

UNIVERSIDADE FEDERAL DO RIO DE JANEIRO

ROGÉRIO SERAFINI DOS SANTOS

CLASSIFICAÇÃO DE VEÍCULO EM SISTEMA VISUAL DE MONITORAÇÃO DE
TRÂNSITO

Rio de Janeiro

2008

S237 Santos, Rogério Serafini dos.

Classificação de veículo em sistema visual de monitoração de trânsito / Rogério Serafini dos Santos. - Rio de Janeiro, 2008. 133 f.: il.

Dissertação (Mestrado em Informática)–Universidade Federal do Rio de Janeiro, Instituto de Matemática, Núcleo de Computação Eletrônica, 2008.

Orientador: Antonio Carlos Gay Thomé

1. Reconhecimento de Padrões - Teses. 2. Reconhecimento de Veículo - Teses. 3. Redes Neurais – Teses. 4. Processamento de Imagens – Teses. I. Antonio Carlos Gay Thomé (Orient.). II. Universidade Federal do Rio de Janeiro. Instituto de Matemática. Núcleo de Computação Eletrônica. III. Título.

CDD:

ROGÉRIO SERAFINI DOS SANTOS

CLASSIFICAÇÃO DE VEÍCULO EM SISTEMA VISUAL DE MONITORAÇÃO DE
TRÂNSITO

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Informática, Instituto de Matemática, Núcleo de Computação Eletrônica, Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Mestre em Informática.

Orientador: Antonio Carlos Gay Thomé

Rio de Janeiro

2008

Rogério Serafini dos Santos

CLASSIFICAÇÃO DE VEÍCULO EM SISTEMA VISUAL DE MONITORAÇÃO DE
TRÂNSITO

Dissertação de Mestrado apresentada ao Programa Pós-Graduação em Informática, Instituto de Matemática, Núcleo de Computação Eletrônica, Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Mestre em Informática.

Aprovada em: 24 de junho de 2008.

Antonio Carlos Gay Thomé – Ph.D. IM/NCE – Orientador

Adriano Joaquim de Oliveira Cruz – Ph.D. NCE

Josefino Cabral Melo Lima – Ph.D. IM/DCC

Ricardo Cordeiro Farias – Dr. Sc. COPPE

AGRADECIMENTOS

Agradeço.

Em primeiro lugar a Deus, o Arquiteto do Universo.

À minha esposa, Carla, pelo amor, confiança e coragem em mudar o rumo da nossa Vida. Com muito carinho e força ela manteve nossos pequenos filhos, Miguel e Marina, como uma família feliz e unida durante a minha ausência temporária e distante.

À minha Mãe, por seu amor e carinho incondicionais.

Ao meu orientador, professor Thomé, pela sua dedicação e paciência ao longo da dissertação, e também pela experiência profissional no Projeto Kapta.

Aos professores do Mestrado, não somente pelos conhecimentos transmitidos, mas também pelos seus conselhos e experiências.

Ao Marcelo, o “Peruano”, um colega dedicado durante as aulas e desafios do Mestrado, mas também um amigo eterno para a Vida.

Aos órgãos patrocinadores pela grande oportunidade, suporte financeiro e recursos brindados para realizar esta pesquisa.

RESUMO

SANTOS, Rogério Serafini dos. **Classificação de veículo em sistema visual de monitoração de trânsito**. Rio de Janeiro, 2008. Dissertação (Mestrado em Informática)- Instituto de Matemática, Núcleo de Computação Eletrônica, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2008.

No Brasil a frota veicular é de 45 milhões e a malha rodoviária de 1.6 milhões Km. Isso é consequência do crescimento populacional e também da maior oferta de veículos. Estes geram um tráfego intenso tanto em vias urbanas assim como em rodovias. Uma solução seria a construção de novas pistas, porém isso está se tornando inviável devido aos altos custos e ao pouco espaço disponível. Soluções contemporâneas enfatizam a melhoria do planejamento e controle para um uso mais eficiente da infra-estrutura já existente. Porém, a grande extensão da malha rodoviária, existente nas grandes cidades e nas rodovias estaduais ou federais, inviabiliza o emprego de uma vigilância presencial que seja abrangente, efetiva e em tempo integral. A monitoração por vídeo vem se tornando uma grande aliada do poder público, apesar de ainda ser realizada de forma manual na sua grande maioria. A automatização da monitoração será o próximo passo tecnológico. O objetivo deste trabalho é construir um modelo computacional capaz de realizar o reconhecimento da categoria e da cor predominante de um veículo a partir de sua imagem digital. No escopo deste trabalho os veículos são classificados 5 categorias, 11 subcategorias e 8 cores predominantes. Os classificadores utilizados são do tipo neural, treinados com algoritmo de retro-propagação do erro, com resultados alcançados em bateria de testes com 7 mil veículos está próximo aos 99% para categoria e 80% para cor.

ABSTRACT

SANTOS, Rogério Serafini dos. **Classificação de veículo em sistema visual de monitoração de trânsito**. Rio de Janeiro, 2008. Dissertação (Mestrado em Informática)- Instituto de Matemática, Núcleo de Computação Eletrônica, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2008.

In Brazil the vehicular fleet is of 45 million and the net of highways of 1.6 million Km. That is consequence of the population growth and also of the largest offer of vehicles. These generate a so much intense traffic in urban roads as well as in highways. A solution would be the construction of new tracks, however that is being turned unviable due to the high costs and to the little available space. Contemporary solutions already emphasize the improvement of the planning and control for a more efficient use of the infrastructure existent. However, the great extension net of highways, existent in the great cities and in the state highways or you federate, it makes unfeasible the employment of a present surveillance that is including, effective and in integral time. The surveillance for video is becoming a big one allied of the public power, in spite of still to be accomplished in a manual way in your great majority. The automation of the surveillance will be the next technological step. The objective of this work is to build a model in computer capable to accomplish the recognition of the category and of the predominant color of a vehicle starting from your digital image. In the mark of this work the vehicles are classified in 5 categories, 11 subcategories and 8 predominant colors. The used classifiers are of the type neural, with algorithm of retro-propagation of the error, with results reached in battery of tests with 7 thousand vehicles it is close to the 99% for category and 80% for color.

SUMÁRIO

1.	Introdução	17
1.1.	Objetivos.....	21
1.2.	Organização da dissertação	22
2.	ITS - Sistemas de Transportes Inteligentes	23
2.1.	Sistemas de ITS	23
2.2.	Pesquisa em ITS	25
2.3.	Processamento de Imagens em ITS	27
2.3.1.	Sistemas de monitoração de tráfego veicular	28
2.3.2.	Sistemas de Apoio ao Motorista.....	29
2.4.	Sensores	29
2.4.1.	Vídeo	30
2.4.2.	Laço Indutivo.....	30
2.4.3.	<i>Transponder</i>	31
2.4.4.	Infravermelho	31
2.4.5.	Microondas	32
2.4.6.	Ultra-som	32
3.	Técnicas para construção de um Sistema Visual de Monitoração.....	34
3.1.	Aquisição	35
3.1.1.	Visão de câmera.....	36
3.1.2.	Calibração de contexto	39
3.2.	Pré-processamento.....	43
3.2.1.	Redução de complexidade	43
3.2.2.	Correção de imagem.....	44
3.3.	Segmentação	44
3.3.1.	Estática Individual	45
3.3.2.	Estática Temporal	47
3.3.3.	Dinâmica Individual	49
3.4.	Representação e Descrição	50
3.5.	Reconhecimento e interpretação.....	51
3.5.1.	Classificação de Categoria.....	52
3.5.2.	Confirmação de presença.....	59
3.5.3.	Rastreamento de trajetória	61
3.5.4.	Interpretação de eventos	62
3.6.	Base de Conhecimento	63
4.	Proposta e avaliação de desempenho do modelo computacional.....	64
4.1.	Aquisição e segmentação.....	65
4.2.	Descritores de Categoria.....	68
4.2.1.	Geométricos.....	68
4.2.2.	Regionais	70
4.3.	Descritores de Cor	75
4.3.1.	Modelo RGB	76
4.3.2.	Modelo HSI	77
4.3.3.	Modelo YCbCr	78
4.3.4.	Modelo Lab	79
4.4.	Classificação Neural	80
4.5.	Estudo de Caso	82
4.5.1.	Categorias e Subcategorias.....	85
4.5.2.	Cores.....	88
4.6.	Configuração e Treinamento dos Classificadores	90

