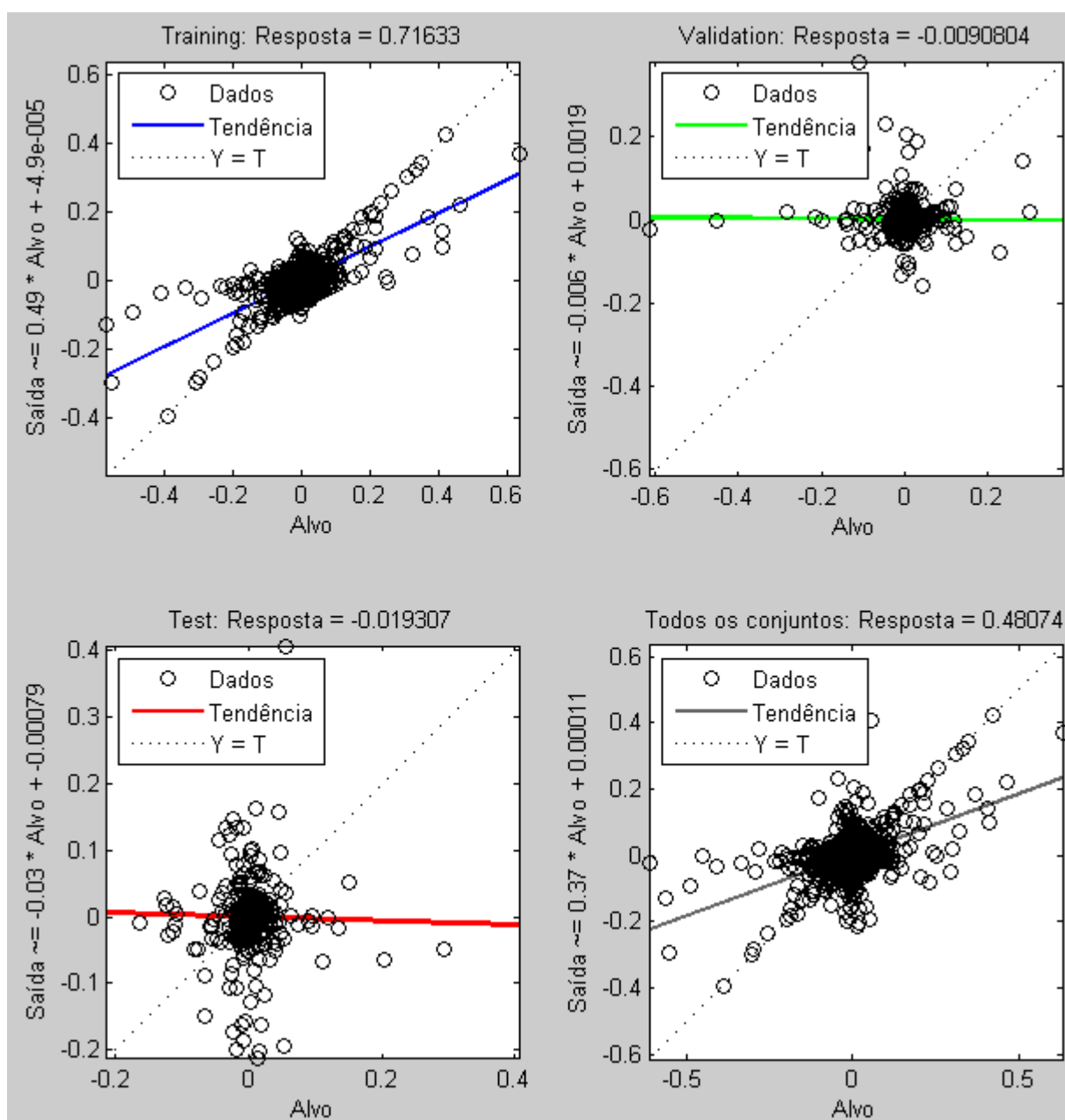


Figura B.8: Regressão das saídas alvo 10% para a segunda abordagem RGB.

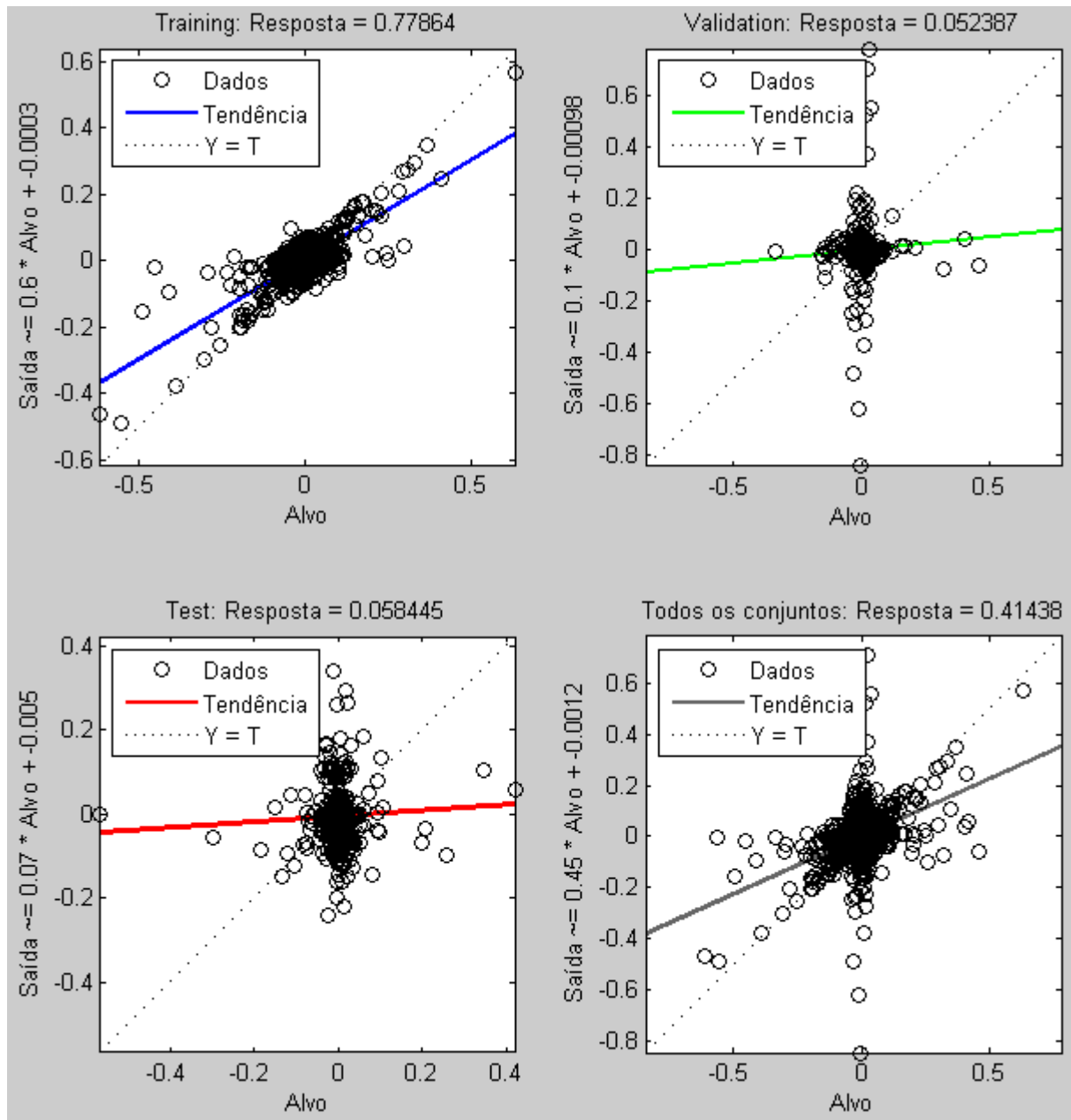


Figura B.9: Regressão das saídas alvo 20% para a segunda abordagem RGB.

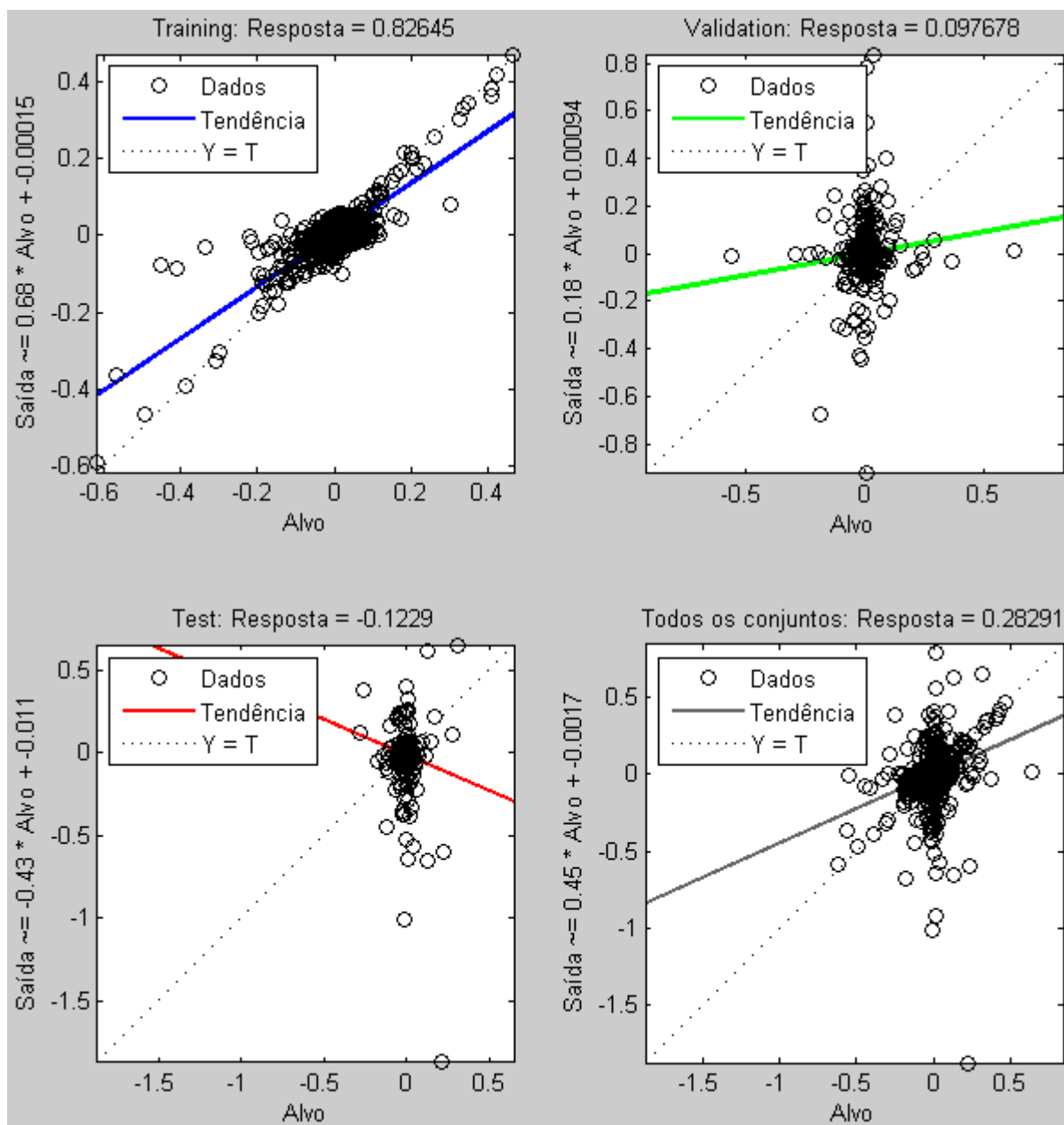


Figura B.10: Regressão das saídas alvo 30% para a segunda abordagem RGB.

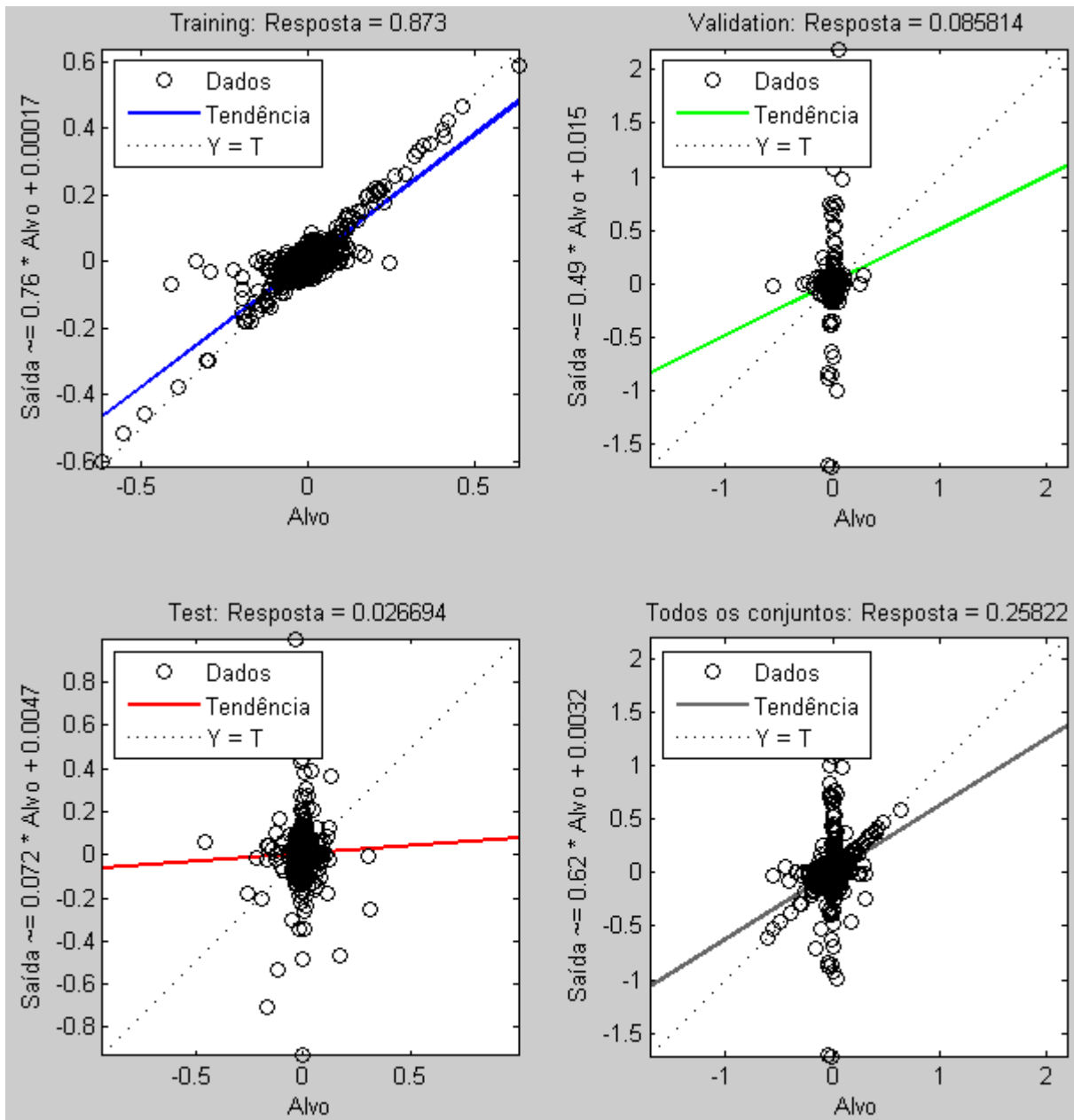


Figura B.11: Regressão das saídas alvo 40% para a segunda abordagem RGB.

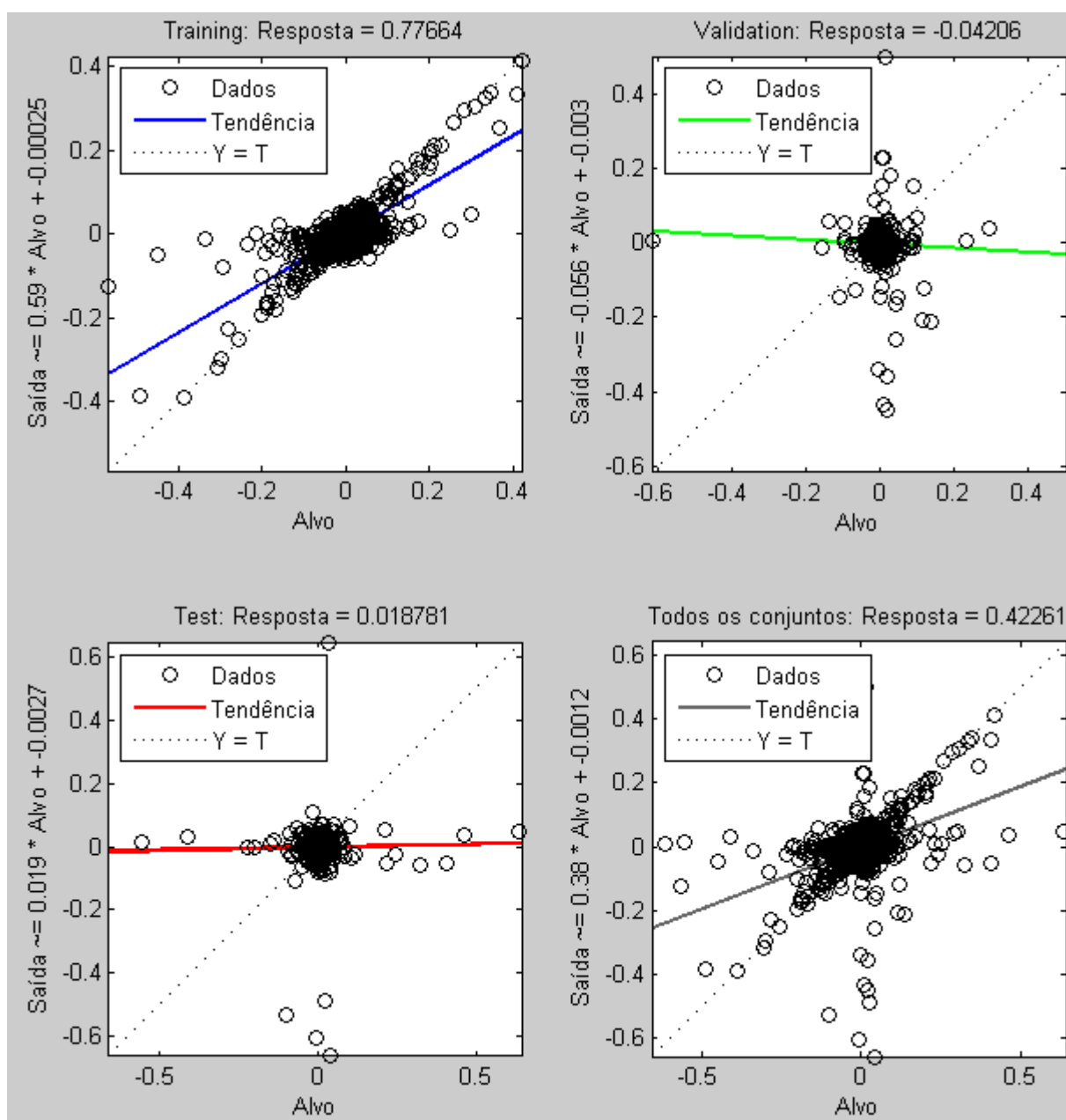


Figura B.12: Regressão das saídas alvo 50% para a segunda abordagem RGB.

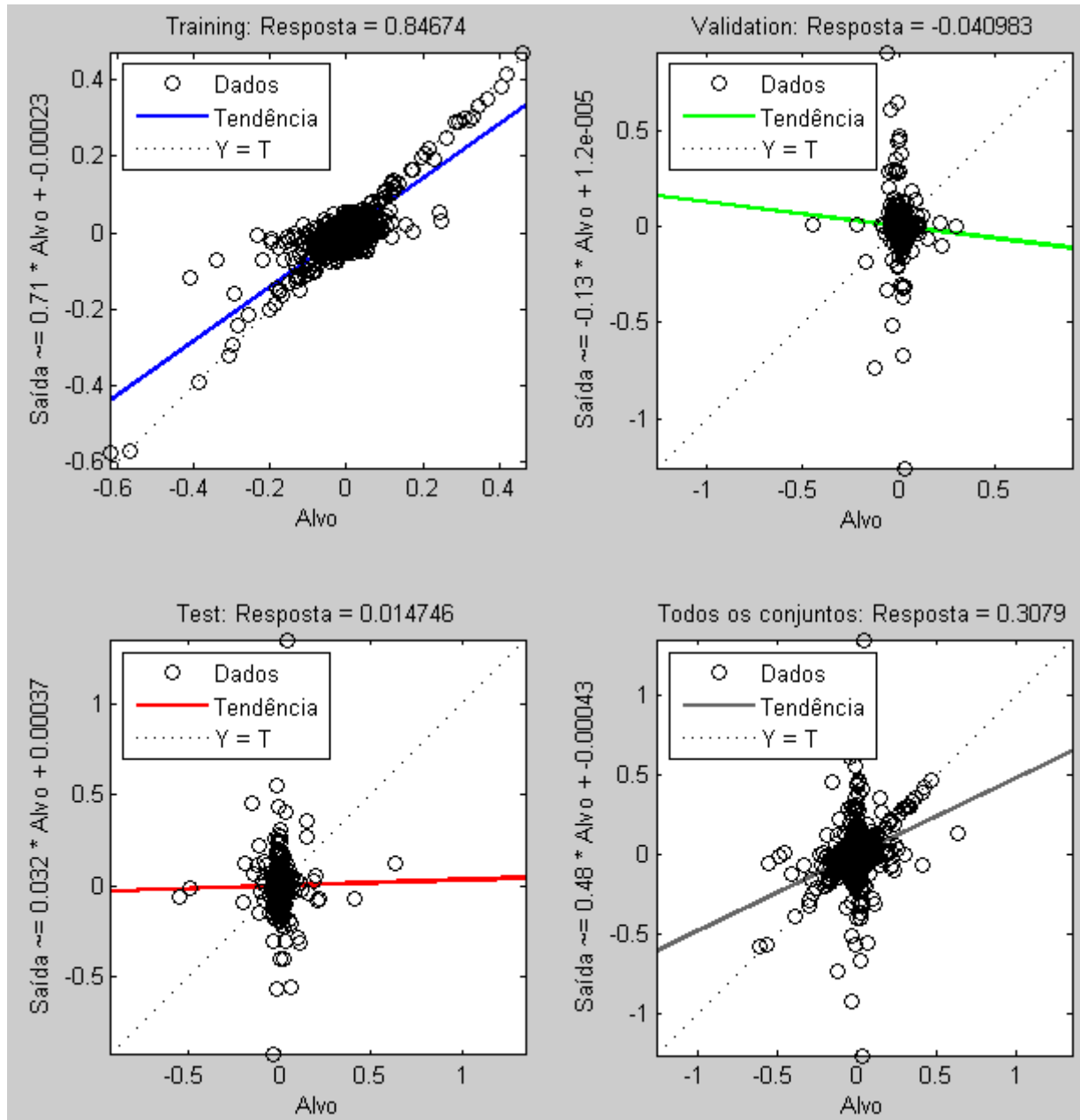


Figura B.13: Regressão das saídas alvo 75% para a segunda abordagem RGB.

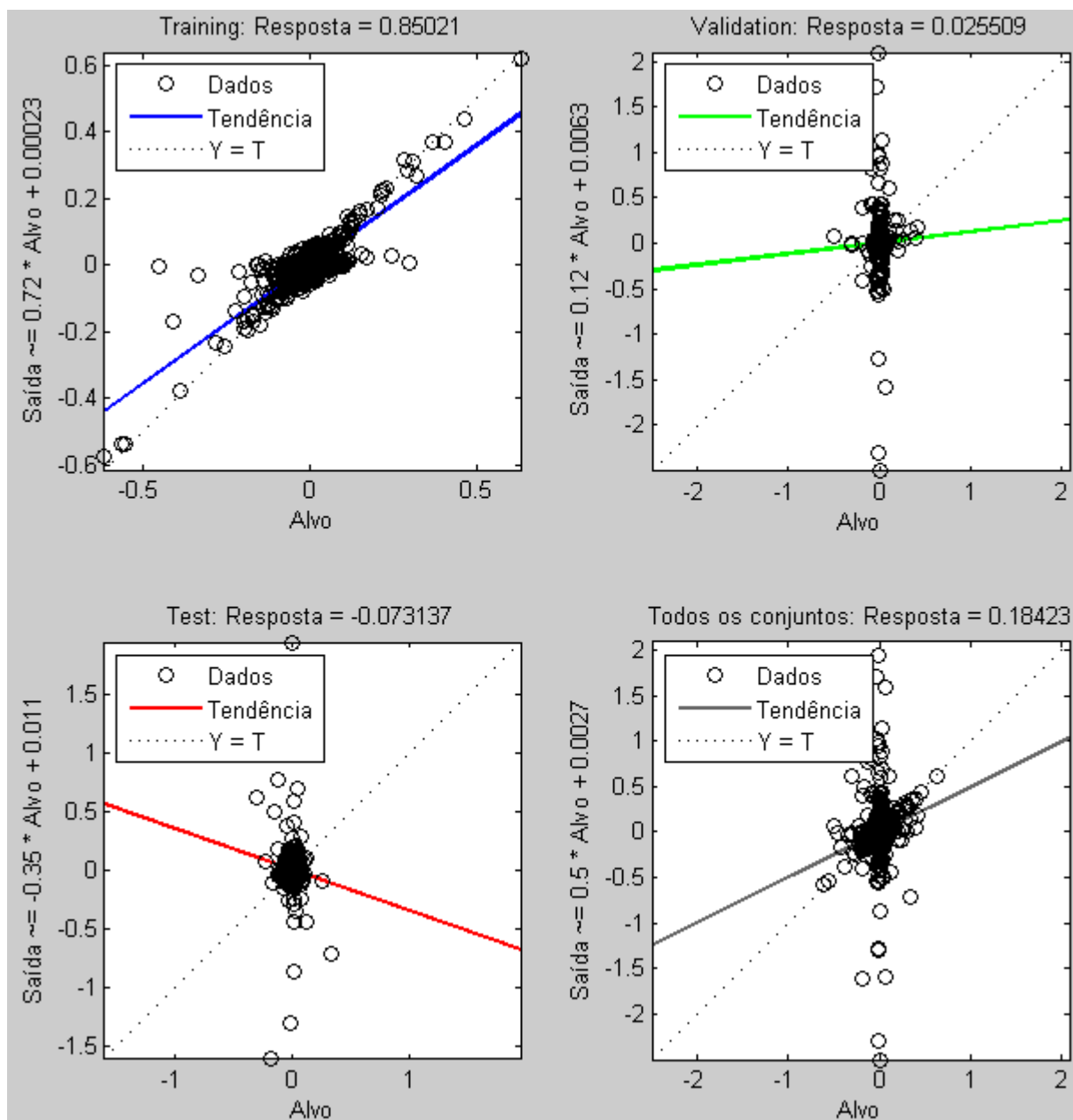


Figura B.14: Regressão das saídas alvo 100% para a segunda abordagem RGB.

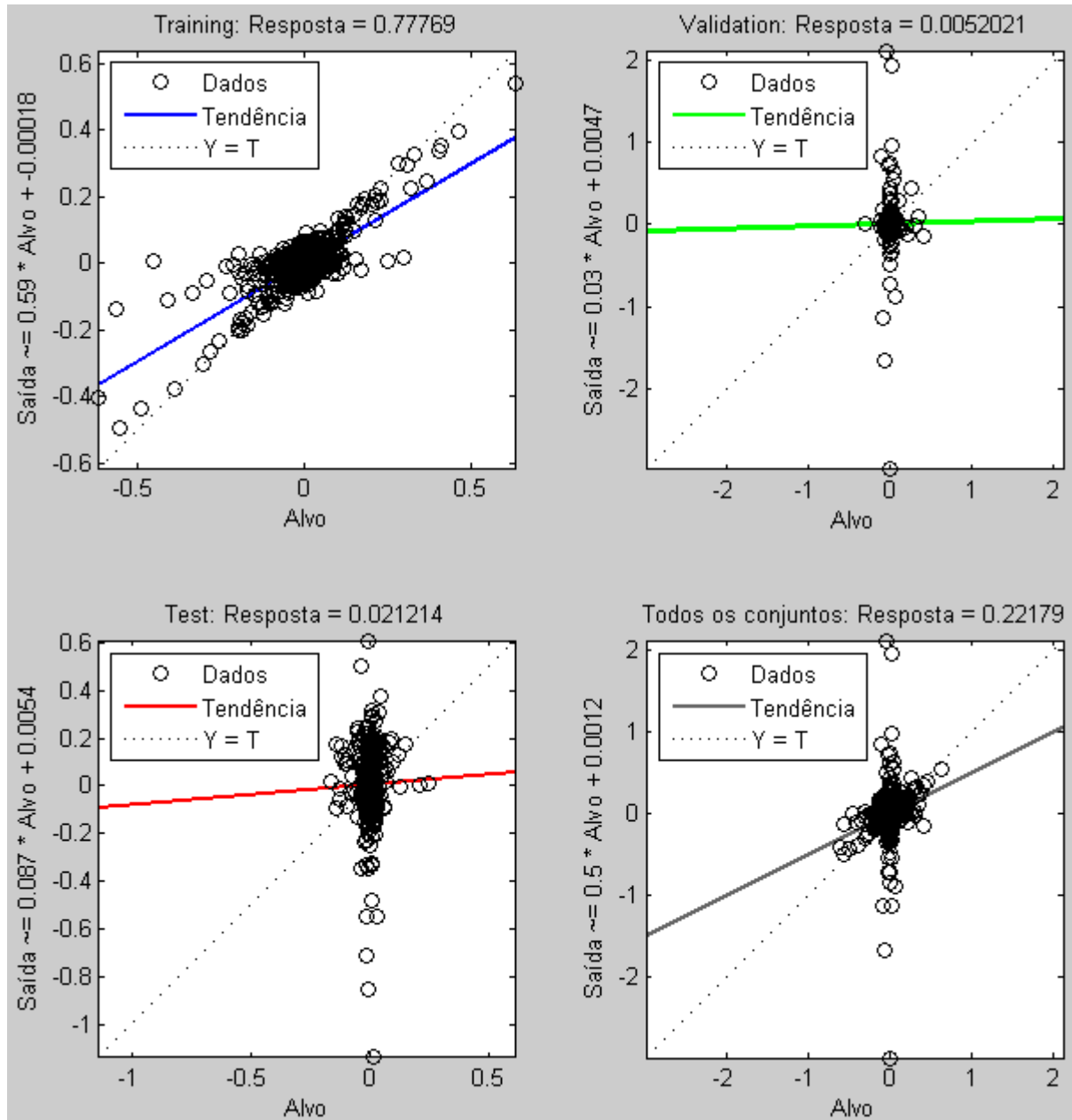


Figura B.15: Regressão das saídas alvo 10% para a segunda abordagem RGB.

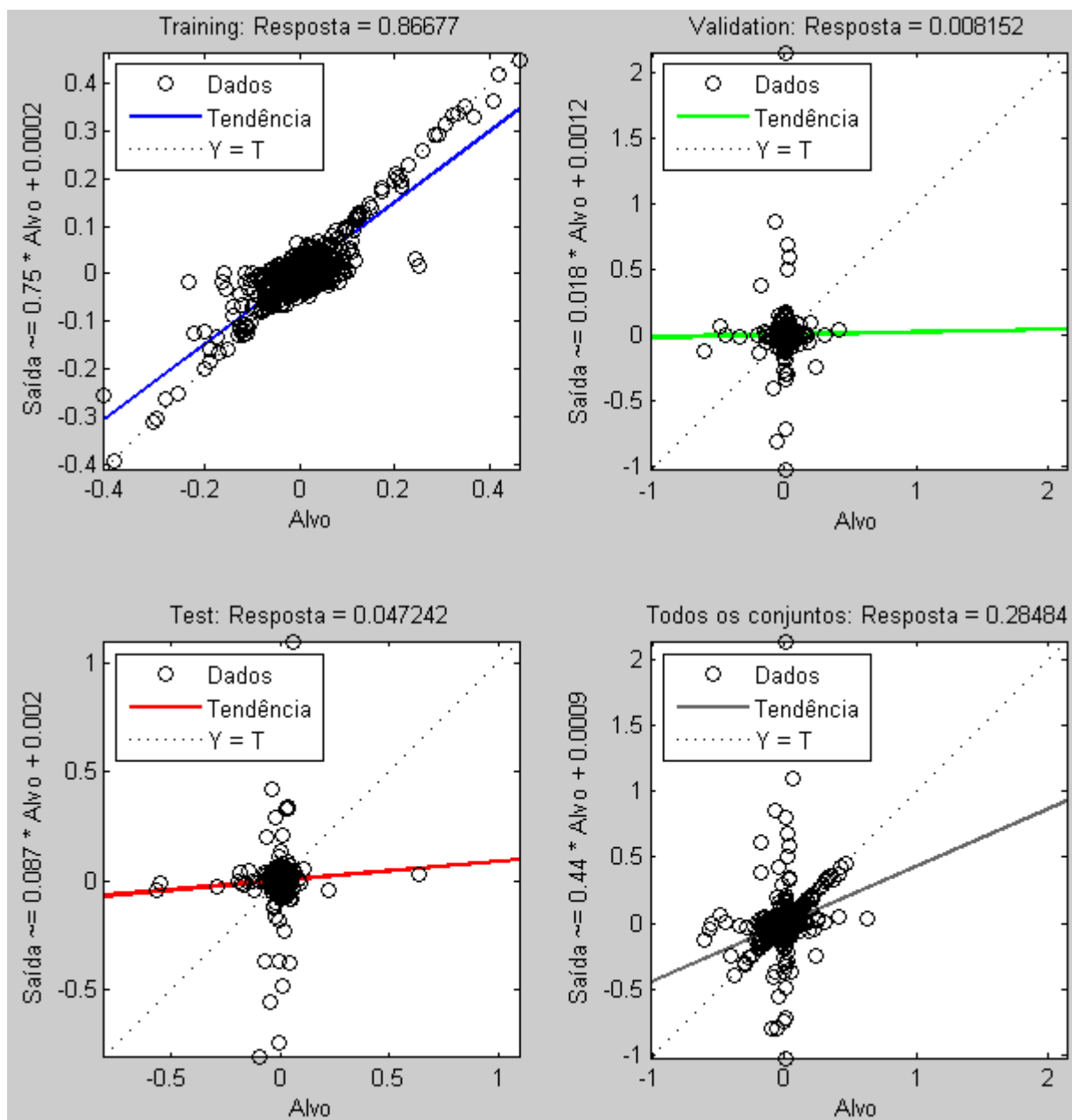


Figura B.16: Regressão das saídas alvo 20% para a segunda abordagem RGB.

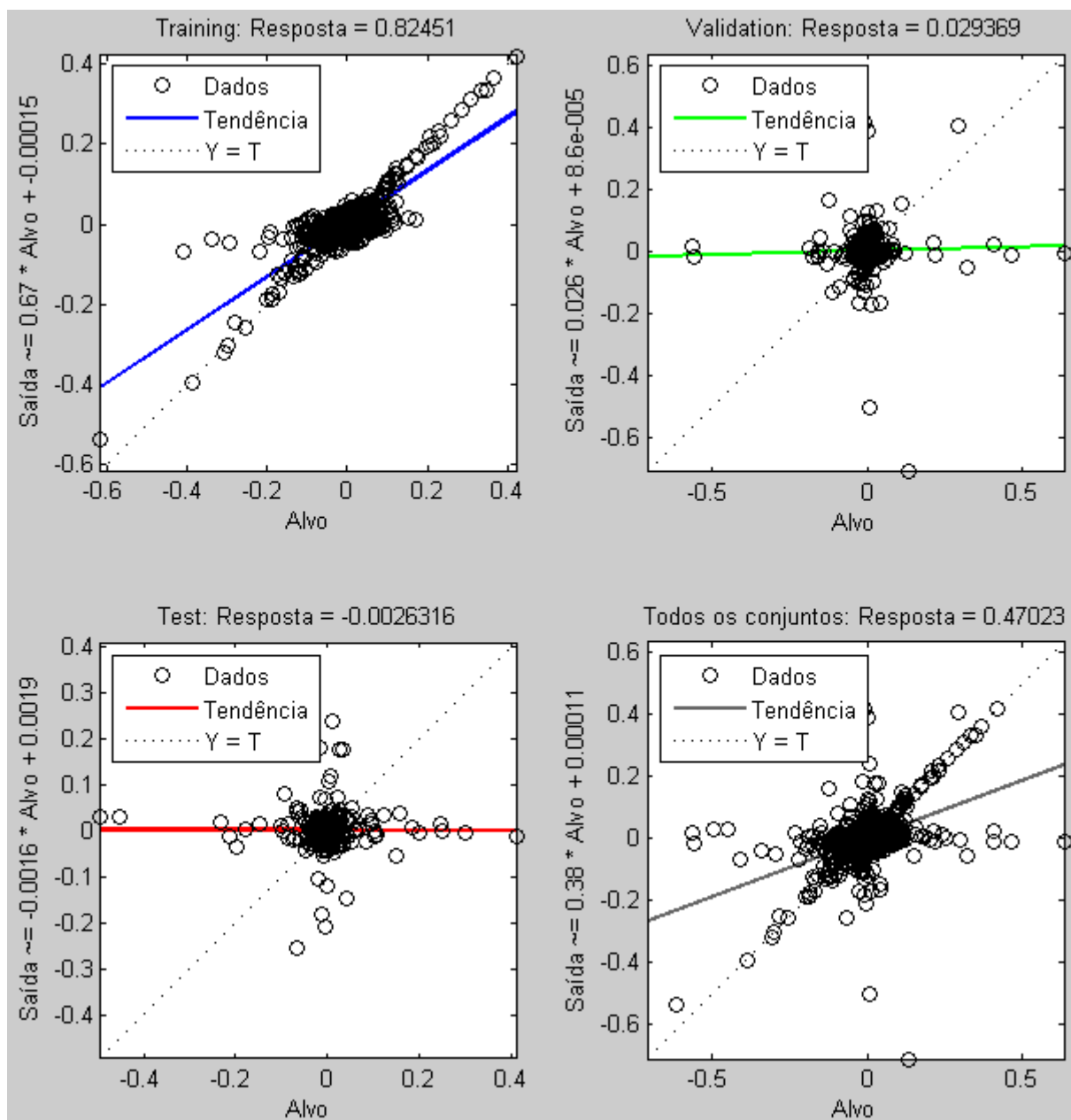


Figura B.17: Regressão das saídas alvo 30% para a segunda abordagem RGB.

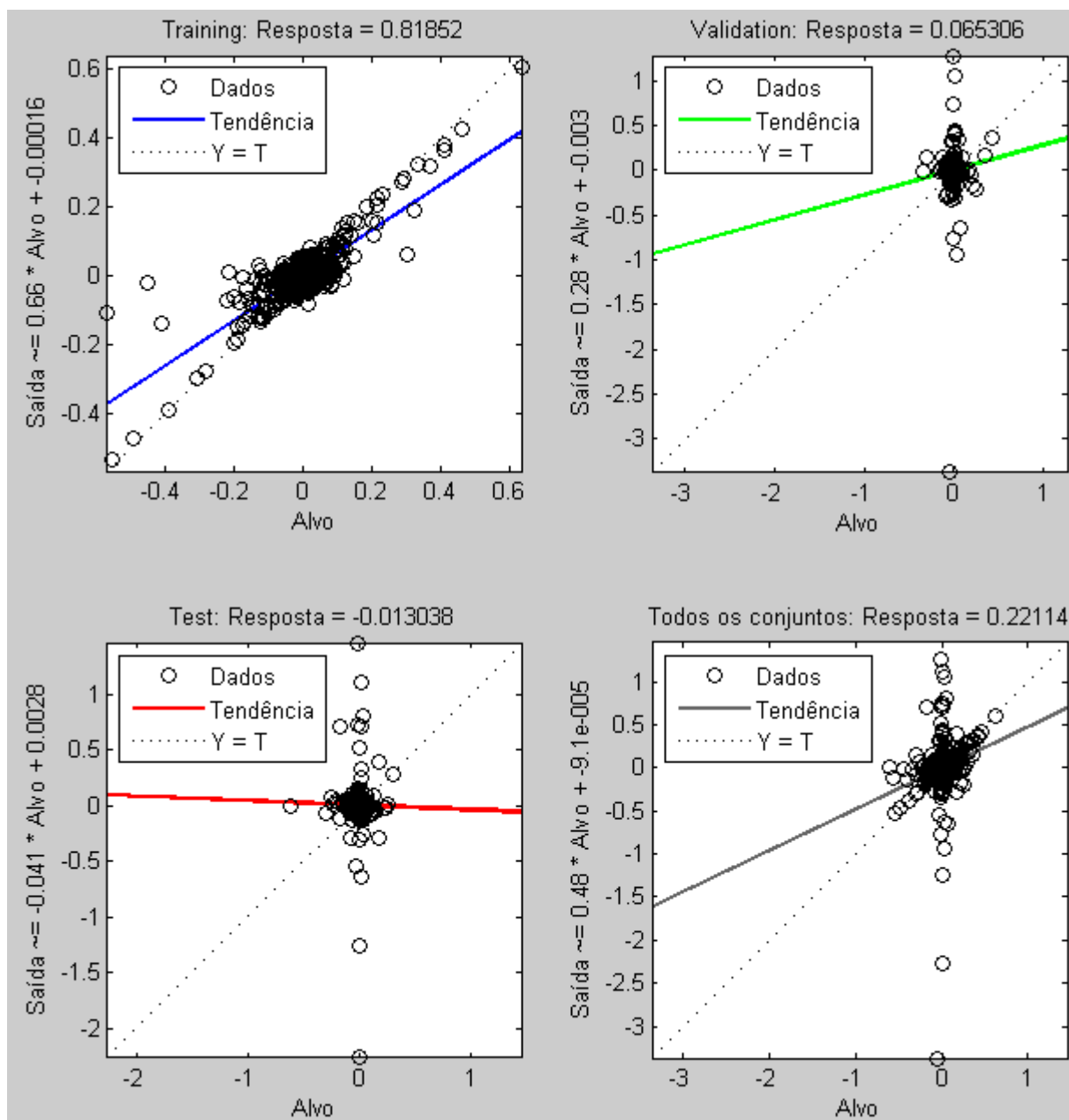


Figura B.18: Regressão das saídas alvo 40% para a segunda abordagem RGB.