4.6.1.	Classificação de Categoria.....	93
4.6.2.	Classificação de Cor Predominante	95
4.7.	Testes e Resultados.....	96
4.7.1.	Classificação de Categoria.....	96
4.7.2.	Classificação de Subcategoria	99
4.7.3.	Classificação de Cor	101
5.	Análise e Conclusões.....	103
5.1.	Análise	103
5.2.	Conclusões.....	105
5.3.	Trabalhos futuros	106
6.	Apêndice 1 - Rotinas	107
7.	Apêndice 2 - Resultados de Classificação de Categoria.....	111
8.	Apêndice 3 - Resultados de Classificação de Cor	117
9.	Apêndice 4 – Redes Neurais.....	119
10.	Apêndice 5 – Visualizações em conjunto.....	122
11.	Referências	126

LISTA DE SIGLAS

AG	Algoritmos Genéticos
CET-RIO	Companhia de Engenharia de Tráfego – Rio de Janeiro
CNT	Confederação Nacional de Transporte
DENATRAN	Departamento Nacional de Trânsito
HSI	<i>Hue, Saturation, Intensity</i>
IEEE	<i>Institute of Electrical and Electronics Engineers</i>
ITS	<i>Intelligent Transportation Systems</i>
ITSS	<i>Intelligent Transportation Systems Society</i>
MLP	<i>Multi Layer Perceptron</i>
PCA	Análise de Componente Principal
RBF	<i>Radial basis function</i>
RGB	<i>Red, Green, Blue</i>
RNA	Redes Neurais Artificiais
SVM	<i>Support Vector Machines</i>
WTA	<i>Winner Takes All</i>

LISTA DE ILUSTRAÇÕES

Figura 1: Câmera de vídeo em monitoração de trânsito.....	17
Figura 2: Exemplos de câmeras da CET-Rio no trânsito.	19
Figura 3: Tipos de sensores incorporados no sistema VSAM.....	21
Figura 4: Veículo individual.....	21
Figura 5: Grupos de Sistemas Visuais.....	27
Figura 6: Modelo geral de análise da imagem.....	35
Figura 7: Ângulos de aquisição: horizontal θ e vertical ϕ	36
Figura 8: Veículos individuais e com imagem próxima em ângulos variados.....	37
Figura 9: Imagens em cruzamento de pistas.....	38
Figura 10: Imagens em pistas múltiplas.	38
Figura 11: Visão múltipla, traseira e variável em câmera a bordo de veículo.	38
Figura 12: Calibração de regiões incluídas e excluídas em cruzamento de trânsito.	40
Figura 13: Correspondência entre o plano da pista e o sistema global.....	40
Figura 14: Ferramenta de calibração por linhas.	40
Figura 15: Calibração de cruzamento múltiplo de vias urbanas.....	41
Figura 16: Calibração de regiões esperadas para detecção.....	41
Figura 17: Modelo de veículo com projeção de sombra, pára-brisa, teto e capô.	42
Figura 18: Ponto de fuga representado pela interseção das linhas projetadas ao infinito.	42
Figura 19: Calibração a partir de um par de linhas paralelas.	42
Figura 20: Ambientes variados com iluminação natural ou artificial.	44
Figura 21: Amostras para segmentação por cor.	46
Figura 22: Segmentação de veículos em monitoração.	46
Figura 23: Passos para segmentação estática.	46
Figura 24: Passos de segmentação temporal em monitoração.	48
Figura 25: Detecção de linhas principais de veículo.	49
Figura 26: Passos da segmentação dinâmica.	49
Figura 27: Passos da detecção de pista.....	50
Figura 28: Segmentação de ambiente 3D.	50
Figura 29: Modelos poligonais de sete lados.	53
Figura 30: Exemplos de veículos em visão lateral.	53
Figura 31: Diferença entre retângulo de candidato e modelo.....	54
Figura 32: Passos da construção de esqueleto morfológico estatístico.	54
Figura 33: Pistas com faixas múltiplas.....	55
Figura 34: Veículos em visão frontal, superior e distante.	56
Figura 35: Passos de reconhecimento de modelo de veículo.	57
Figura 36: Veículos em entrada de estacionamento.	57
Figura 37: Passos de classificação por modelo 3D.	58
Figura 38: Categorias em modelos 3D: carro, microônibus, van e van alta.....	58
Figura 39: Seqüência em curva de estacionamento.....	58
Figura 40: Categorias em modelos 3D de 12 medidas.	58
Figura 41: Veículos centralizados em imagem 128x128 e alinhamento de faróis.	59
Figura 42: Exemplos de imagens traseiras	59
Figura 43: Geração multi-resolução de regiões candidatas.	60
Figura 44: Imagens reconstruídas através da seleção de autovetores (av) específicos.	61
Figura 45: Imagens reconstruídas através da seleção automatizada de autovetores.	61
Figura 46: Diagrama de módulos funcionais.....	64
Figura 47: Imagem de veículo em via de acesso à UFRJ.....	64
Figura 48: Segmentação interativa de veículo.....	65

Figura 49: Aquisição de seqüência de 3 imagens.....	66
Figura 50: Veículo em trânsito.	66
Figura 51: Recorte retangular colorido com fundo colorido.	66
Figura 52: Tela de marcação manual de cor, categoria e contorno de veículo.....	68
Figura 53: Imagem Convexa.	69
Figura 54: Descritor cinza.	70
Figura 55: Descritor binário com fundo preto.	71
Figura 56: Descritor de Bordas com fundo preto.	71
Figura 57: Anulação do fundo cinza do recorte.	72
Figura 58: Descritor Silhueta.....	72
Figura 59: Exemplos de recortes originais.	73
Figura 60: Exemplos de redimensionamento sem relação de aspecto.	74
Figura 61: Exemplos de redimensionamento com relação de aspecto.	74
Figura 62: Recortes cinza redimensionados sem relação de aspecto.	74
Figura 63: Recortes cinza redimensionados com relação de aspecto.....	75
Figura 64: Recortes de outros descritores redimensionados sem proporção em 6,25%	75
Figura 65: Modelo RGB.	76
Figura 66: Modelo HSI baseado em planos triangulares.....	77
Figura 67: Componentes HSI referentes ao cubo RGB.....	78
Figura 68: Imagens de veículos nas pistas do Posto 3 de acesso à UFRJ.	83
Figura 69: Posição da câmera 1 (círculo) relativa ao laço magnético (retângulo na pista).....	83
Figura 70: Pistas com múltiplos veículos e oclusão entre eles.....	84
Figura 71: Mudança de escala devido a variação de distância da câmera.....	84
Figura 72: Deformação de veículos devido ao deslocamento de trajetória.	84
Figura 73: Variações de cor na mesma categoria.	85
Figura 74: Exemplos de categorias de veículos.....	86
Figura 75: Subcategorias de veículos, de acordo com os seus respectivos códigos.....	88
Figura 76: Imagens descartadas por falta de qualidade.	88
Figura 77: Distribuição de referência para cores.....	89
Figura 78: Distribuição de referência para cores no grupo Global.....	89
Figura 79: Gráfico da função de transferência logsig.....	90
Figura 80: Representação visual do conjunto da Pista 3, descritor cinza com fundo.....	104
Figura 81: Representação visual do conjunto da Pista 3, descritor cinza com fundo e normalização.....	105
Figura 82: Representação visual do conjunto da Pista 3, descritor cinza com fundo e PCA.	105
Figura 83: Neurônio artificial de McCulloch / Pitts.	120
Figura 84: Rede em modo de alimentação progressiva (feed-forward).	120
Figura 85: Representação visual do conjunto da Pista 3, descritor cinza sem fundo.	122
Figura 86: Representação visual do conjunto da Pista 3, descritor de bordas sem fundo.	122
Figura 87: Representação visual do conjunto da Pista 3, descritor binário sem fundo.	123
Figura 88 – Representação visual do conjunto da Pista 3, descritor silhueta sem fundo.	123
Figura 89: Representação visual do conjunto da Pista 3, descritor retângulo proporcional... ..	124
Figura 90- Representação visual do conjunto da Pista Global, descritor cinza sem fundo....	125

LISTA DE TABELAS

Tabela 1: Brasil: Totalização da frota de veículos e da população (x milhão).....	18
Tabela 2: Grupos de pesquisa internacionais.	27
Tabela 3: Grupos de pesquisa nacionais.....	27
Tabela 4: Acerto de classificação em SVM por PCA e <i>Eigen-vehicles</i>	52
Tabela 5: Matriz de confusão para modelos poligonais e minimização <i>metropoli</i>	53
Tabela 6: Matriz de confusão em proporções entre medidas de contorno.	54
Tabela 7: Matriz de confusão em retângulo de candidato e modelo.	54
Tabela 8: Matriz de confusão em esqueleto morfológico estatístico e modelos 3D.	55
Tabela 9: Matriz de confusão em <i>convex hull</i> e modelos 3D.....	55
Tabela 10: Desempenho por tamanho e linearidade, com 1 ou múltiplos quadros.....	56
Tabela 11: Desempenho de testes por classificador individual.....	56
Tabela 12: Desempenho de testes por combinação dos classificadores.....	56
Tabela 13: Taxa de acerto em veículos estacionados.....	57
Tabela 14: Taxa de acerto em veículos em entrada de estacionamento.	57
Tabela 15: Comparação de desempenho entre classificadores.....	59
Tabela 16: Redução de quantidade de exemplos de carros.	86
Tabela 17: Total de exemplos por pista e categoria (número de exemplos).	86
Tabela 18: Totais de subcategorias no grupo global	87
Tabela 19: Valores de alvos binários.....	92
Tabela 20: Conjunto de treinamento para categoria em 70% do total (número de exemplos). 94	
Tabela 21: Conjunto de teste para categoria em 30% do total (número de exemplos).	94
Tabela 22: Dimensões de teste	94
Tabela 23: Número de neurônios na saída de acordo com a classe.....	95
Tabela 24: Conjunto de treinamento para cores em 70% do total (número de exemplos).....	95
Tabela 25: Conjunto de teste para cores em 30% do total (número de exemplos).....	95
Tabela 26: Desempenho para categorias sem relação de aspecto e com fundo preto.	96
Tabela 27: Matriz de confusão para descritor cinza deformado para categoria na pista 3.....	96
Tabela 28: Desempenho para categorias sem relação de aspecto e com fundo original.	97
Tabela 29: Desempenho para categorias sem relação de aspecto.	97
Tabela 30: Desempenho para categorias com relação de aspecto e com fundo preto.....	97
Tabela 31: Desempenho para categorias com relação de aspecto e com fundo original.....	97
Tabela 32: Desempenho para categorias com relação de aspecto.	98
Tabela 33: Desempenho para categorias para descritor Geométrico.	98
Tabela 34: Matriz de confusão para descritor geométrico para categorias na pista 3.....	98
Tabela 35: Teste de Robustez da rede de cada Pista com dados das outras pistas, com fundo preto.	98
Tabela 36: Teste de Robustez da rede de cada Pista com dados das outras pistas, com fundo original.....	99
Tabela 37: Teste de Robustez da rede de cada Pista com dados das outras pistas, geométricos.	99
Tabela 38: Desempenho comparativo entre descritores para categorias no grupo global.....	99
Tabela 39: Matriz de confusão para descritor cinza sem fundo no grupo global.....	99
Tabela 40: Matriz de confusão para descritor geométrico no grupo global.	99
Tabela 41: Desempenho comparativo entre descritores para subcategorias no grupo pista 3.	100
Tabela 42: Matriz de confusão para descritor cinza em subcategorias da pista 3.....	100
Tabela 43: Desempenho comparativo entre descritores para subcategorias na grupo global. 100	
Tabela 44: Matriz de confusão para descritor geométrico em subcategorias do grupo global.	101

Tabela 45: Desempenho comparativo entre classificação direta e em cascata de subcategoria.	101
Tabela 46: Resultados para os descritores de cores com PCA.	102
Tabela 47: Resultados para os descritores de cores sem PCA.	102
Tabela 48: Acertos na classificação de categorias por outros autores.	103
Tabela 49: Resultados para o descritor cinza deformado para categoria em todas as pistas.	111
Tabela 50: Resultados para o descritor cinza deformado sem permutação na pista 3.	112
Tabela 51: Resultados para o descritor cinza deformado sem PCA na pista 3.	112
Tabela 52: Resultados para o descritor cinza deformado com redução do conjunto de treino na pista 3.	112
Tabela 53: Resultados do descritor cinza deformado com redução do conjunto de treino na pista global.	112
Tabela 54: Resultados do descritor cinza deformado com seleção de captura na pista 3.	112
Tabela 55: Resultados do descritor cinza não deformado na pista 3.	112
Tabela 56: Resultados para o descritor cinza deformado em subcategorias na pista 3.	113
Tabela 57: Resultados para o descritor cinza deformado em subcategorias na pista global.	113
Tabela 58: Resultados para o descritor binário deformado para categoria na pista 3.	113
Tabela 59: Resultados para o descritor binário deformado para categoria na pista global.	113
Tabela 60: Resultados do descritor binário não deformado na pista 3.	113
Tabela 61: Resultados para o descritor binário deformado em subcategorias na pista 3.	113
Tabela 62: Resultados para o descritor binário deformado em subcategorias na pista global.	114
Tabela 63: Resultados para o descritor de bordas deformado para categoria em todas na pista 3.	114
Tabela 64: Resultados do descritor de bordas não deformado na pista 3.	114
Tabela 65: Resultados do descritor de bordas não deformado no grupo global.	114
Tabela 66: Resultados para o descritor de bordas deformado em subcategorias na pista 3.	114
Tabela 67: Resultados para o descritor de bordas deformado em subcategorias na pista global.	114
Tabela 68: Resultados do descritor silhueta não deformado na pista 3.	115
Tabela 69: Resultados do descritor de retângulo não deformado na pista 3.	115
Tabela 70: Resultados para o descritor de retângulo deformado para categoria na pista global.	115
Tabela 71: Resultados para o descritor de retângulo deformado em subcategorias na pista 3.	115
Tabela 72: Resultados para o descritor de retângulo deformado em subcategorias na pista global.	115
Tabela 73: Resultados para o descritor geométrico para categorias na pista 3.	115
Tabela 74: Resultados para o descritor geométrico sem PCA em todas as pistas.	116
Tabela 75: Resultados para o descritor geométrico em subcategorias na pista 3 e global.	116
Tabela 76: Resultados para o descritor momentos deformado para categoria na pista 3.	116
Tabela 77: Resultados para o descritor momentos não deformado para categoria na pista global.	116
Tabela 78: Resultados para o descritor momentos deformado para subcategoria na pista 3.	116
Tabela 79: Resultados para o descritor momentos não deformado para subcategoria na pista global.	116
Tabela 80: Resultados para o descritor RGB para cores na pista 3 e global.	117
Tabela 81: Resultados para o descritor RGB para cores sem PCA na pista 3.	117
Tabela 82: Resultados para o descritor HSI para cores na pista 3 e global.	117
Tabela 83: Resultados para o descritor HSI para cores sem PCA na pista 3.	117