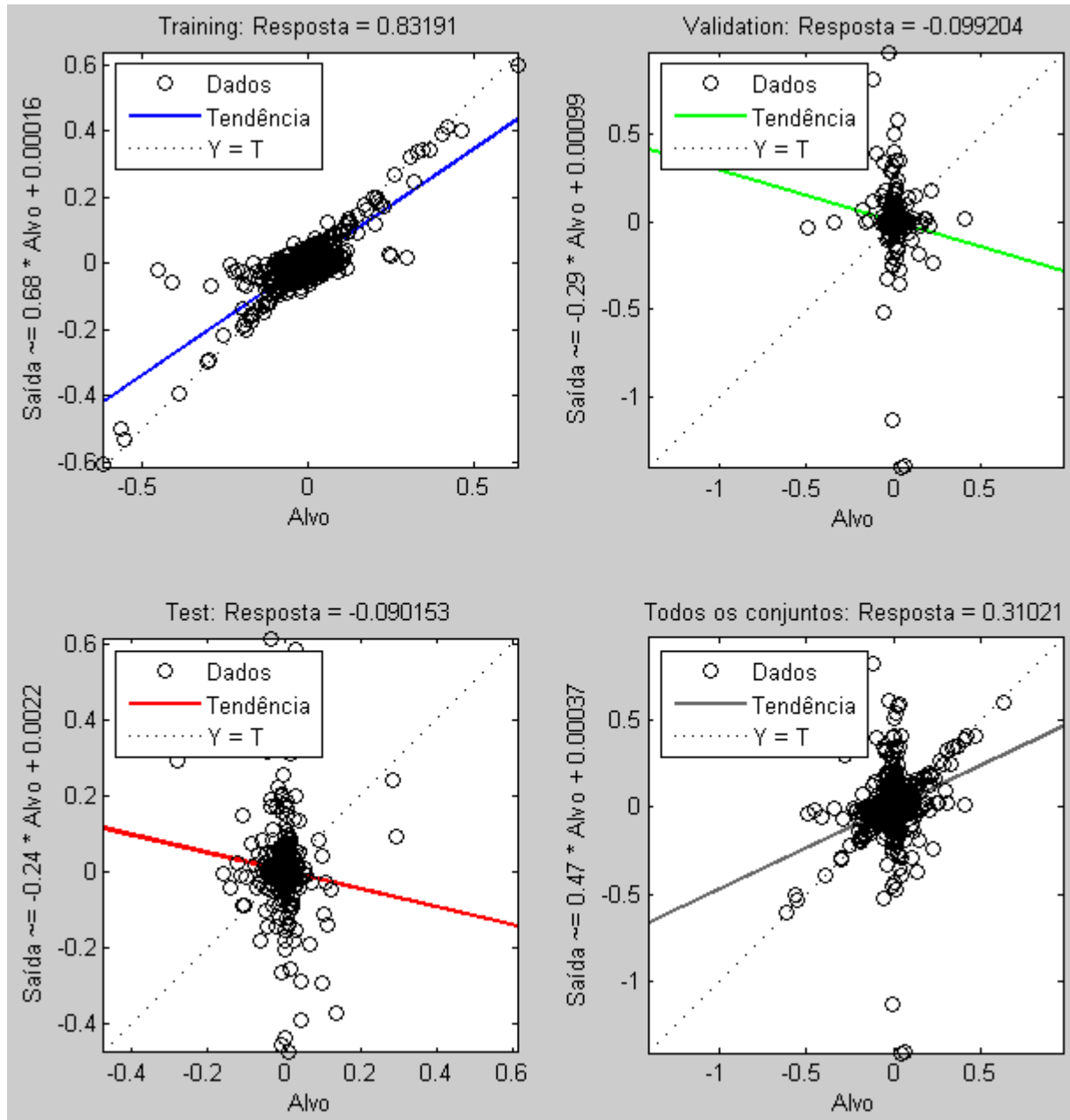


Figura B.19: Regressão das saídas alvo 50% para a segunda abordagem RGB.

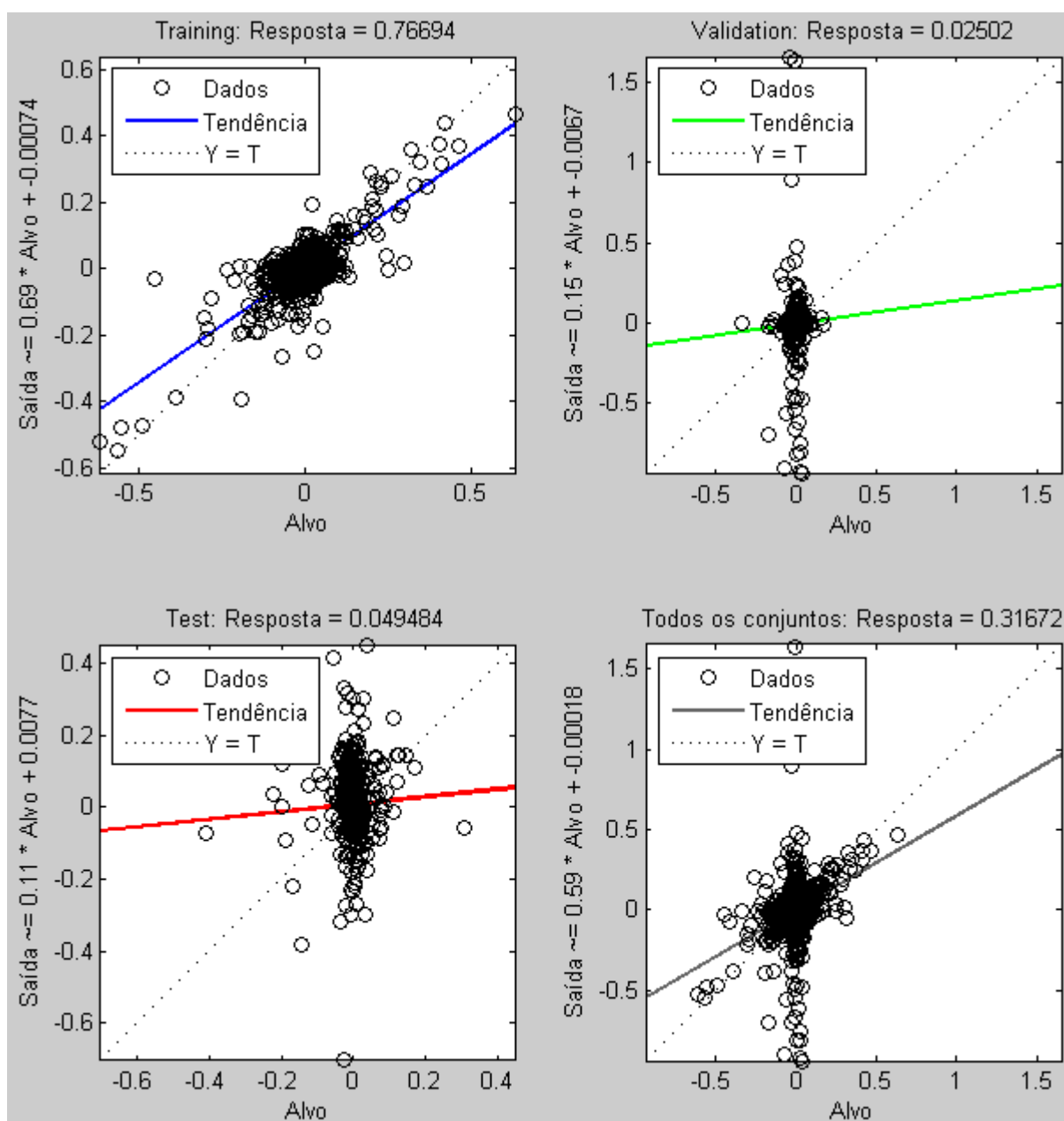


Figura B.20: Regressão das saídas alvo 75% para a segunda abordagem RGB.

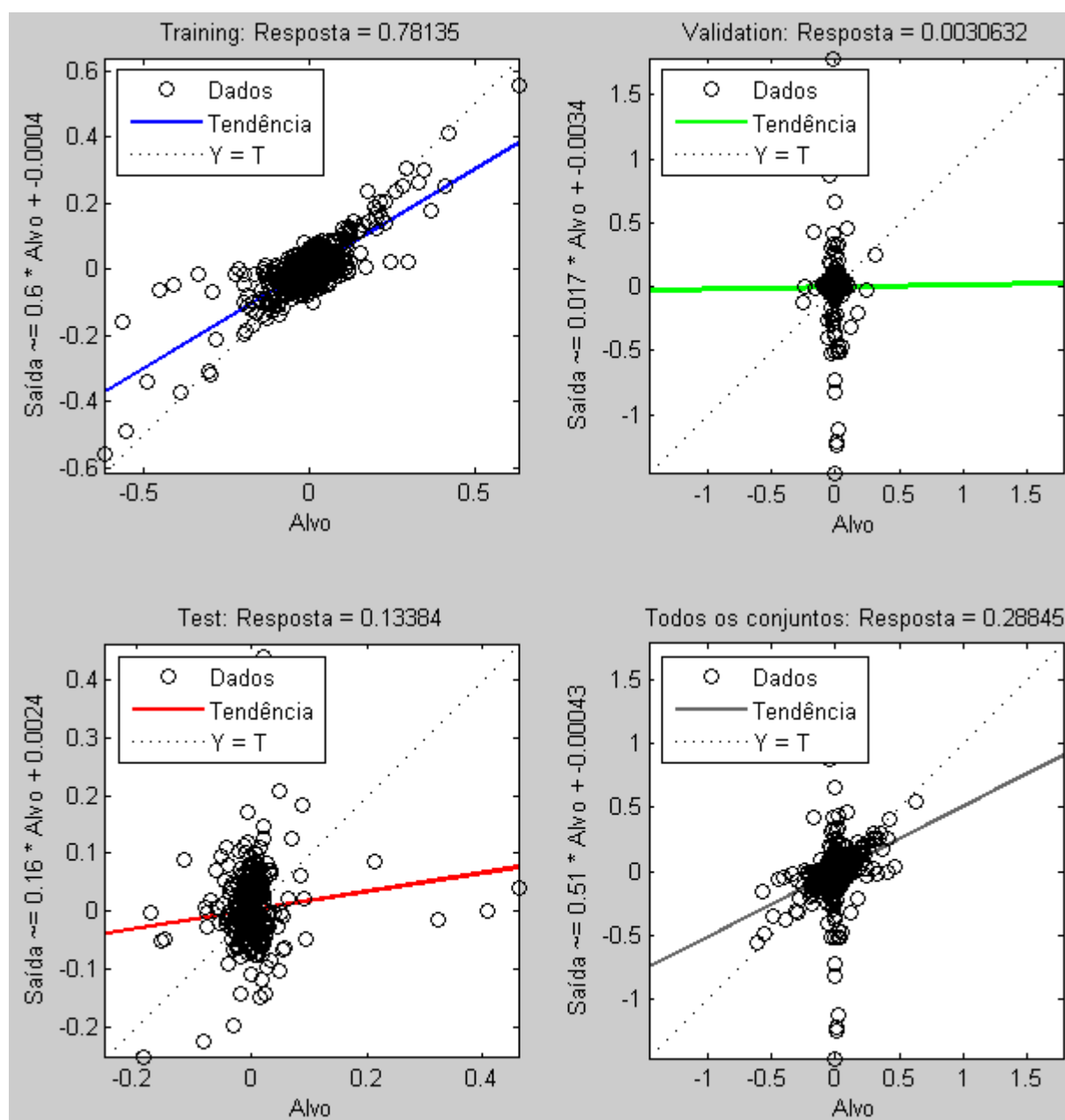


Figura B.21: Regressão das saídas alvo 100% para a segunda abordagem RGB.

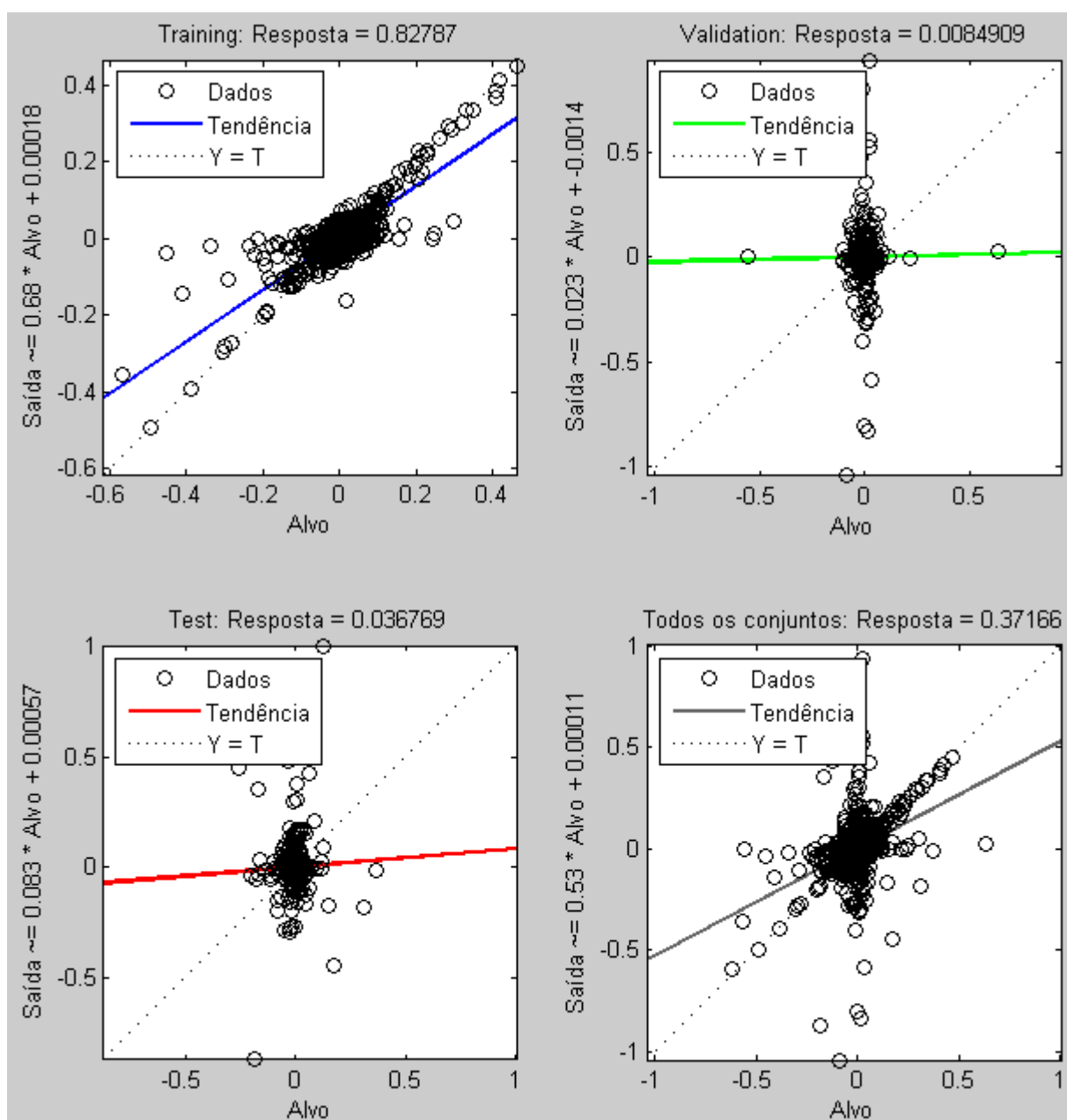


Figura B.22: Regressão das saídas alvo 10% para a segunda abordagem RGB.

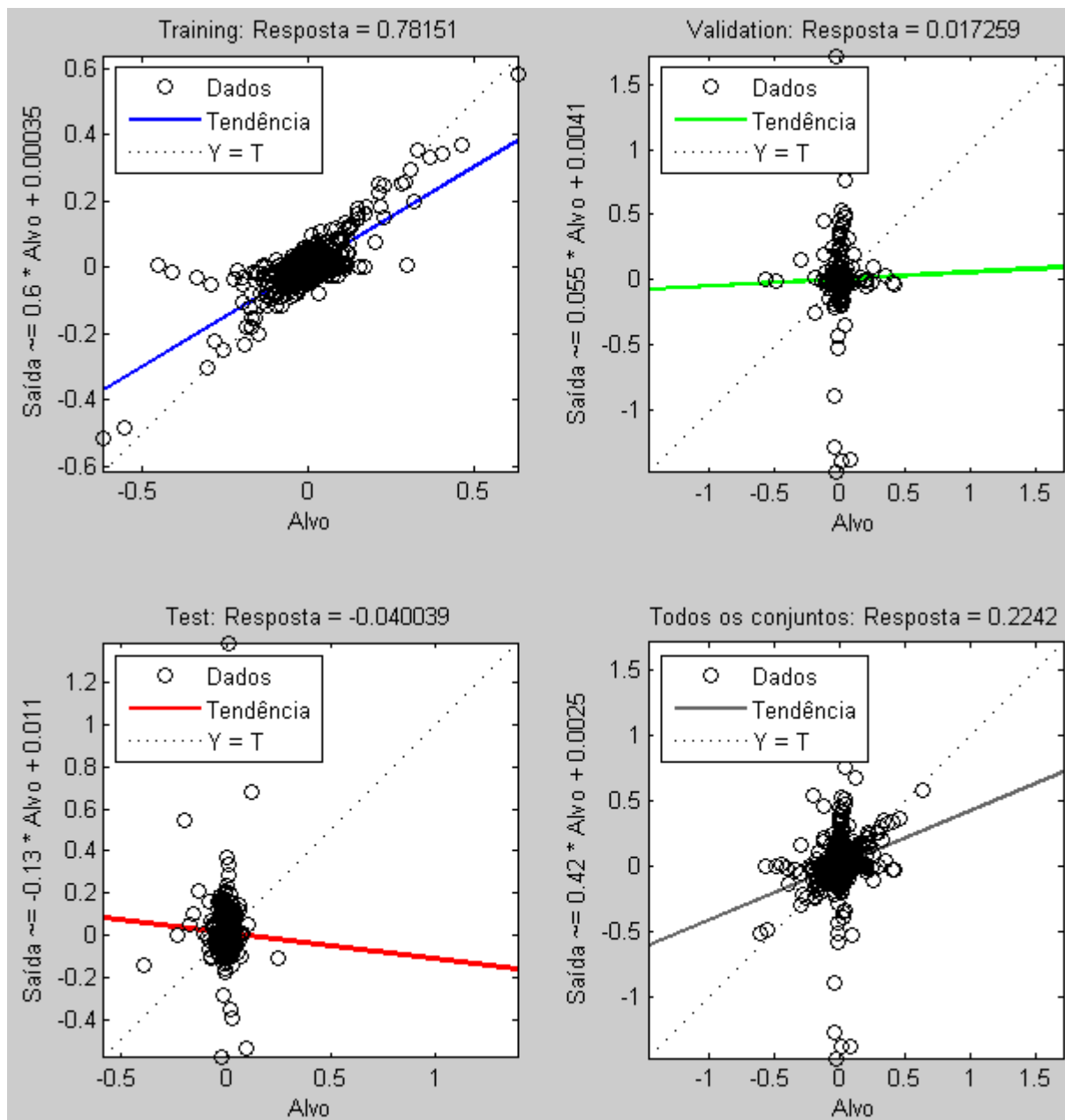


Figura B.23: Regressão das saídas alvo 20% para a segunda abordagem RGB.

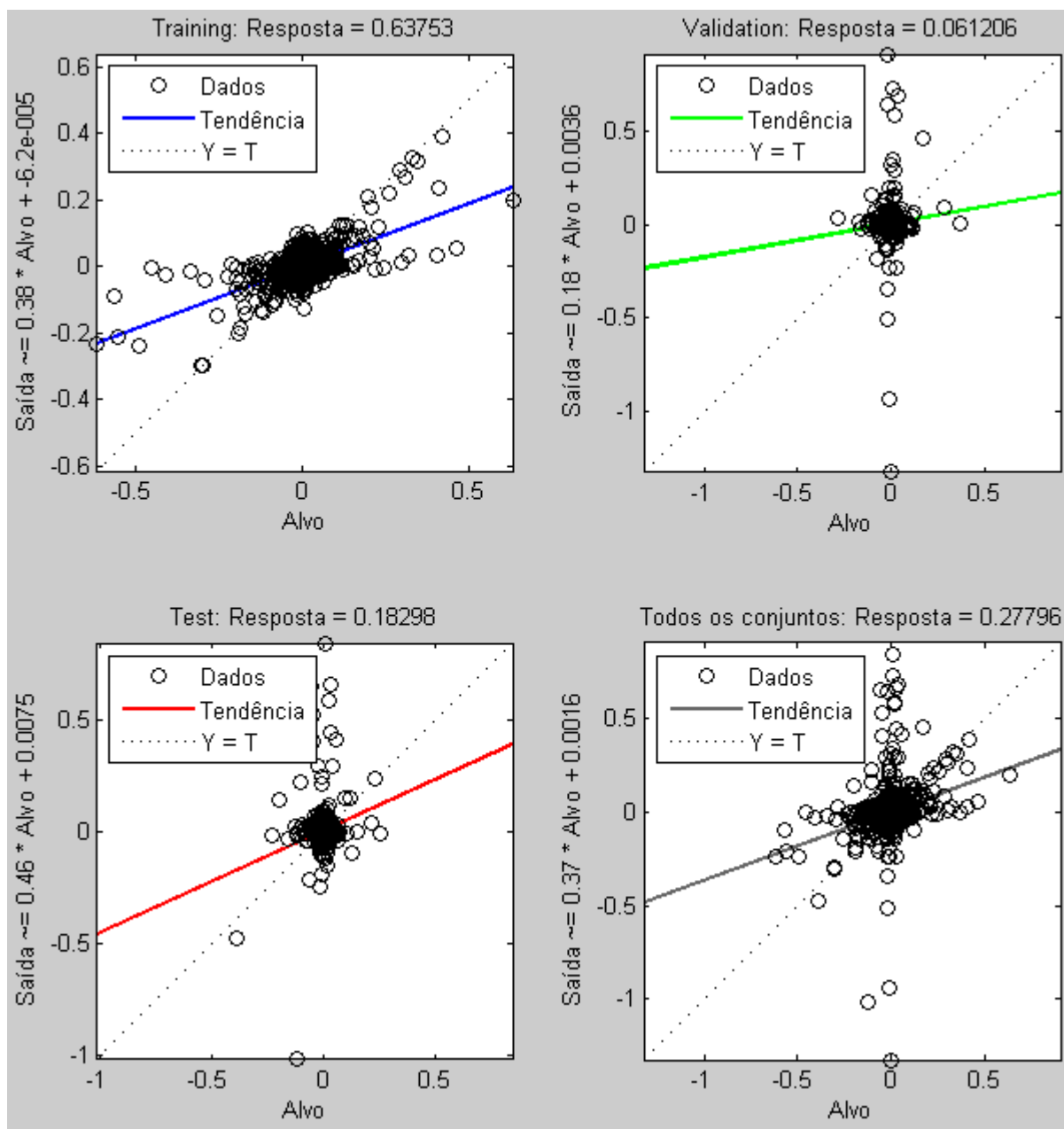


Figura B.24: Regressão das saídas alvo 30% para a segunda abordagem RGB.

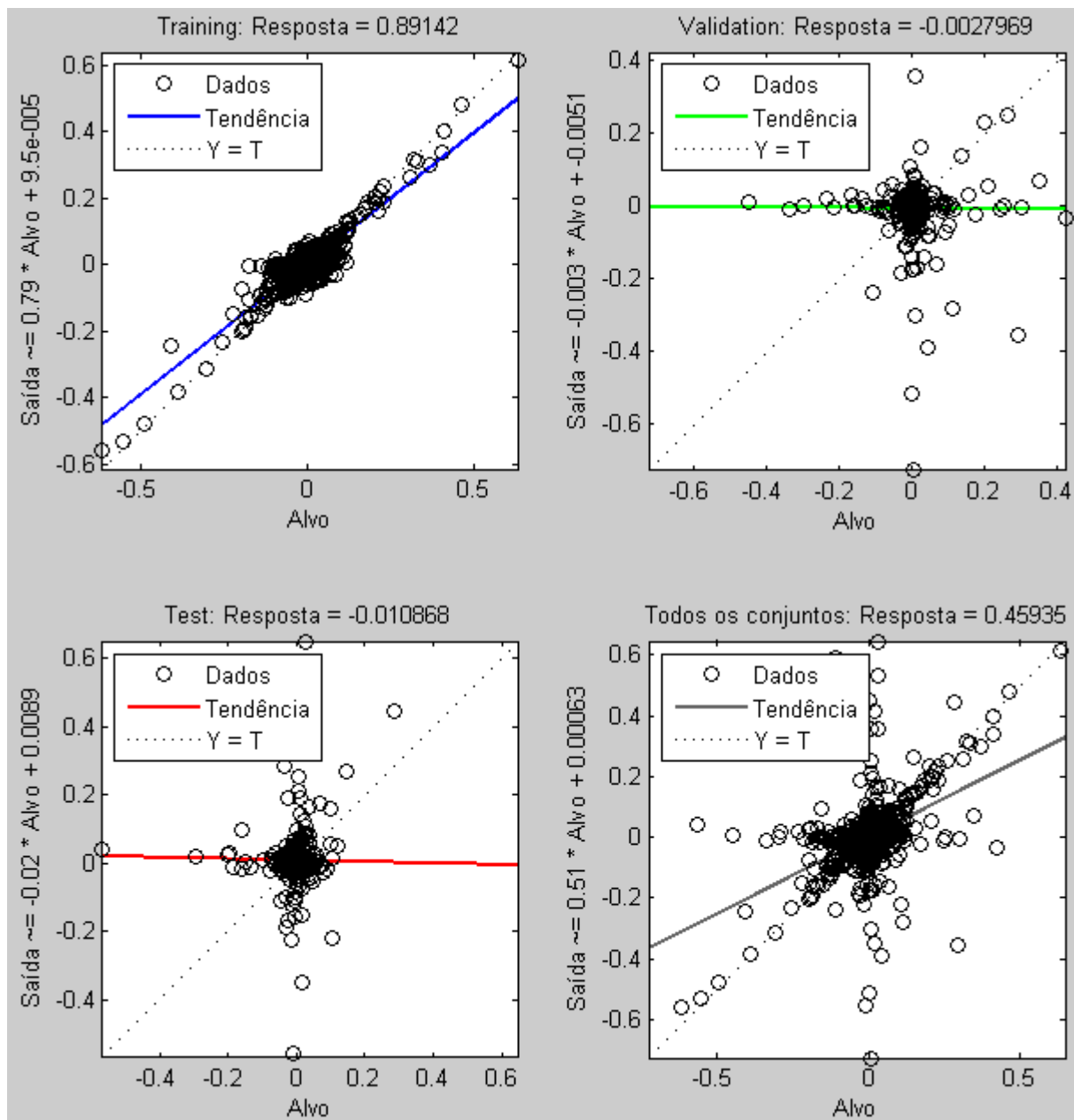


Figura B.25: Regressão das saídas alvo 40% para a segunda abordagem RGB.

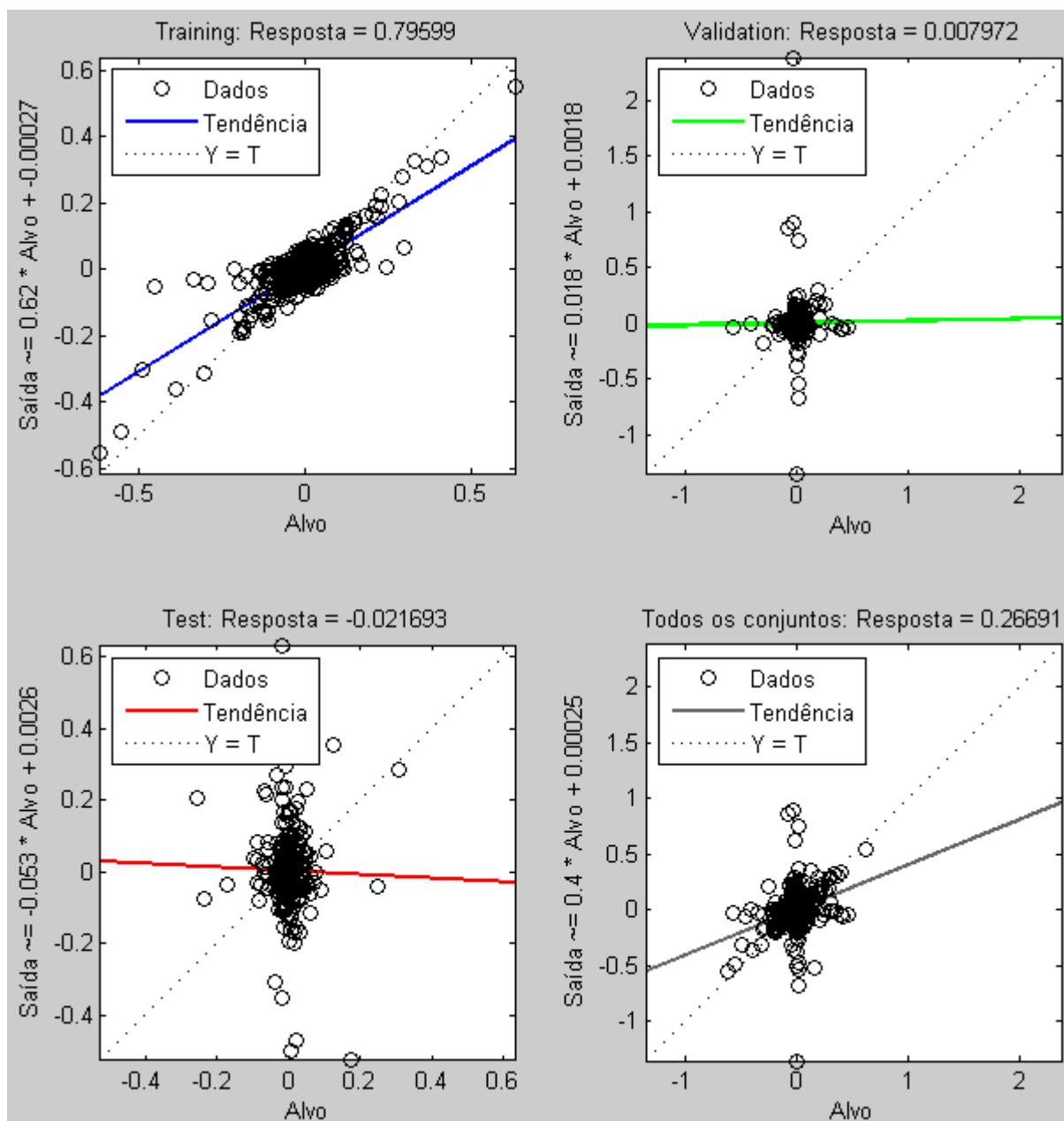


Figura B.26: Regressão das saídas alvo 50% para a segunda abordagem RGB.

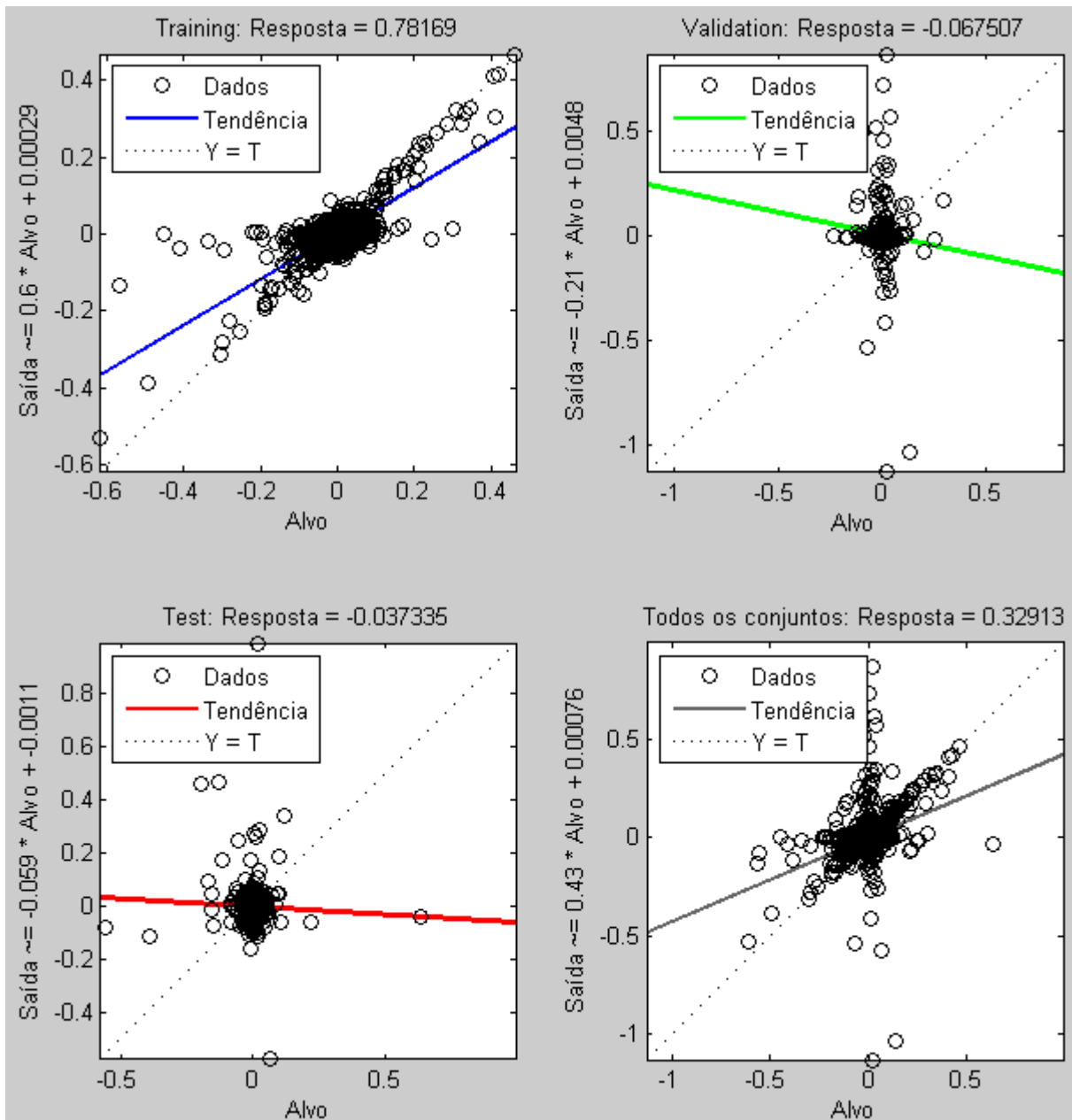


Figura B.27: Regressão das saídas alvo 75% para a segunda abordagem RGB.

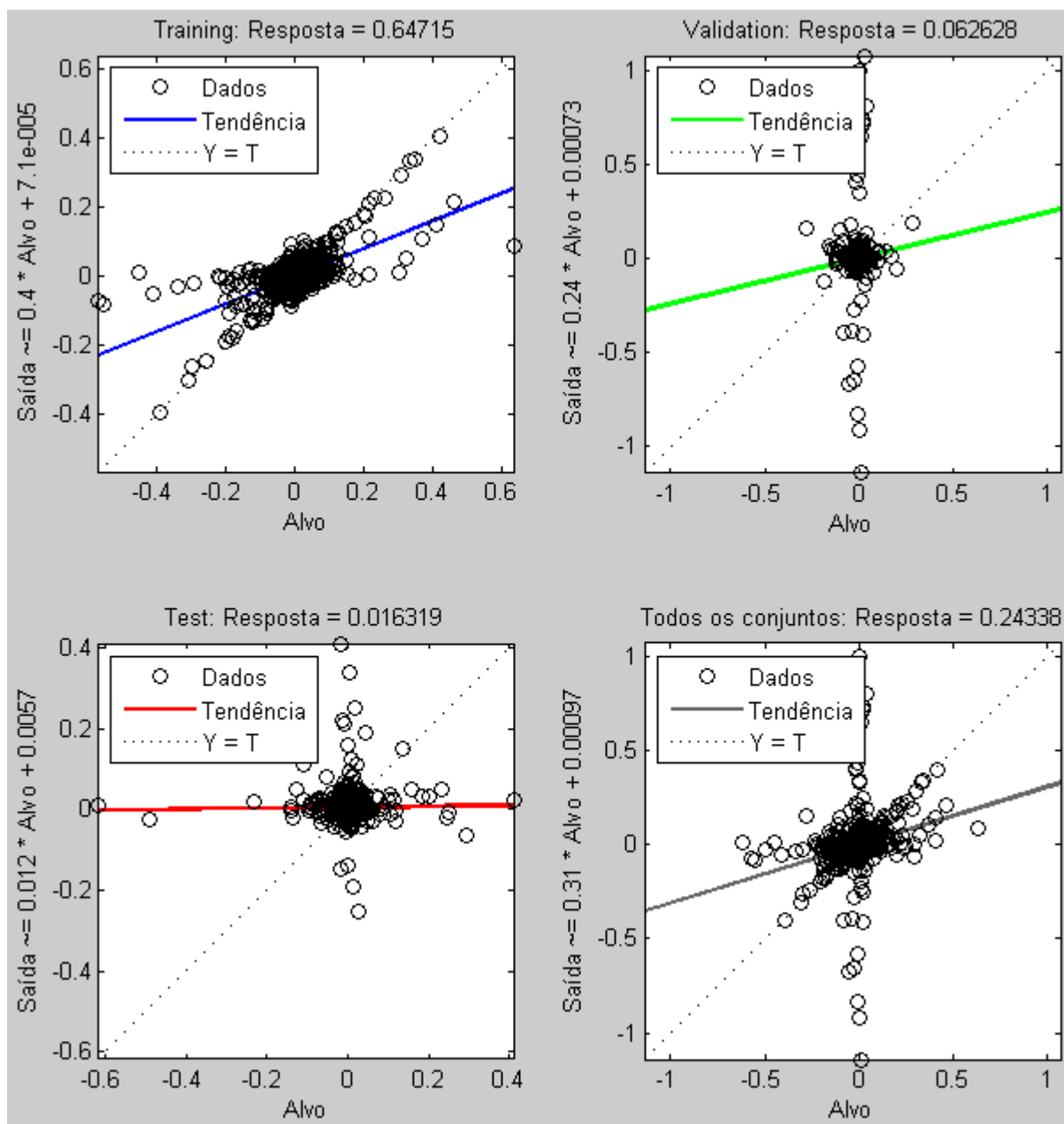


Figura B.28: Regressão das saídas alvo 100% para a segunda abordagem RGB.

APÊNDICE C. EQM/MSE do classificador

RNA para as duas abordagens

Os gráficos abaixo denotam a relação entre o erro quadrático médio (*Mean Squared Error* MSE/EQM) e quantidade de épocas na obtenção do melhor índice de desempenho durante a validação dos dados para com os conjuntos de treinamento e teste. Quanto menor o EQM, melhor é o desempenho do classificador.

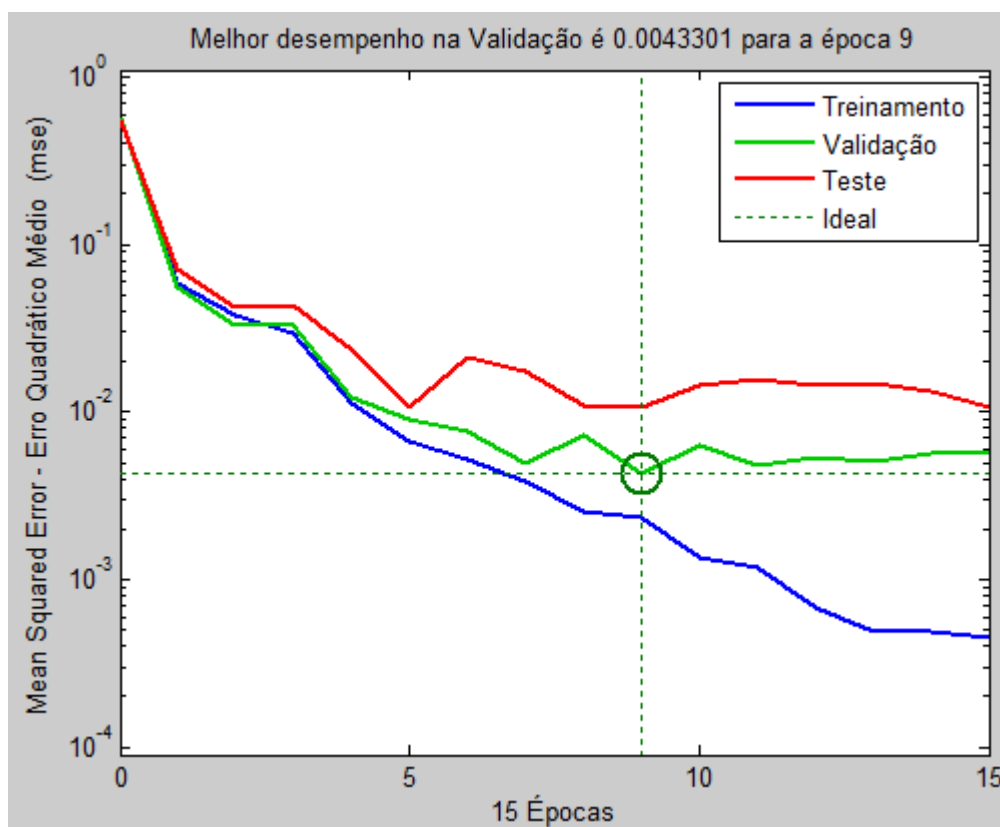


Figura C.1: MSE 10% para a primeira abordagem RGB.