Tabela 84: Resultados para o descritor Lab para cores na pista 3 e global.	118
Tabela 85: Resultados para o descritor Lab para cores sem PCA na pista 3.....	118
Tabela 86: Resultados para o descritor Ycbcr para cores na pista 3 e global.....	118
Tabela 87: Resultados para o descritor Ycbcr para cores sem PCA na pista 3.	118

LISTA DE EQUAÇÕES

Equação 1: Componente H do modelo HSI.....	78
Equação 2: Componente S do modelo HSI.	78
Equação 3: Componente I do modelo HSI.	78
Equação 4: Componente Y do modelo YCbCr.	79
Equação 5: Componente Cb do modelo YCbCr.....	79
Equação 6: Componente Cr do modelo YCbCr.	79
Equação 7: Conversão de RGB para XYZ.	79
Equação 8: Componente L do modelo Lab.	80
Equação 9: Componentes A do modelo Lab.	80
Equação 10: Componentes B do modelo Lab.....	80
Equação 11: Transformação $f(w)$ para componentes A e B do modelo Lab.....	80
Equação 12: Normalização Z-score em <i>prestd</i>	81
Equação 13: Função de transferência logsig.	90

1. Introdução

Atualmente existe uma grande necessidade da engenharia de tráfego automotivo em conseguir informações rápidas e precisas sobre os veículos em trânsito. Isso com o objetivo de monitoração para detectar infrações, levantar estatísticas, realizar ações de controle e planejamento. Nas grandes metrópoles brasileiras a monitoração é realizada a partir de fotos capturadas nas vias públicas (Figura 1), porém na maioria ainda é feita manualmente e não em tempo real. Com um sistema automatizado o ganho é a imediata disponibilidade da informação para uso em outras aplicações.



Figura 1: Câmera de vídeo em monitoração de trânsito

No Brasil o crescimento da frota veicular vem aumentando consideravelmente. Segundo o [DENATRAN, 2006] no ano de 1990 a frota nacional era de 18 milhões de veículos e evoluiu para 45 milhões até 2006 (Tabela 1, Figura 1), sendo que 61% destes são automóveis. Isso representa um crescimento de aproximadamente 124% em 17 anos. Isso não é somente consequência do crescimento populacional, mas também da maior oferta de veículos, assim como pode ser evidenciado na queda da proporção entre população e frota (Tabela 1)(Gráfico 1). Essa grande quantidade de veículos gera um tráfego intenso tanto em vias urbanas assim como em vias rodoviárias. Outro motivo para a monitoração é a grande extensão da malha rodoviária que possui 1.6 milhões Km [CNT, 2006] de vias a serem mantidas e monitoradas. Esta grande extensão inviabiliza uma vigilância abrangente e presencial por agentes fiscalizadores pessoais.

Tabela 1: Brasil: Totalização da frota de veículos e da população (x milhão).
Fontes: Frota Nacional - Anfavea (anuário 2000) e Denatran (2006); População - IBGE (censo 1996 e 2000; estimativa 2006).

Ano	População	Frota	População / Frota
1950	52	0,4	130
1960	70	0,8	88
1970	93	3	29
1980	119	11	10
1990	147	18	8
1996	157	28	6
2000	170	30	6
2006	187	45	4

A solução ideal para o problema da expansão do tráfego seria a pavimentação de novas pistas para evitar os congestionamentos. Porém isso está se tornando inviável devido aos altos custos e também a inexistência de espaço físico disponível. Soluções contemporâneas enfatizam a melhoria do planejamento e controle para o uso mais eficiente da infra-estrutura já existente. A monitoração do fluxo de veículos e eventos relacionados pode ser utilizada para estes objetivos.

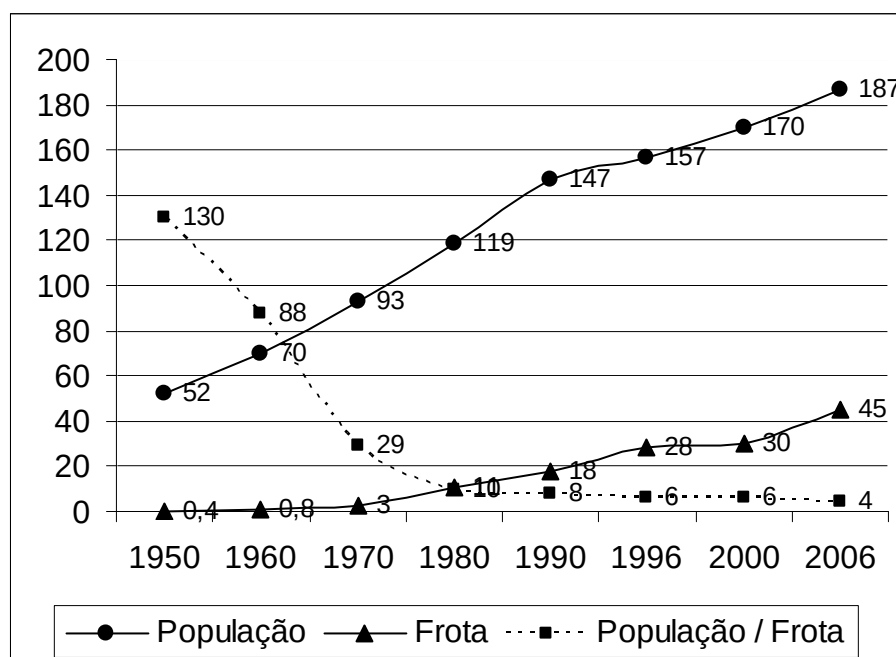


Gráfico 1: Evolução da proporção entre população e frota (x milhão).

A necessidade crescente de monitoração constante do fluxo de veículos vem motivando a intensificação da pesquisa e do emprego de tecnologia nesta área. Um sinal visível deste fato é a instalação cada vez mais freqüente de câmeras de monitoração nas vias, assim como de outros equipamentos eletrônicos. Um exemplo disso é a Prefeitura do Rio de Janeiro que disponibiliza na Internet boletins sobre as condições de trânsito [TRANSITORIO, 2006] obtidas através da análise visual nas câmeras da CET-Rio

[TRANSITO ONLINE, 2006] (Figura 2). Esses boletins fornecem o fluxo e a densidade do trânsito em determinados trechos (93 câmeras). Adicionalmente existem 32 radares fixos para aferição de velocidade em conjunto com câmeras para registro de veículos infratores. Também existem três radares móveis e 42 lombadas eletrônicas utilizados para este mesmo objetivo.



Figura 2: Exemplos de câmeras da CET-Rio no trânsito.

A implementação de monitoração resulta no uso de sensores de ambiente para aquisição de informações do tráfego. Alguns destes são: vídeo, laço indutivo, *transponder*, infravermelho, microondas e ultra-som. Uma técnica promissora é a monitoração através de processamento de imagens de vídeo, o qual pode fornecer informações tradicionais de trânsito como detecção, classificação, fluxo e velocidade, assim como outras adicionais como o rastreamento da trajetória de veículos. Além disso, câmeras oferecem economia em quantidade de equipamentos instalados. Isso porque apenas uma destas pode oferecer uma visão abrangente para várias pistas e veículos, inclusive em direções e sentidos distintos (cruzamento de trânsito). Também é vantagem a mobilidade das câmeras que podem ser facilmente trocadas de posição.

A riqueza visual adicionada por imagem fornece vários detalhes que não são encontrados em outros sensores. Um detalhe fornecido apenas por imagem é a cor do veículo. Assim como a possibilidade de identificação de veículo através da leitura do número da placa, que pode ser usado para controlar o acesso a locais restritos ou identificar veículos com situação irregular. Além disso, permite reconhecer um veículo através de um símbolo ou texto nele estampado como, por exemplo, um logotipo da empresa a qual ele pertence, e assim monitorar a passagem de tais veículos por um determinado local.