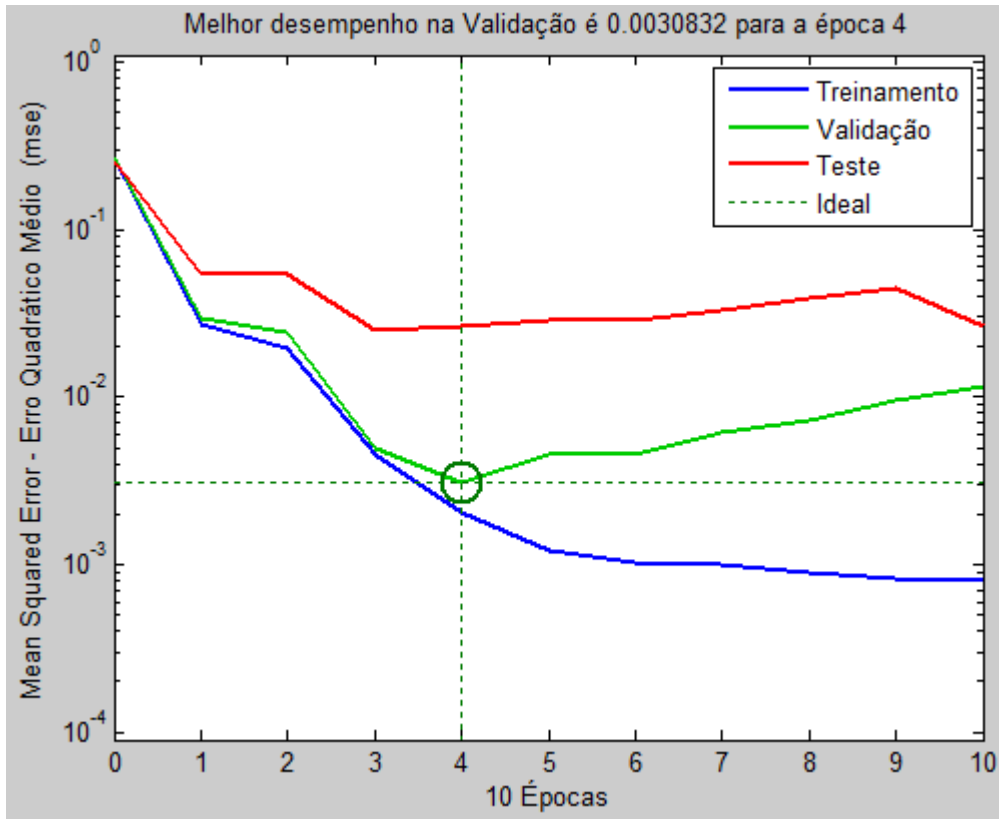


Figura C.2: MSE 20% para a primeira abordagem RGB.

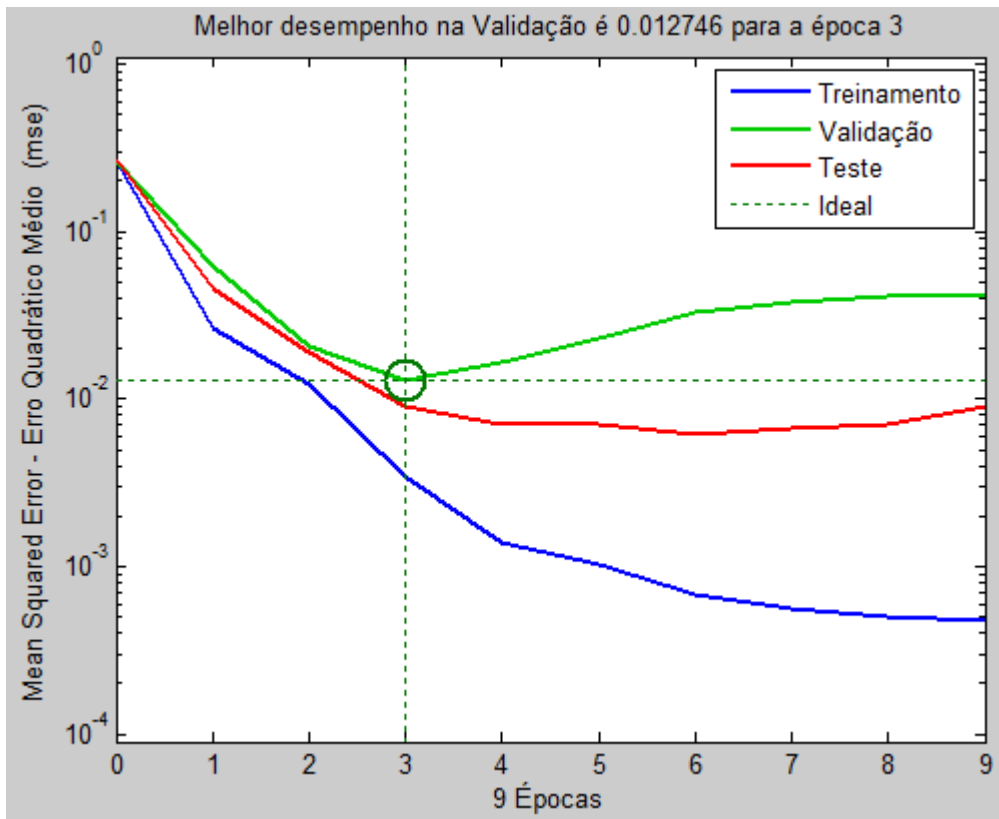


Figura C.3: MSE 30% para a primeira abordagem RGB.

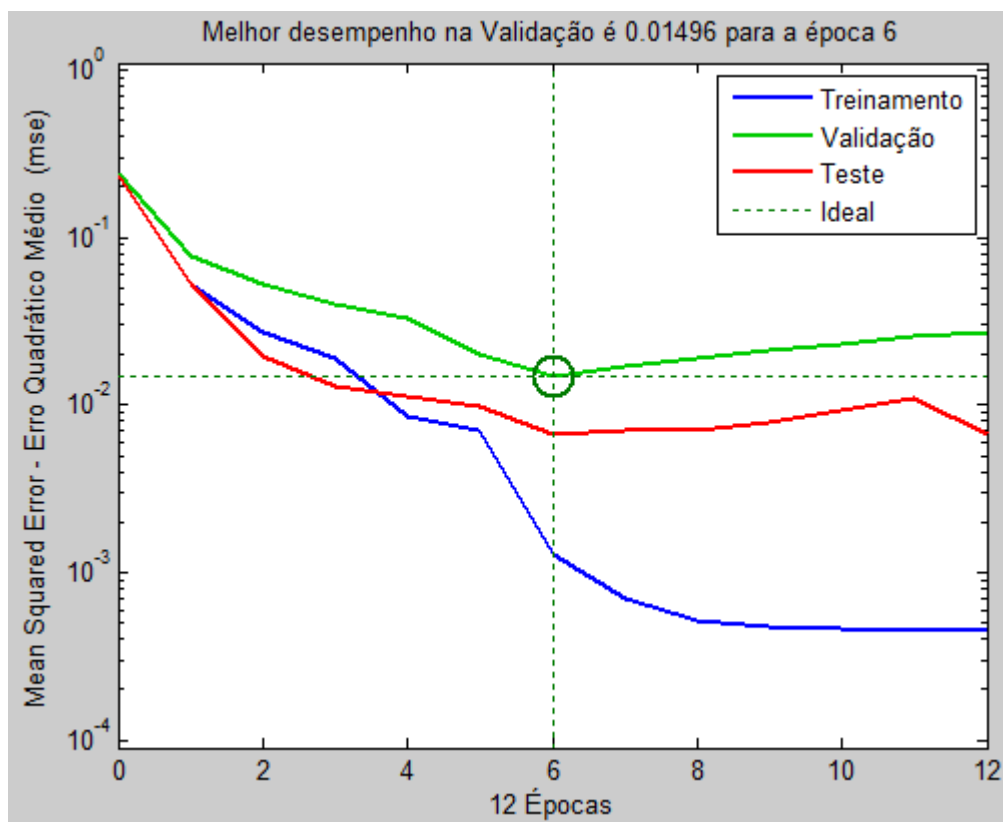


Figura C.4: MSE 40% para a primeira abordagem RGB.

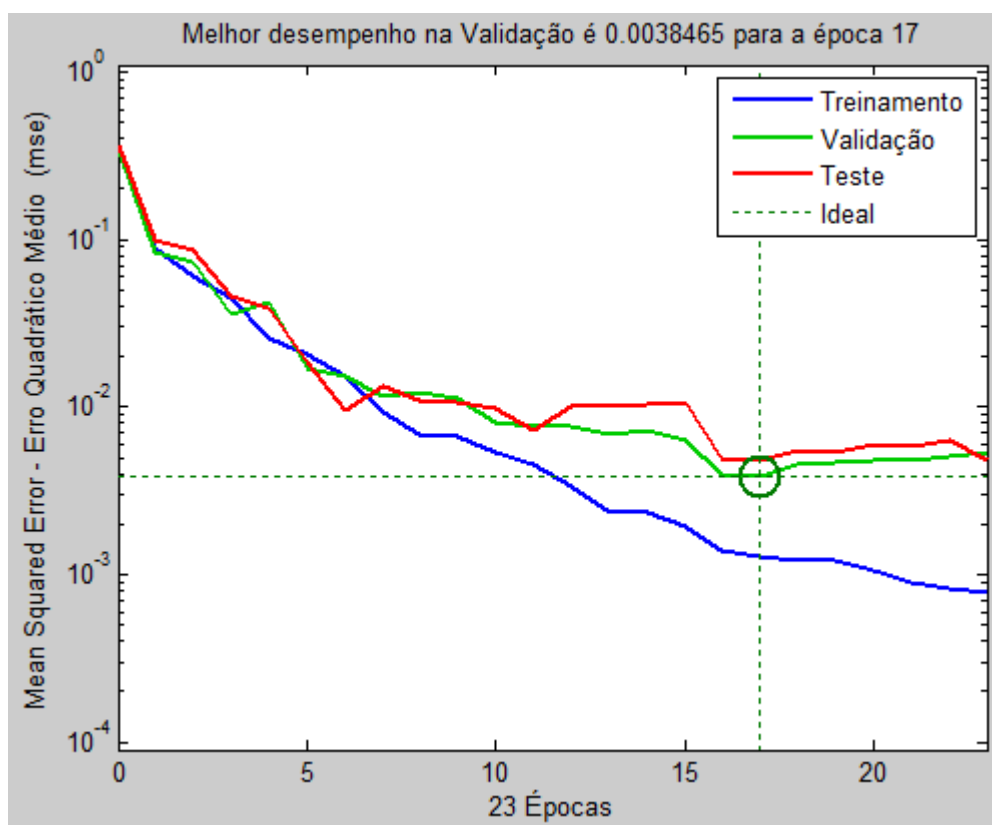


Figura C.5: MSE 50% para a primeira abordagem RGB.

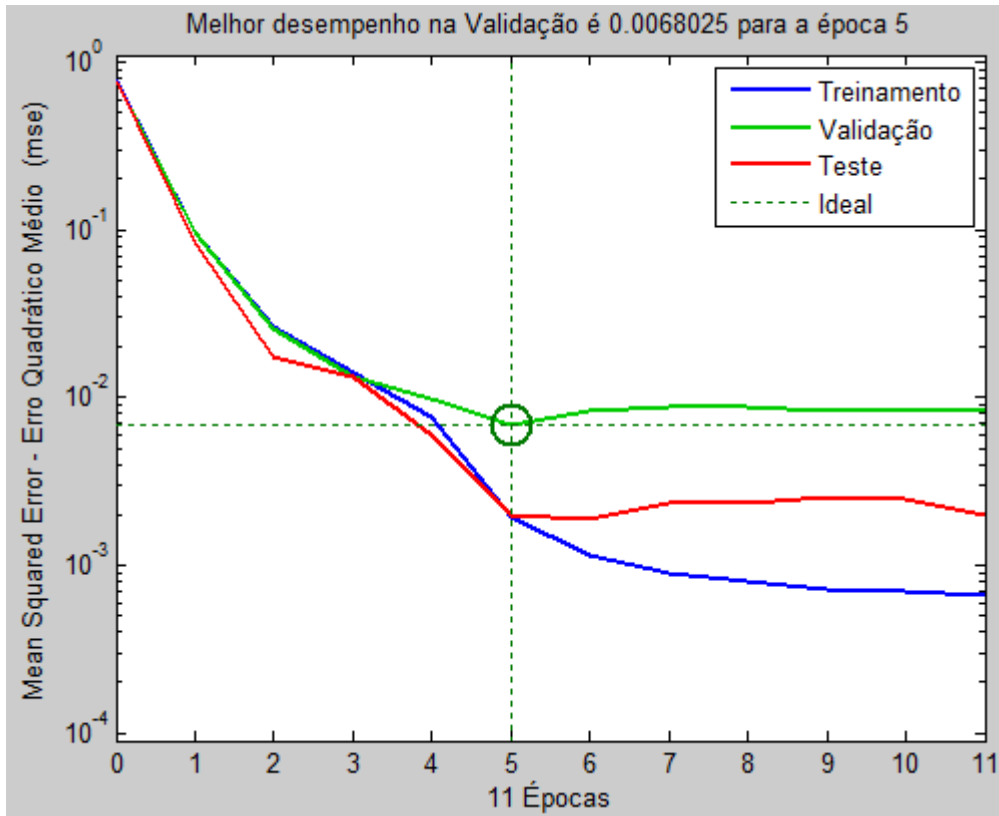


Figura C.6: MSE 75% para a primeira abordagem RGB.

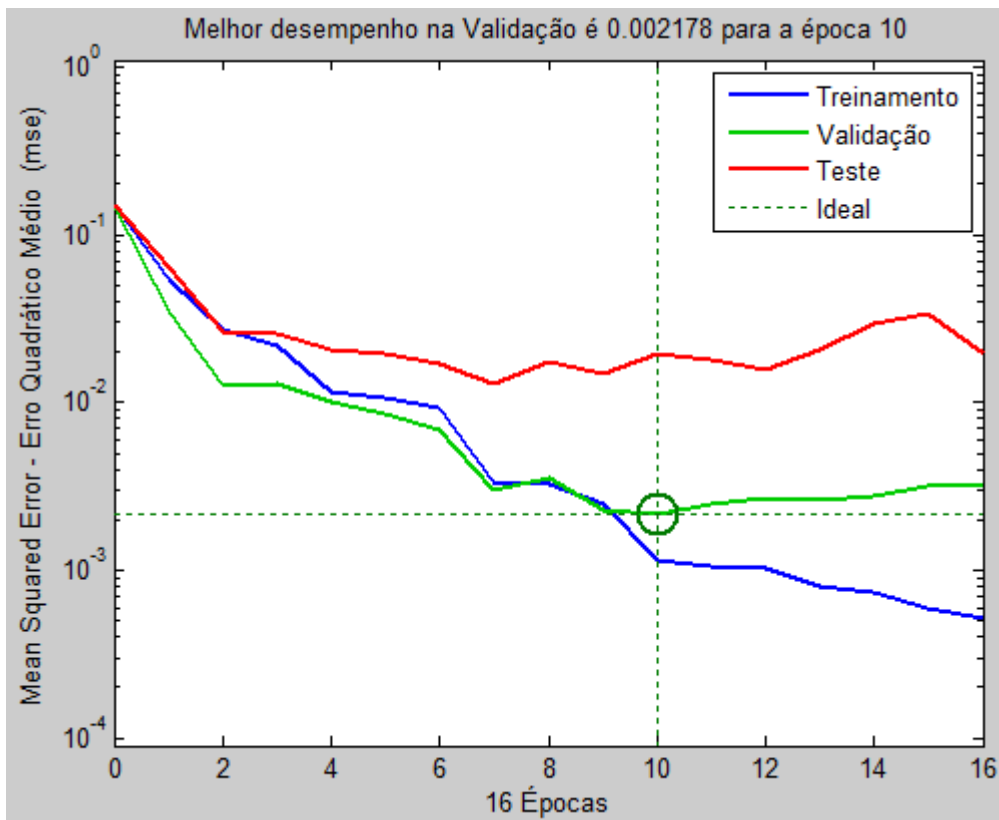


Figura C.7: MSE 100% para a primeira abordagem RGB.

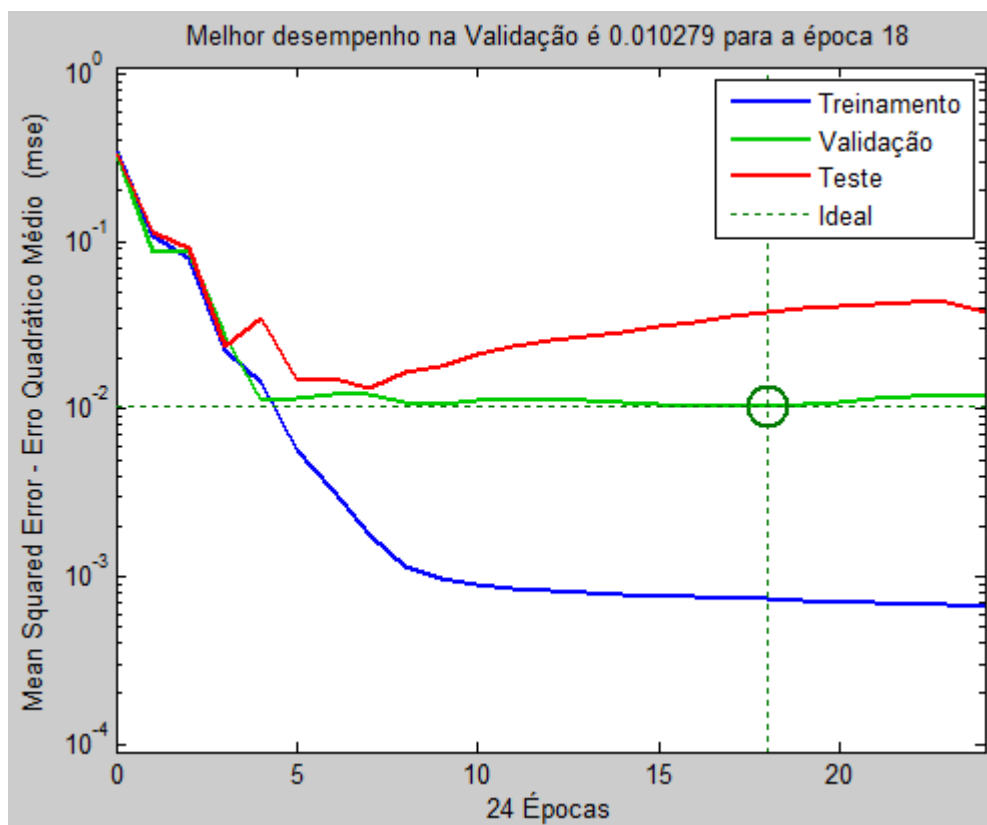


Figura C.8: MSE 10% para a primeira abordagem HSV.

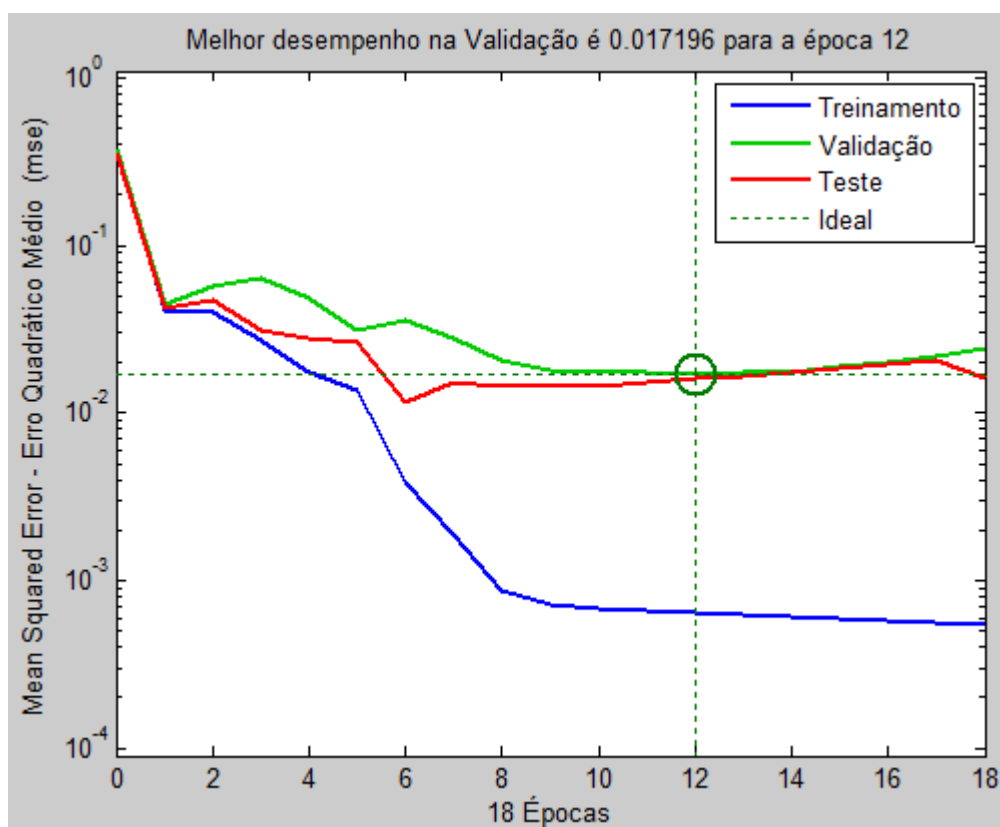


Figura C.9: MSE 20% para a primeira abordagem HSV.

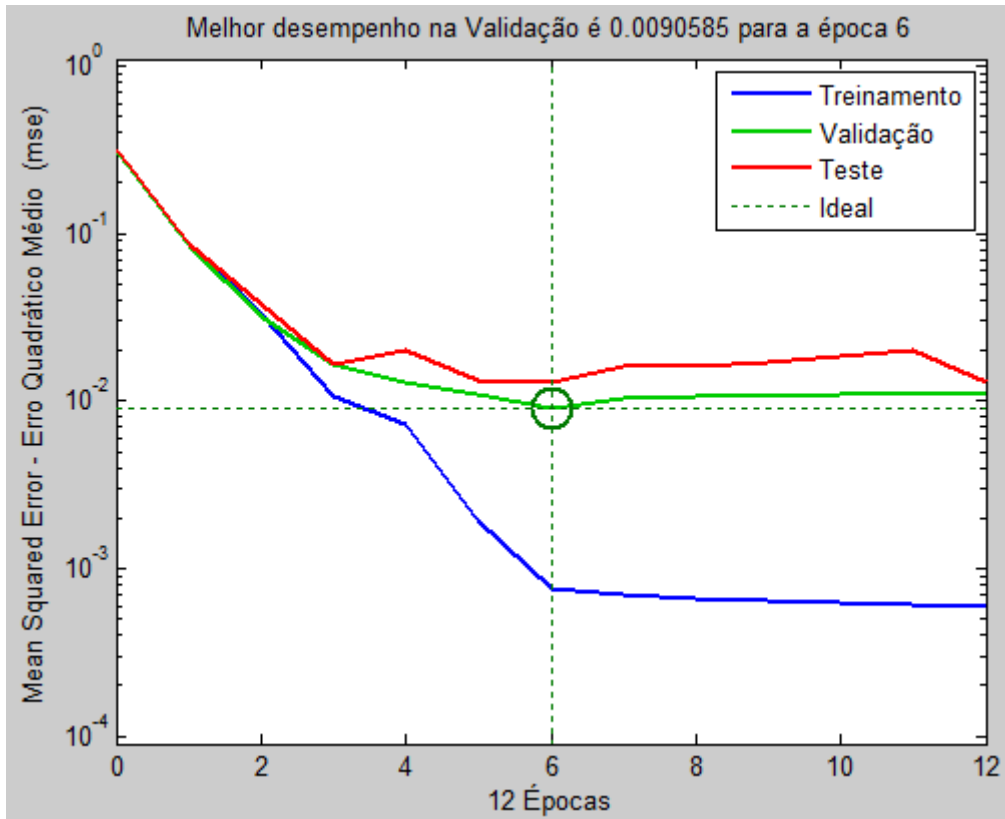


Figura C.10: MSE 30% para a primeira abordagem HSV.

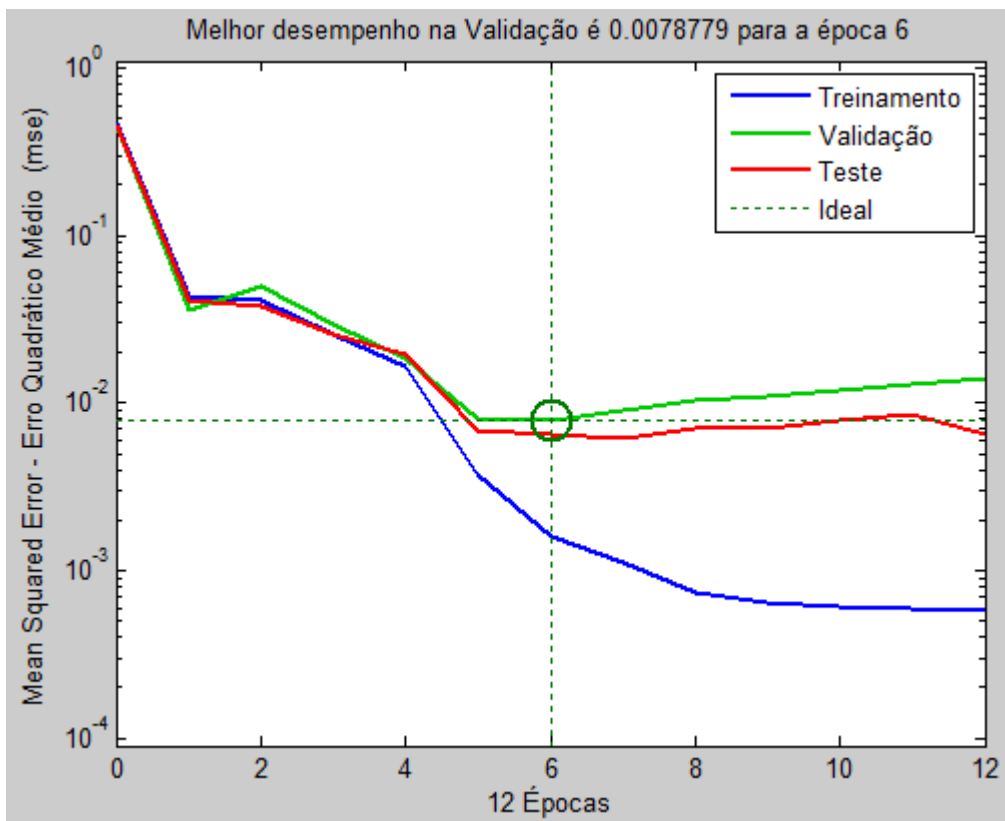


Figura C.11: MSE 40% para a primeira abordagem HSV.

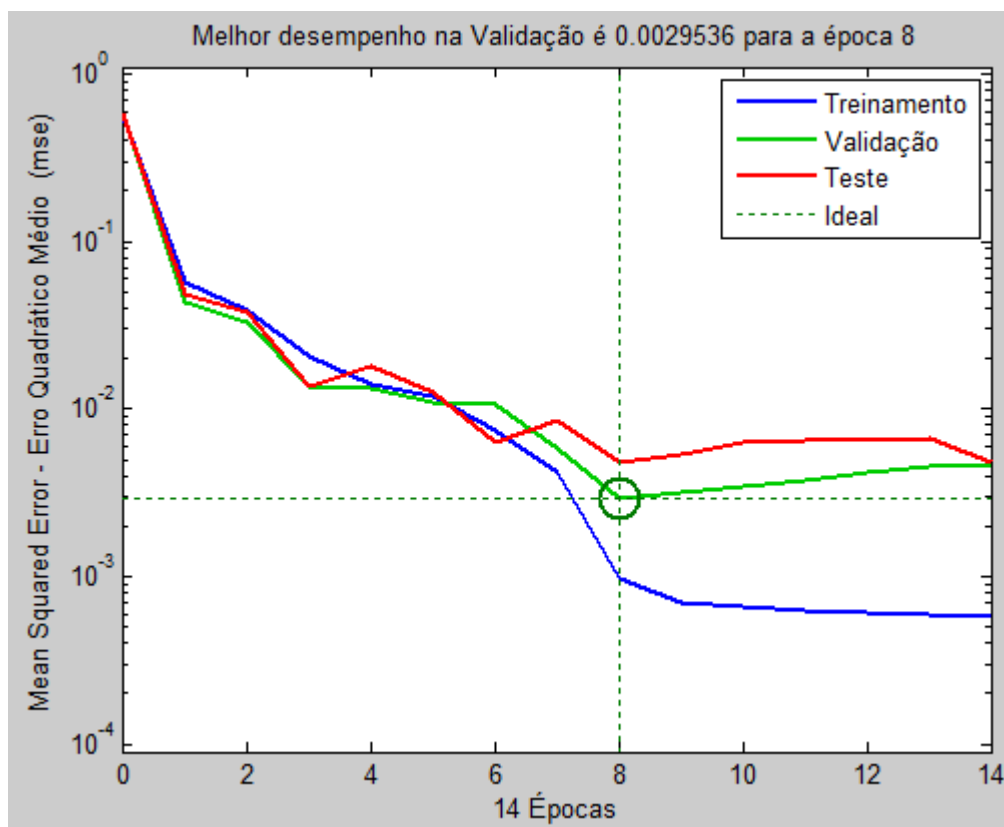


Figura C.12: MSE 50% para a primeira abordagem HSV.

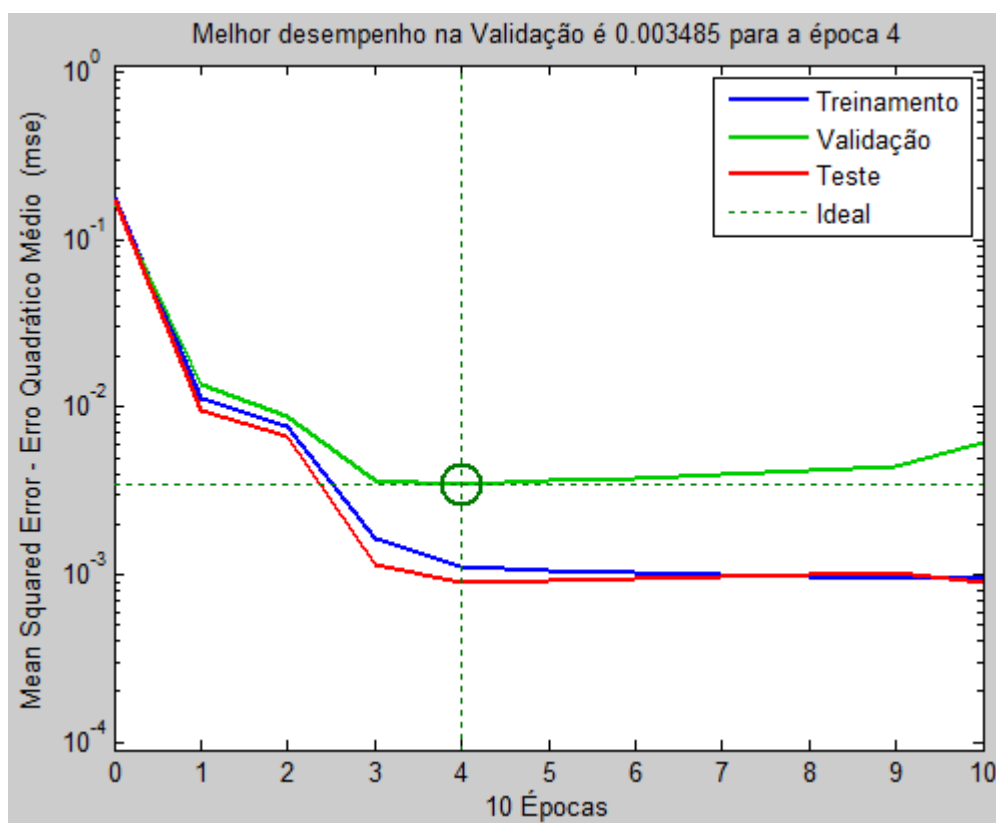


Figura C.13: MSE 75% para a primeira abordagem HSV.

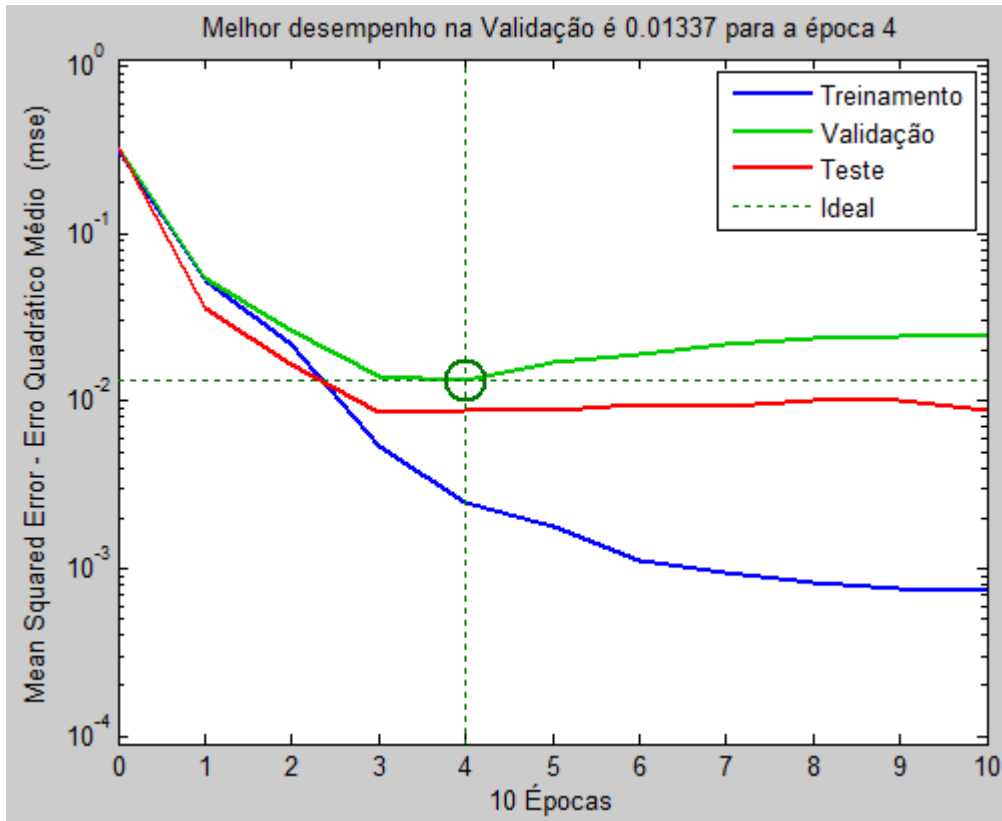


Figura C.14: MSE 100% para a primeira abordagem HSV.

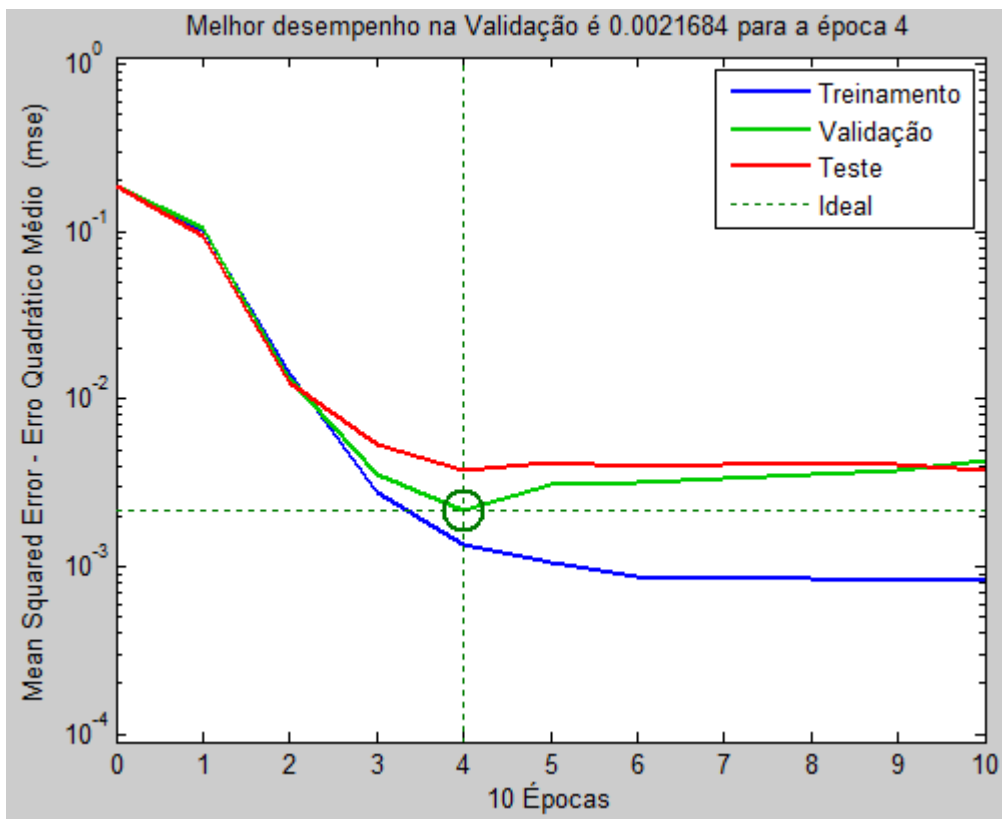


Figura C.15: MSE 10% para a segunda abordagem RGB.

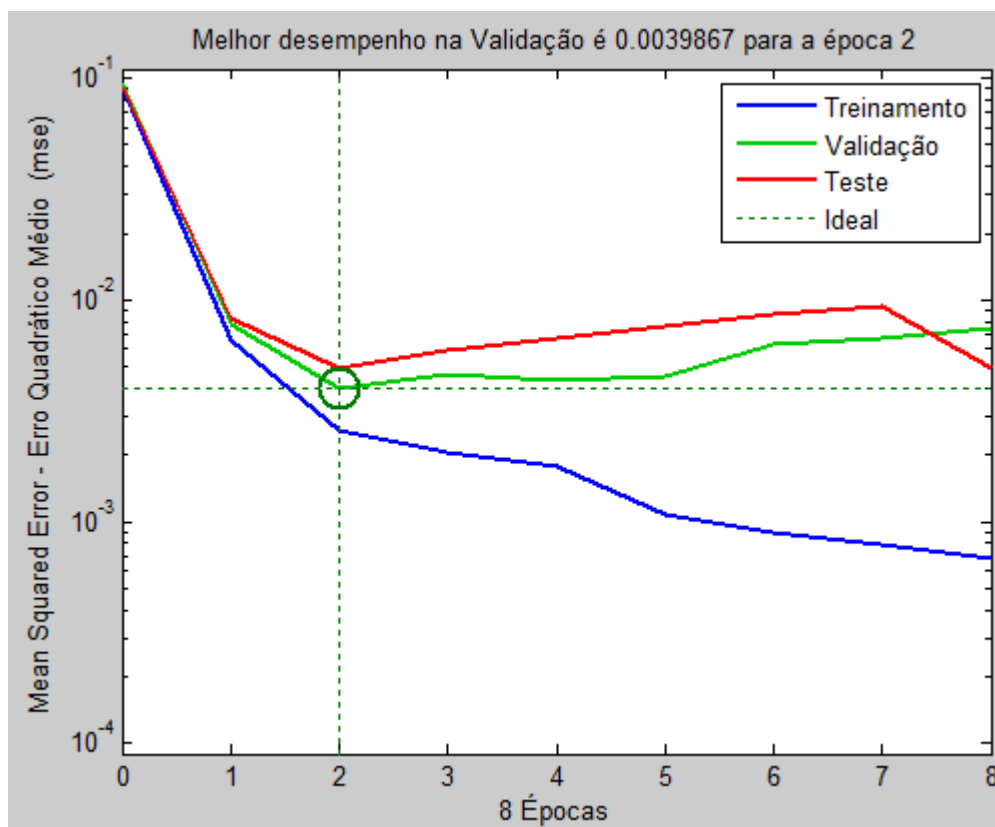


Figura C.16: MSE 20% para a segunda abordagem RGB.

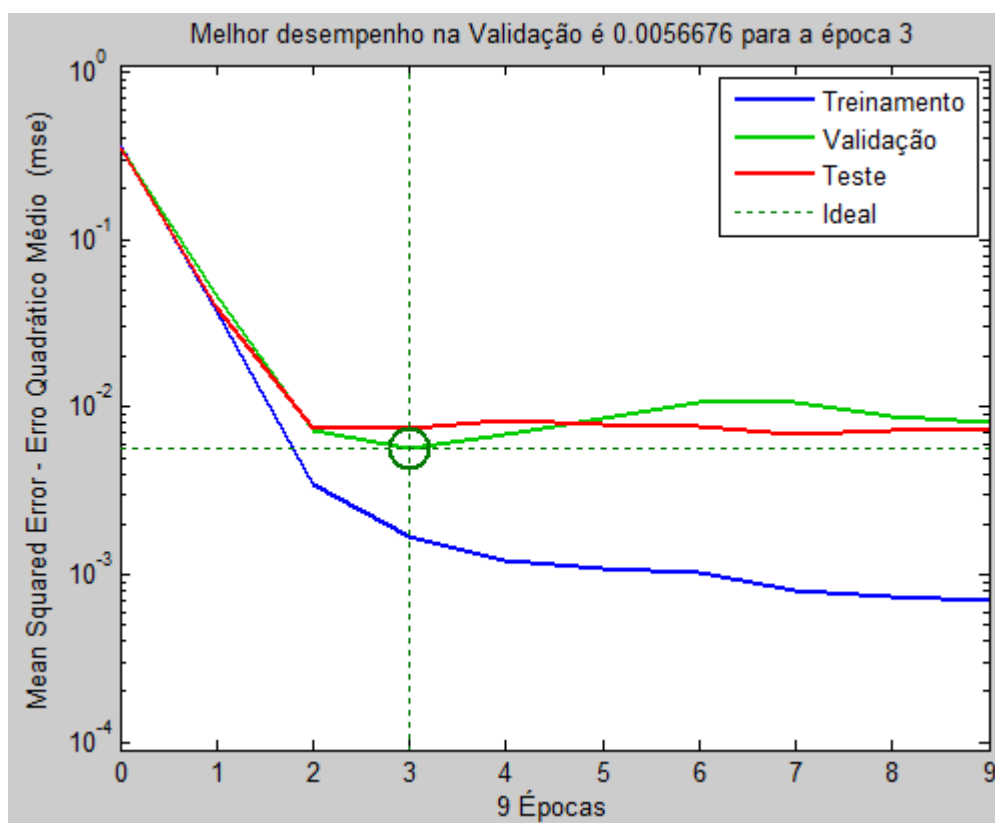


Figura C.17: MSE 30% para a segunda abordagem RGB.

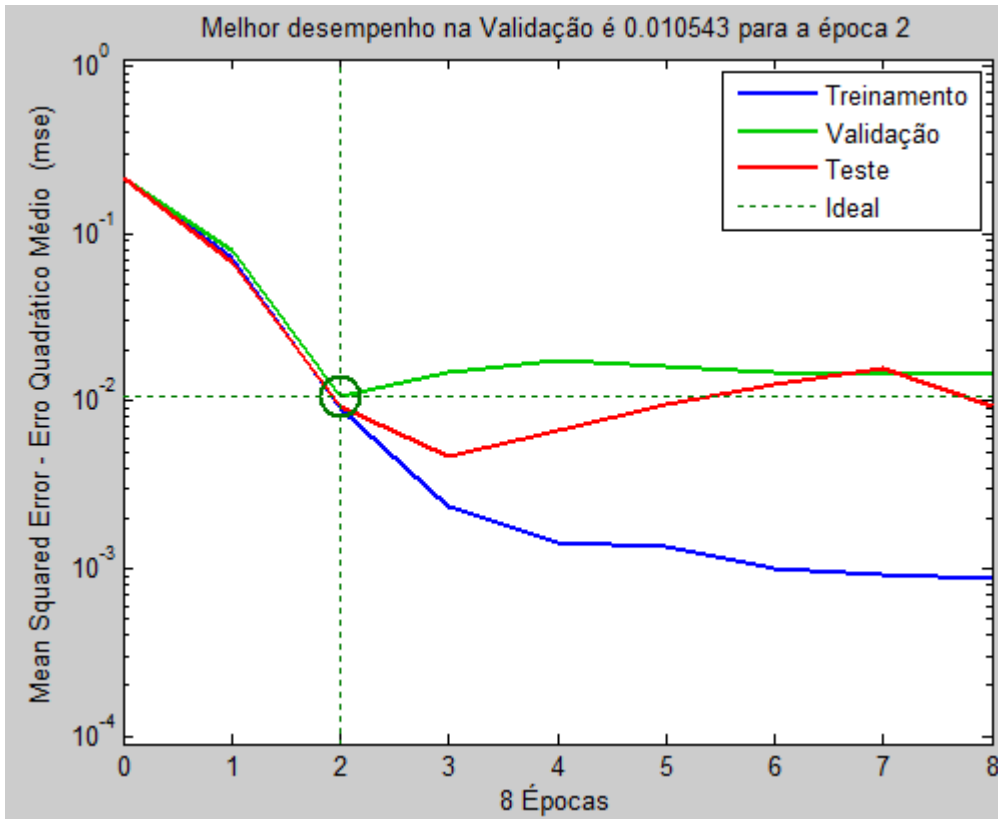


Figura C.18: MSE 40% para a segunda abordagem RGB.

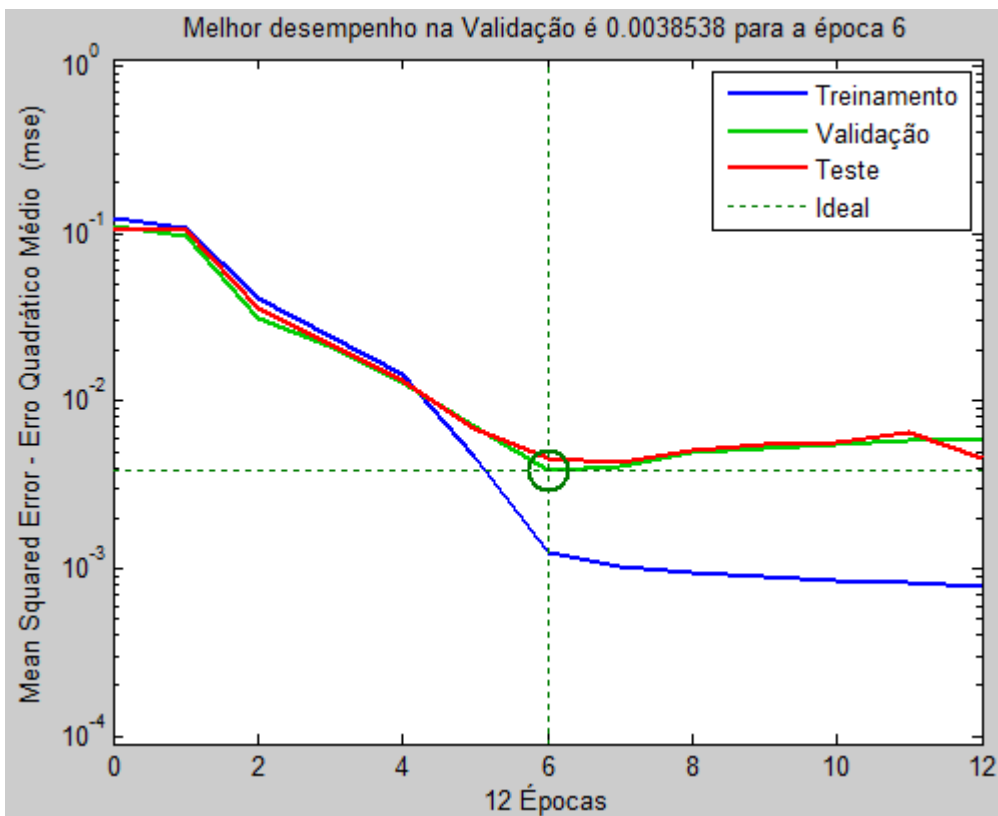


Figura C.19: MSE 50% para a segunda abordagem RGB.

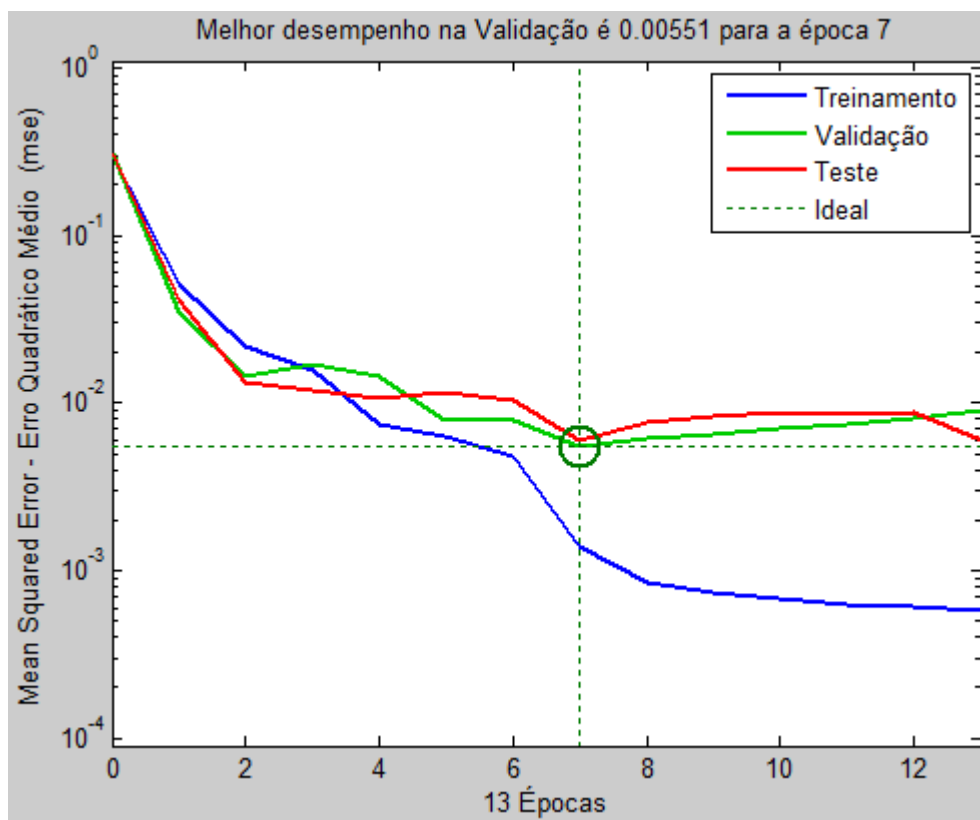


Figura C.20: MSE 75% para a segunda abordagem RGB.

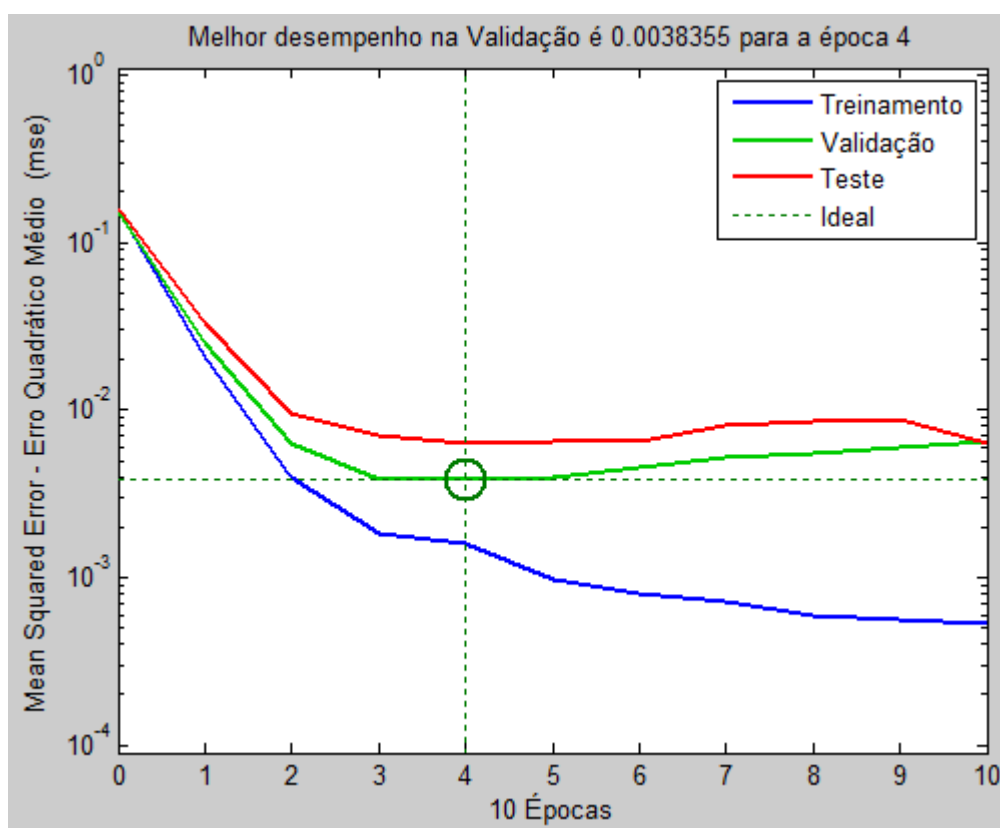


Figura C.21: MSE 100% para a segunda abordagem RGB.

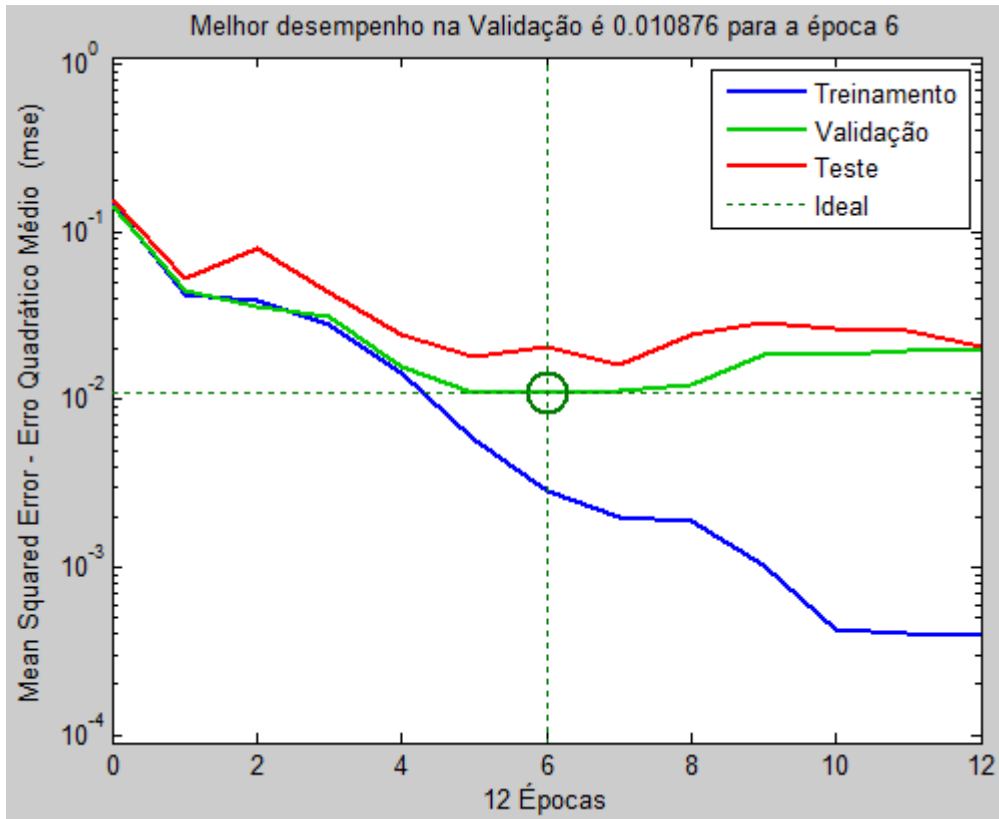


Figura C.22: MSE 10% para a segunda abordagem HSV.

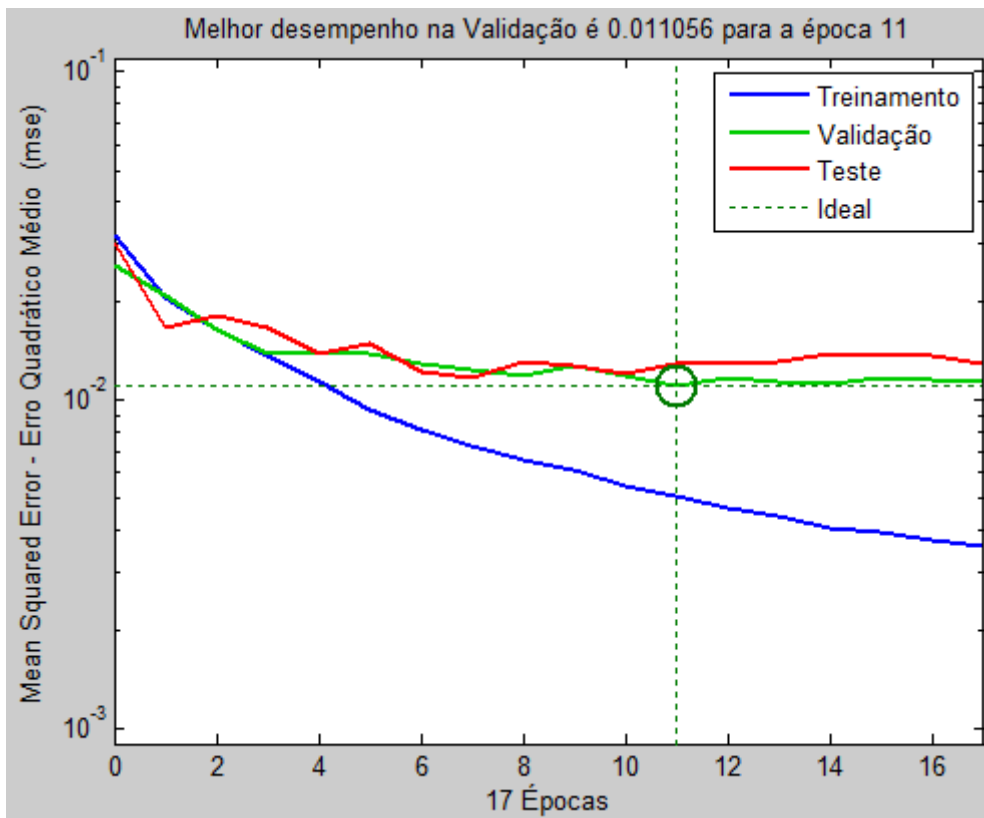


Figura C.23: MSE 20% para a segunda abordagem HSV.

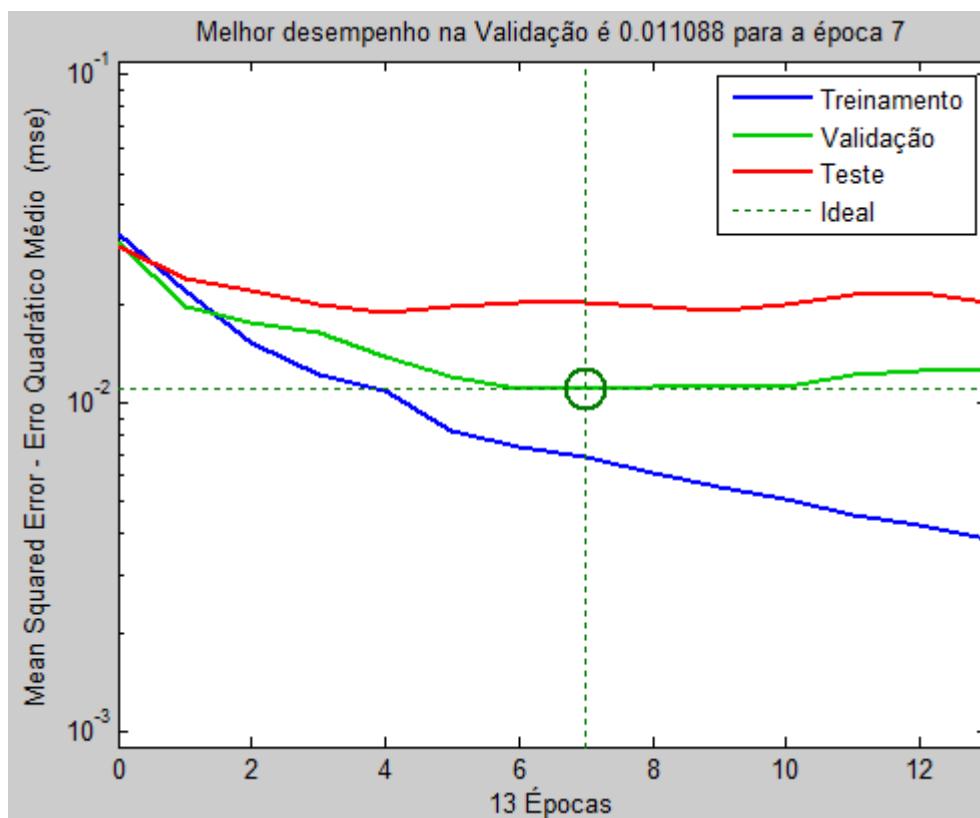


Figura C.24: MSE 30% para a segunda abordagem HSV.

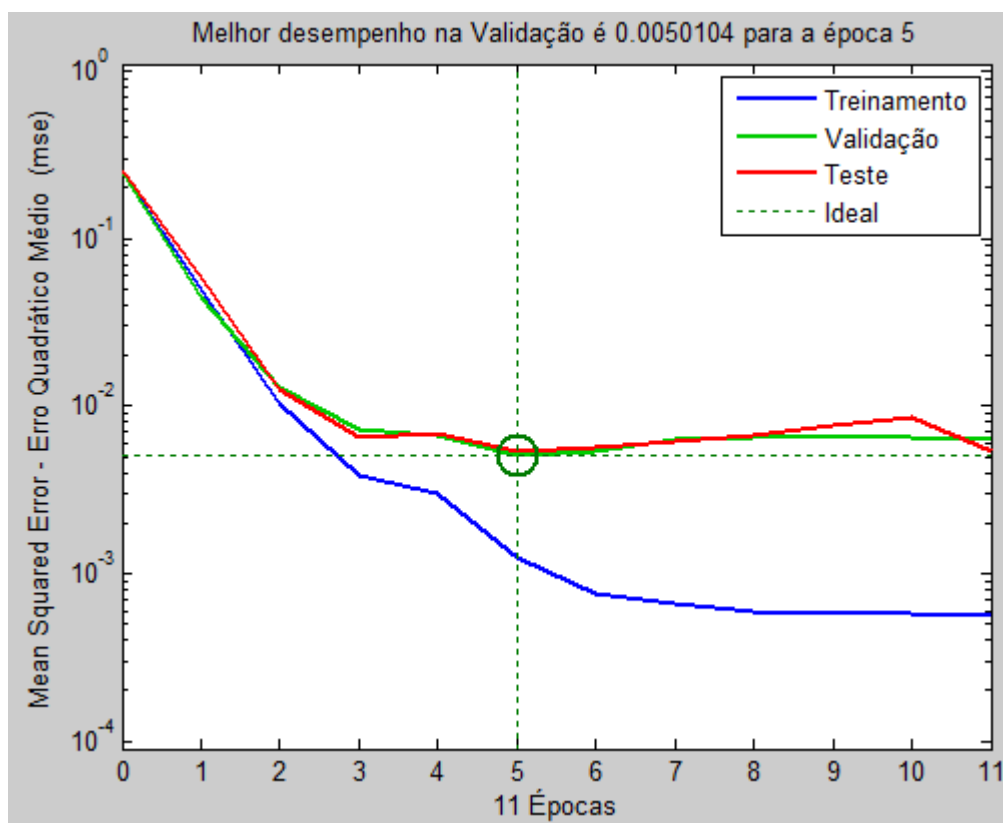


Figura C.25: MSE 40% para a segunda abordagem HSV.

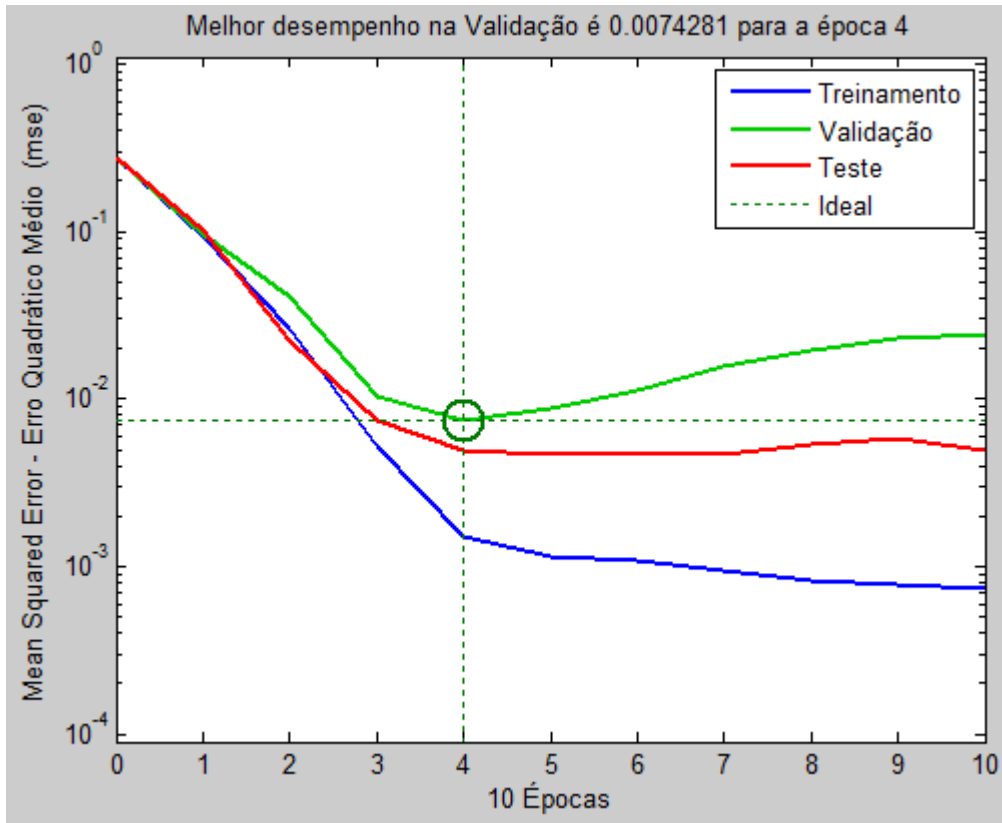


Figura C.26: MSE 50% para a segunda abordagem HSV.

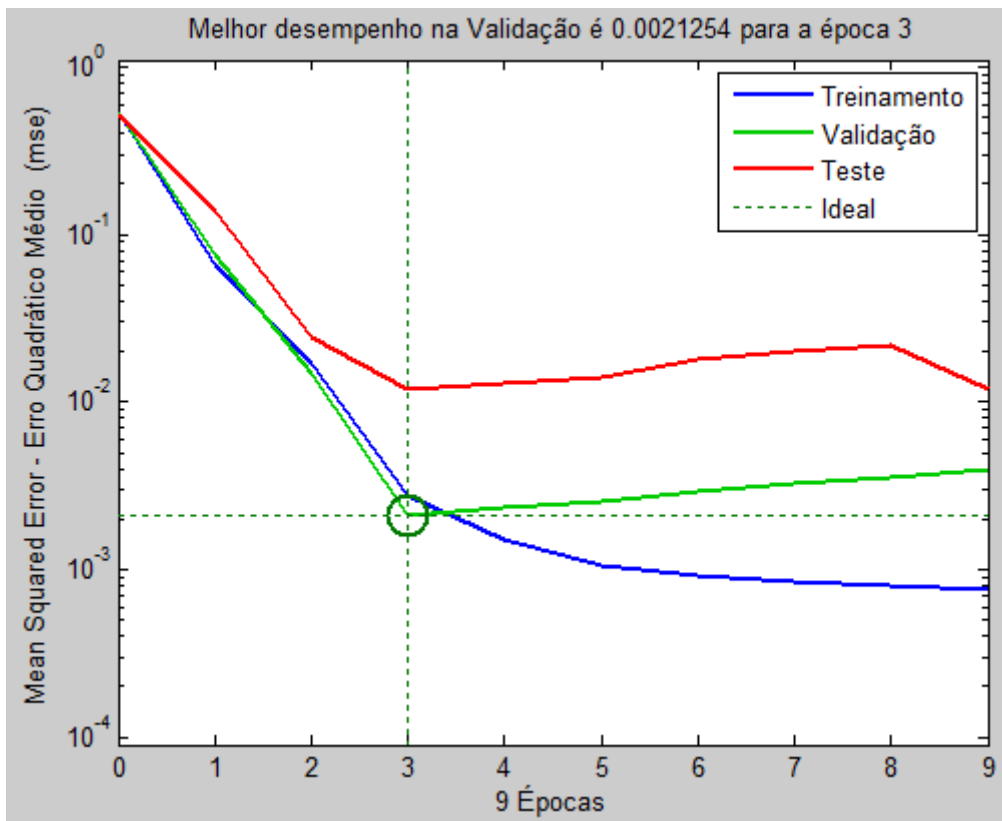


Figura C.27: MSE 75% para a segunda abordagem HSV.

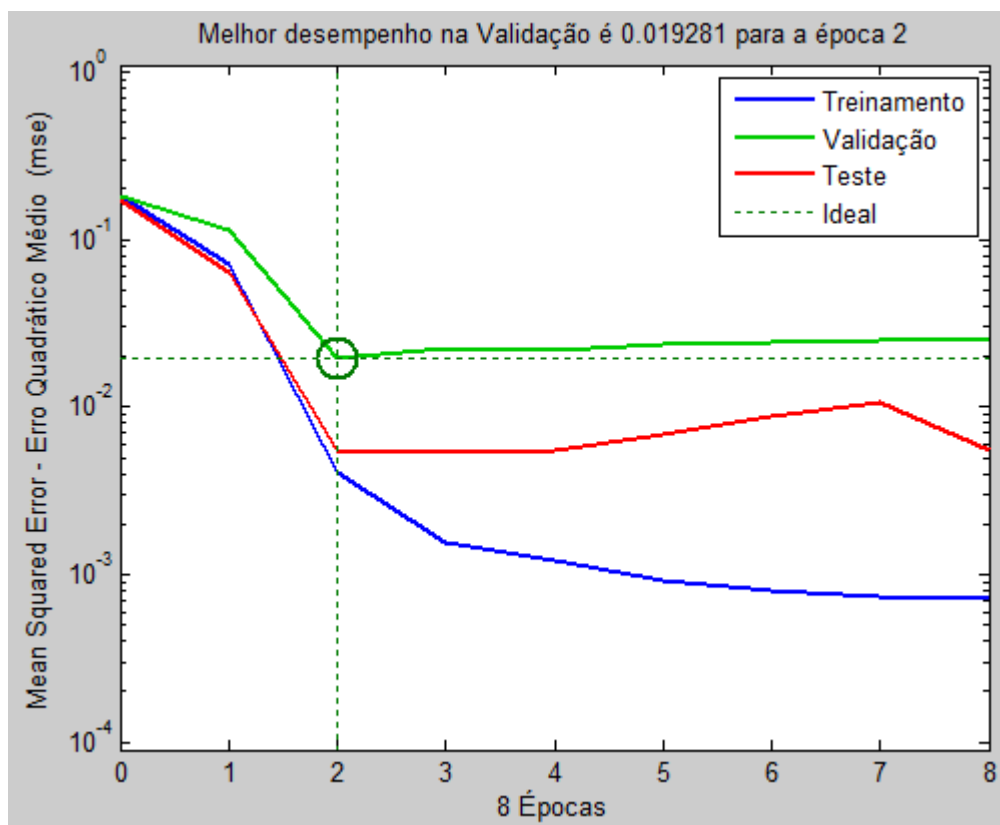


Figura C.28: MSE 100% para a segunda abordagem HSV.

APÊNDICE D. Gráficos de superfície dos estados de inferência Fuzzy

Na primeira abordagem são inferidas 24 combinações entre os estados de entrada *RDDL*, *RDEL*, *RDDF* e *RDDF*. Alguns estados, por apresentarem mesmo padrão, são ilustrados abaixo relacionando seus desvios de inferência para cada combinação.

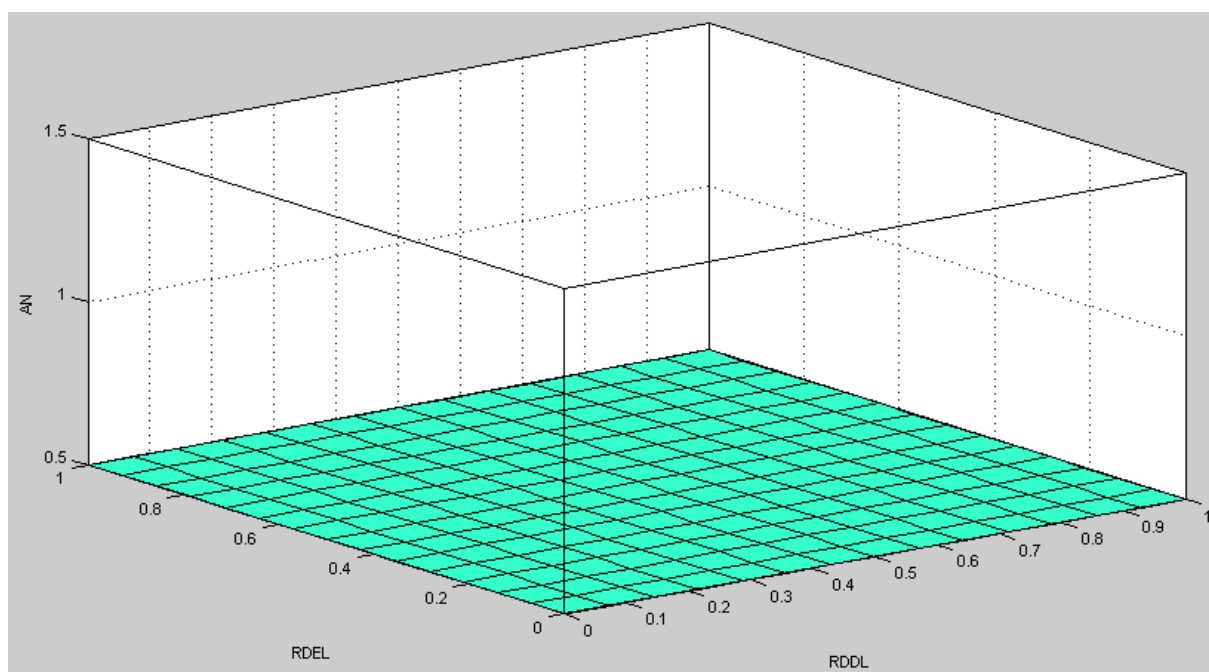


Figura D.1: Gráfico de superfície para as combinações *RDDL/RDEL/AN*, *RDDL/RDEF/AN*, *RDDL/RDEF/AnN*, *RDEL/RDDL/AN*, *RDEL/RDDF/AN*, *RDEL/RDDF/AnN*, *RDEL/RDEF/AN*, *RDDF/RDEL/AN*, *RDDF/RDEL/AnN*, *RDDF/RDEF/AnN*, *RDEF/RDDL/AN*, *RDEF/RDDL/AnN*, *RDEF/RDDF/AnN*.

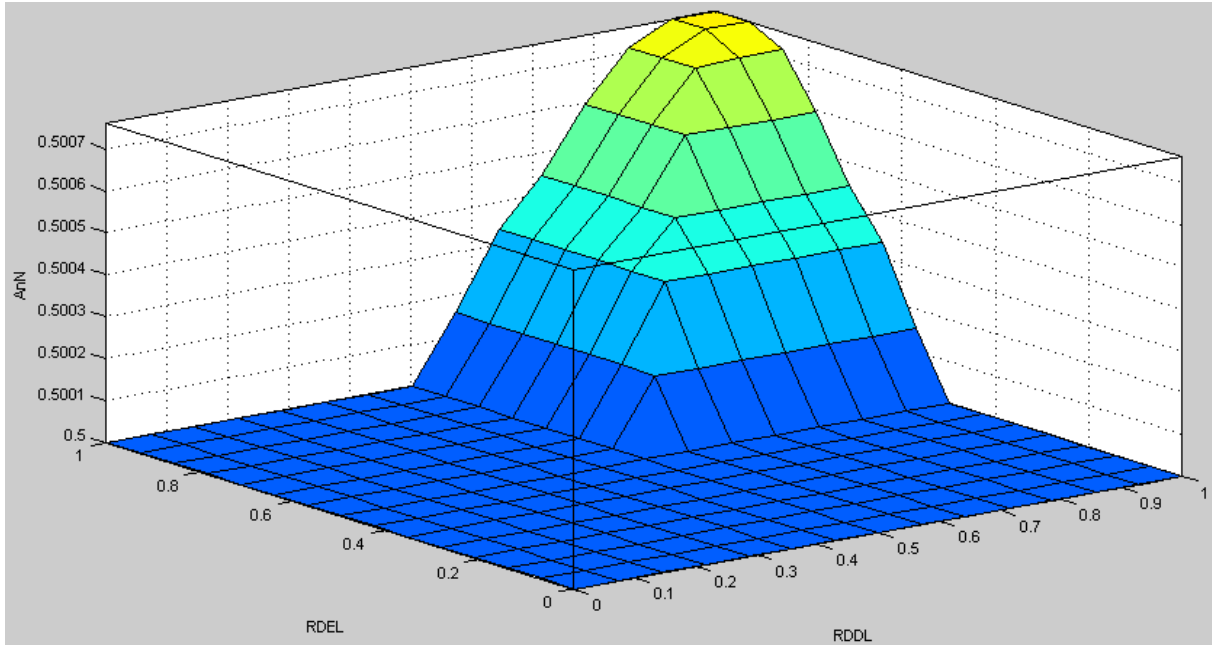


Figura D.2: Gráfico de superfície para as combinações $RDDL/RDEL/AnN$ e $RDEL/RDDL/AnN$.

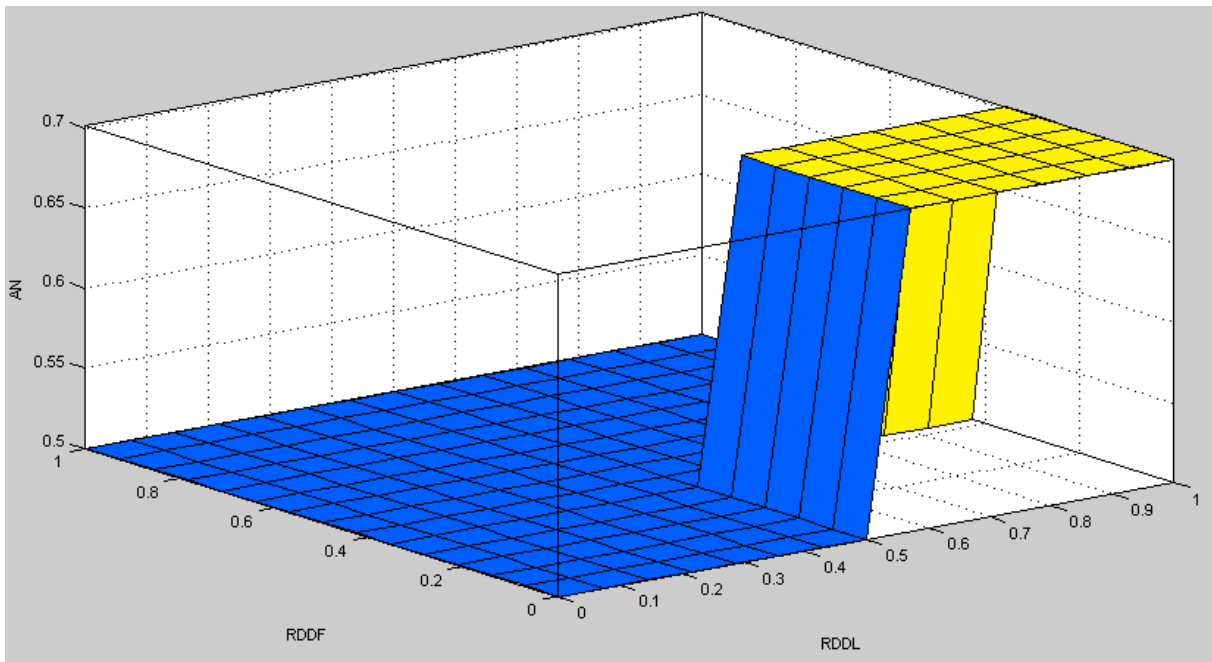


Figura D.3: Gráfico de superfície para a combinação $RDDL/RDDF/AN$.

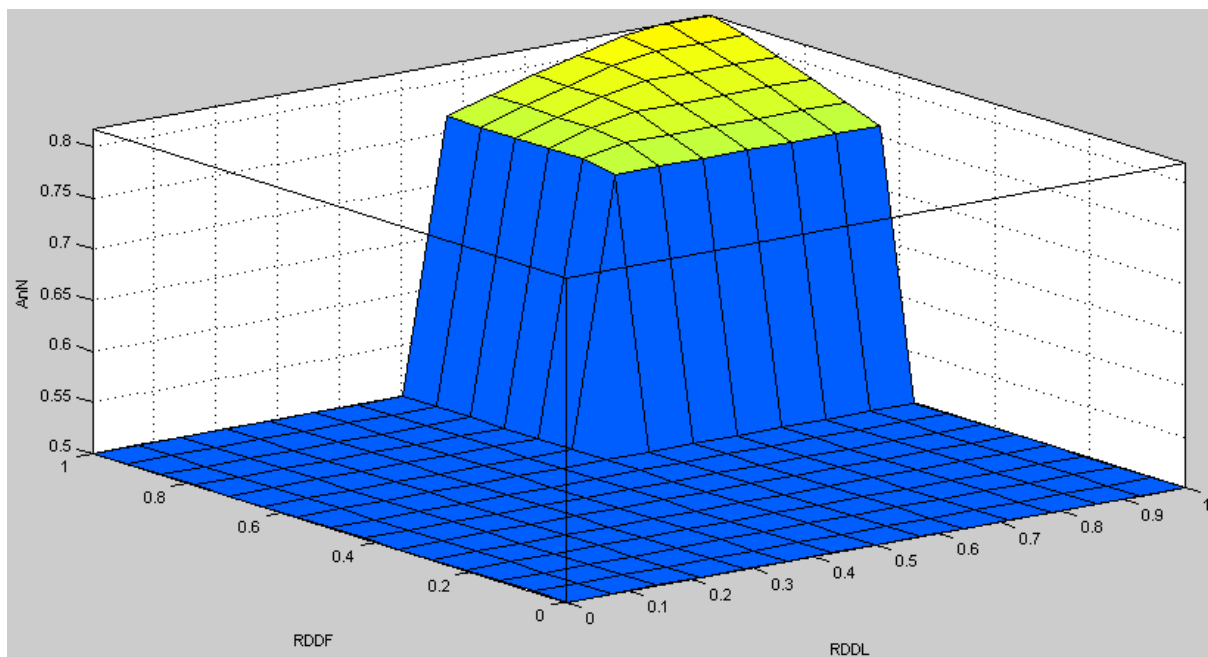


Figura D.4: Gráfico de superfície para as combinações $RDDL/RDDF/AnN$ e $RDDF/RDDL/AnN$.