As informações adicionais obtidas pela trajetória do veículo incluem o acompanhamento de movimento (rastreamento) no espaço de visão da câmera, o que permite a medição de velocidade e aceleração. Além disso, a trajetória percorrida revela eventos como a troca de pistas, paradas, ultrapassagens e outras manobras que o motorista possa vir a realizar. Esses podem ser utilizados para detectar ações impróprias dos motoristas como

excesso de velocidade, ultrapassagem proibida, tráfego em sentido incorreto, avanço de sinal vermelho, retorno em local proibido ou entrada em zona restrita.

Também é possível medir parâmetros globais no trânsito, como a densidade do fluxo do trânsito e a medição de filas em pontos de parada temporária (semáforos). Além disso, também são usadas para detectar eventos interativos com o meio, assim como acidentes, atropelamentos, assaltos, mudanças de trajetória, alterações bruscas de velocidade, e condições anormais de visibilidade. Adicionalmente, podem ser detectados obstáculos indevidos na pista como, por exemplo, a presença de pedestres, animais e outros objetos estranhos.

Além da monitoração, outra aplicação especial do processamento de imagens no trânsito é a instalação de vídeo a bordo de veículos, tanto para o auxílio à navegação do motorista, quanto para condução autônoma (piloto automático). Nestes, além dos parâmetros como detecção e trajetória de veículos, o vídeo também pode ser usado para identificar regiões e objetos externos à pista, assim como vegetação, céu e construções. Assim como componentes adicionais como faixa de pedestres, e sinais de trânsito. Isso inclusive com a projeção e reconstrução do ambiente e seus objetos em espaço virtual 3D.

Pesquisas envolvendo processamento de imagens aplicado ao trânsito têm se desenvolvido muito nas últimas duas décadas devido à evolução computacional e tecnológica. Esse crescimento originou a área denominada Sistemas de Transportes Inteligentes (ITS – *Intelligent Transportation Systems*) que engloba várias tecnologias aplicadas ao controle e monitoração de trânsito.

No Brasil existem poucos trabalhos publicados em ITS. Porém existem vários grupos de pesquisa em Processamento de Imagens com condições técnicas de eventualmente desenvolver trabalhos nesta área. Entretanto em âmbito mundial vários podem ser encontrados em monitoração de trânsito e auxílio ao motorista.

Um exemplo de monitoração é o projeto *Video Surveillance and Monitoring (VSAM)* [COLLINS, 2000], no qual foi desenvolvido um sistema para vigilância e monitoramento por vídeo. Nesse sistema são monitoradas as interações de pessoas e veículos em um ambiente de muitas câmeras. Esses são classificados em categorias semânticas (humano, carro, e caminhão) através de análise de forma, cor e consistência temporal. Também é realizada classificação de atividade humana, como caminhar e correr. Geo-localizações destas

categorias são determinadas usando um conjunto estéreo de duas ou mais câmeras fotográficas com sobreposição ou interseção de visão. Estas tecnologias foram demonstradas em conjunto com sensores de calor, sensores de movimento e câmera *omni* (Figura 3).



Figura 3: Tipos de sensores incorporados no sistema VSAM
a) câmera fixa; b) sensor térmico; c) câmera em furgão; d) câmera em avião; e) câmera *omni*;

1.1.Objetivos

O objetivo deste trabalho inclui o estudo, concepção, implementação e avaliação de um modelo computacional para efetuar o reconhecimento da categoria e da cor predominante de um veículo a partir de sua imagem digital. Isso deve ser realizado através de técnicas de processamento de imagens e reconhecimento de padrões (Figura 4).



Figura 4: Veículo individual

Além do objetivo principal, existem outros módulos que precedem este e são implementados ou avaliados neste trabalho, como os de detecção, localização e extração de características do veículo. Também é objetivo deste trabalho realizar uma revisão abrangente do estado da arte em técnicas de processamento de imagens aplicadas na resolução de desafios relacionados à monitoração de trânsito e auxílio ao motorista, os quais são inseridos no contexto mais amplo de ITS.

A área de apoio ao motorista não é objetivo direto deste trabalho, porém a sua revisão é importante por incluir técnicas de processamentos de imagens que também podem ser aplicadas em monitoração de trânsito. Isso porque ambas as áreas analisam imagens de trânsito e apenas se diferenciam na forma de aquisição, a qual na monitoração é realizada a partir de câmera fixa e no auxílio ao motorista a partir de câmera móvel a bordo do veículo.

Para o desenvolvimento deste trabalho assumiu-se as seguintes restrições de contexto e escopo:

1. as imagens são adquiridas em ambiente externo com iluminação natural;
2. as imagens são adquiridas com veículo em movimento e sob visão frontal com pequeno índice de inclinação (Figura 4);
3. as imagens são adquiridas por quatro diferentes câmeras fixadas nas laterais da rodovia;
4. as imagens são em cores e tomadas com o emprego de um sensor de presença tipo laço magnético;
5. a aquisição e segmentação automática das imagens não fazem parte do escopo deste trabalho;
6. o modelo a ser proposto deve ser robusto quanto a pequenas variações de luminosidade (sombra, reflexo e ângulo solar), padrões (variações na mesma categoria), posição (trajetória), tonalidades (cores) e ao estado de conservação do veículo.

1.2.Organização da dissertação

No capítulo 2 faz-se uma breve revisão sobre a área de ITS e de algumas das suas aplicações. Faz-se também uma descrição sucinta dos sensores comumente utilizados na monitoração de trânsito e das principais aplicações utilizando processamento de imagens.

O capítulo 3 contém uma revisão abrangente do estado da arte em técnicas e aplicações relacionadas à monitoração de trânsito e apoio ao motorista. Isso inclui trabalhos que utilizam em conjunto processamento de imagens e reconhecimento de padrões.

O capítulo 4 é dedicado à descrição do modelo proposto de classificação de veículos. Este modelo é descrito de forma detalhada em seus módulos funcionais e é avaliado em estudo de caso. Também inclui os resultados e comentários sobre os testes realizados.

No capítulo 5 estão incluídos os comentários, as avaliações finais e algumas sugestões para trabalhos futuros.

2. ITS - Sistemas de Transportes Inteligentes

As aplicações que objetivam a monitoração e o controle de tráfego veicular podem ser incluídas na área de Sistemas de Transportes Inteligentes (ITS – *Intelligent Transportation Systems*). O desenvolvimento em ITS é um campo de pesquisa multidisciplinar que envolve desde estudos na área de engenharia de transportes até técnicas computacionais.

2.1. Sistemas de ITS

As aplicações de ITS podem ser divididas em sistemas que utilizam vários tipos de tecnologias, as quais podem ou não incluir ferramentas computacionais. Não existe um padrão único em relação aos possíveis sistemas a serem implementados. Logo abaixo são descritos alguns dos principais sistemas de acordo com o modelo do *USA Department of Transportation's* [ITS-DOT, 2006], o qual criou em 1991 o *ITS Program* [ITS-DOT-OVERVIEW, 2006].

Administração de vias arteriais: administram o fluxo de tráfego ao longo de estradas arteriais, empregando detectores de tráfego e semáforos, e alterando dinamicamente o sentido de pistas reversíveis e a velocidade máxima permitida.

Administração de auto-estradas: realizam medidas em vias de acesso a auto-estradas para otimização das velocidades de viagem e dos tempos de espera nos seus acessos (filas), distribuição de tráfego e uso máximo da capacidade de fluxo das pistas. Também são usados na administração de eventos especiais como, por exemplo, no controle do impacto de congestionamentos no acesso a estádios ou centros de convenções.

Administração de trânsito: incluem vigilância, roteamento, monitoração e desativação remotas de veículo.

Administração de incidente: visam reduzir os efeitos de congestionamentos relacionados a incidentes, diminuindo assim o tempo para detectá-los, para a chegada dos veículos de emergência e o tempo requerido para o tráfego voltar às suas condições normais.