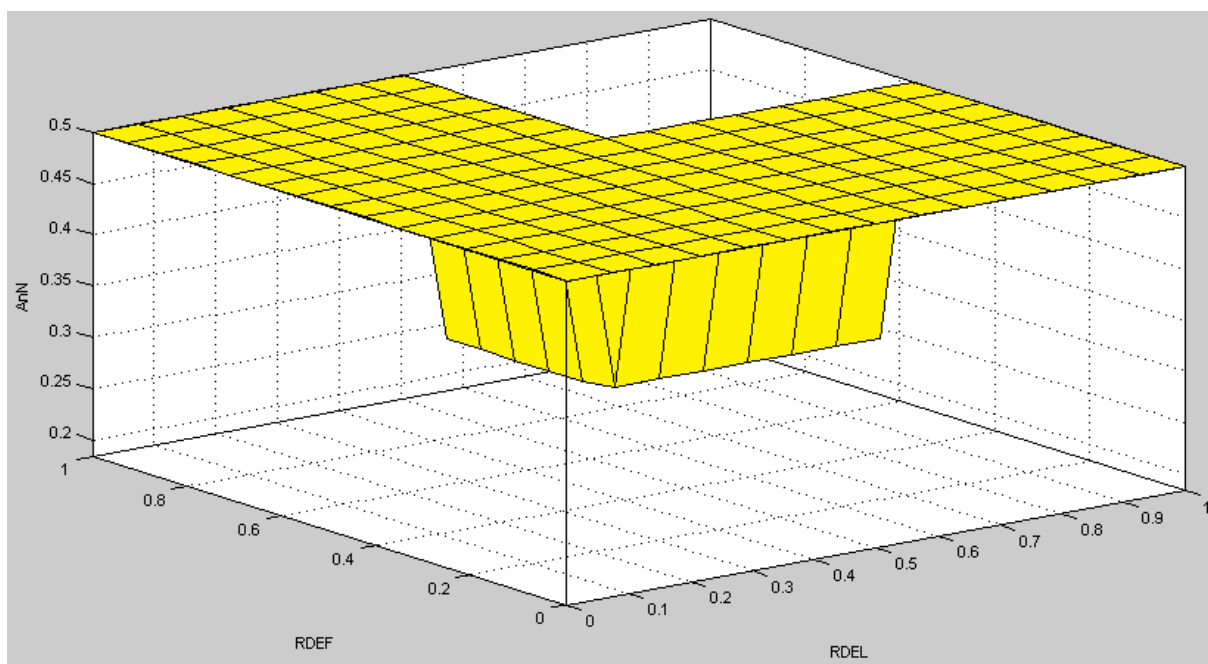


Figura D.5: Gráfico de superfície para as combinações $RDEL/RDEF/AnN$, $RDEF/RDEL/AnN$.

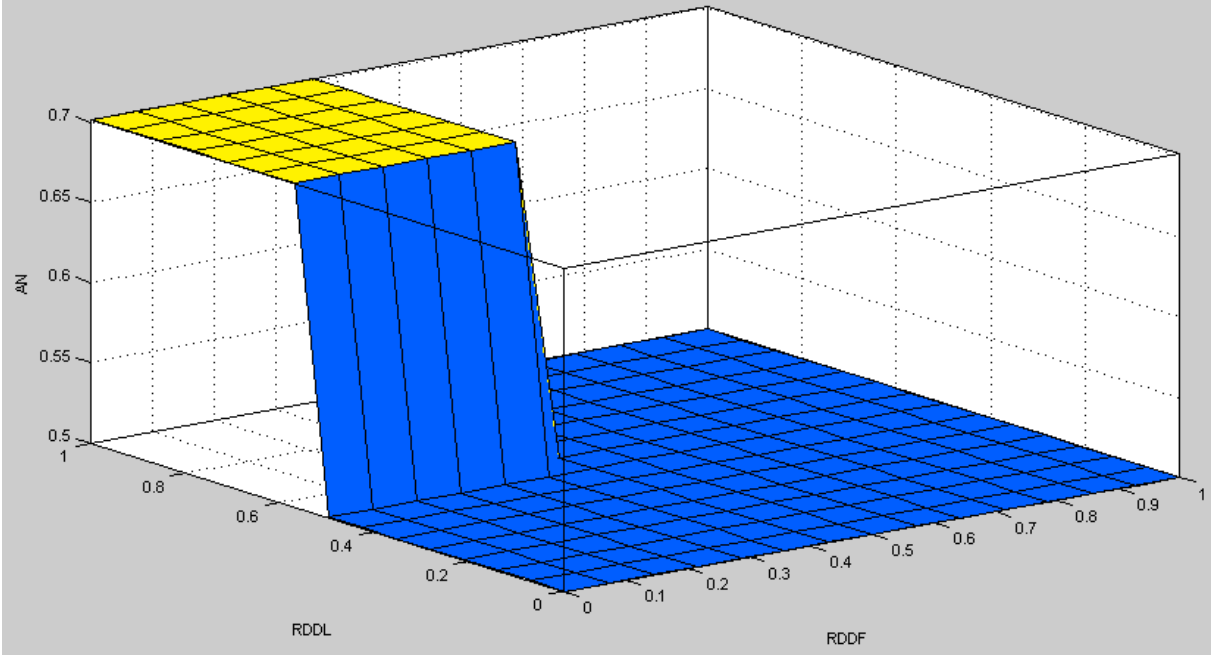


Figura D.6: Gráfico de superfície para a combinação *RDDF/RDDL/AN*.

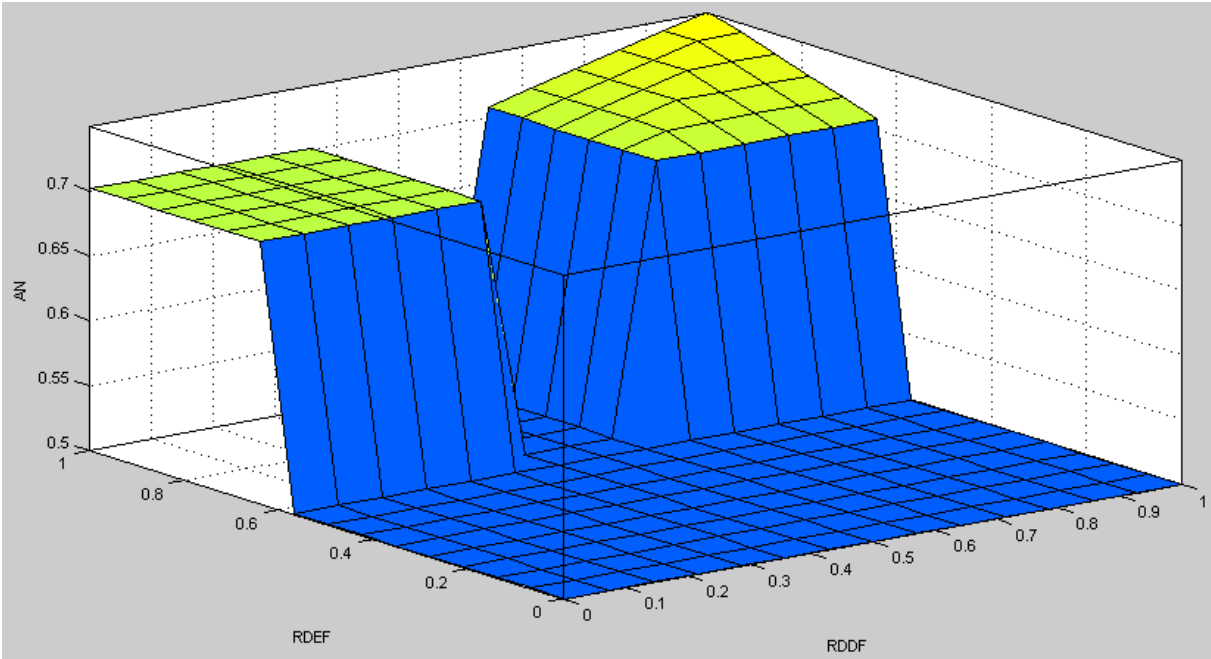


Figura D.7: Gráfico de superfície para a combinação *RDDF/RDEF/AN*.

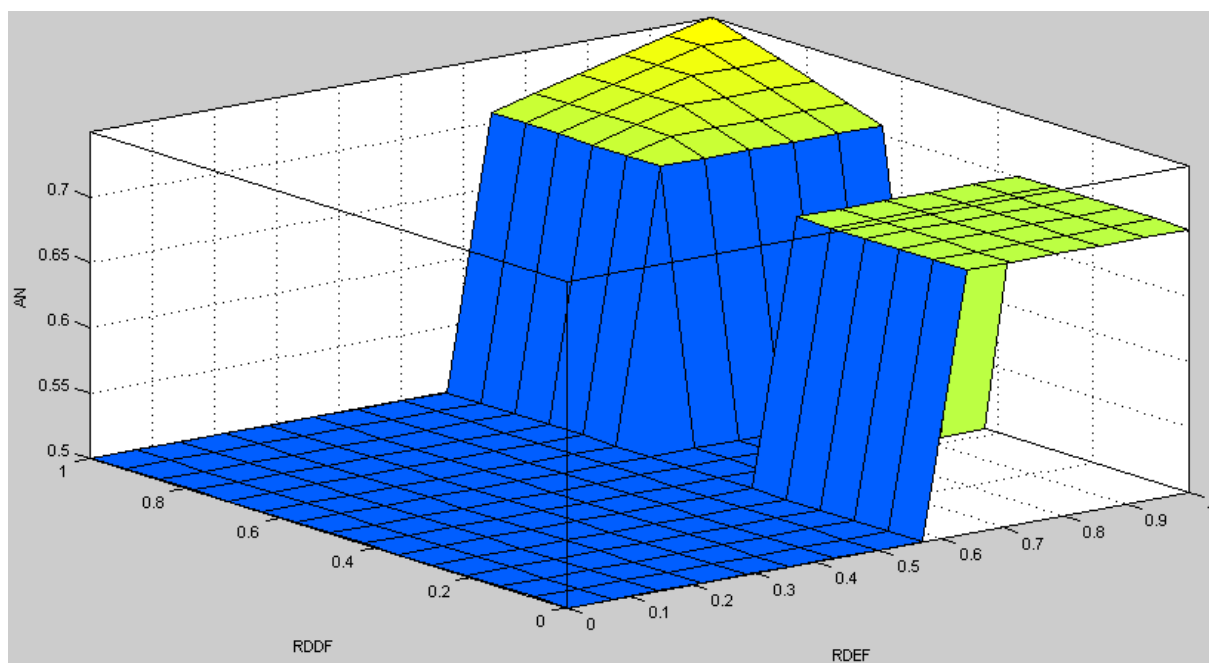


Figura D.8: Gráfico de superfície para a combinação *RDEF/RDDF/AN*.

Na segunda abordagem são inferidas 22 combinações. Os estados *CA100*, *CA200*, *AA300*, *FL* e *CF* são direcionados à localização de frutos. Os estados *LA100*, *LA300*, *GM* e *CG* são direcionados à localização de gramíneas, e os estados *AIM*, *AIB*, *FRG*, *SVR* e *FC* são direcionados aos objetos intermediários (folhas, ramos, caules e galhos).

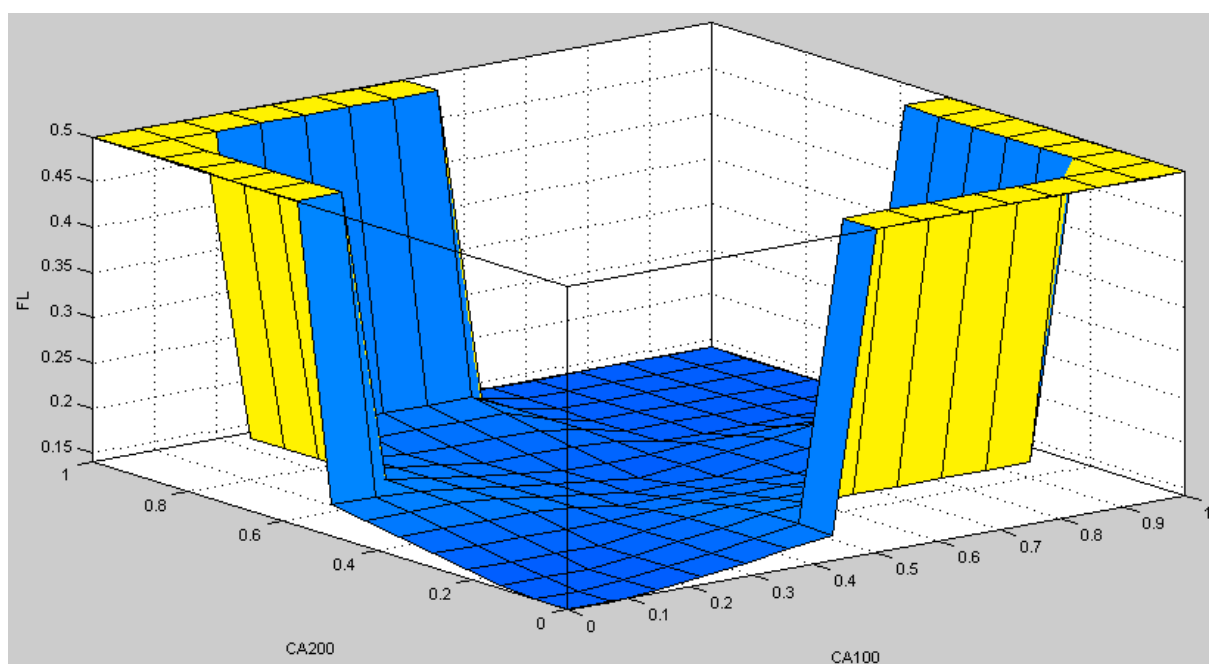


Figura D.9: Gráfico de superfície para as combinações *CA100/CA200/FL*, *CA100/CA200/CF*, *CA200/CA100/FL* e *CA200/CA100/CF*.

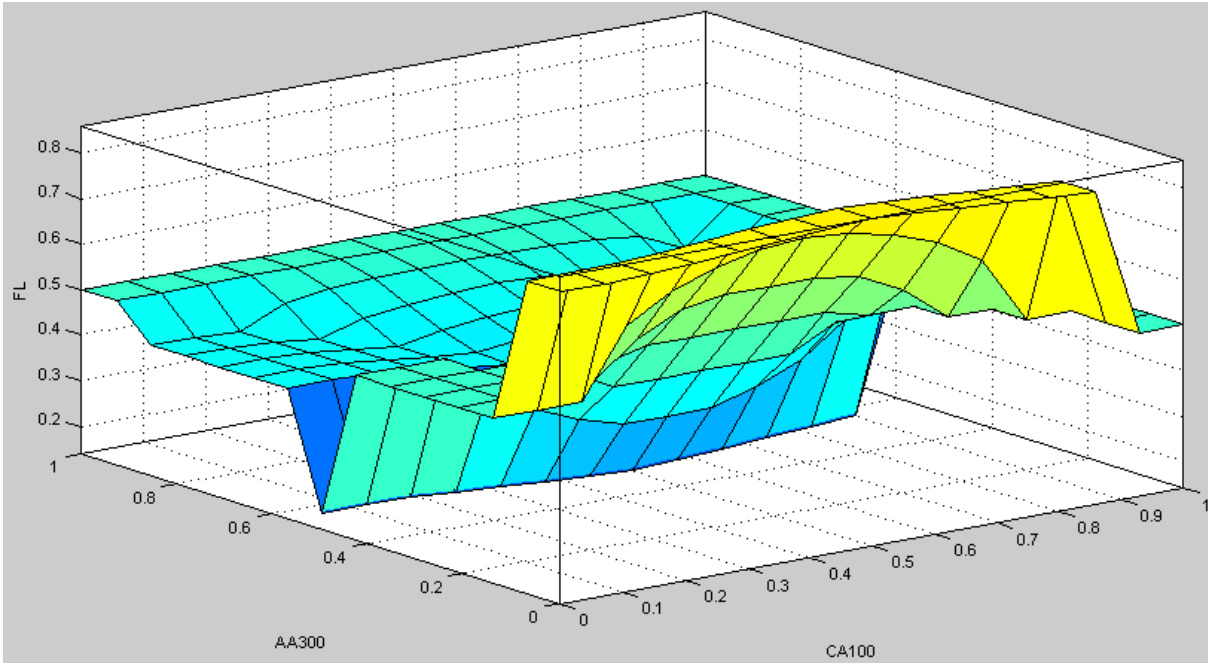


Figura D.10: Gráfico de superfície para as combinações *CA100/AA300/FL*, *CA100/AA300/CF*, *CA200/AA300/FL* e *CA200/AA300/CF*.

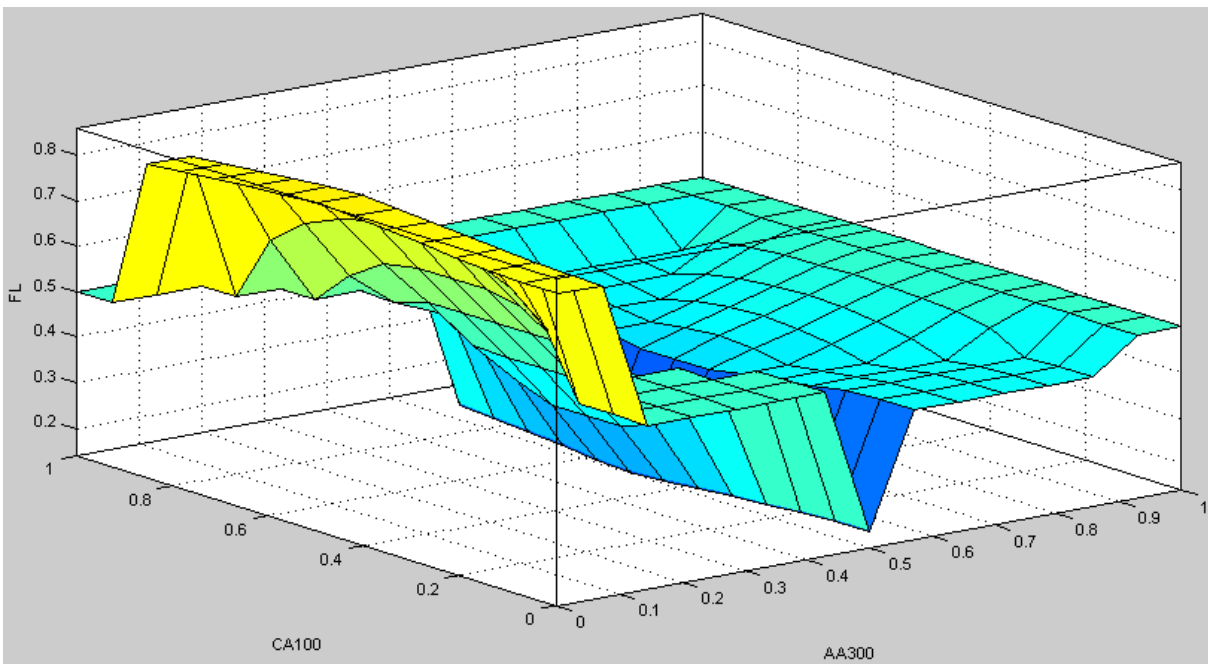


Figura D.11: Gráfico de superfície para as combinações *AA300/CA100/FL*, *AA300/CA100/CF*, *AA300/CA200/FL* e *AA300/CA200/CF*.

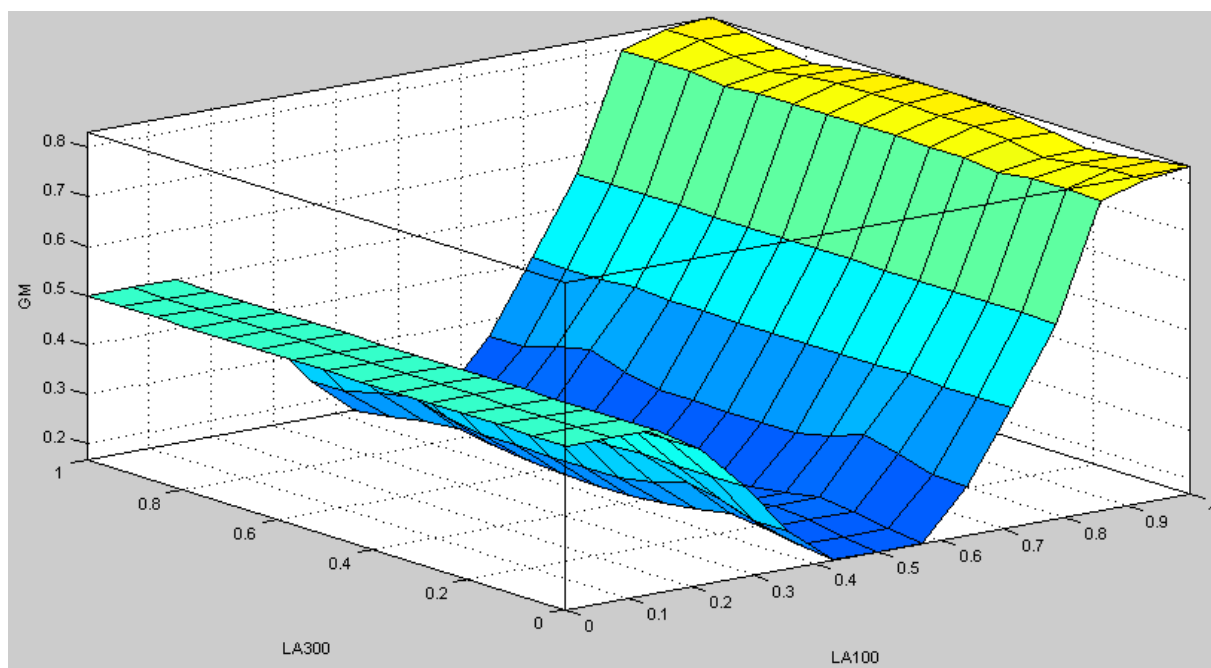


Figura D.12: Gráfico de superfície para as combinações *LA100/LA300/GM* e *LA100/LA300/CG*.

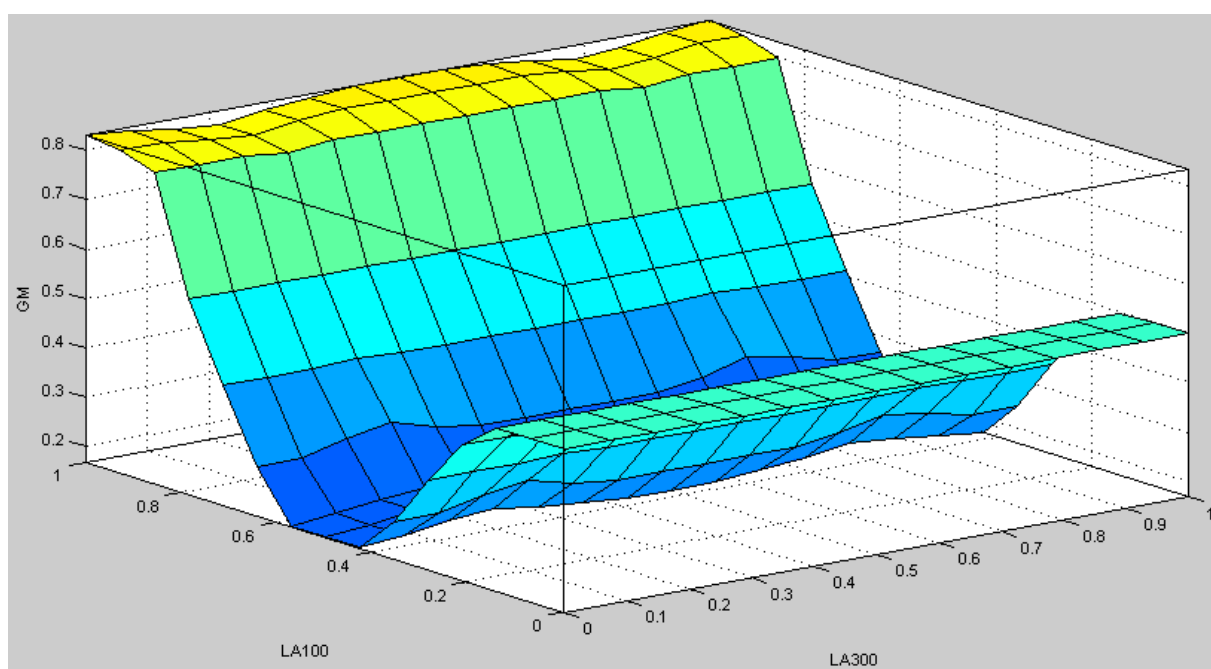


Figura D.13: Gráfico de superfície para as combinações *LA300/LA100/GM* e *LA300/LA100/CG*.

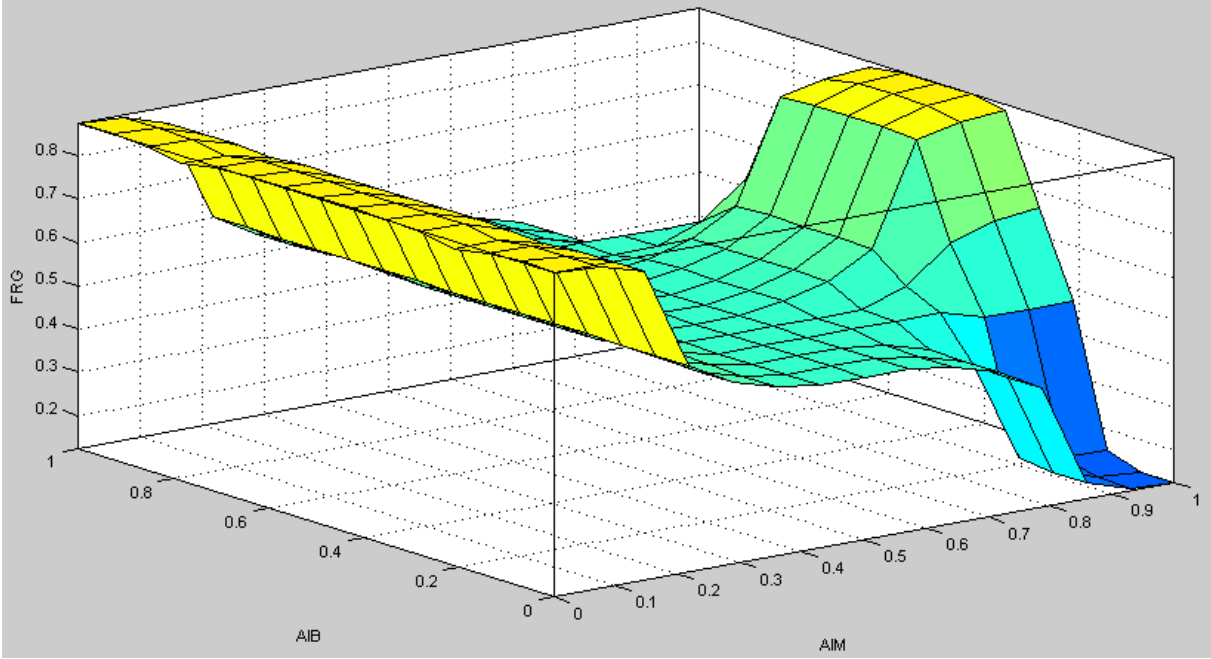


Figura D.14: Gráfico de superfície para a combinação AIM/AIB/FRG.

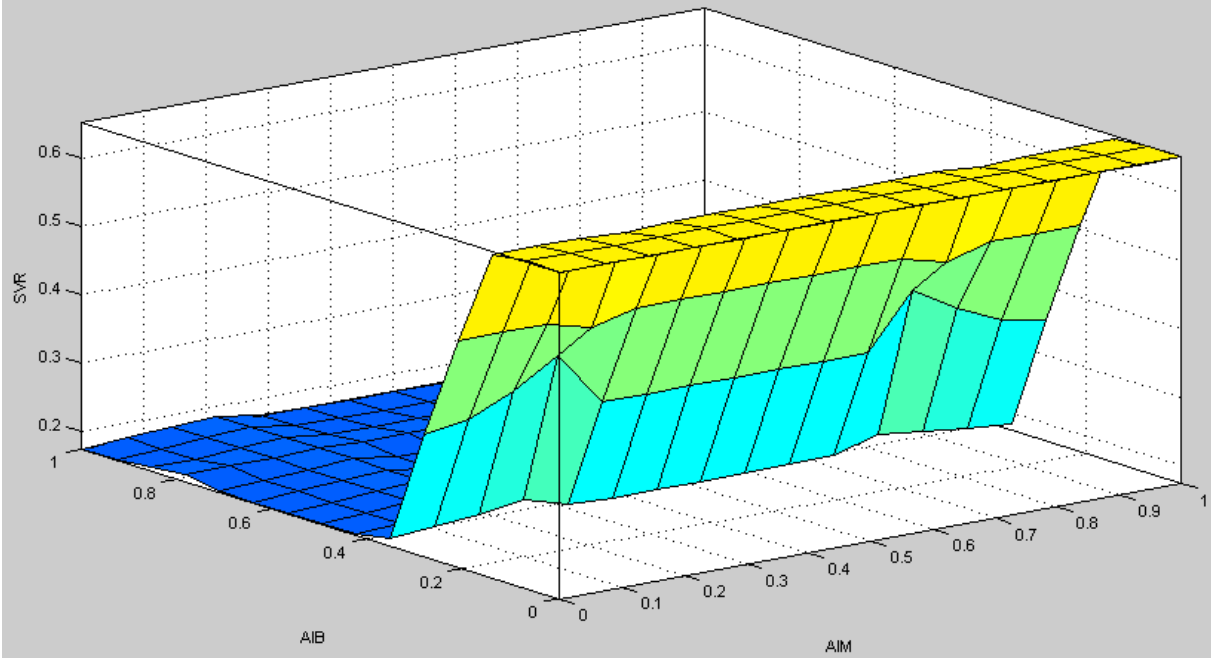


Figura D.15: Gráfico de superfície para a combinação AIM/AIB/SVR.

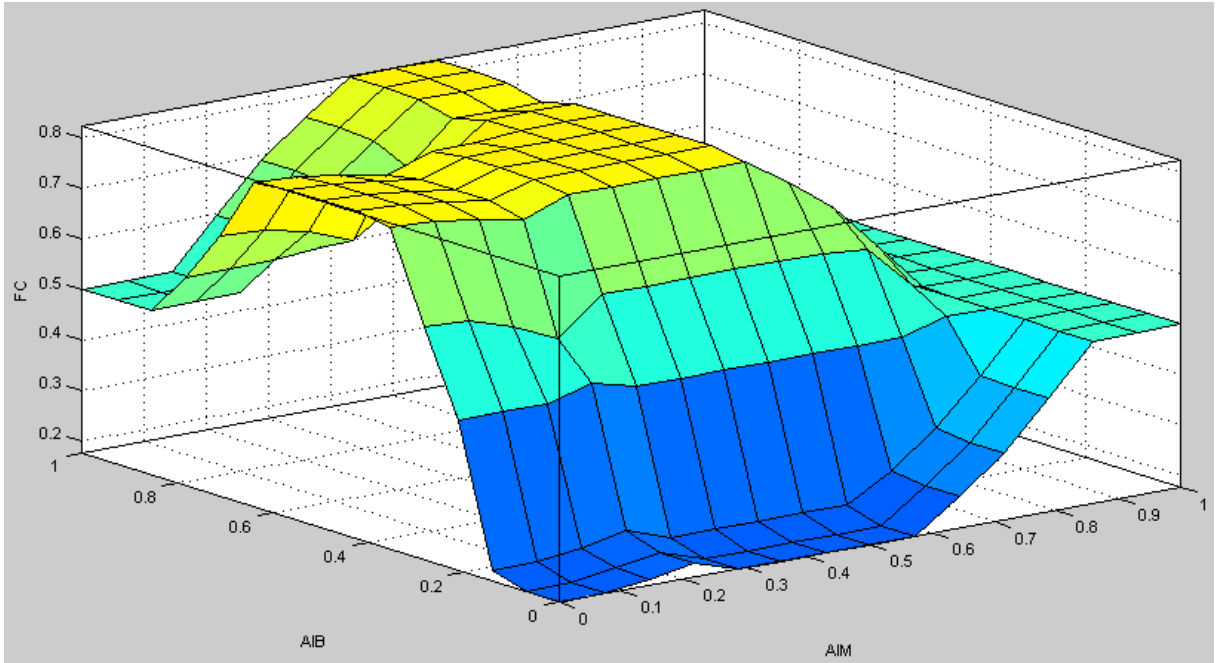


Figura D.16: Gráfico de superfície para a combinação *AIM/AIB/FC*.

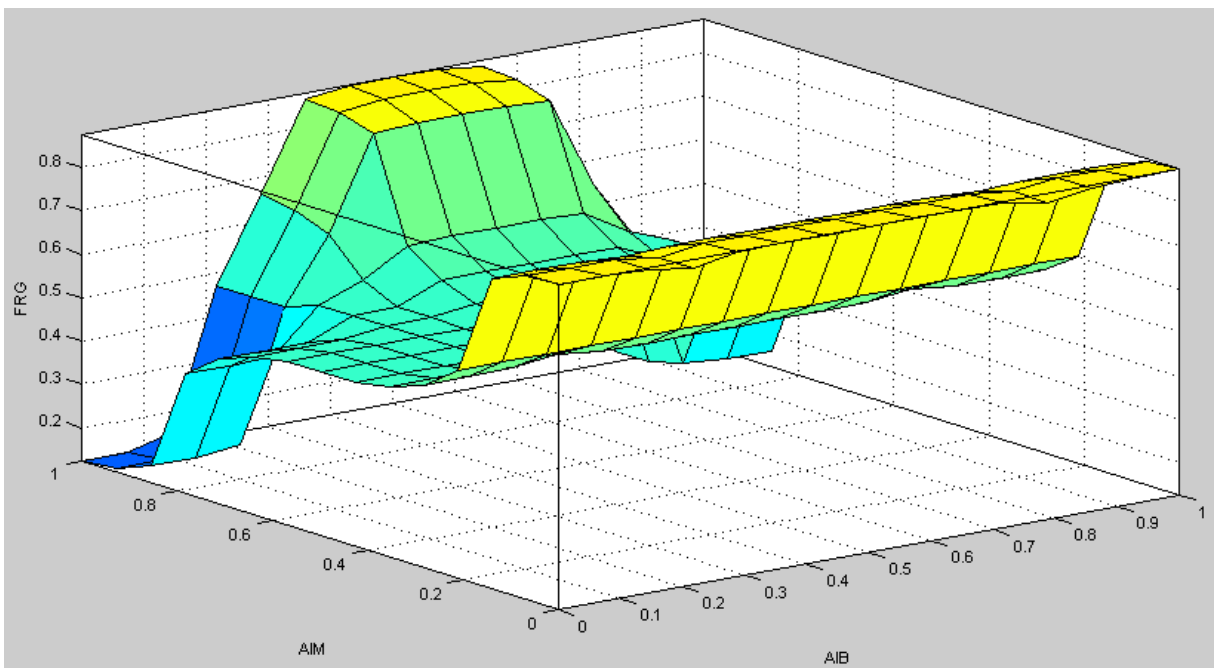


Figura D.17: Gráfico de superfície para a combinação *AIB/AIM/FRG*.

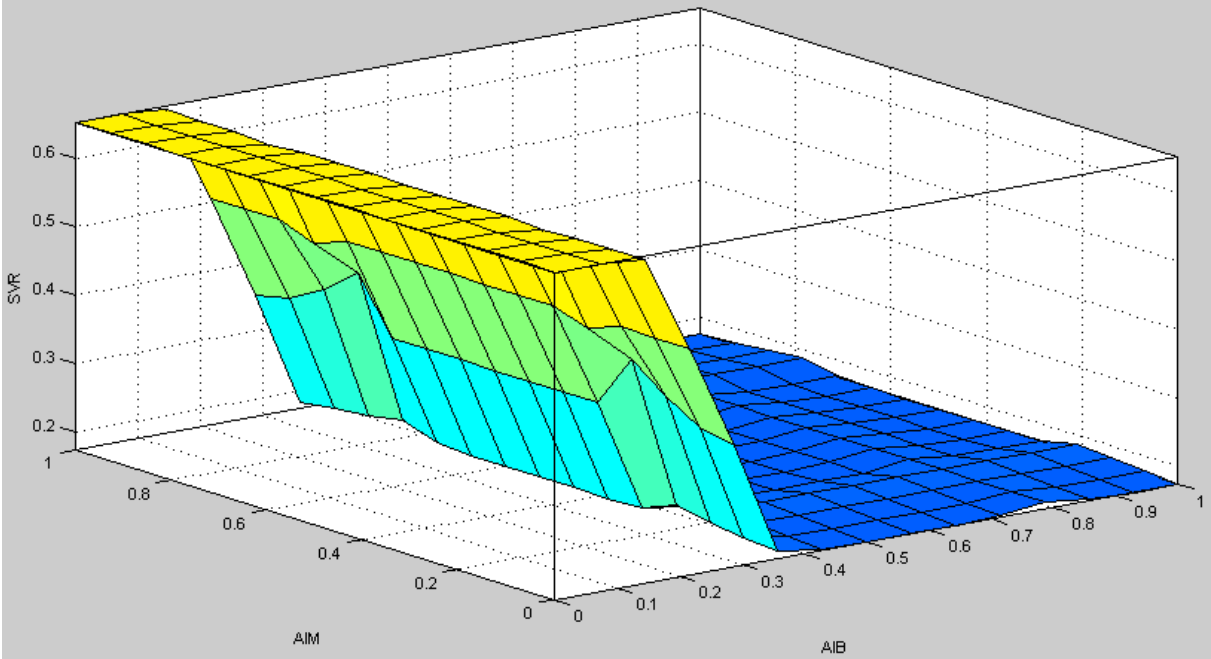


Figura D.18: Gráfico de superfície para a combinação *AIB/AIM/FRG*.

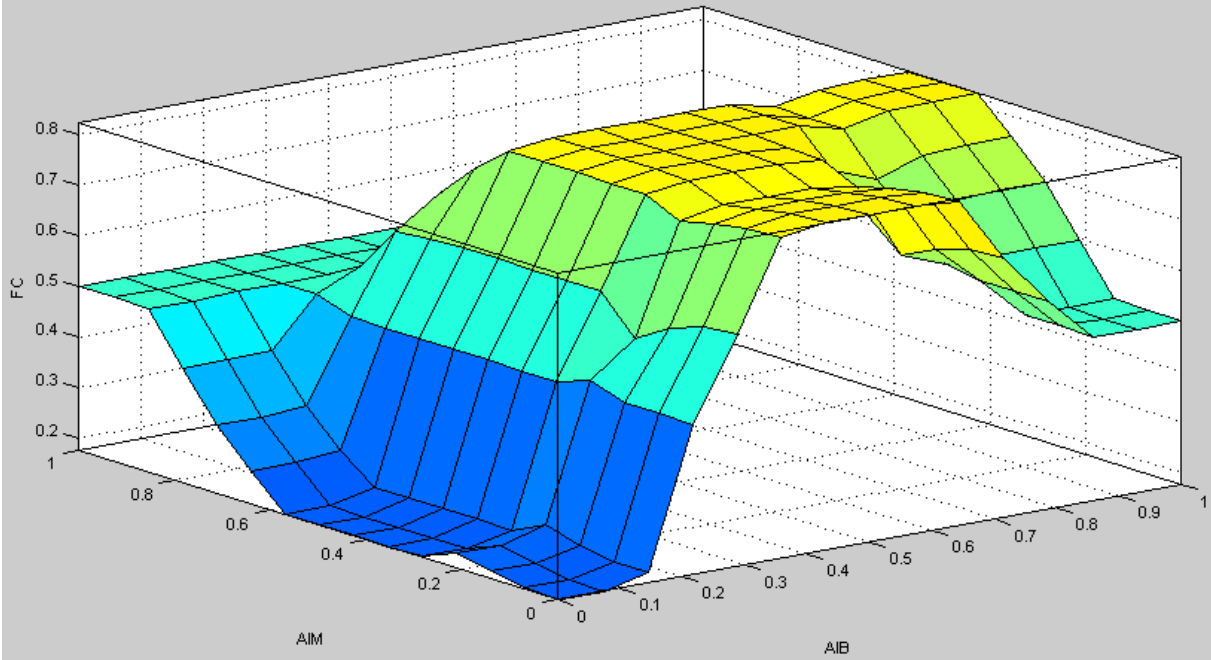


Figura D.19: Gráfico de superfície para a combinação *AIB/AIM/FRG*.

APÊNDICE E. Mapas de classes para as duas abordagens

Laranja – imagem 02

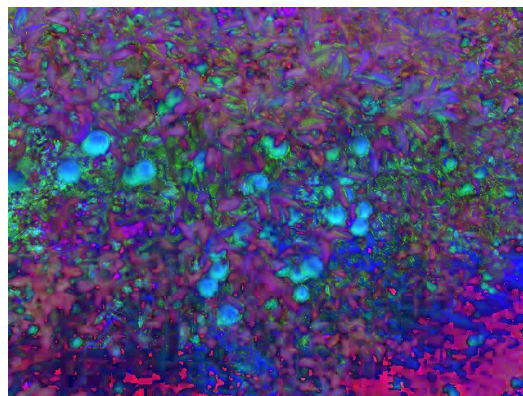


Figura E.1: Imagem original RGB e HSV (laranja 02).



Figura E.2: Mapa de classe 2 (Bayes RGB e HSV).

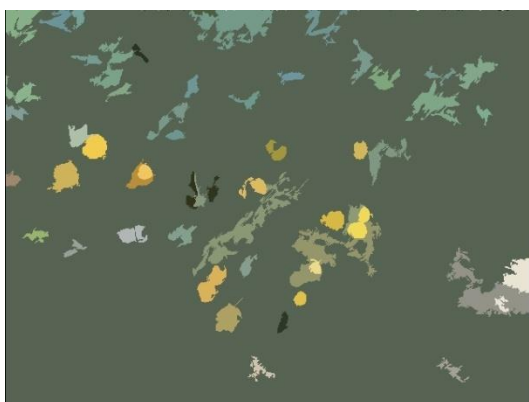


Figura E.3: Mapa de classe 2 (Bayes ingênuo RGB e HSV).

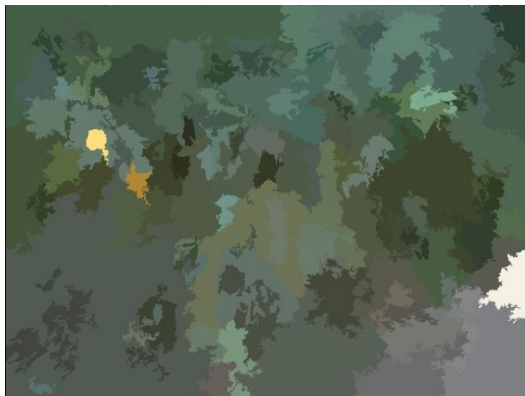


Figura E.4: Mapa de classe 2 (Fuzzy preeditivo RGB e HSV).

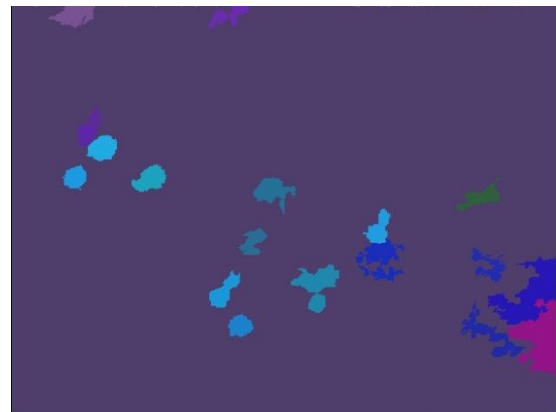


Figura E.5: Mapa de classe 2 (RNA RGB e HSV).

Laranja – imagem 03

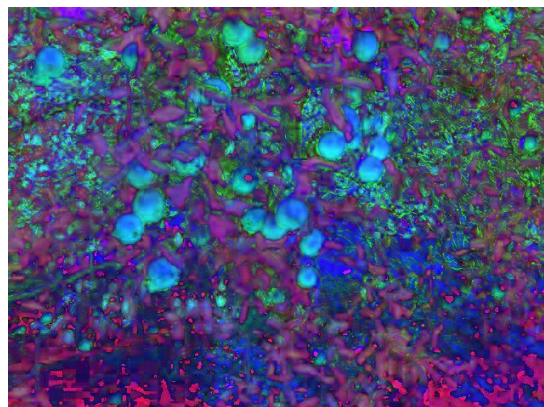


Figura E.6: Imagem original RGB e HSV (laranja 03).



Figura E.7: Mapa de classe 3 (Bayes RGB e HSV).

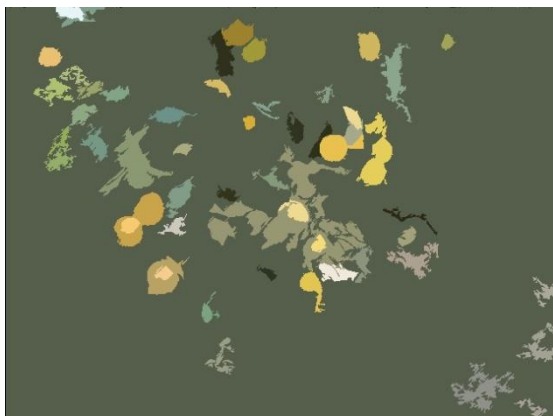


Figura E.8: Mapa de classe 3 (Bayes ingênuo RGB e HSV).

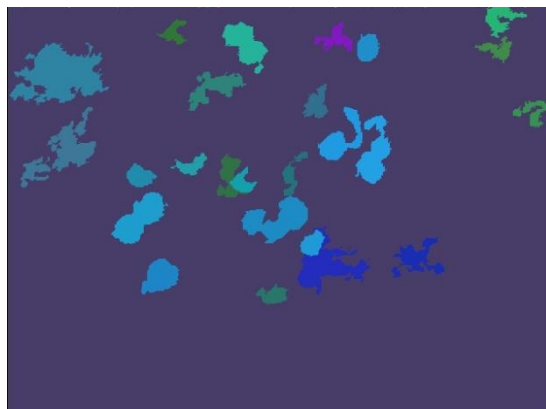
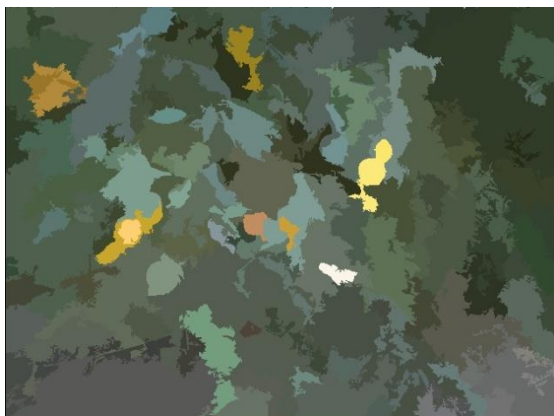


Figura E.9: Mapa de classe 3 (Fuzzy preditivo RGB e HSV).



Figura E.10: Mapa de classe 3 (RNA RGB e HSV).

Laranja – imagem 04

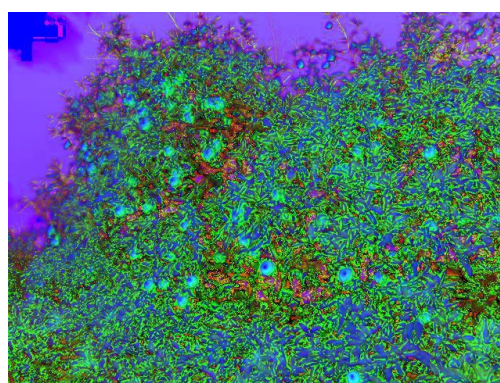


Figura E.11: Imagem original RGB e HSV (laranja 04).



Figura E.12: Mapa de classe 4 (Bayes RGB e HSV).

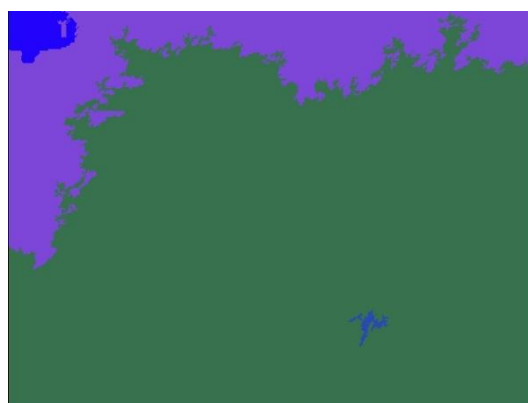
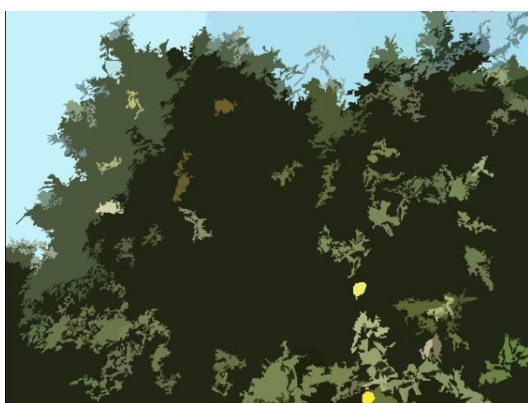


Figura E.13: Mapa de classe 4 (Bayes ingênuo RGB e HSV).



Figura E.14: Mapa de classe 4 (Fuzzy preditivo RGB e HSV).

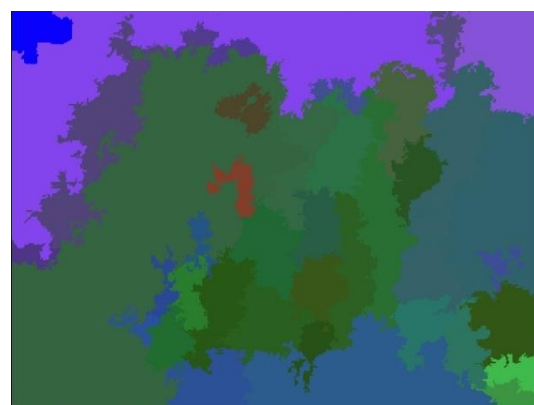


Figura E.15: Mapa de classe 4 (RNA RGB e HSV).

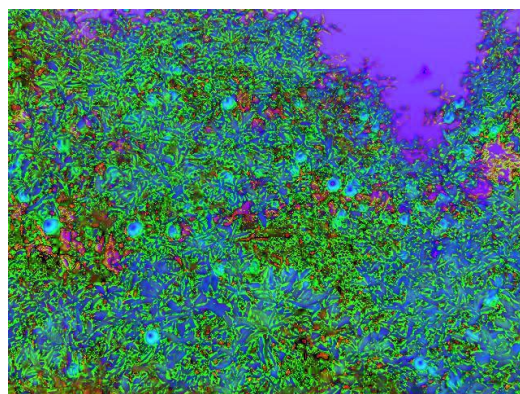
Laranja – imagem 05

Figura E.16: Imagem original RGB e HSV (laranja 05).

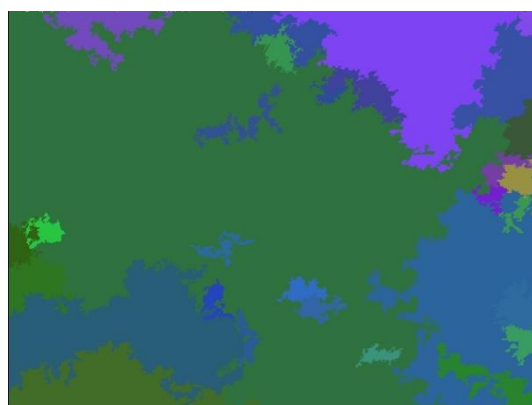
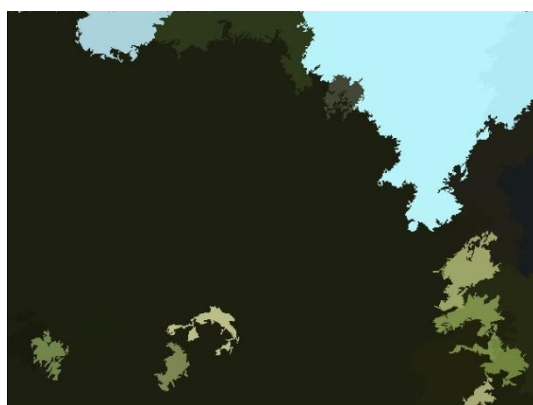


Figura E.17: Mapa de classe 5 (Bayes RGB e HSV).

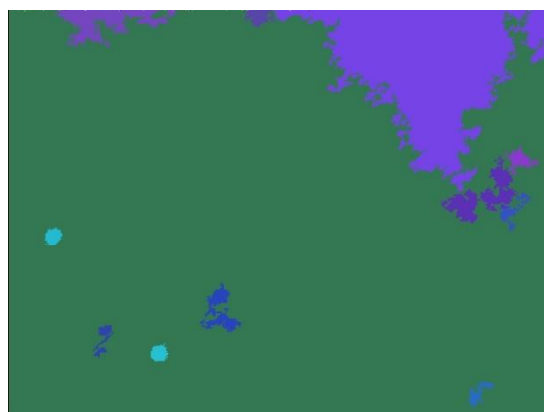


Figura E.18: Mapa de classe 5 (Bayes ingenuo RGB e HSV).

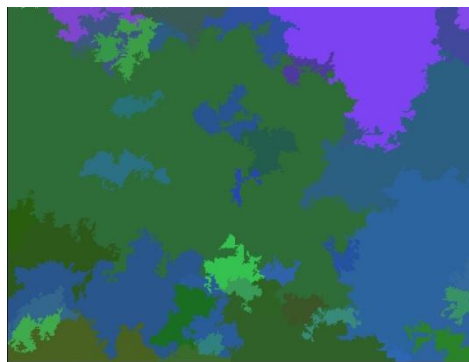
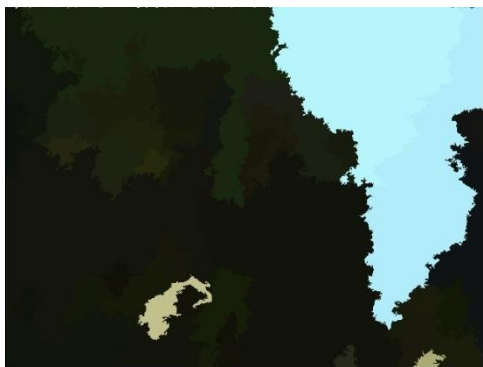


Figura E.19: Mapa de classe 5 (Fuzzy preditivo RGB e HSV).

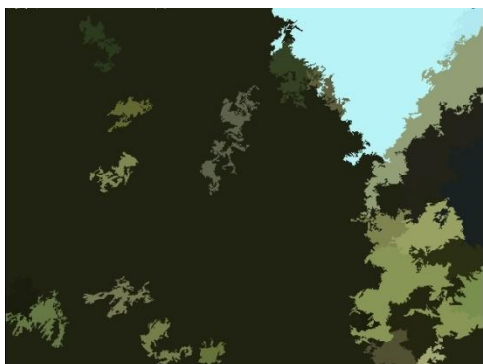


Figura E.20: Mapa de classe 5 (RNA RGB e HSV).

Linhas de plantio – Pomares 02

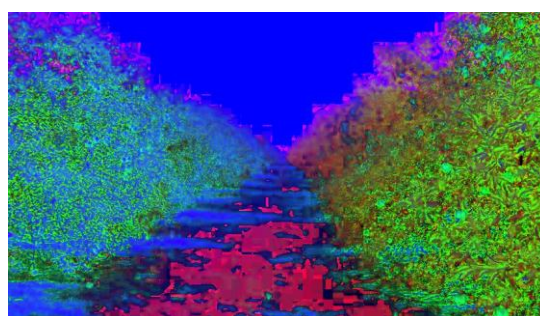


Figura E.21: Imagem original RGB e HSV (pomares 02).



Figura E.22: Mapa de classe 2 (Bayes RGB e HSV).



Figura E.23: Mapa de classe 2 (Bayes ingênuo RGB e HSV).



Figura E.24: Mapa de classe 2 (Fuzzy preditivo RGB e HSV)

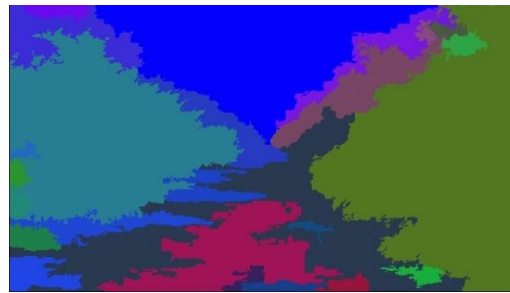


Figura E.25: Mapa de classe 2 (RNA RGB e HSV).

Linhas de plantio – Pomares 03

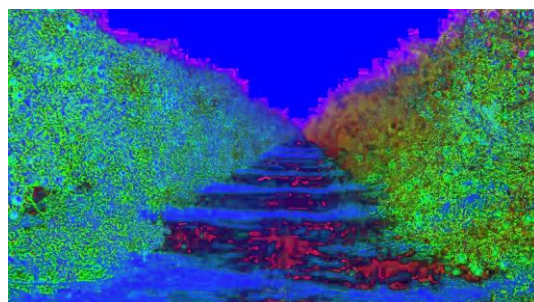


Figura E.26: Imagem original RGB e HSV (pomares 03).

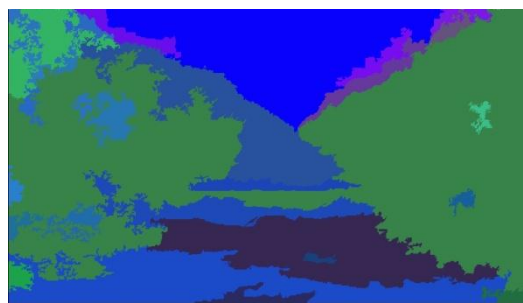


Figura E.27: Mapa de classe 3 (Bayes RGB e HSV).



Figura E.28: Mapa de classe 3 (Bayes ingênuo RGB e HSV).

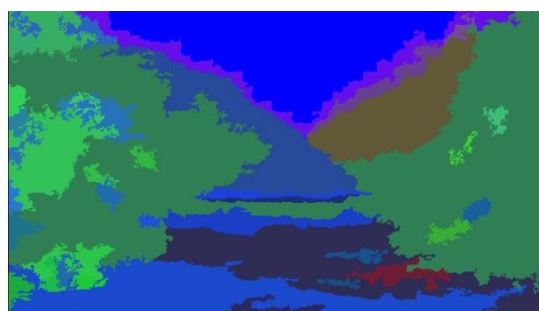


Figura E.29: Mapa de classe 3 (Fuzzy preditivo RGB e HSV).

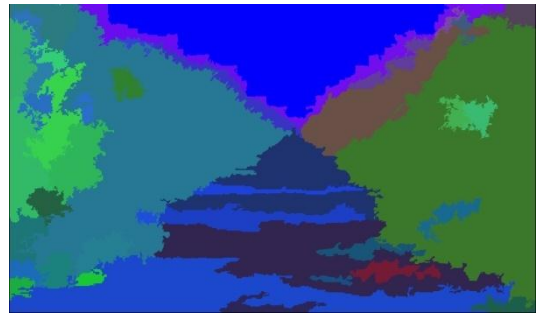


Figura E.30: Mapa de classe 3 (RNA RGB e HSV).

Linhas de plantio – Pomares 04

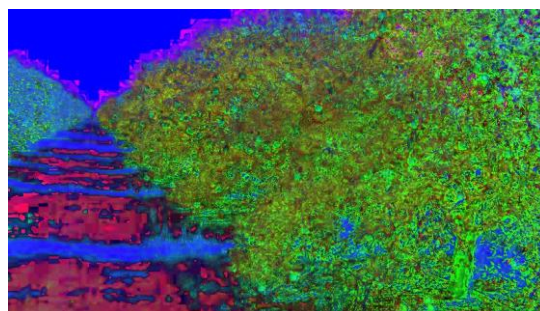


Figura E.31: Imagem original RGB e HSV (pomares 04).



Figura E.32: Mapa de classe 4 (Bayes RGB e HSV).



Figura E.33: Mapa de classe 4 (Bayes ingênuo RGB e HSV).



Figura E.34: Mapa de classe 4 (Fuzzy preditivo RGB e HSV).



Figura E.35: Mapa de classe 4 (RNA RGB e HSV).

Linhas de plantio – Pomares 05

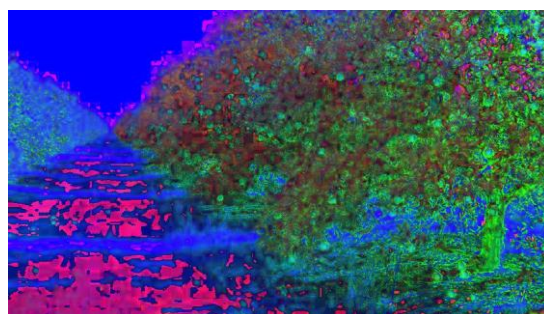


Figura E.36: Imagem original RGB e HSV (pomares 05).



Figura E.37: Mapa de classe 5 (Bayes RGB e HSV).



Figura E.38: Mapa de classe 5 (Bayes ingênuo RGB e HSV).

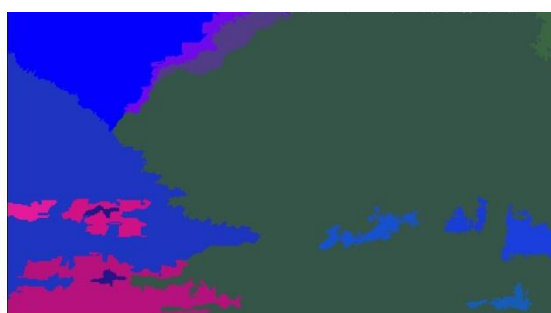


Figura E.39: Mapa de classe 5 (Fuzzy preditivo RGB e HSV).

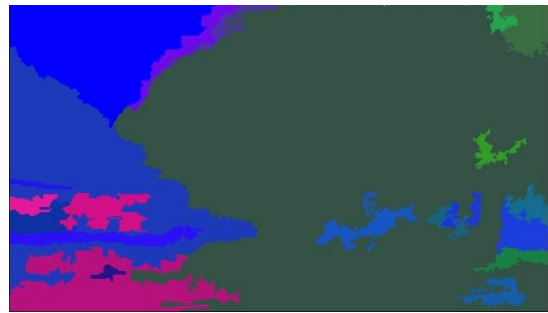
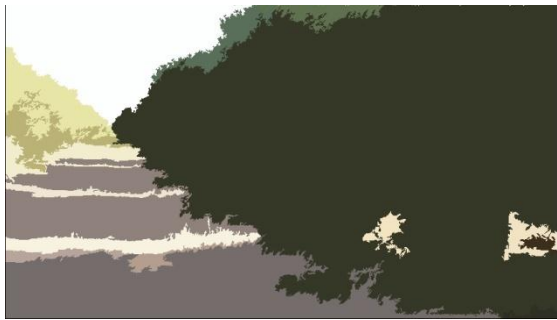


Figura E.40: Mapa de classe 5 (RNA RGB e HSV).

Linhas de plantio – Milharais 02

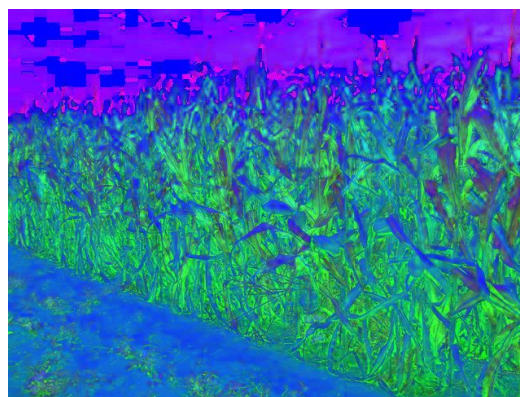


Figura E.41: Imagem original RGB e HSV (milharais 02).

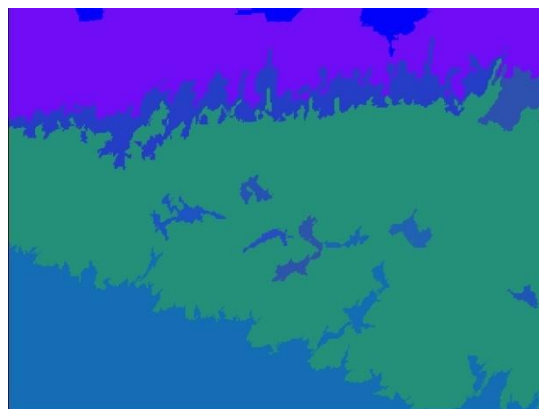


Figura E.42: Mapa de classe 2 (Bayes RGB e HSV).



Figura E.43: Mapa de classe 2 (Bayes ingênuo RGB e HSV)

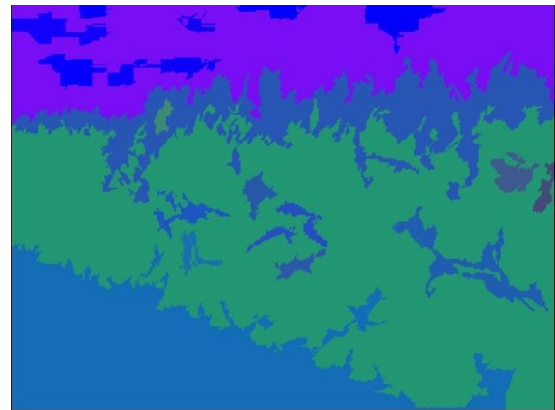


Figura E.44: Mapa de classe 2 (Fuzzy preditivo RGB e HSV)

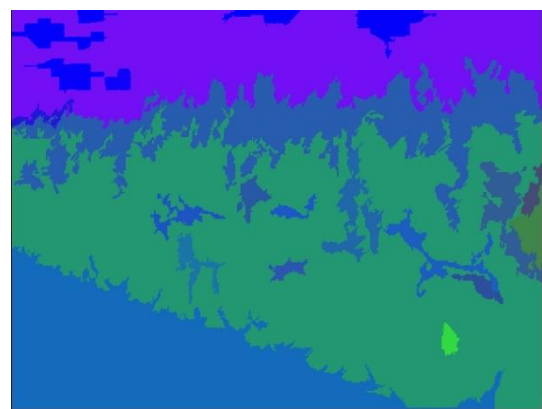


Figura E.45: Mapa de classe 2 (RNA RGB e HSV).

Linhas de plantio – Milharais 03

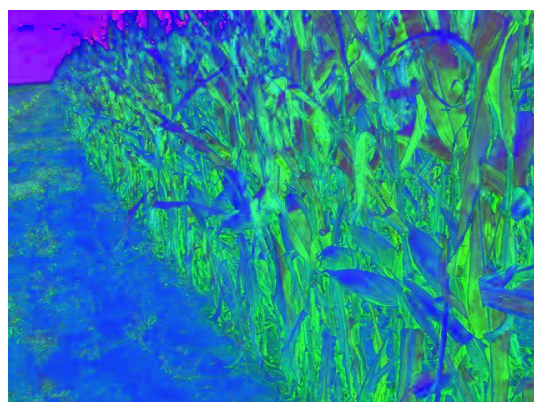


Figura E.46: Imagem original RGB e HSV (milharais 03).

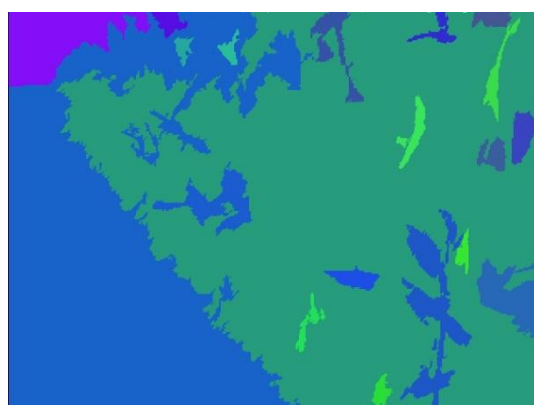


Figura E.47: Mapa de classe 3 (Bayes RGB e HSV).



Figura E.48: Mapa de classe 3 (Bayes ingênuo RGB e HSV).

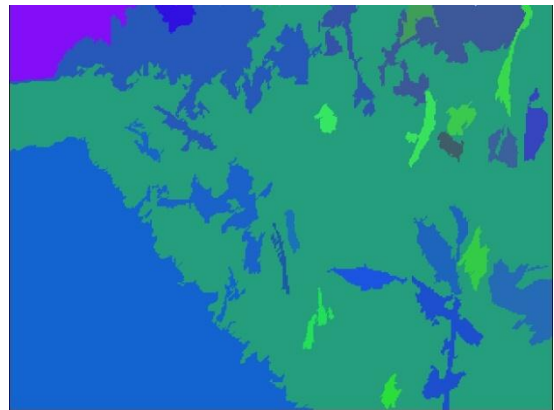
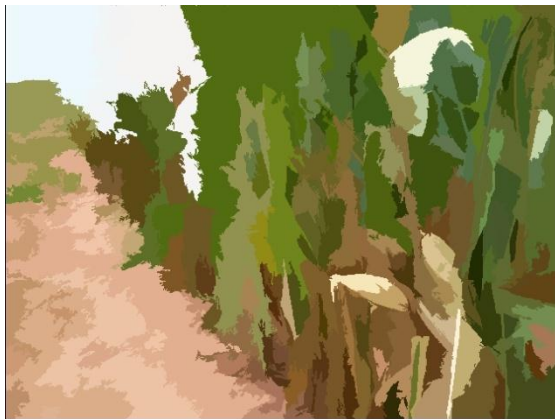


Figura E.49: Mapa de classe 3 (Fuzzy preditivo RGB e HSV).

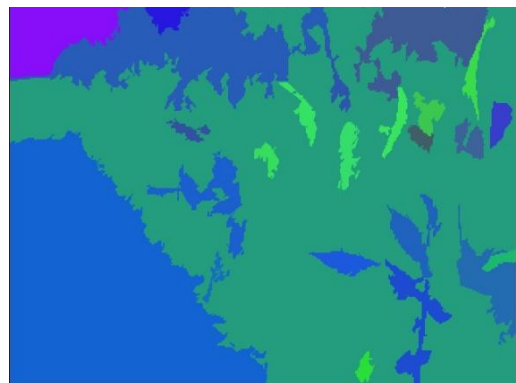


Figura E.50: Mapa de classe 3 (RNA RGB e HSV).

Linhas de plantio – Milharais 04

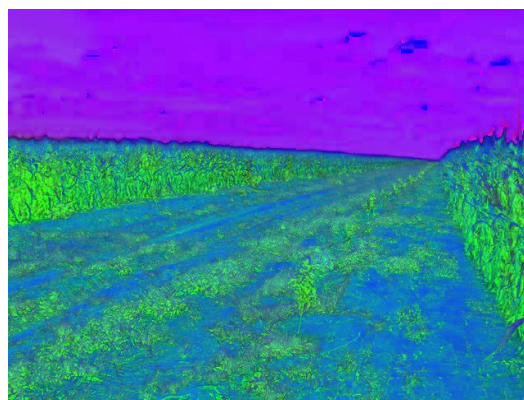


Figura E.51: Imagem original RGB e HSV (milharais 04).



Figura E.52: Mapa de classe 4 (Bayes RGB e HSV).



Figura E.53: Mapa de classe 4 (Bayes ingênuo RGB e HSV).

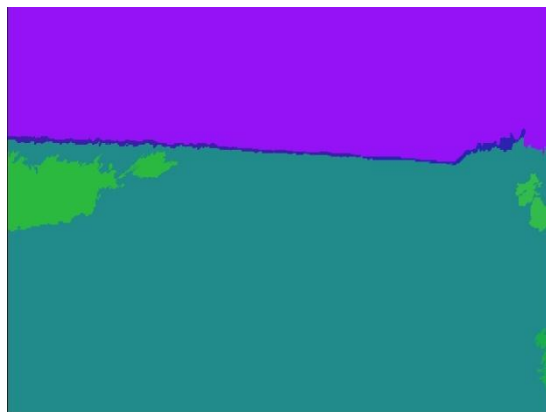
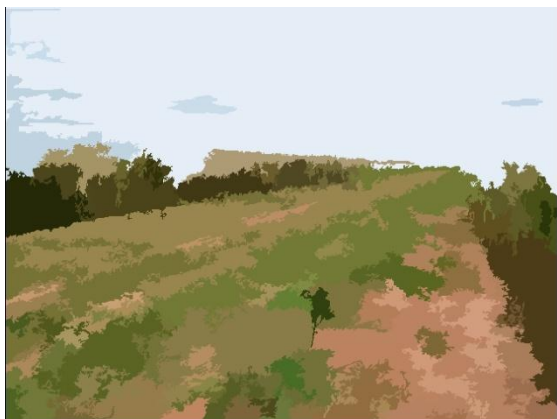


Figura E.54: Mapa de classe 4 (Fuzzy preditivo RGB e HSV).

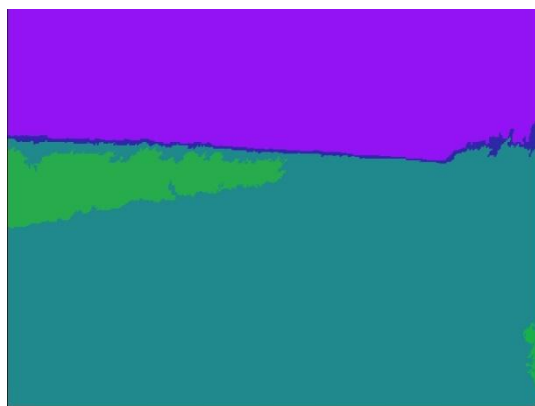


Figura E.55: Mapa de classe 4 (RNA RGB e HSV).

Linhas de plantio – Milharais 05

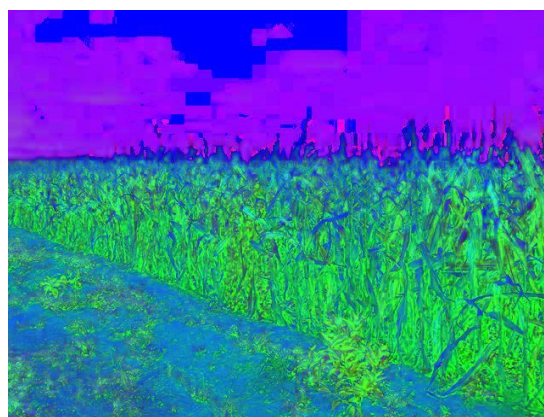


Figura E.56: Imagem original RGB e HSV (milharais 05).

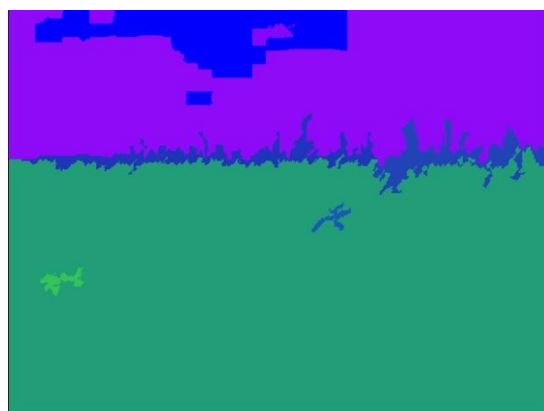


Figura E.57: Mapa de classe 5 (Bayes RGB e HSV).



Figura E.58: Mapa de classe 5 (Bayes ingênuo RGB e HSV).

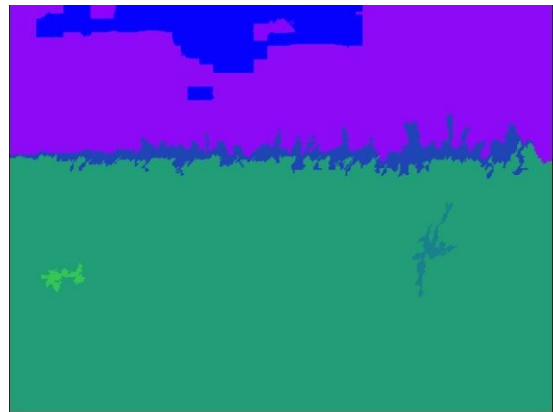
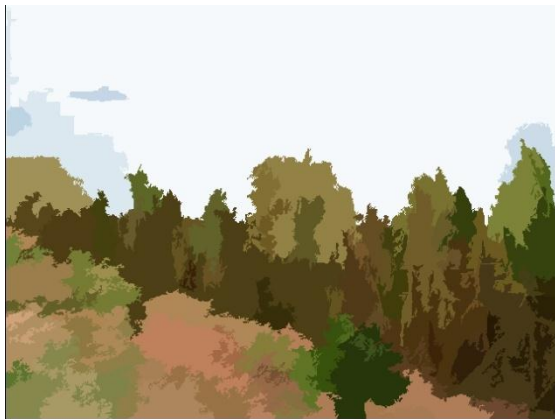


Figura E.59: Mapa de classe 5 (Fuzzy preditivo RGB e HSV).

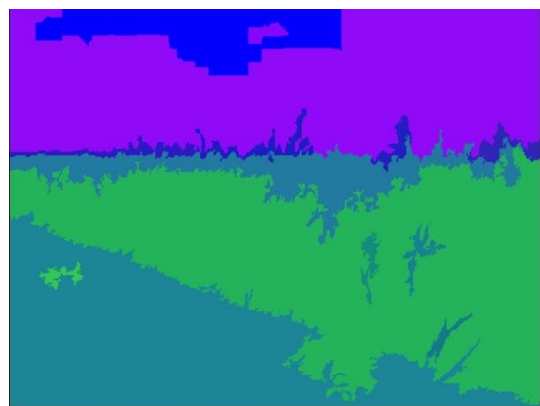


Figura E.60: Mapa de classe 5 (RNA RGB e HSV).

Linhas de plantio – Canaviais 02

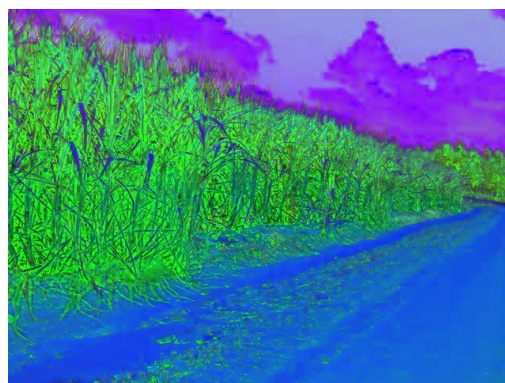


Figura E.61: Imagem original RGB e HSV (canaviais 02).



Figura E.62: Mapa de classe 2 (Bayes RGB e HSV).

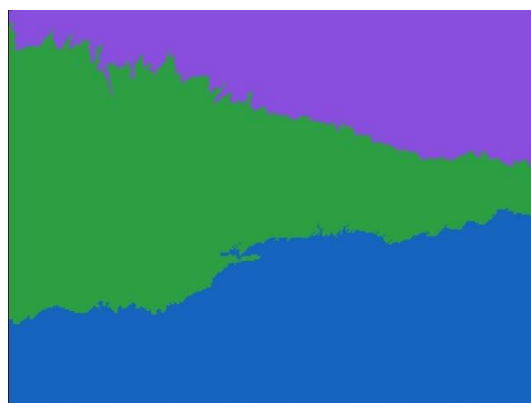


Figura E.63: Mapa de classe 2 (Bayes ingênuo RGB e HSV).

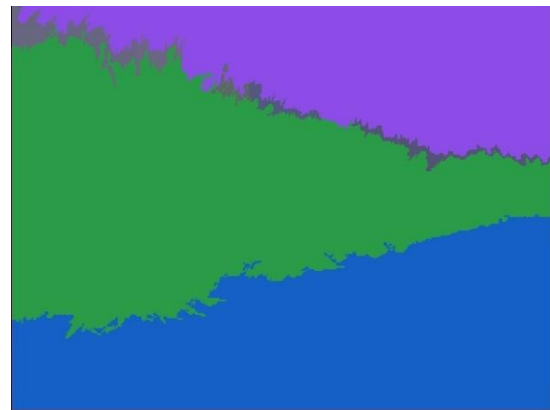


Figura E.64: Mapa de classe 2 (Fuzzy preditivo RGB e HSV).

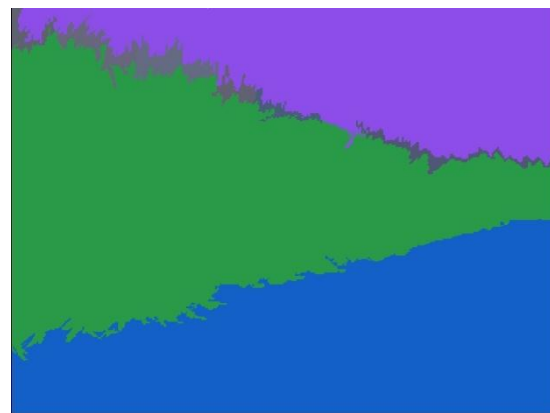


Figura E.65: Mapa de classe 2 (RNA RGB e HSV).

Linhas de plantio – Canaviais 03

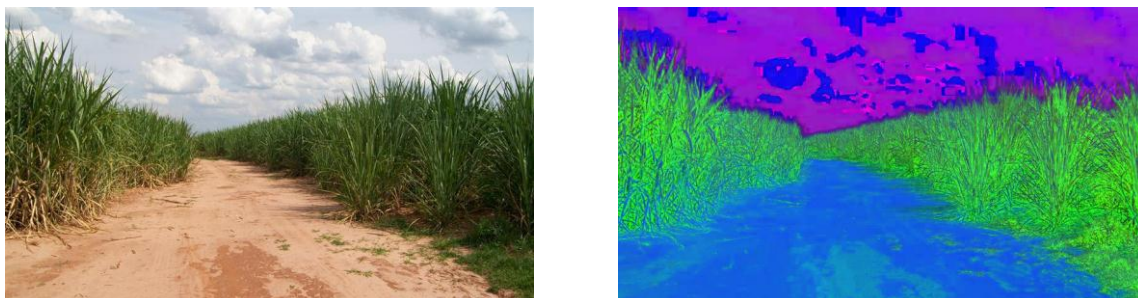


Figura E.66: Imagem original RGB e HSV (canaviais 03).

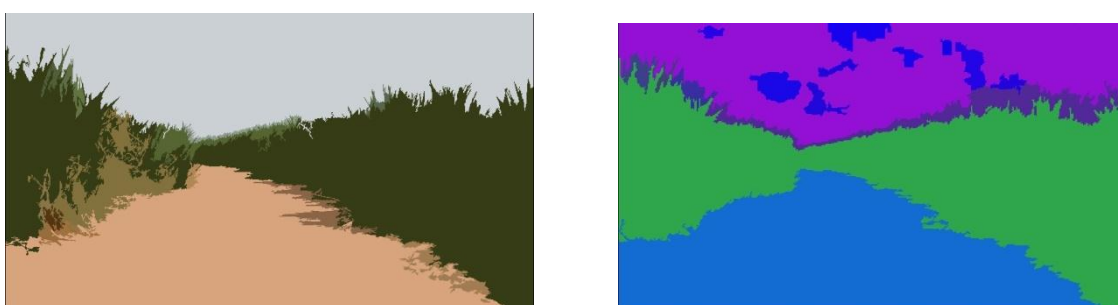


Figura E.67: Mapa de classe 3 (Bayes RGB e HSV).

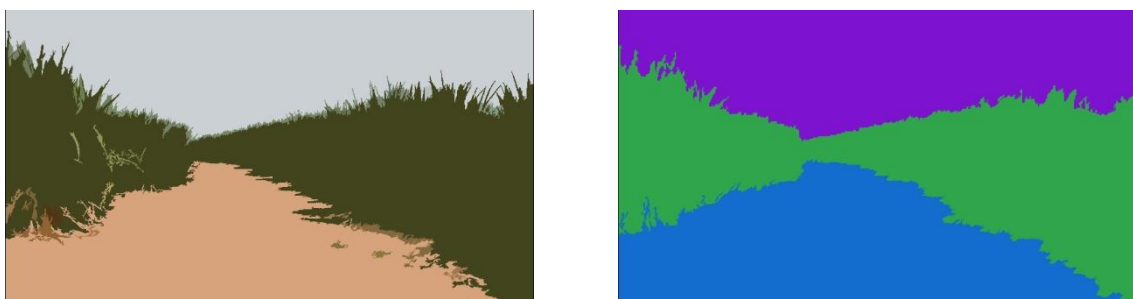


Figura E.68: Mapa de classe 3 (Bayes ingênuo RGB e HSV).

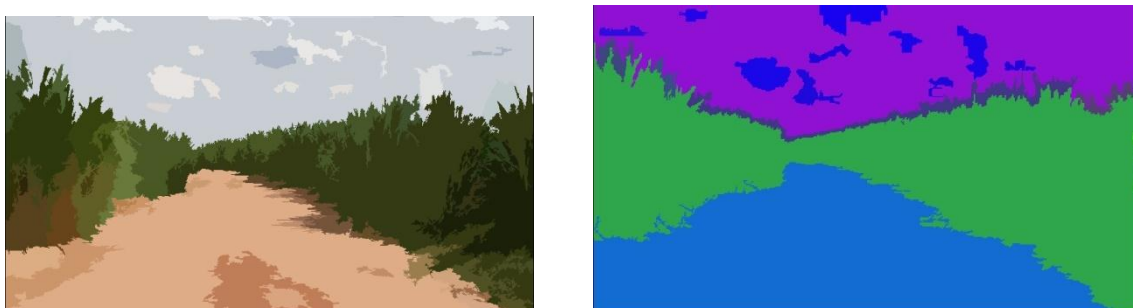


Figura E.69: Mapa de classe 3 (Fuzzy preditivo RGB e HSV).