Administração de emergência: incluem administração de materiais perigosos (ex., combustíveis), serviço médico de emergência, resposta à emergência em grande e pequena escala, e operação de evacuação.

Pagamento eletrônico: facilitam o pagamento de pedágios e tarifas de trânsito.

Informação ao viajante: utilizam letreiros luminosos dinâmicos, Internet, telefone móvel, linhas diretas de telefone, televisão e rádio, informando aos motoristas as partidas de viagem, condições, rotas e modos de viagem.

Administração de informação: apóiam o arquivamento e recuperação de dados gerados por sistemas de ITS. Sistemas de apoio de decisão, previsões, e monitoração de desempenho são algumas das aplicações possíveis através da análise das informações. Esses sistemas podem ajudar no planejamento de transporte, pesquisa, e atividades de administração de segurança.

Segurança e prevenção de colisões: emitem alertas em locais perigosos ou que necessitam de atenção maior como curvas perigosas, acessos de saída, viadutos, cruzamentos e interseções. Além disso, advertem em relação à presença de objetos indevidos na pista como materiais, substâncias, pedestres, ciclistas e animais. Também fornecem advertência levando em conta as características da categoria de veículo e a velocidade máxima indicada para ele em determinadas condições de estrada e visibilidade.

Operações de manutenção: focam em administração integrada de frotas de manutenção, veículos de serviço especializados, remediação de condições perigosas de estrada, e zonas de serviço para mobilidade e segurança.

Condições climáticas: as aplicações ajudam na monitoração e previsão de condições atmosféricas e de estrada, controle de limites variáveis de velocidade de acordo com as essas condições, assim como atividades de manutenção de efeitos climáticos (deslizamento de terra; alagamento; acúmulo de neve).

Veículos comerciais: projetam operações para aumentar a comunicação entre os transportes motorizados e as agências reguladoras. Exemplos incluem inscrição eletrônica e programas de autorização, troca eletrônica de dados entre agências para a melhoria de inspeções, e sistemas de blindagem eletrônica.

Trânsito intermodal: facilita a conexão eficiente, segura, e sem intermediação de fretes entre diferentes modais de transporte como ferroviária, aquaviária, rodovia e aerovia. Aplicações acompanham fretes e ativos particionados transportados como *containers* e

chassis, e melhoram a eficiência de processos nos terminais de frete, operações de *drayage*, e travessias de fronteiras internacionais.

Advertência anticolisão: são usados para melhorar a habilidade dos motoristas em evitar acidentes. Monitoram o ambiente em torno do veículo e alertam o motorista sobre condições que poderiam conduzir a uma colisão. Exemplos disso incluem advertências de colisão frontal, colisão traseira, colisão em cruzamento de pistas, colisão em troca de pista (ultrapassagem), detecção de obstáculos, rompimentos na pista (buracos), e inclinação do veículo com tendência à capotagem.

Assistência ao motorista: ajudam o motorista a dirigir seguramente o veículo. Sistemas estão disponíveis para ajudar na navegação como, por exemplo, melhoria de visibilidade, controle inteligente de velocidade, diagnóstico automatizado de veículo, monitoração de condições de carga, comunicação entre motoristas, controle de estabilidade e inclinação. Esses sistemas facilitam a direção segura durante condições adversas. Outros sistemas ajudam com tarefas motrizes difíceis como trânsito e ancoramento (estacionamento) de veículo de grande porte.

Notificação pós-colisão: melhoram o tempo de resposta para salvar vidas através de sistemas para detectar e informar o local e a severidade de incidentes a agências e serviços responsáveis para coordenar ações apropriadas de resposta à emergência. Estes sistemas podem ser ativados manualmente, ou automaticamente com notificação de colisão automática, e sistemas avançados podem transmitir informação sobre o tipo de impacto, número de passageiros, e a probabilidade de danos.

2.2. Pesquisa em ITS

No campo internacional de pesquisas a respeitada organização IEEE, acompanhando as tendências acadêmicas, criou em 1998 a *Intelligent Transportation Systems Society (ITSS)* [ITSS, 2006]. Esta Sociedade se interessa em aspectos teóricos, experimentais e operacionais de engenharia elétrica e eletrônica, assim como as tecnologias de informação aplicadas em sistemas de ITS. As áreas de interesse e temas relacionados destacados pela ITSS em eventos, como o *International IEEE Conference on Intelligent Transportation Systems 2007*, são as seguintes:

Viagem e Administração de Tráfego: Informação e Orientação para Viagem; Escolha e Reserva de Rotas; Informação de Serviços ao viajante; Controle de tráfego; Administração de incidentes; Administração de Demanda de viagem; Teste e Redução de emissões; Interseção de malha de rodovia.

Administração de transporte público: Administração de transporte público; Informação de Trânsito em rota; Trânsito Público personalizado; Segurança de Viagem pública.

Operações de Veículos comerciais: Liberação Eletrônica de Veículo comercial; Inspeção automatizada de Segurança em margem de estrada; Monitoração de Segurança à bordo; Veículo comercial Administrativo; Processos; Resposta a Incidente com Material perigoso; Administração comercial Rápida.

Pagamentos Eletrônicos: Serviços de Pagamentos Eletrônicos.

Administração de emergência: Notificação de emergência e Segurança Pessoal; Administração de Veículo de emergência; ITS e Segurança Nacional.

Sistemas avançados de Segurança de Veículo: Prevenção de colisão; Melhoria de visão; Sistemas avançados de Segurança; Operação Automatizada de Veículo.

Modelagem e Análise de ITS: Mineração de dados e Análise; Comportamento de viagem; Simulação e Modelagem; Teoria de tráfego; Modelagem estatística; Otimização e Controle; Teoria e Modelagem; Sistemas de Informação geográfica.

Outros Tópicos: Captura e Análise de Imagem; Fusão de Multi-sensores; Sistemas Cooperativos; Serviços de Ajuda ao Motorista; Interface de Homem-máquina; Arquiteturas e Padrões; ITS para Deficientes Físicos; Tecnologias modernas.

Outra sociedade de destaque é a *ITSA (Intelligent Transportation Society of America)* que há quinze anos desenvolve a área nos EUA, incorporando membros de corporações privadas, agências públicas, instituições acadêmicas e centros de pesquisa [ITSA, 2006]. Também existem outros países que estão se destacando em ITS, como Austrália [ITS-AUSTRALIA], Canadá [ITS-CANADA], Chile [ITS-CHILE], Singapura [ITS-SINGAPORE], África do Sul [ITS-SAS], Coreia [ITS-KOREA] e também os países da União Européia.

No ambiente acadêmico, várias universidades internacionais se destacam em avançadas pesquisas em processamento de imagens e reconhecimento de padrões (Tabela 2). Porém no Brasil estas áreas são emergentes e encontradas em poucas instituições de ensino.

Tabela 2: Grupos de pesquisa internacionais.

Chinese Academy of Sciences	National Laboratory of Pattern Recognition Institute of Automation
Czech Technical University	The Center for Machine Perception
University of California Berkeley	Computer Vision Group
Carnegie Mellon University	Robotics Institute
University of Nevada Reno	Computer Vision Laboratory
Vienna University of Technology	Pattern Recognition and Image Processing Group
University of Southern Califórnia	Computer Vision Laboratory
University of Pennsylvania	Laboratory of Robotics Research and Education

Tabela 3: Grupos de pesquisa nacionais.