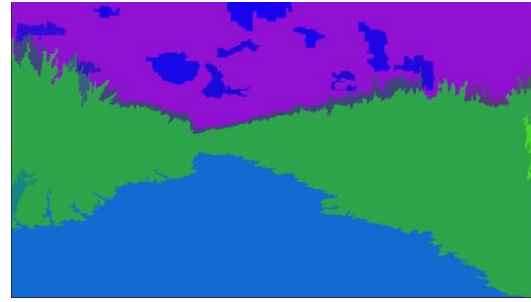
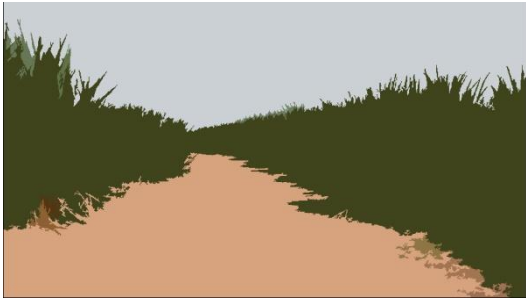


Figura E.70: Mapa de classe 3 (RNA RGB e HSV).

Linhas de plantio – Canaviais 04

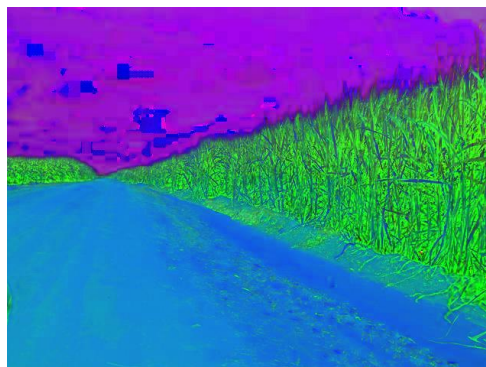


Figura E.71: Imagem original RGB e HSV (canaviais 04).

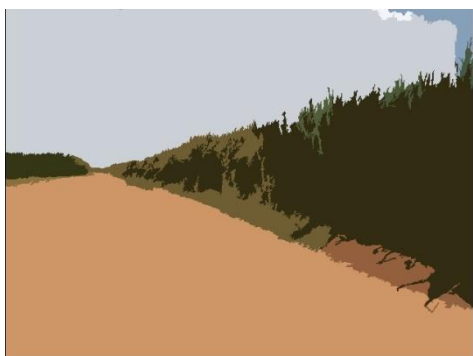


Figura E.72: Mapa de classe 4 (Bayes RGB e HSV).



Figura E.73: Mapa de classe 4 (Bayes ingênuo RGB e HSV)



Figura E.74: Mapa de classe 4 (Fuzzy preditivo RGB e HSV).



Figura E.75: Mapa de classe 4 (RNA RGB e HSV).

Linhas de plantio – Canaviais 05

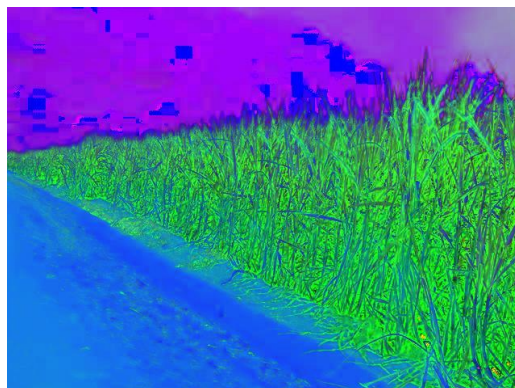


Figura E.76: Imagem original RGB e HSV (canaviais 05).

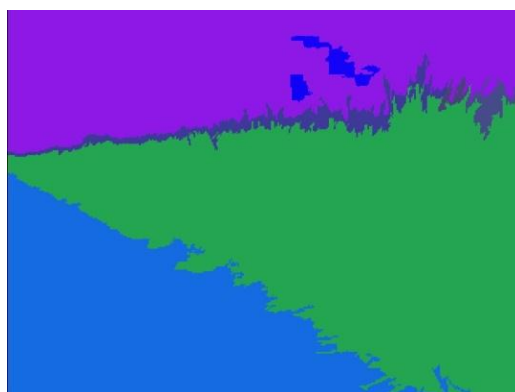


Figura E.77: Mapa de classe 5 (Bayes RGB e HSV).



Figura E.78: Mapa de classe 5 (Bayes ingênuo RGB e HSV).

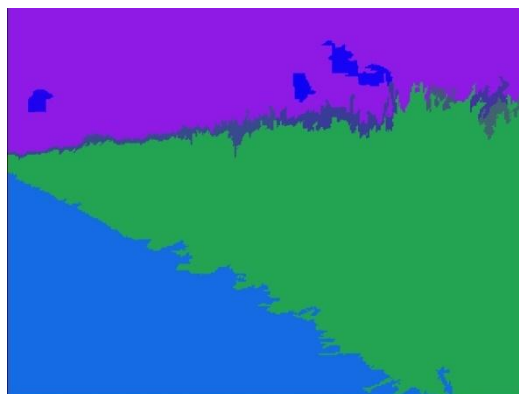


Figura E.79: Mapa de classe 5 (Fuzzy preditivo RGB e HSV).

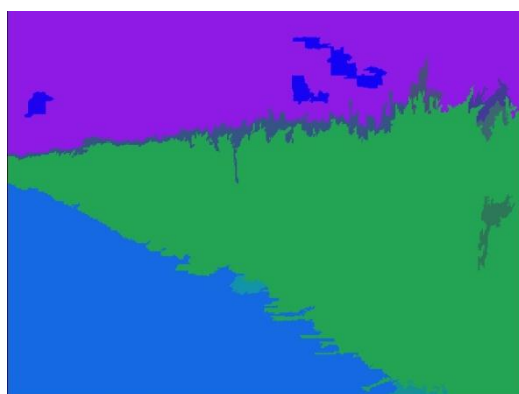
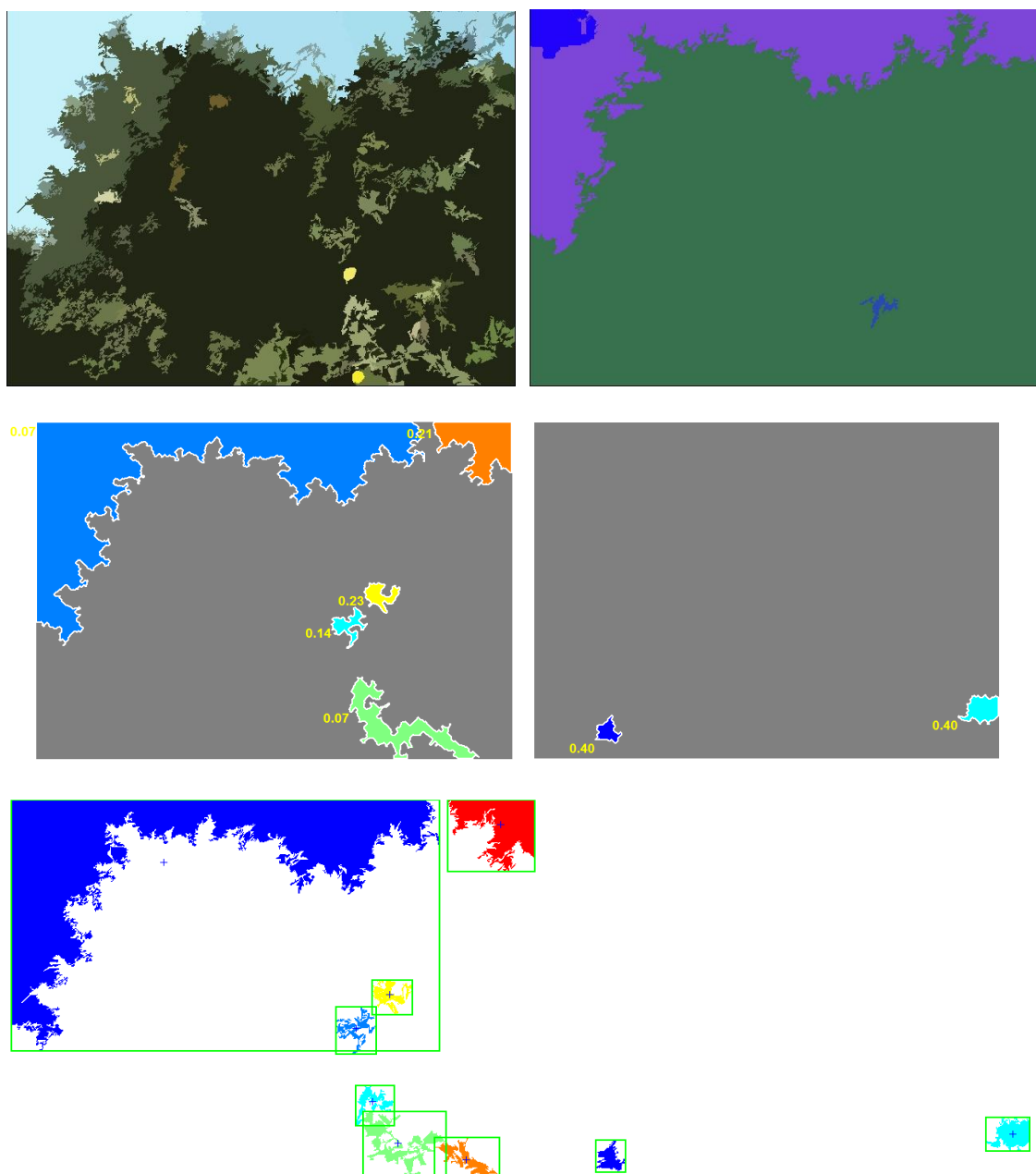


Figura E.80: Mapa de classe 5 (RNA RGB e HSV).

APÊNDICE F. Gráficos de quantificação e localização de objetos

Para os casos RGB e HSV são apresentadas as imagens abaixo, da segunda abordagem, de localização de frutos.



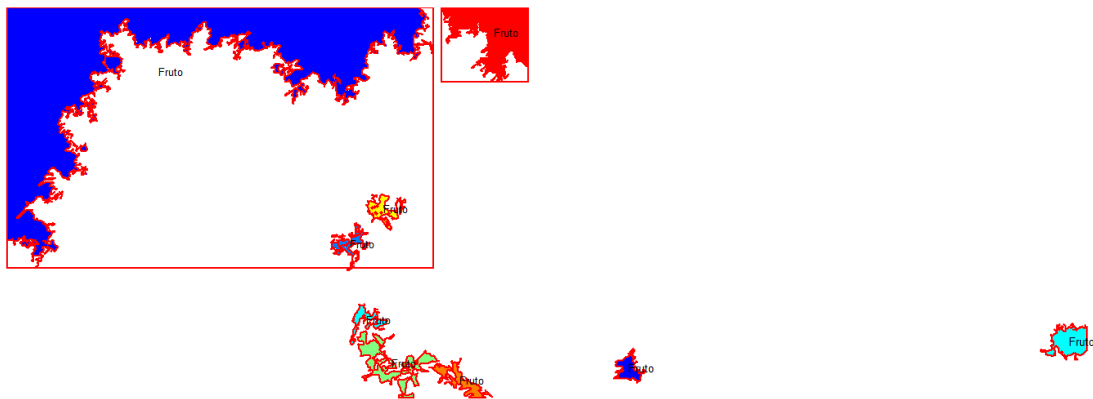


Figura F.1: Quantificação 03 – mapas de classe, métricas circulares, centróides e localização de frutos (RGB – esquerda; HSV – direita).

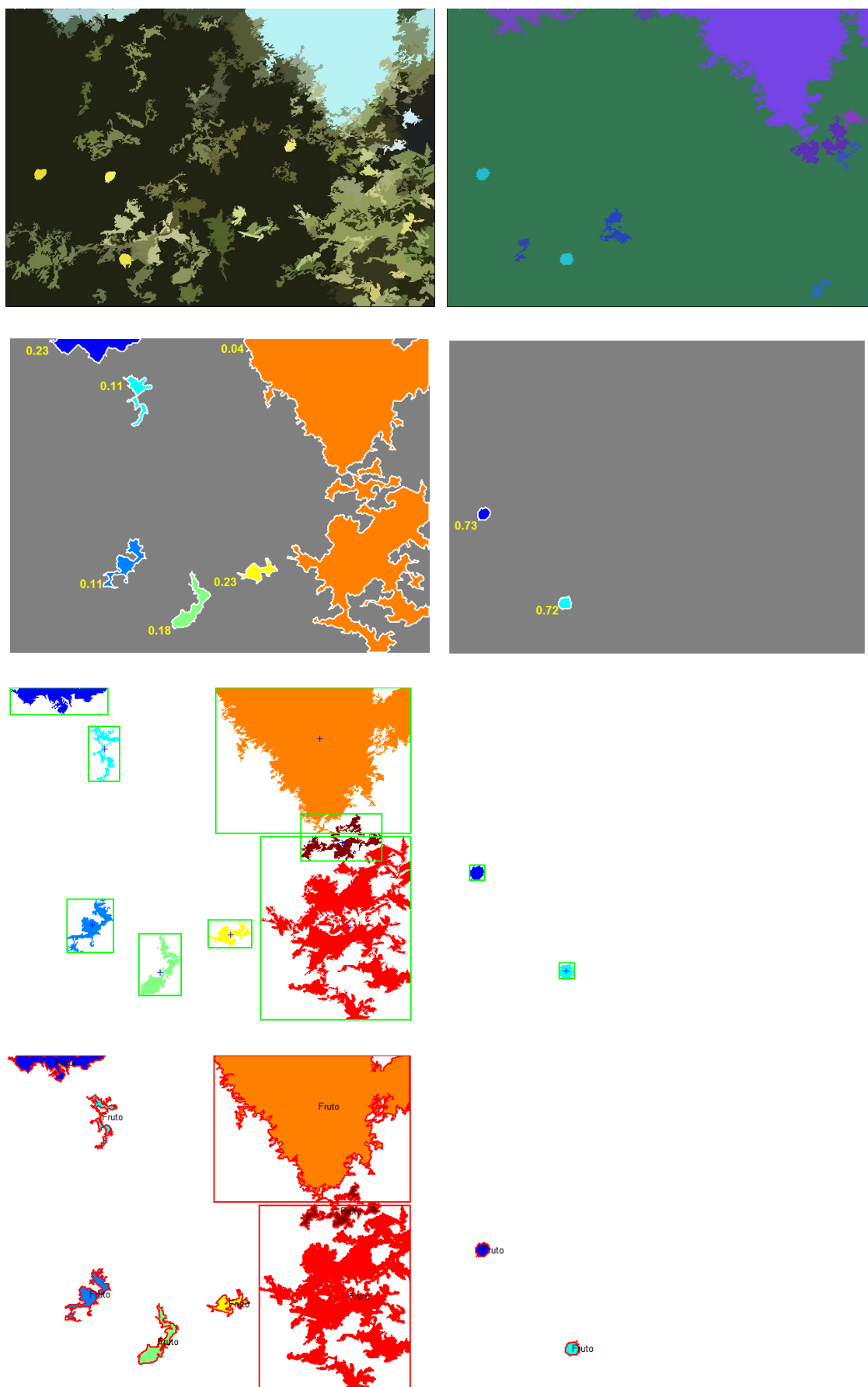


Figura F.2: Quantificação 04 – mapas de classe, métricas circulares, centróides e localização de frutos (RGB – esquerda; HSV – direita).

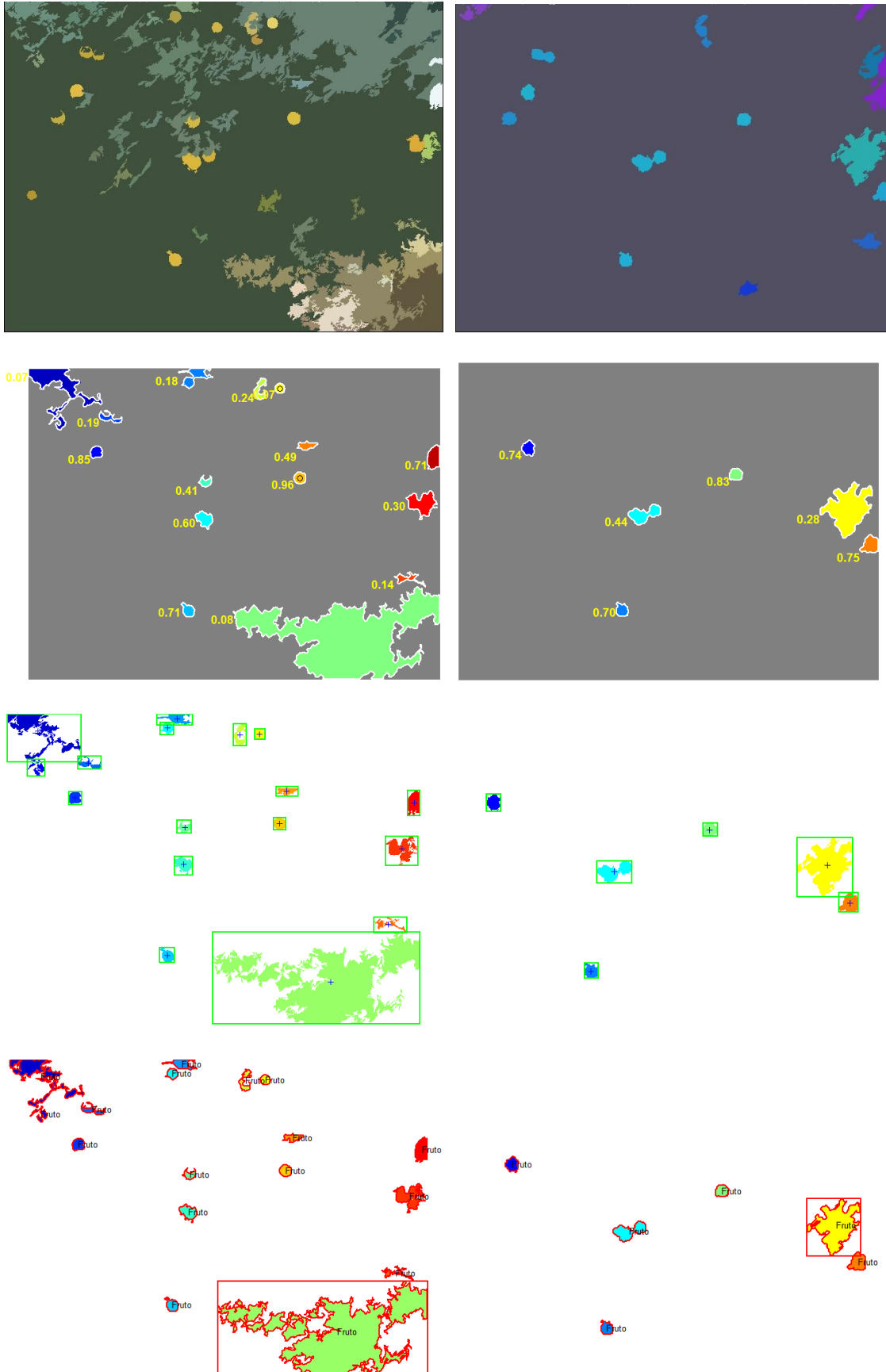


Figura F.3: Quantificação 05 – mapas de classe, métricas circulares, centróides e localização de frutos (RGB – esquerda; HSV – direita).

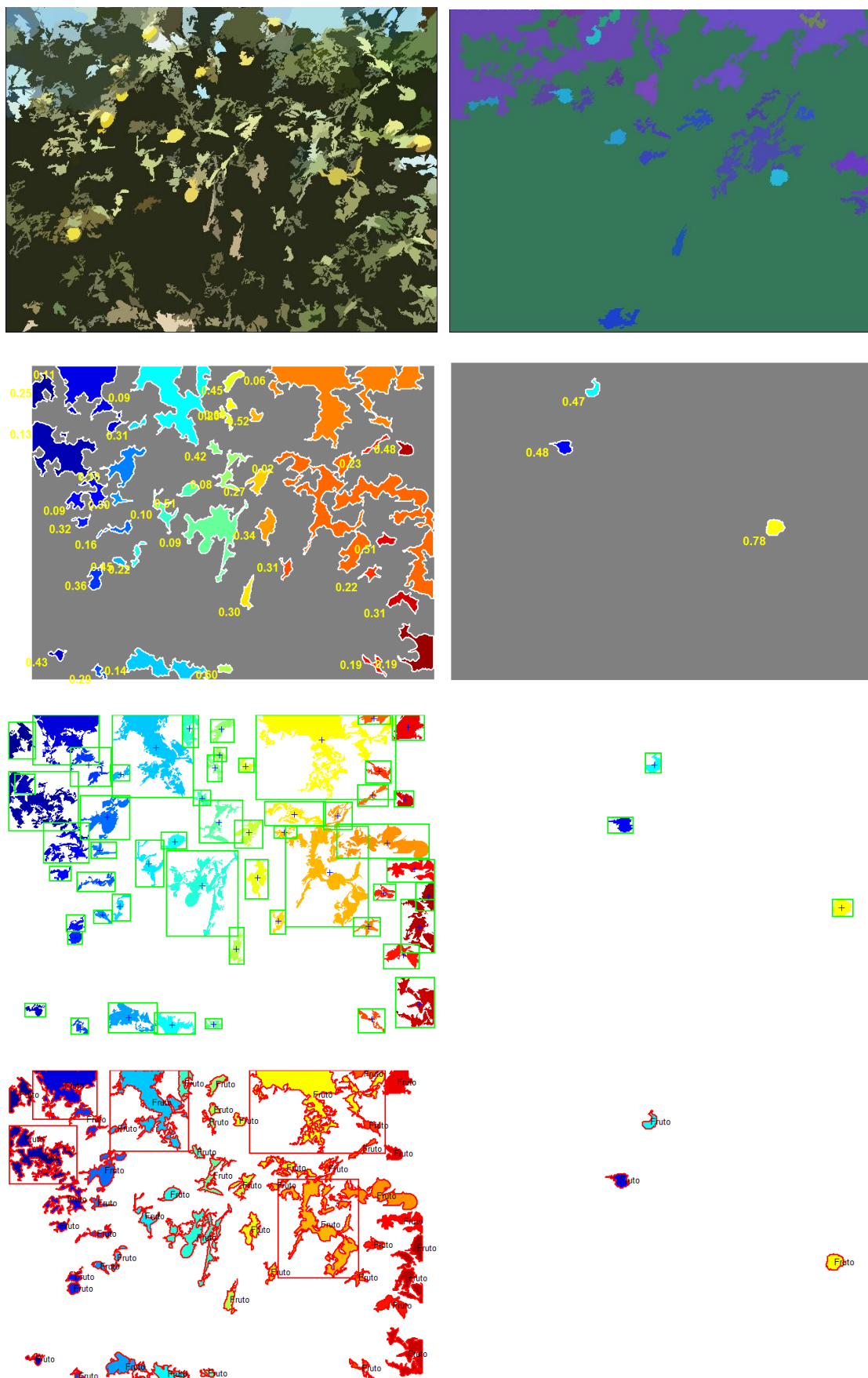


Figura F.4: Quantificação 06 – mapas de classe, métricas circulares, centróides e localização de frutos (RGB – esquerda; HSV – direita).

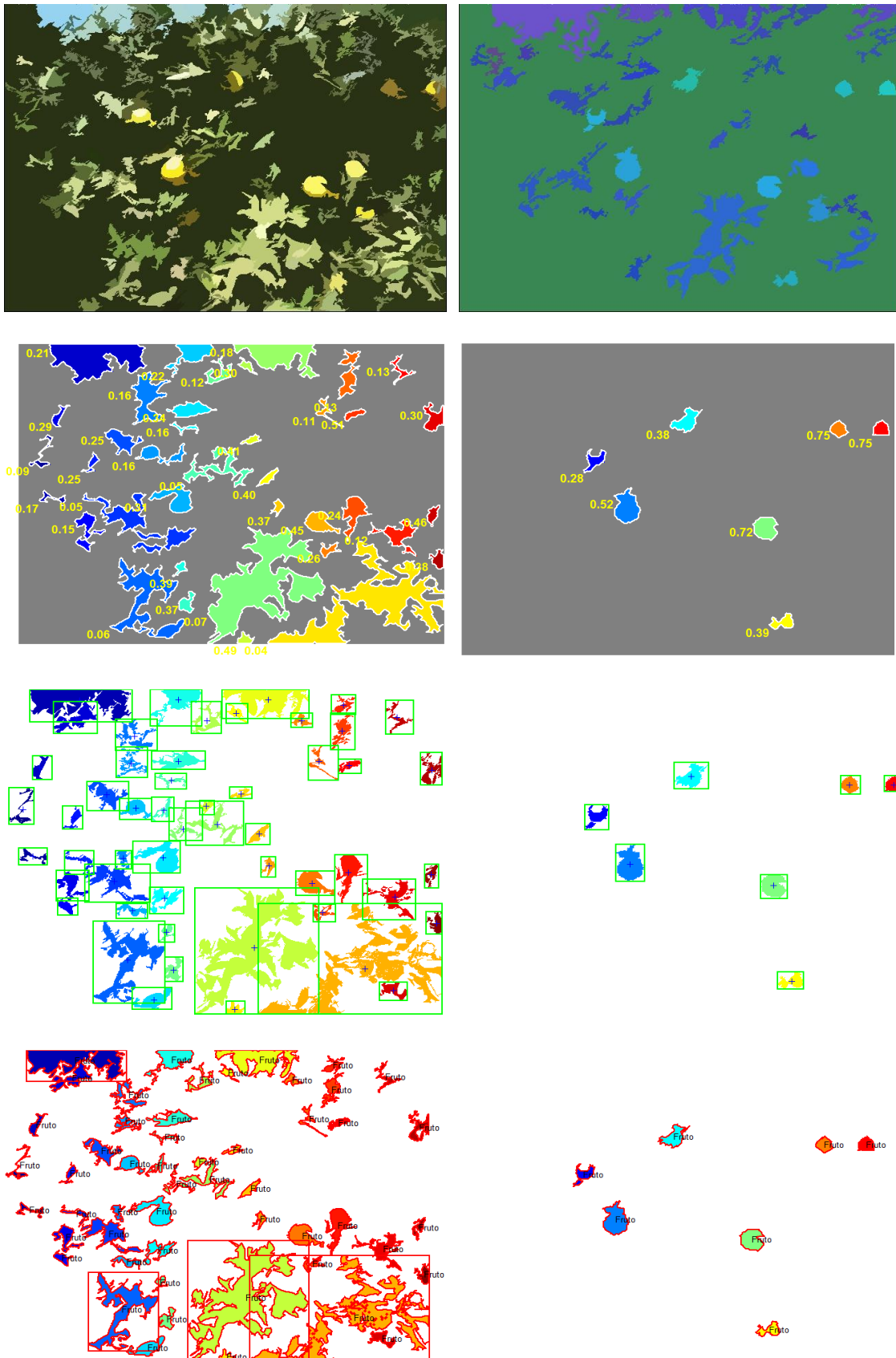


Figura F.5: Quantificação 07 – mapas de classe, métricas circulares, centróides e localização de frutos (RGB – esquerda; HSV – direita).

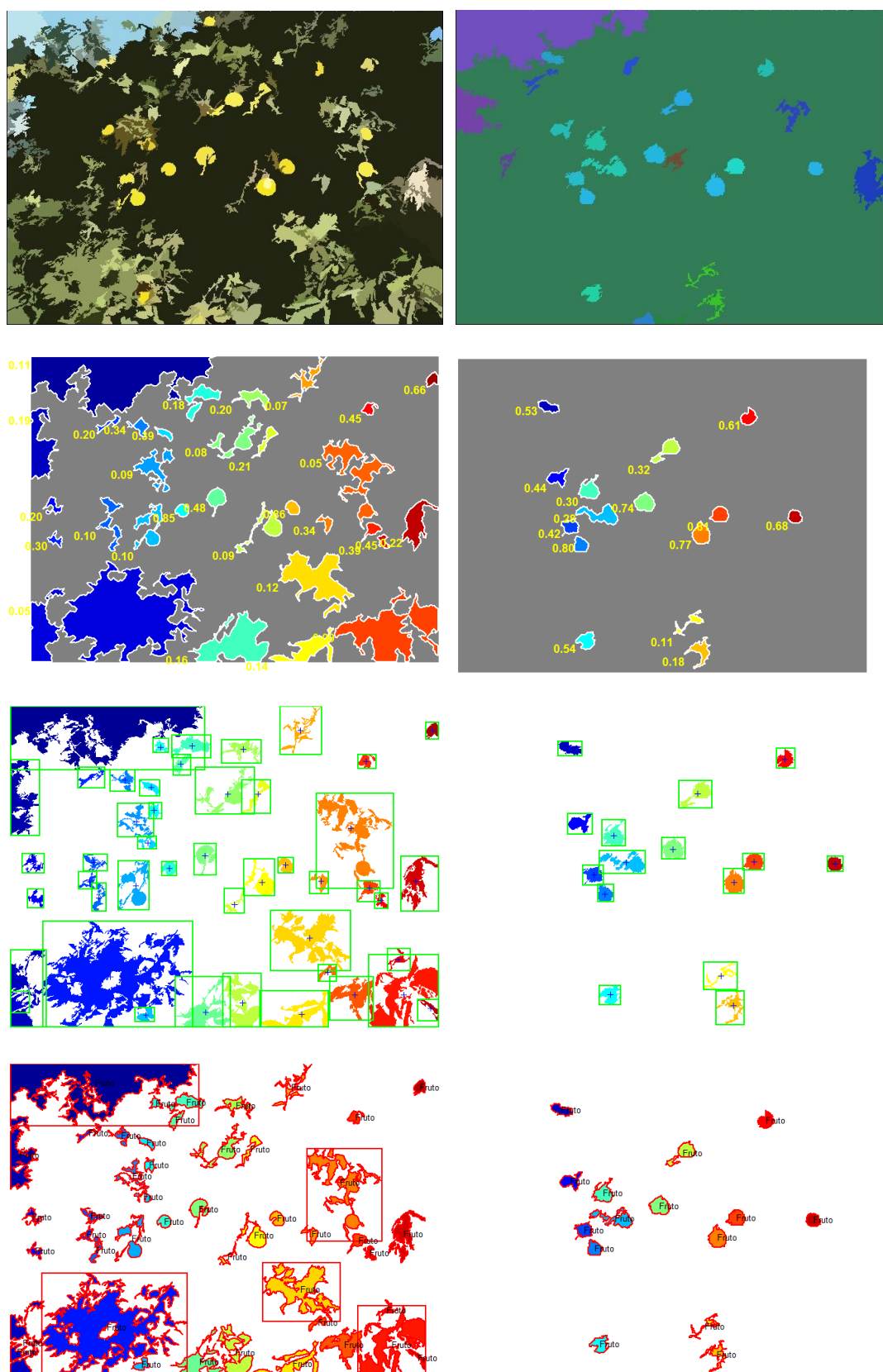


Figura F.6: Quantificação 08 – mapas de classe, métricas circulares, centróides e localização de frutos (RGB – esquerda; HSV – direita).

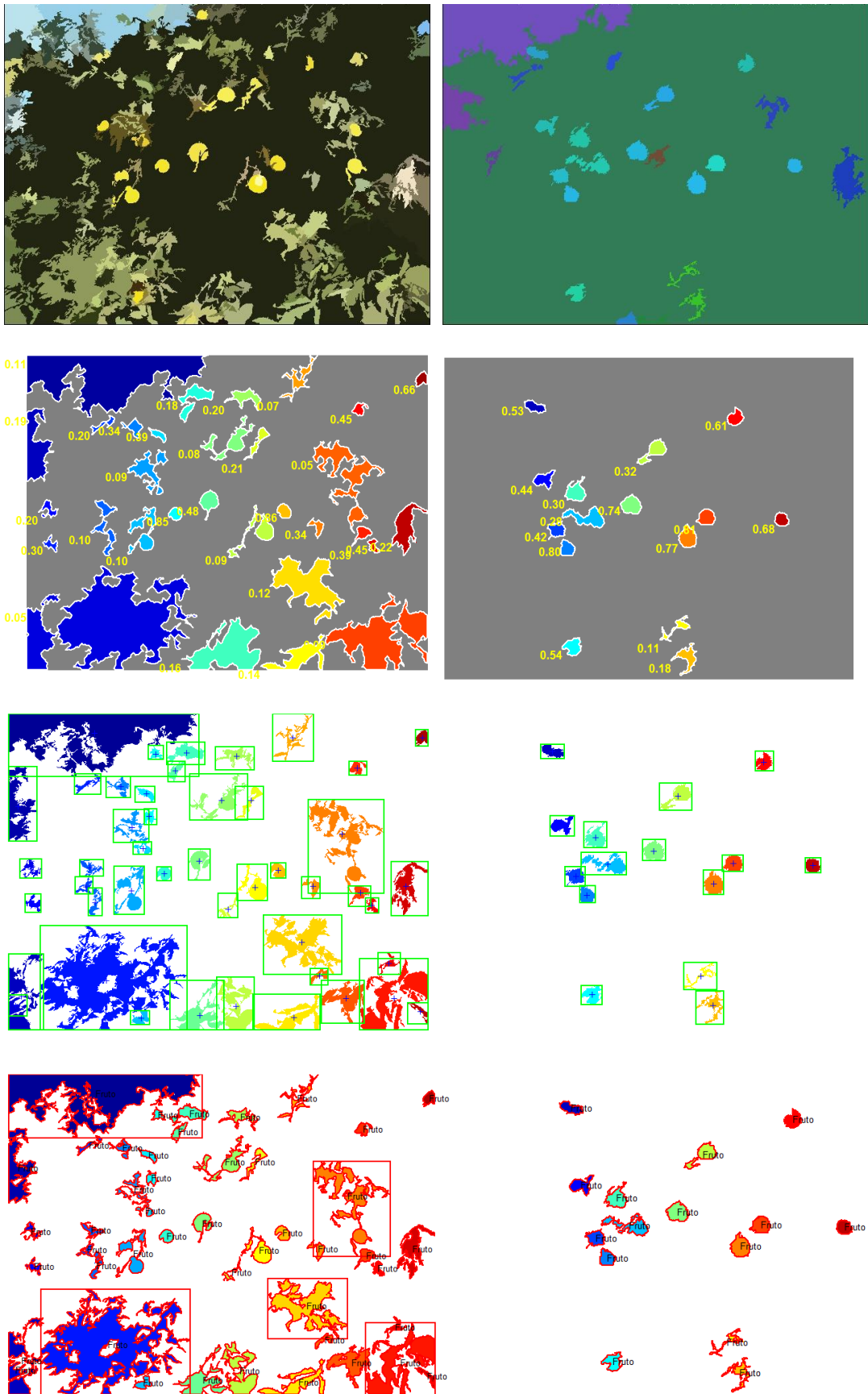


Figura F.7: Quantificação 09 – mapas de classe, métricas circulares, centróides e localização de frutos (RGB – esquerda; HSV – direita).

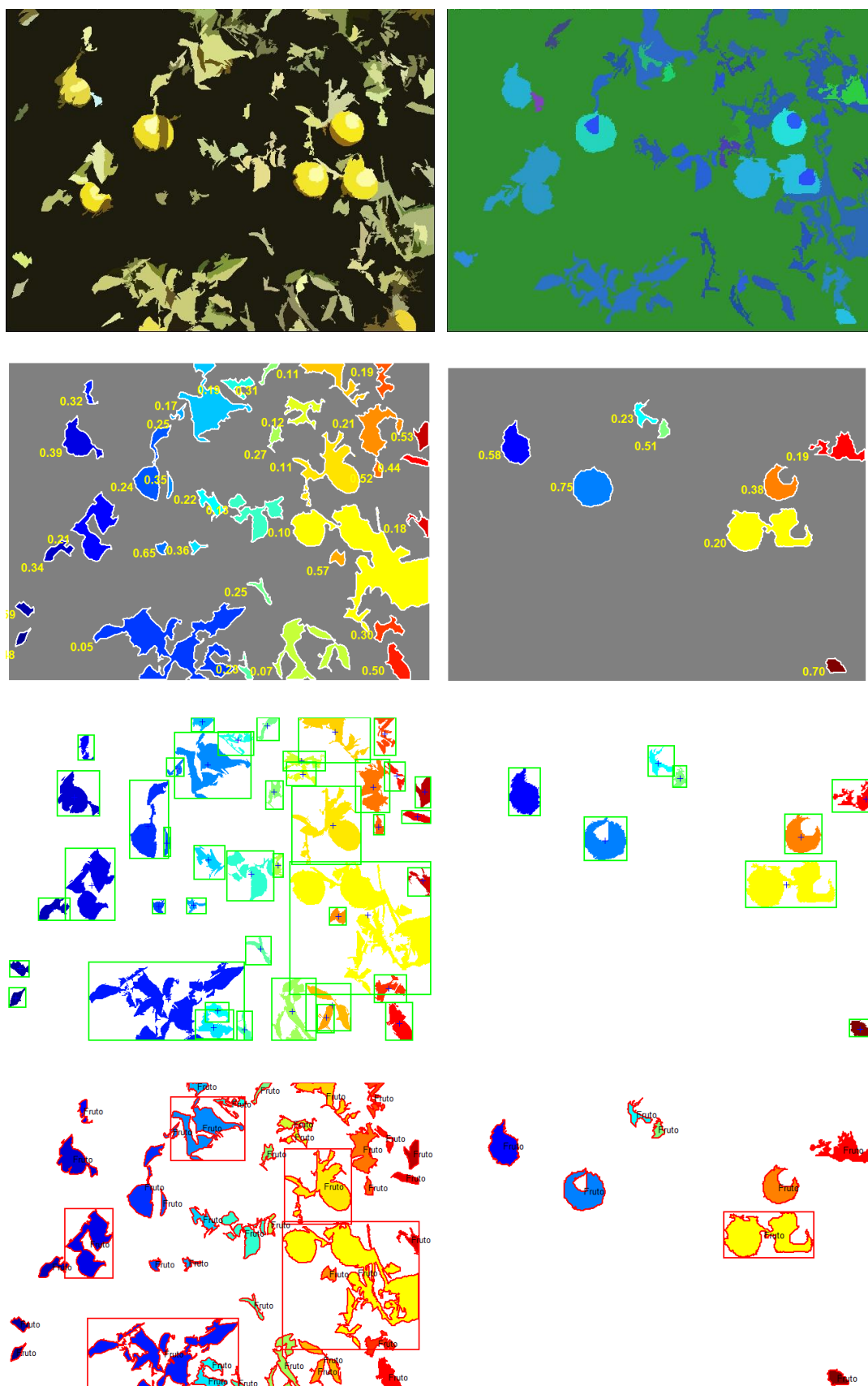


Figura F.8: Quantificação 10 – mapas de classe, métricas circulares, centróides e localização de frutos (RGB – esquerda; HSV – direita).

APÊNDICE G. Publicações científicas

Os seguintes artigos científicos referentes a este trabalho foram publicados abaixo:

- LULIO, L. C.; Prof. Dr. PORTO, A. J. V.; Prof. Dr. TRONCO, M. L. **Omnidirectional Vision System for Agricultural Mobile Robot Navigation**. In: VI Congresso Nacional de Engenharia Mecânica (ABCM/CONEM 2010), 2010, Campina Grande, PB - Brasil. Controle & Automação (Impresso), 2010.
- LULIO, L. C.; Prof. Dr. TRONCO, M. L.; Prof. Dr. PORTO, A. J. V. **JSEG-based Image Segmentation in Computer Vision for Agricultural Mobile Robot Navigation**. In: CIRA 2009 - IEEE International Symposium on Computational Intelligence in Robotics and Automation, 2009, Daejeon. Proceedings of IEEE IS CIRA 2009.
- LULIO, L. C.; Prof. Dr. TRONCO, M. L.; Prof. Dr. PORTO, A. J. V. **Mobile Robot Navigation with JSEG-based Image Segmentation Technique in Computer Vision**. In: 30th CILAMCE - Iberian-Latin-American Congress on Computational Methods in Engineering, 2009, Armação dos Búzios - RJ. Proceedings of ABMEC (Brazilian Association for Computational Methods in Engineering), 2009.
- LULIO, L. C.; Prof. Dr. PORTO, A. J. V.; Prof. Dr. TRONCO, M. L. **Image Segmentation-based Techniques in Computer Vision for Agricultural Mobile Robots Navigation** (aceito para publicação). In: 7 Congresso Brasileiro de Agroinformática (SBIAgro), 2009, Viçosa, MG - Brasil. Revista Brasileira de Agroinformática, 2009.
- LULIO, L. C.; Prof. Dr. PORTO, A. J. V.; Prof. Dr. TRONCO, M. L. **Statistical ANN-based Image Recognition for Autonomous Mobile Robot Navigation** - IEEE International Conference on Control, Automation, Robotics and Vision (ICARCV'10). IEEE Computer Society, 2010.
- LULIO, L. C.; Prof. Dr. TRONCO, M. L.; Prof. Dr. PORTO, A. J. V. **JSEG-based Image Segmentation and Statistical-ANN merging technique for pattern recognition in Agricultural Mobile Robot Navigation** - IEEE/IAPR International Conference on Pattern Recognition (ICPR'10). IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010.
- LULIO, L. C.; Prof. Dr. PORTO, A. J. V.; Prof. Dr. TRONCO, M. L. **Mobile Robot Navigation with JSEG-based Image Segmentation Technique and Statistical ANN Pattern Recognition** - AIMI International Conference on Image Analysis and Recognition (ICIAR'10)/AIMI International Conference on Autonomous and Intelligent Systems (ICAIS'10) - Springer LNCS. Lecture Notes in Computer Science, 2010.
- LULIO, L. C.; Prof. Dr. TRONCO, M. L.; Prof. Dr. PORTO, A. J. V. **ANN Statistical Image Recognition Method for Computer Vision in Agricultural Mobile Robot**

Navigation - IEEE International Conference on Mechatronics and Automation (ICMA'10)/IEEE International Conference on Advanced Mechatronics (ICAM'10). IEEE Transactions on Robotics and Automation, 2010.