Universidade Federal do Rio de Janeiro	LabIC – Laboratório de Inteligência Computacional
Universidade Federal do Paraná	IMAGO - Grupo de Pesquisa em Visão Computacional, Computação Gráfica e Processamento de Imagens
Universidade Estadual de São Paulo	Creative Vision Research Group
Universidade Federal do Rio Grande do Sul	Laboratório de Processamento de Sinais e Imagens

2.3.Processamento de Imagens em ITS

Visão é o principal sentido que usamos para perceber a estrutura do ambiente que nos cerca. Devido à importância da informação visual para o ser humano, torna-se claro o papel eminente da Visão Computacional para projetos de ITS. Com esse objetivo os sistemas visuais podem ser divididos em dois grupos: Sistemas de Monitoração de Tráfego Veicular (Figura 5-a) e Sistemas de Apoio ao Motorista (Figura 5-b).



Figura 5: Grupos de Sistemas Visuais
(a) Sistemas de Monitoração de Tráfego Veicular; (b) Sistemas de Apoio ao Motorista.

As aplicações de Processamento de Imagens em ITS incluem as seguintes:

- Monitoração em visão múltipla;
- Detecção de presença de veículo;
- Reconhecimento de Características de veículo: Categoria, Cor, Tamanho;
- Leitura de informações estampadas no
- Fluxo de trânsito: volume, densidade;
- Observação ambiental: pista, céu, vegetação, construções, sinais de trânsito, marcações de pista, faixa de segura;
- Detecção de Condições visuais: fumaça,

- | | |
|---|--|
| veículo: placa, símbolo, texto; | chuva, neblina, sombra, luminosidade; |
| • Detecção de obstáculos na pista:
pedestres, animais, veículos, resíduos; | • Análise de eventos: troca de pista,
ultrapassagem, avanço de sinal, aceleração |
| • Rastreamento de trajetória de veículo:
direção, sentido, velocidade; | excesso de velocidade, sentido incorreto,
retorno, parada, acidente, atropelamento. |

2.3.1. Sistemas de monitoração de tráfego veicular

O objetivo da monitoração [DIEHL, 2000] é observar um determinado ambiente, relatar informação sobre atividades relevantes e, se necessário, ativar as reações apropriadas. Sistemas de monitoração por vídeo usam sensores óticos para observar um ambiente. Sendo assim, um desses sistemas básicos consiste de câmeras fixas (ângulo de visão constante) ou dinâmico (ângulo variável).

Na visão constante os ângulos da câmera são ajustados apenas uma vez no momento de sua instalação, e não são mais alterados. Porém, na visão dinâmica a posição da câmera é fixa, porém possui uma base articulada com ativadores mecânicos (*pan-tilt*) que possibilitam a alteração dos ângulos por controle remoto.

A partir das câmeras são coletadas seqüências de imagens que constituem os fluxos de vídeos que são transmitidos a um local central. Estes são monitorados por vigilantes para descobrir alguma atividade que necessite de uma ação como resposta. Dado que tais eventos podem acontecer raramente, a descoberta deles requer observação focalizada durante períodos longos de tempo.

Sistemas de monitoração automatizados objetivam reduzir o trabalho do usuário detectando automaticamente eventos por vídeo. Para isso proporcionam avaliações oportunas de atividades pertinentes que permitam ao usuário alguma ação em contrapartida. A interpretação em tempo-real é obrigatória para que a informação produzida seja útil ao usuário e também para ativar possíveis reações automáticas do sistema.

2.3.2. Sistemas de Apoio ao Motorista

Os Sistemas de Apoio ao Motorista [RODRIGUES, 2002] devem comportar-se como um co-piloto cooperativo, ou seja, devem ajudar ao motorista em atividades de rotina que geralmente provocam desatenção.

O objetivo principal desses sistemas é aumentar a segurança do tráfego. Assim, devem fornecer informações que facilitem o raciocínio humano nas decisões tomadas sobre as ações imediatas no trânsito. Para isso eles podem comunicar-se com o motorista visando aprender sobre suas metas correntes e informá-lo sobre o que julgar importante. Assim como monitorar a situação de tráfego no ambiente imediato ao veículo e nos próximos segmentos ao longo da rota.

Diversas técnicas são investigadas para a automação completa ou parcial de tarefas relacionadas com a direção de veículos. As técnicas de Processamento de Imagens possibilitam auxílio em tarefas como: seguir uma rota, manter o veículo na pista correta, evitar obstáculos, permanecer à distância segura de outros veículos, regular a velocidade de acordo com as condições de tráfego e estrada, estacionar dentro de ambientes urbanos e fornecer informações sobre sinalização.

Outra tarefa viável é monitorar continuamente o veículo e suas reações aos comandos que o motorista transmite quando opera atuadores como o volante, freio e acelerador. A partir disso pode avaliar tais atividades do motorista e operar automaticamente os atuadores sob certas condições pré-definidas e seguras. Estão relacionados com este item os sistemas de direção automática.

2.4.Sensores

Vários tipos de sensores podem ser utilizados para extrair informações do ambiente de tráfego de veículos, os quais particularmente possuem seus princípios de operação que determinam suas capacidades e limitações [RITCHIE, 2005]. De acordo com as suas características os sensores podem ser usados individualmente ou em conjunto (fusão) para fornecer o resultado necessário ao sistema.

Tecnologias de sensores incluem o infravermelho, magnético, radar, microondas, ultra-sônico, acústico, e vídeo. Todos, com exceção do ultra-sônico e o acústico, utilizam

alguma forma de energia eletromagnética para descobrir a presença de um veículo. Processadores de imagem de vídeo usam imagem de espectro visível em comprimentos de onda mais curtos; seguidos por infravermelho, microondas, e eletromagnéticos. Devido a isso, o processador de vídeo é o único sensor afetado por variação de iluminação.

Sensores passivos não transmitem energia, mas sim detectam a energia que é emitida pelos veículos e a estrada, ou é refletida por eles. Sensores ativos transmitem energia, sendo que uma porção da qual é refletida ou se dissipa a partir do veículo ou estrada retornando ao sensor.

2.4.1. Vídeo

Detectores baseados em vídeo usam um microprocessador para analisar a entrada de imagem. Aproximações diferentes são usadas por sensores de detecção por vídeo. Alguns analisam a imagem de uma área designada no pavimento. A mudança na imagem na área designada, como um veículo que a atravessa, é processada. Outra aproximação identifica quando um veículo alvo entra no campo de visão e rastreia seu percurso nesta área.

Capacidades: colecionar volume, velocidade, presença, ocupação, densidade, comprimento de fila, tempo de permanência, intervalo de tempo entre veículos, movimentos de retorno, aceleração, mudança de pista, e classificação.

Limitações: condições ambientais que afetam a qualidade da imagem de vídeo podem reduzir o desempenho do sistema. Tais condições incluem névoa, chuva, pó, congelamento, condensação, sujeira na lente, condições de iluminação adversas, clarão de farol em pavimento molhado, luz solar de baixo ângulo, pouco contraste entre veículo e estrada, e reflexão de farol em curvas da estrada. Configuração e calibração apropriadas são fatores críticos para coletar dados mais precisos e alcançar um desempenho satisfatório em condições pobres de iluminação.

2.4.2. Laço Indutivo

Os detectores de veículo mais comuns [TARTARI, 2002] são aqueles que utilizam o sistema indutivo. Nestes um rolo de fio de cobre é enterrado sob a pista para criar um laço indutivo. Este normalmente é instalado em um corte de abertura retangular, próximo à superfície da pista por onde se desloca o veículo a ser detectado.

Há dois tipos principais de magnetômetros [MN/DOT, 2001]: de indução ou de eixo dual. Magnetômetros de indução medem mudanças magnéticas no campo induzido causadas por componentes de metal do veículo em movimento, especialmente o motor ou o alternador. Magnetômetros de eixo dual detectam mudanças nos componentes horizontais e verticais do campo magnético da Terra, causadas pela passagem ou presença de um veículo.

Capacidades: detectam volume, dimensão, intervalo de tempo entre veículos, presença e velocidade, através de análise do sinal de um ou mais sensores.

Limitações: sensores magnéticos presos à superfície da estrada estão sujeitos a deslocação ou dano causado por tráfego normal ou equipamentos de obras. Sensores montados debaixo do pavimento requerem perfuração para instalar canais e laços. Não podem descobrir veículos parados.