- LULIO, L. C.; Prof. Dr. PORTO, A. J. V.; Prof. Dr. TRONCO, M. L. **Pattern Recognition Structured Heuristics Methods for Image Processing in Mobile Robot Navigation** - IEEE International Conference on Robotics Systems (IROS'10). IEEE Robotics & Automation Magazine, 2010.
- LULIO, L. C.; Prof. Dr. TRONCO, M. L.; Prof. Dr. PORTO, A. J. V. **Customized Statistical ANN-based method Classification for Image Segmentation in a Mobile Robot Vision System** - IEEE World Congress on Computational Intelligence (WCCI'10). IEEE Computational Intelligence Magazine, 2010.
- LULIO, L. C.; Prof. Dr. TRONCO, M. L.; Prof. Dr. PORTO, A. J. V. **JSEG Algorithm and Statistical ANN Image Segmentation Techniques for Natural Scenes** - Capítulo de Livro – *Image Segmentation* – INTECH Publishing. ISBN 978-953-307-228-9. 2011.

6. Referências Bibliográficas

ALEIXOS, N., BLASCO, J., NAVARRÓN, F. e MOLTÓ, E. (2002). **Multispectral inspection of citrus in real-time using machine vision and digital signal processors.** Computers and Electronics in Agriculture, v 33, n 2, p 121-137.

ANDREASSON, H. e DUCKETT, T. (2004). **Topological localization for mobile robots using omnidirectional vision and local features.** In: Proceedings of IAV 2004 - The 5th IFAC Symposium on Intelligent Autonomous Vehicles.

ANNAMALAI, P., LEE, W. S. e BURKS, T. F. (2004). **Color vision system for estimating citrus yield in real-time.** In: ASAE/CSAE ANNUAL INTERNATIONAL MEETING, 2004, Ottawa, Ontário, Canadá.

AUERNHAMMER, H. e SPECKMANN, H. (2006). **Dedicated communication systems and standards for agricultural applications.** Chapter 7 Communication Issues and Internet Use, CIGR Handbook of Agricultural Engineering, v. 7, p. 435-452.

BACZYK, R., KASIŃSKI, A. e SKRZYPCZYŃSKI, P. (2003). **Vision-based mobile robot localization with simple artificial landmarks.** 7th IFAC Symposium on Robot Control.

BAIS, A., SABLATNIG, R., GU, J., KHAWAJA, M., USMAN, M., HASAN, G. e IQBAL, M. (2008). **Stereo vision based self-localization of autonomous mobile robots.** Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), v 4931 LNCS, p 367-380, 2008, Robot Vision - Second International Workshop, RobVis 2008, Proceedings.

BAKKER, T., WOUTERS, H., van ASSELT, K., BONTSEMA, J., TANG, L., MÜLLER, J. e van STRATEN, G. (2008). **A vision based row detection system for sugar beet.** Computers and Electronics in Agriculture, v 60, n 1, p 87-95, January 2008.

BAYES, T. e PRICE, M. (1763). **An Essay towards solving a Problem in the Doctrine of Chances.** Philosophical Transactions of the Royal Society of London 53: 370-418. doi:10.1098/rstl.1763.0053.

BESL, P. J. e MACKAY, N. D. (1992). **A method for registration of 3-D shapes.** IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 14 (2), p.239 - 256.

BITTENCOURT de, A. (1998). **Fotogrametria.** Ed. SBEE. Curitiba, Brasil.

BLASCO, J., ALEIXOS, N., GÓMEZ-SANCHIS, J. e MOLTÓ, E. (2009). **Recognition and classification of external skin damage in citrus fruits using multispectral data and morphological features.** Biosystems Engineering, v 103, n 2, p 137-145.

BLASCO, J., ALEIXOS, N., GOMEZ, J. e MOLTO, E. (2007-a). **Citrus sorting by identification of the most common defects using multispectral computer vision algorithm.** Journal of Food Engineering 83, 384–391.

BLASCO, J., ALEIXOS, N. e MOLTO, E. (2007-b). **Computer vision detection of peel defects in citrus by means of a region oriented segmentation algorithm.** Journal of Food Engineering 81, 535–543.

BONATO, V. (2007). **Proposta de uma arquitetura de hardware em FPGA implementada para SLAM com multi-câmeras aplicada à robótica móvel.** Tese (Doutorado) - Instituto de Ciências Matemáticas e Computação, Universidade de São Paulo (ICMC/USP) - São Carlos, SP, Brasil.

BOYLE, W. S. e SMITH, G. E. (1970). **Charge Coupled Semiconductor Devices.** Bell Sys. Tech. J. 49 (4): 587–593.

BRADSKI, G. e KAEHLER, A. (2008). **Learning OpenCV: Computer Vision with the OpenCV Library.** O'Reilly Media; 1st edition (September 24, 2008).

BUENO, J. M. (2001). **Suporte à recuperação de imagens médicas baseada em conteúdo através de histogramas métricos.** 162 p. (Tese Doutorado). Instituto de Ciências Matemáticas e de Computação/Universidade de São Paulo (ICMC/USP). São Carlos, São Paulo - Brasil.

BULANON, D. M., BURKS, T. F. e ALCHANATIS, V. (2009). **Fruit visibility analysis for robotic citrus harvesting.** Transactions of the ASABE, v 52, n 1, p 277-283, January-February 2009.

BULANON, D. M., KATAOKA, T., OKAMOTO, H. e HATA, S. (2004). **Determining the 3-d location of the apple fruit during harvest.** In: Proceedings of AUTOMATION TECHNOLOGY FOR OFF-ROAD EQUIPMENT, Kyoto, Japão.

BURT, D. J. (1974). **Basic Operation of the Charge Coupled Device.** International Conference on Technology and Applications of Charge Coupled Devices. September 1974. Edinburgh: University of Edinburgh, Centre for Industrial Consultancy and Liaison.

CASTANÓN, C. A. B. e TRAINA, A. J. M. (2003). **Caracterização de imagens através de análise multiresolução por wavelets para sistemas cbir.** In: XXIX CONFERÊNCIA LATINOAMERICANA DE INFORMÁTICA.

CAVANI, F. A. (2007). **Análise de cenas de pomares de laranjeiras através de segmentação de imagens e reconhecimento de padrões.** . Dissertação (Mestrado). Escola de Engenharia de São Carlos, Universidade de São Paulo. p. 287. São Carlos, SP Brasil.

CHANG, C., WANG, H. e FU, S. (2010). **Texture image classification using modular radial basis function neural networks.** Journal of Electronic Imaging, v 19, n 1, 2010.

CHEN, J., PAPPAS, N., MOJSILOVIC, A. e ROGOWITZ, B. E. (2005). **Adaptive perceptual color-texture image segmentation.** IEEE Transactions on Image Processing, v. 14, n. 10, p. 1524–1536, out. ISSN 1057-7149.

CHEN, K. e CHEN, S. (2002). **Color texture segmentation using feature distributions**. Pattern Recognition Letters, v. 23, n. 7, p. 755–771.

CHEN, T. e CHAN, E. (2009). **Optimal in-place algorithms for 3-d convex hulls and 2-d segment intersection**. Proceedings of the Annual Symposium on Computational Geometry, p 80-87.

CHENG, H. D., JIANG, X. H., SUN, Y. e WANG, J. (2001). **Color image segmentation: advances and prospects**. Pattern Recognition, v. 34, p. 2259–2281.

CHENG, H. D., LIU, B., HUANG, J., TIAN, J., TANG, X. e LIU, J. (2010). **Fully automatic and segmentation-robust classification of breast tumors based on local texture analysis of ultrasound images**. Pattern Recognition, v 43, n 1, p 280-298, January 2010.

CHENG, H. D., ZHANG, M. e ZHANG, L. (2010). **A neutrosophic approach to image segmentation based on watershed method**. Signal Processing, v 90, n 5, p 1510-1517, May 2010.

COMPANY, E. K. (2003). **Charge-coupled device CCD image sensors**. Disponível em: <Erro! A referência de hiperlink não é válida.>. Acesso em: 20/06/2009.

COSTA, L. F. e CESAR JÚNIOR, R. M. (2001). **Shape analysis and classification - Theory and Practice**. 1.ed. Boca Raton, Florida, EUA: CRC Press LLC. ISBN 0-8493-3493-4.

CRUVINEL, P. E. e MINATEL, E. R. (2002). **Image processing in automated pattern classification of oranges**. In: (Proceedings) WORLD CONGRESS OF COMPUTERS IN AGRICULTURE AND NATURAL RESOURCES, Foz do Iguaçu, Brasil. ASAE, p.56–61.

CYGANEK, B. e SIEBERT, P. (2009). **An Introduction to 3D Computer Vision Techniques and Algorithms**. John Wiley & Sons, Ltd. ISBN: 978-0-470-01704-3.

DAINESE, R. C., PUGLIESE, P. B., JORGE, L. A. C., O.;, T. J. e JR.;, C. (2004). **Avaliação da incidência de plantas invasoras por meio de segmentação de imagens e redes neurais**. In: CONGRESSO BRASILEIRO DE AGRICULTURA DE PRECISÃO (Anais), Piracicaba, São Paulo, Brasil: USP/Esalq.

DALSA, D. L. (2005). **CMOS vs. CCD: Maturing Technologies, Maturing Markets**. Reprinted from August 2005 issue of PHOTONICS SPECTRA - Laurin Publishing. Disponível em: <http://www.dalsa.com/corp/markets/ccd_vs_cmos.aspx>. Acesso em: 09/07/2009.

DECCÓ, C. C. G. (2004). **Construção de mapas de ambiente para navegação de robôs móveis com visão omnidirecional estéreo**. Dissertação (mestrado). Escola Politécnica - Universidade de São Paulo. 202p. São Paulo, SP - Brasil.

DENG, Y., KENNEDY, S., MOORE, M. S. e MANJUNATH, B. S. (1999b). **Peer group filtering and perceptual color image quantization**. Proc. IEEE International Symposium on Circuits and Systems VLSI (ISCAS'99), Orlando, FL, vol. 4, pp. 21-24.

DENG, Y. e MANJUNATH, B. S. (1999c). **An efficient low-dimensional color indexing scheme for region based image retrieval**. Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP-99), Phoenix, Arizona, March 1999.

DENG, Y. e MANJUNATH, B. S. (2001). **Unsupervised segmentation of color-texture regions in images and video**. IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI '01).

DENG, Y., MANJUNATH, B. S. e SHIN, H. (1999a). **Color Image Segmentation**. Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition CVPR'99, Fort Collins, CO, vol. 2, pp. 446-51.

DeSOUZA, G. N. e KAK, A. C. (2002). **Vision for mobile robot navigation: a survey**. IEEE Transactions on Pattern Analysis and Machine Intelligence, New York (USA), v. 24, n. 2, p. 237–267, February, 2002.

DIAZ, R., GIL, L., SERRANO, C., BLASCO, J. e MOLTO, E. (2004). **Comparison of three algorithms in the classification of table olives by means of computer vision**. Journal of Food Engineering 61, 101–107.

DIJKSTRA, E. W. (1959). **A note on two problems in connexion with graphs**. Numerische Mathematik , 1 (1), 269-271.

DUDA, R. O. e HART, P. E. (1970). **Pattern Classification and Scene Analysis**. John Wiley & Sons, New York.

EASTMAN KODAK, C. (2001). **CCD Primer**. MTD/PS-0218.

EGMONT-PETERSEN, M., RIDDER, D. e HANDELS, H. (2002). **Image processing with neural networks - a review**. Pattern Recognition, v. 35, n. 10, p. 2279–2301.

FARMER, M. E. (2007). **Chaotic Phenomena from Motion in Image Sequences**. EEE International Conference on Neural Networks - p 835-840. The 2007 International Joint Conference on Neural Networks (IJCNN), Orlando, Florida, USA. .

FIGUEIREDO, M. A. T. e JAIN, A. K. (2002). **Unsupervised learning of finite mixture models**. IEEE Transactions on Pattern Analysis and Machine Intelligence, v. 24, n. 3, p. 381–396. ISSN 0162-8828.

FLORCZYK, S. (2005). **Robot Vision: Video-based Indoor Exploration with Autonomous and Mobile Robots**. WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim ISBN: 3-527-40544-5.

FORSYTH, D. A. e PONCE, J. (2002). **Computer Vision: A Modern Approach**. Prentice Hall - US ed., 693 p. - ISBN: 0130851981.

GAO, H., XU, W., SUN, J. e TANG, Y. (2010). **Multilevel thresholding for image segmentation through an improved quantum-behaved particle swarm algorithm**. IEEE Transactions on Instrumentation and Measurement, v 59, n 4, p 934-946.

GERSHO, A. e GRAY, R. M. (1992). **Vector quantization and signal compression**. Kluwer Academic, Norwell, MA.

GNEDENKO, B. V. (1982). **The theory of probability**. Moscow: MIR, p.392.

GODOY, E. P. (2007). **Desenvolvimento de uma ferramenta de análise de desempenho de redes CAN (Controller Area Network) para aplicações em sistemas agrícolas**. Dissertação (Mestrado) - Escola de Engenharia de São Carlos, Universidade de São Paulo. 93p. São Carlos, SP - Brasil.

GOLDBERG, K., SONG, D. e KHOR, Y. (2002). **Collaborative online teleoperation with spatial dynamic voting and a human "tele-actor"**. Proceedings - IEEE International Conference on Robotics and Automation, v 2, p 1179-1184.

GOMIDE, F. A. C. e GUDWIN, R. R. (1994). **Modelagem, Controle, Sistemas e Lógica Fuzzy**. SBA Controle & Automação, Vol. 4,Nº 3, Campinas, Brasil.

GONZAGA, A. (2010). **Notas de aula em Visão Computacional (Disciplina Pós-Graduação)**. Universidade de São Paulo, Escola de Engenharia de São Carlos, Departamento de Engenharia Elétrica - Laboratório de Visão Computacional (USP/EESC/SEL/LAVI) - São Carlos, São Paulo - Brasil.

GONZALEZ, R. C. e WOODS, R. E. (2007). **Digital Image Processing**. 3.ed. New Jersey, EUA: Prentice-Hall Inc. ISBN 013168728X.

GOZDOWSK, D. e SAMBORSKI, S. (2007). **Book Reviews - Precision Agriculture**. Journal: Communications in Biometry and Crop Science. v. 2, p. 90-94.

GRITAI, A., RAO, C., SHAH, M. e SHEIKH, Y. (2009). **Matching trajectories of anatomical landmarks under viewpoint, anthropometric and temporal transforms**. International Journal of Computer Vision, v 84, n 3, p 325-343.

GUIMARÃES, A. A. (2003). **Análise da norma ISO11783 e sua utilização na implementação do barramento do implemento de um monitor de semeadora**. . Dissertação (Mestrado). Escola Politécnica, Universidade de São Paulo. p. 98. São Paulo, SP Brasil. .

HAN, Z., ZHAO, Y. e YANG, J. (2010). **Detection of embryo based on independent components for kernel RGB images in maize**. Nongye Gongcheng Xuebao/Transactions of the Chinese Society of Agricultural Engineering, v 26, n 3, p 222-226, March 2010.

HASTIE, T., TIBSHIRANI, R. e FRIEDMAN, J. (2003). **The elements of statistical learning**. 1.ed. New York, EUA: Springer. ISBN 0-387-95284-5.

HAYKIN, S. (2008). **Neural Networks and Learning Machines**. 3.ed. McMaster University, Canada: Prentice-Hall. ISBN 0131471392.

HECKBERT, P. (1990). **A seed fill algorithm**. Graphics Gems, Andrew Glassner, ed., Academic Press, Boston, 1990, pp. 275-277 721-722.

HELMERT, F. R. (1875). **Über die Bestimmung des wahrscheinlichen Fehlers aus einer endlichen Anzahl wahrer Beobachtungsfehler.** Z. Math. Phys. 20, 300-3.

HELMERT, F. R. (1876a). **Über die Wahrscheinlichkeit der Potenzsummen der Beobachtungsfehler und über einige damit in Zusammenhang stehende Fragen.** Z. Math. Phys. 21, 192-218.

HELMERT, F. R. (1876b). **Die Genauigkeit der Formel von Peters zur Berechnung des wahrscheinlichen Beobachtungsfehlers director Beobachtungen gleicher Genauigkeit** Astron. Nachr. 88, 113-32.

HERATH, D., KODAGODA, S. e DISSANAYAKE, G. (2008). **New framework for simultaneous localization and mapping: Multi map SLAM.** Proceedings of IEEE International Conference on Robotics and Automation, p 1892-1897, 2008 - ICRA'2008.

HILL, P. R. (2002). **Wavelet based texture analysis and segmentation for image retrieval and fusion.** Tese (Doutorado) — University of Bristol, Bristol, Inglaterra.

IGEL, C. e HÜSKEN, M. (2000). **Improving the Rprop learning algorithm.** In: Proceedings of the Second International ICSC Symposium on Neural Computation (ICSC Academic Press), p. 115–121.

JAFARI, A., MOHTASEBI, H., EGHBALI, J. e OMID, M. (2004). **Color feature extraction by means discriminant analysis for weed segmentation.** In: (Proceedings) ASAE ANNUAL INTERNATIONAL MEETING, 2004, Ottawa, Ontário, Canadá.

JANESICK, J. (2001). **Scientific charge-coupled devices.** SPIE Press. p. 4. ISBN 9780819436986.

JANESICK, J. (2002). **Dueling detectors. CCD or CMOS?** SPIE OE Magazine, February, 30–33.

JIANG, G. e ZHAO, C. (2010). **A vision system based crop rows for agricultural mobile robot.** 2010 International Conference on Computer Application and System Modeling (ICCASM/IEEE) - 978-1-4244-7237-6.

JIANYUAN, D. e CHONGYANG, H. (2009). **3D fast level set image segmentation based on Chan-Vese model.** 3rd International Conference on Bioinformatics and Biomedical Engineering, iCBBE 2009.

JING, F., LIZ, M., ZHAN, H. J. e B., Z. (2003). **Unsupervised image segmentation using local homogeneity analysis.** In: (Proceedings) INTERNATIONAL SYMPOSIUM ON CIRCUITS AND SYSTEMS, IEEE, v. 2, p. 456–459. ISBN 0-7803-7761-3.

KAESS, M., DELLAERT, F. e NI, K. (2009). **Flow separation for fast and robust stereo odometry.** Proceedings - IEEE International Conference on Robotics and Automation (ICRA 2009), p. 3539-3544.

- KENJI, K. e NAOKI, S. (2009). **Obstacle detection based on occupancy grid maps from virtual disparity image**. ICCAS-SICE 2009 - ICROS-SICE International Joint Conference 2009, p. 268.
- KISE, M. e ZHANG, Q. (2008). **Development of a stereovision sensing system for 3D crop row structure mapping and tractor guidance**. Biosystems Engineering, v 101, n 2, p 191-198, October 2008.
- KISE, M., ZHANG, Q. e ROVIRA, M. (2005). **A stereovision-based crop row detection method for tractor-automated guidance**. Biosystems Engineering, v 90, n 4, p 357-367, April 2005.
- KORN, F., PAGEL, B.-U. e FALOUTSOS, C. (2001). **On the 'dimensionality curse' and the 'self-similarity blessing'**. IEEE Transactions on Knowledge and Data Engineering, IEEE Educational Activities Department, Piscataway, NJ, USA, v. 13, n. 1, p. 96–111, 2001. ISSN 1041-4347.
- LAND, A. H. e DOIG, A. G. (1960). **An automatic method of solving discrete programming problems**. Econometrica 28 (3): pp. 497-520. doi:10.2307/1910129.
- LEE, G., WILFRED, O., YEOM, J. e CHO, B. (2009). **Extraction of Key Postures using Shape Contexts**. International Conference on Advanced Communication Technology, ICACT, v 2, p 1311-1314.
- LEKHAL, F., NASRI, M. e HITMY, M. (2008). **A general scheme for texture images classification, by a genetic method, for artificial vision quality control**. Proceedings of the 2008 International Conference on Image Processing, Computer Vision, and Pattern Recognition, IPCV 2008, p 435-439.
- LEVINE, M. D. (1985). **Vision in man and machine**. 1. ed. Nova Iorque, EUA: McGraw-Hill, Inc. ISBN 0-07-037446-5.
- LI, J., RAO, X., YING, Y., MA, B. e GUO, J. (2009). **Background and external defects segmentation of navel orange based on mask and edge gray value compensation algorithm**. Transactions of the Chinese Society of Agricultural Engineering, v 25, n 12, p 133-137.
- LI, Z., HONG, T., WANG, W. e SONG, S. (2007). **Automatic detection of growing orange fruits by machine vision**. ASABE Annual International Meeting, Technical Papers, v 8 BOOK, 2007.
- LIMING, X. e TANCHAO, Z. (2010). **Automated strawberry grading system based on image processing**. Computers and Electronics in Agriculture, v 71, n SUPPL. 1, p S32-S39.
- LIU, Y., ZHANG, J., RICHARDS, M., PHAM, B., ROE, P. e CLARKE, A. (2009). **Towards continuous surveillance of fruit flies using sensor networks and machine vision**. Proceedings - 5th International Conference on Wireless Communications, Networking and Mobile Computing, WiCOM 2009.

LOPEZ, J. J., AGUILERA, E. e COBOS, M. (2009). **Defect detection and classification in citrus using computer vision**. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), v 5864 LNCS, n PART 2, p 11-18, 2009, Neural Information Processing - 16th International Conference, ICONIP 2009.

LÜ, K., CAI, J., ZHAO, J., WANG, F. e TANG, M. (2010). **Real-time recognition of citrus on trees in natural scene**. Transactions of the Chinese Society of Agricultural Machinery, v 41, n 2, p 185-188+170.

LUCCHESI, L. e MITRA, S. K. (2001). **A new method for color image equalization**. IEEE International Conference on Image Processing, v 1, p 133-136.

LUCCHESI, L. e MITRA, S. K. (2004). **A new class of chromatic filters for color image processing. Theory and applications**. IEEE Transactions on Image Processing, v 13, n 4, p 534-548.

LULIO, L. C., MEGDA, P. T., TRONCO, M. L., PORTO, A. J. V., CAURIN, G. A. P. e BECKER, M. (2010b). **Controle de Trajetórias para Robôs Móveis utilizando Lógica Fuzzy**. Congresso Nacional de Engenharia Mecânica (VI CONEM). Campina Grande, PB – Brasil.

LULIO, L. C., TRONCO, M. L. e PORTO, A. J. V. (2010a). **Omnidirectional Vision System for Agricultural Mobile Robot Navigation**. Congresso Nacional de Engenharia Mecânica (VI CONEM). Campina Grande, PB – Brasil.

MACARTHUR, D. K., SCHUELLER, J. K., LEE, W., CRANE, C. e PARSONS, L. (2006). **Remotely-piloted helicopter citrus yield map estimation**. ASABE Annual International Meeting, 2006.

MALYSZCO, D. e STEPANIUK, J. (2009). **Rough entropy based K-means clustering**. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), v 5908 LNAI, p 406-413, 2009, Rough Sets, Fuzzy Sets, Data Mining and Granular Computing - 12th International Conference, RSFDGrC 2009.

MALYSZCO, D. e STEPANIUK, J. (2010). **Adaptive multilevel rough entropy evolutionary thresholding**. Information Sciences, v 180, n 7, p 1138-1158, April 1, 2010.

MATHWORKS, T. (2010). **MATLAB R2010a**. Version 7.10.0.499 - The Language of technical computing.

McCULLOCH, W. e PITTS, W. (1943). **A logical calculus of the ideas immanent in nervous activity**. Bulletin of Mathematical Biophysics, 7:115 - 133.

MENEGATTI, L. A. A. e MOLIN, J. P. (2004). **A cana e a Agricultura de Precisão**. IDEA NEWS 4.

MEURIE, C., RUCHEK, Y., COHEN, A. e MARAIS, J. (2010). **A hybrid and adaptive segmentation method using color and texture information**. Proceedings of SPIE - The

International Society for Optical Engineering, v 7538, 2010, Proceedings of SPIE-IS and T Electronic Imaging - Image Processing: Machine Vision Applications III.

MEWADA, H. e PATNAIK, S. (2009). **Variable Kernel based Chan-Vese model for image segmentation**. Proceedings of INDICON 2009 - An IEEE India Council Conference.

MORIMOTO, T., TAKEUCHI, T., MIYATA, H. e HASHIMOTO, Y. (2000). **Pattern recognition of fruit shape based on the concept of chaos and neural networks**. Computers and Electronics in Agriculture, v 26, n 2, p 171-186.

NAKAMURA, T. e ISHIGURO, H. (2002). **Automatic 2-D map construction using special catadioptric sensor**. Proceeding of the IEEE Int. Conference on Intelligent Robots and Systems, Oct. 2002.

NANDEDKAR, A. (2010). **Supervised colour image segmentation using granular reflex fuzzy min-max neural network**. Proceedings of SPIE - The International Society for Optical Engineering, v 7546, 2010, 2nd International Conference on Digital Image Processing.

NIKOLAU, M. (2001). **Model predictive controllers: A critical synthesis of theory and industrial needs**. Advances in Chemical Engineering, Academic Press, 2001, Volume 26, Pages 131-204.

NISSEN, H. J. e HENNINGER, G. (2008). **ISOBUS Status of Standardization and Implementation**. Workshop ISOBUS Brasil. *Nissen is the Supervisor Systems Engineering of John Deere AMS Europe.* .

ORLANDO, R. C., PINTO, F. A. C. e QUEIROZ, D. M. (2004). **Identificação automática de plantas daninhas e milho utilizando técnicas de visão artificial**. In: CONGRESSO BRASILEIRO DE AGRICULTURA DE PRECISÃO, Piracicaba, São Paulo, Brasil.

PAALANEN, P., KAMARAINEN, J., LLONEN, J. e KALVIAINEN, H. (2006). **Feature representation and discrimination based on gaussian mixture model probability densities – practices and algorithms**. Pattern Recognition, v. 39, p.1346–1358.

PARK, S. B., LEE, J. W. e KIM, S. K. (2004). **Content-based image classification using a neural network**. Pattern Recognition Letters, v. 3, n. 25, p. 287–300, fev. ISSN 0167-8655.

PEDRO, L. M. (2008a). **Fixação de objetos por garras de robôs: aplicações de redes neurais e proposta de auto-aprendizagem para os casos 2D e 3D**. (Dissertação Mestrado). Escola de Engenharia de São Carlos/Universidade de São Paulo (EESC/USP). São Carlos, São Paulo - Brasil.

PEDRO, L. M. e CAURIN, G. A. P. (2008c). **Algoritmo de Rápido Processamento Baseado em Redes Neurais Artificiais para Fixação de Objetos Desconhecidos por Garras Industriais**. Revista Minerva, v. 4, p. 1-14, 2008.

PEDRO, L. M., DIAS, A. L., MASSARO, L. C. e CAURIN, G. A. P. (2008b). **Dynamic Modelling and hardware-in-the-loop Simulation applied to a Mechatronic Project**. In: ABCM Symposium Series in Mechatronics, vol.3, 2008.

PEREIRA, R. R. D. (2008). **Protocolo ISO 11783: Procedimentos para comunicação serial de dados com o Controlador de Tarefas**. Dissertação (Mestrado) - Escola de Engenharia de São Carlos, Universidade de São Paulo. 188p. São Carlos, SP - Brasil.

PLEBE, A. e GRASSO, G. (2001). **Localization of spherical fruits for robotic harvesting**. Machine Vision and Applications, v. 13, p. 70–79.

PORTILLO, J., CABANES, I., MARCOS, M. e ESTEVEZ, E. (2006). **CANopen Network for μ controller-based Real Time Distributed Control Systems**. IEEE, p. 4644-4649. .

PYDIPATI, R., F., B. T. e LEE, W. S. (2005). **Statistical and neural network classifiers for citrus disease detection using machine vision**. Transactions of the American Society of Agricultural Engineers, v 48, n 5, p 2007-2014.

QINGZHONG, L. (2002). **Computer vision based system for apple surface defect detection**. Computer and Electronics in Agriculture 36, 215–223.

RAO, X., Q. e YING, Y. (2009). **Grading a fruit by it's surface color**. Journal of Zhejiang University (Engineering Science), v 43, n 5, p 869-871.

RAO, X. Q. e YING, Y. (2009). **Grading a fruit by it's surface color**. Proceedings of SPIE - The International Society for Optical Engineering, v 5996 (Optical Sensors and Sensing Systems for Natural Resources and Food Safety and Quality).

RATLE, F., CAMPS, V. e WESTON, J. (2010). **Semisupervised neural networks for efficient hyperspectral image classification**. IEEE Transactions on Geoscience and Remote Sensing, v 48, n 5, p 2271-2282, May 2010.

REGUNATHAN, M. e LEE, W. S. (2005). **Citrus fruit identification and size determination using machine vision and ultrasonic sensors**. In: ASAE/CSAE ANNUAL INTERNATIONAL MEETING (Proceedings), Tampa, Flórida, EUA.

RICHTER, E., LINDNER, P., TAKAGI, K. e ISAGAI, A. (2009). **Advanced occupancy grid techniques for lidar based object detection and tracking**. IEEE Conference on Intelligent Transportation Systems, Proceedings (ITSC 2009), p. 450-454.

RONCATTI, M. A. (2008). **Avaliação de métodos ótimos e subótimos de seleção de características de texturas em imagens**. (Dissertação de Mestrado). Instituto de Ciências Matemáticas e Computação/Universidade de São Paulo (ICMC/USP). São Carlos, São Paulo - Brasil.

ROSSELOT, D., AULL, M. e HALL, E. (2010). **Predictive vision from stereo video: Robust object detection for autonomous navigation using the unscented Kalman filter on streaming stereo images**. Proceedings of SPIE - The International Society for Optical Engineering, v 7539, 2010, Proceedings of SPIE-IS and T Electronic Imaging - Intelligent Robots and Computer Vision XXVII: Algorithms and Techniques.

RUSSELL, S. e NORVIG, P. (2009). **Artificial Intelligence: A Modern Approach**. Prentice-Hall, 3rd edition - USA. ISBN: 0136042597.

SANTOS, D. P. (2006). **Seleção de características: abordagem via redes neurais aplicada a segmentação de imagens**. (Dissertação de Mestrado) - Instituto de Ciências Matemáticas e de Computação/Universidade de São Paulo (ICMC/USP). São Carlos, São Paulo - Brasil.

SAVAKAR, D. e ANAMI, B. S. (2009). **Recognition and classification of food grains, fruits and flowers using machine vision**. International Journal of Food Engineering, v 5, n 4, 2009.

SCATENA, J. M. (2003). **Navegação de robôs móveis baseados em computação reconfigurável**. (Dissertação Mestrado). Instituto de Ciências Matemáticas e de Computação/Universidade de São Paulo (ICMC/USP). São Carlos, São Paulo - Brasil.

SEBE, N., COHEN, I., GARG, A. e HUANG, T. S. (2005). **Machine Learning in Computer Vision (Computational Imaging and Vision)**. Springer - 1. ed., 242 p. - ISBN: 1402032745.

SERFLING, R. (2004). **Depth Functions in Nonparametric Multivariate Inference**. DIMACS Series in Discrete Mathematics and Theoretical Computer Science.

SHI, Y., DINOV, I. e MIO, W. (2010). **A computational model of multidimensional shape**. International Journal of Computer Vision, v 89, n 1, p 69-83.

SHIIGI, T., KONDO, N., KURITA, M. e KAMATA, J. (2008). **Strawberry harvesting robot for fruits grown on table top culture**. American Society of Agricultural and Biological Engineers Annual International Meeting 2008, v 5, p 3139-3148.

SHOU, Z., WANG, Q., GUI, J. e WANG, Y. (2009). **Fruit shape detection by optimizing Chan-Vese model**. Proceedings of SPIE - The International Society for Optical Engineering, v 7495, 2009, MIPPR 2009 - Automatic Target Recognition and Image Analysis.

SI, Y., LIU, G. e GAO, R. (2009). **Segmentation algorithm for green apples recognition based on k-means algorithm**. Transactions of the Chinese Society of Agricultural Machinery, v 40, n SUPPL. 1, p 100-104.

SIBLEY, G., MATTHIES, L. e SUKHATME, G. (2008). **A sliding window filter for incremental SLAM**. Lecture Notes in Electrical Engineering, v 8 LNEE, p 103-112, 2008, Unifying Perspectives in Computational and Robot Vision.

SIEGWART, R., NOURNAKSH, I. e SCARAMUZZA, D. (2011). **Introduction to Autonomous Mobile Robots**. The MIT Press - Massachusetts Institute of Technology Cambridge. ISBN: 0-262-19502-X.

SILVA, L. L. (2008). **Classificação e caracterização de ambientes para navegação de robôs móveis baseada em mapas**. Dissertação (Mestrado) - Instituto de Biociências, Letras e Ciências Exatas - Universidade Estadual Paulista "Julio de Mesquita Filho" (IBILCE/UNESP).

SIMION, G., GUI, V. e OTESTEANU, M. (2009). **Hand posture recognition using compositional techniques**. Proceedings - 2009 5th International Symposium on Applied Computational Intelligence and Informatics, SACI 2009, p 435-439.

SIRAKOV, N. e USHKALA, K. (2009). **An integral active contour model for convex hull and boundary extraction**. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), v 5876 LNCS, n PART 2, p 1031-1040, 2009, Advances in Visual Computing - 5th International Symposium, ISVC 2009.

SLAUGHTER, D. C., STAAB, E. S., GILES, D. K., DOWNEY, D. e ZHANG, Y. (2009). **Precision automated weed control using hyperspectral vision identification and heated oil**. American Society of Agricultural and Biological Engineers Annual International Meeting 2009, v 6, p 3567-3583.

SMITH, R., SELF, M. e CHEESEMAN, P. (1988). **A stochastic map for uncertain spatial relationships**. In: Fourth International Symposium of Robotics Research, Santa Clara (USA): MIT Press, 1988, p. 467–474.

SMITH, R., SELF, M. e CHEESEMAN, P. (1990). **Estimating uncertain spatial relationships in robotics**. Autonomous robot vehicles, New York (USA): Springer-Verlag, p. 167–193.

SOUZA, R. V. (2002). **CAN (Controller Area Network): Uma abordagem para automação e controle na área agrícola**. . Dissertação (Mestrado). Escola de Engenharia de São Carlos, Universidade de São Paulo. p. 84. São Carlos, SP Brasil. .

SOUZA, R. V. (2007). **Robô agrícola móvel (RAM): uma arquitetura baseada em comportamentos hierárquicos e difusos para sistemas autônomos de guiagem e navegação**. Tese (Doutorado). Escola de Engenharia de São Carlos, Universidade de São Paulo (EESC/USP).

STEWART, B. L., TIAN, L. F. e TANG, L. (2004). **Reduced-dimension clustering for vegetation segmentation**. Transactions of the ASAE, v. 47, p. 609–616.

STORY, D., KACRIA, M., KUBOTA, C., AKOGLU, A. e AN, L. (2010). **Lettuce calcium deficiency detection with machine vision computed plant features in controlled environments**. Computers and Electronics in Agriculture.

SUVINEN, A. e SAARILAHTI, M. (2006). **Measuring the mobility parameters of forwarders using GPS and CAN bus techniques**. Journal of Terramechanics, v. 43, p. 237-252.

TANG, L., TIAN, L. e STEWARD, B. L. (2003). **Classification of broadleaf and grass weeds using gabor wavelets and an artificial neural network**. Transactions of the ASAE, v. 46, p. 1247–1254.

TANG, L., TIAN, L. F. e STEWARD, B. L. (2005). **At early growth stages: Machine vision-based automated corn plant population and spacing measurement**. Engineering and Technology for Sustainable World, v 12, n 7, p 15.

TANGERINO, G. T. (2009). **Sistemas de sensoriamento embarcado para uso em controle de aplicações agrícolas à taxa variável** Dissertação (Mestrado) - Escola de Engenharia de São Carlos, Universidade de São Paulo. 102p. São Carlos, SP - Brasil.

THRUN, S. (2002). **Robotic Mapping: a survey**. Technical report - CMU-CS-02-111, School of Computer Science - Carnegie Mellon University.

THRUN, S., BURGARD, W. e FOX, D. (2005). **Probabilistic Robotics (Intelligent Robotics and Autonomous Agents)**. The MIT Press, 667 p. - ISBN: 0262201629.

THRUN, S., LIU, Y., KOLLER, D., NG, A. Y., GHAHRAMANI, Z. e DURRANT-WHYTE, H. (2004b). **Simultaneous localization and mapping with sparse extended information filters**. The International Journal of Robotics Research, London (UK), v. 23, n. 7/8, p. 693–716, July-August, 2004.

THRUN, S., THAYER, S., WHITTAKER, W., BAKER, C. e BURGARD, W. (2004a). **Autonomous exploration and mapping of abandoned mines**. IEEE Robotics and Automation Magazine, New York (USA), n. 4, p. 79–91.

TRONCO, M. L. (1999). **Sistema de reconhecimento de imagens baseado no modelo GSN de rede neural**. (Tese Doutorado). Escola de Engenharia de São Carlos/Universidade de São Paulo (EESC/USP). São Carlos, São Paulo - Brasil.

VERMA, N., GUPTA, P., AGRAWAL, P., HANMANDLU, M., VASIKARLA, S. e CUI, Y. (2009). **Medical image segmentation using improved mountain clustering approach**. ITNG 2009 - 6th International Conference on Information Technology: New Generations, p 1307-1312.

WAN, L. (2009). **Parts-based 2D shape decomposition by convex hull**. IEEE International Conference on Shape Modeling and Applications, SMI 2009, p 89-95.

WANG, J., ZHAO, D., JI, W. e ZHANG, C. (2009). **Apple fruit recognition based on support vector machine using in harvesting robot**. Transactions of the Chinese Society of Agricultural Machinery, v 40, n 1, p 148-151+147.

WANG, X., HUANG, D. e XU, H. (2010). **An efficient local Chan-Vese model for image segmentation**. Pattern Recognition, v 43, n 3, p 603-618, March 2010.

WEI, J., WANG, J. J., ZHAO, D. e TU, J. J. (2009). **Application of support vector machine to apple recognition using in apple harvesting robot**. IEEE International Conference on Information and Automation, ICIA 2009, p 1110-1115.

WIENER, N. (1942). **Extrapolation, Interpolation and Smoothing of Stationary Time Series**. The MIT Press, Cambridge, USA.

WOLF, J., BURGARD, W. e BURKHARDT, H. (2005). **Robust vision-based localization by combining an image-retrieval system with Monte Carlo localization**. IEEE Transactions on Robotics, New York (USA), v. 21, n. 2, p. 208–216, April, 2005.

WU, E., ZHOU, W., DAI, G. e WANG, Q. (2009). **Monocular vision SLAM for large scale outdoor environment**. IEEE International Conference on Mechatronics and Automation, ICMA 2009, p 2037-2041.

WU, M., HUANG, F., WANG, L. e SUN, J. (2009). **Cooperative multi-robot monocular-SLAM using salient landmarks**. Proceedings - 2009 International Asia Conference on Informatics in Control, Automation, and Robotics (CAR 2009), p 151-155.

WURM, K., STACHNISS, C. e GRISSETTI, G. (2010). **Bridging the gap between feature- and grid-based SLAM**. Robotics and Autonomous Systems, v 58, n 2, p. 140-148.

XIA, Y. e FENG, D. (2010). **A general image segmentation model and its application**. Proceedings of the 5th International Conference on Image and Graphics, ICIG 2009, p 227-231.

XUE, J. e JU, W. (2010). **Vision-based guidance line detection in row crop fields**. International Conference on Intelligent Computation Technology and Automation, ICICTA 2010, v 3, p 1140-1143.

YAGI, Y., NISHIZAWA, Y. e YACHIDA, M. (1995). **Map-based navigation for a mobile robot with omnidirectional image sensor COPIS**. IEEE Transactions on Robotics and Automation, vol. 11, n. 5, Oct. 1995.

YE, Z., MOHAMADIAN, H. e YE, Y. (2009). **Intuitive visualization and objective metric evaluation for software programming based advanced image segmentation**. WRI World Congress on Software Engineering, WCSE 2009, v 4, p 38-42.

YIN, J., MAO, H., HU, Y., WANG, X. e CHEN, S. (2008). **Segmentation methods of fruit image and comparative experiments**. Proceedings - International Conference on Computer Science and Software Engineering, CSSE 2008, v 1, p 1098-1102.

YOO, T. S. (2004). **Insight into Images: Principles and Practice for Segmentation, Registration, and Image Analysis**. AK Peters; 1 edition. ISBN: 1568812175.

ZADEH, L. A. (1965). **Fuzzy sets: information and control**. V. 8: 338-353.1965.

ZHANG, F., XU, G., ZHANG, Y., CHENG, Y., WANG, J. e TIAN, J. (2009). **Color image segmentation using combined information of color and texture**. Proceedings of the 2009 Chinese Conference on Pattern Recognition, CCPR 2009, and the 1st CJK Joint Workshop on Pattern Recognition, CJKPR, p 290-293.

ZHANG, F., YING, Y., SHEN, C., JIANG, H. e ZHANG, Q. (2005). **Stereovision-based 3D field recognition for automatic guidance system of off-road vehicle**. Proceedings of SPIE - The International Society for Optical Engineering, v 6000, 2005, Two- and Three-Dimensional Methods for Inspection and Metrology III.

ZHANG, H. (2008). **Image segmentation evaluation: A survey of unsupervised methods**. Computer Vision and Image Understanding, v 110, n 2, p 260-280.

ZHANG, M., ZHANG, L. e CHENG, H. (2010). **A neutrosophic approach to image segmentation based on watershed method**. Signal Processing, v 90, n 5, p 1510-1517.

ZHANG, N., WANG, M. e WANG, N. (2002). **Precision agriculture - a worldwide overview**. Computers and Electronics in Agriculture. v. 36, ed. 2-3, p. 113-132.

ZHANG, X., RAD, A. e WONG, Y. (2008). **A virtual range finder based on monocular vision system in simultaneous localization and mapping**. IFAC Proceedings Volumes (IFAC-PapersOnline), v 17, n 1, 2008, Proceedings of the 17th World Congress, International Federation of Automatic Control.

ZHANG, Y.-j. (2006). **Advances in Image And Video Segmentation**. IRM Press. ISBN: 1591407532.

ZHAO, J., LIU, S., ZOU, X., SHI, J. e YIN, X. (2008). **Recognition of defect Chinese dates by machine vision and support vector machine**. Transactions of the Chinese Society of Agricultural Machinery, v 39, n 3, p 113-115+147.

ZORPETTE, G. (1998). **Seeing the Light: CMOS Image Sensors Are Poised to Take on CCDs**. Scientific American. May 1998 ed.