2.4.3. Transponder

A detecção de veículos é feita através de sinal de rádio. Requer a instalação do dispositivo de comunicação *transponder* (abreviação de *Transmitter-responder*) de identificação, individualmente codificado, em cada veículo.

Quando o veículo se aproxima, o sinal do dispositivo é recebido por uma antena de recepção. O número de identificação do veículo é enviado a um controlador que toma a decisão apropriada.

Capacidades: podem captar dados gravados e emitidos pelo dispositivo, assim como a presença de veículos e o intervalo tempo entre os mesmos.

Limitações: apenas reconhecem veículos que contenham o dispositivo padrão do sistema de recepção.

2.4.4. Infravermelho

O sistema de Infravermelho está baseado na detecção de mudanças tênues de radiação térmica emitidas por qualquer objeto ou corpo que tenha uma temperatura de superfície acima do zero absoluto (-273 C).

Sensores infravermelhos descobrem radiação infravermelha que alcança o detector. Há dois tipos básicos de sensores infravermelhos, passivos e ativos.

O primeiro tipo, **passivo**, descobre as mudanças em energia infravermelha emitida ou refletida de uma determinada área.

Já o **ativo**, irradia laser de baixa energia para uma área designada no pavimento e mede o tempo para o sinal refletido retornar ao sensor.

Capacidades: detectam volume de tráfego, presença e velocidade de veículo. Também são usados para detectar obstáculos e guiar o estacionamento de veículo.

2.4.5. Microondas

Os detectores de microondas emitem sinais que eventualmente são refletidos por um veículo que se move pela área de cobertura.

Esses sensores transmitem radiação de baixa energia a uma área designada no pavimento e então analisam o sinal refletido de volta ao detector. Determinando o tempo do sinal de retorno, calculam a distância até o veículo detectado.

Capacidades: detecta volume de tráfego, presença e velocidade de veículo.

Limitações: não detectam veículos parados. Somente detectam veículos que se movem mais rapidamente além da velocidade mínima pré-definida.

2.4.6. Ultra-som

Detectores de ondas ultra-sônicas emitem pulsos de sinal acústico que ficam além da audição humana e da maioria dos animais. Um veículo que se move pela área de cobertura, reflete os sinais de volta para o detector. Assim, distância entre o detector e o veículo é registrada pela medida de tempo de retorno dos pulsos.

Sensores sônicos ou acústicos **passivos** usam um receptor para descobrir a energia sonora criada pela passagem de veículo para determinar a sua presença, mesmo que ele esteja parado.

Já os **ativos** transmitem ondas a uma área designada no pavimento que são refletidas de volta ao receptor. Este mede o tempo que leva para uma quantidade de energia sonora ser refletida na área designada a um receptor.

Capacidades: sensores acústicos podem detectar dimensão, volume, velocidade, e presença.

Limitações: sensores acústicos estão limitados por condições ambientais que inibem a propagação de ondas de som. Tais condições incluem ventos fortes e precipitação pesada. Veículos altos, que viajam em pistas adjacentes, podem causar falsas leituras.

3. Técnicas para construção de um Sistema Visual de Monitoração

As principais áreas de estudo utilizadas na construção de um Sistema Visual de Monitoração são Processamento Digital de Imagens e Reconhecimento de Padrões.

O Processamento Digital de Imagens consiste da manipulação de uma imagem através de algoritmos implementados em computador de modo que a entrada e a saída do processo sejam imagens ou informações extraídas da imagem [BISHOP, 1995].

Reconhecimento de padrões consiste na manipulação de informações com o objetivo de identificar, descrever, reconhecer ou classificar objetos presentes e que sejam suficientemente similares a padrões desejados e pré-estabelecidos. As três principais linhas de pesquisa que abordam o tema são a estatística (*Bayes*), a sintática (ou estrutural) e as redes neurais [HAYKIN, 1992]. O estudo de Reconhecimento de Padrões em Processamento de Imagens pode ser aplicado em diversas áreas¹, sendo as seguintes as mais pesquisadas²:

- Análise de componentes;
- Classificação de objetos;
- Análise de imagens Biomédicas;
- Análise de Movimento e Trajetória de objetos;
- Análise de cor e textura;
- Inspeção industrial;
- Biometria;
- Identificação de objetos;
- Interpretação de cenas;
- Reconhecimento de caracteres;
- Reconhecimento de objetos;
- Sistemas de Transportes Inteligentes;
- Visão computacional;
- Vigilância inteligente;
- Classificação de eventos e ações;
- Sensoriamento remoto.

A monitoração de trânsito é uma especialização da área mais abrangente da Vigilância Inteligente [SEDKY, 2005; WEIMING HU, 2004] que consiste em técnicas de análise para detecção de objetos em movimento, localização de região, classificação de objeto, identificação de objeto, acompanhamento de movimento, classificação de movimento e de eventos.

¹ Veja maiores detalhes em: Aplicações gerais de reconhecimento de padrões

² Extraídas dos últimos seis anos em conferências relacionadas com reconhecimento de padrões

O reconhecimento dos estímulos visuais é a tarefa central em qualquer sistema de visão [RODRIGUES, 2002]. Qualquer problema de visão ou percepção em geral pode ser visto como um problema de reconhecimento. O reconhecimento faz parte do processo de análise de imagem e, segundo [GONZALES, 2000], as técnicas em análise de imagem podem ser divididas de acordo com três tipos: (1) processamento de baixo nível, com funções de aquisição de imagens e pré-processamento básico; (2) processamento de nível intermediário, com processos de extração e caracterização de componentes em uma imagem; e (3) processamento de alto nível, que envolve os processos de reconhecimento e interpretação (Figura 6).

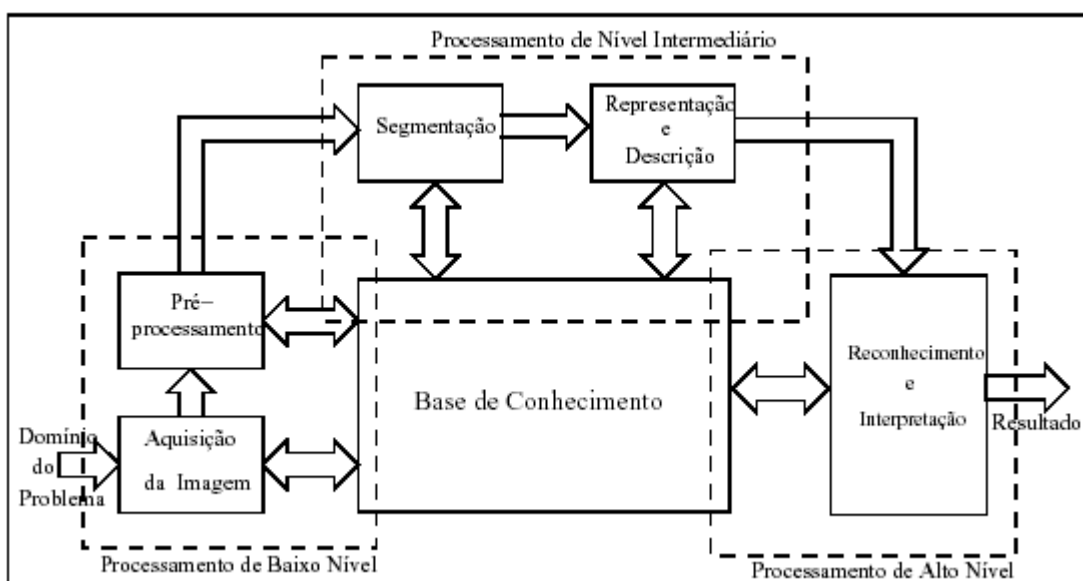


Figura 6: Modelo geral de análise da imagem.

Essas técnicas são agrupadas em módulos básicos que correlacionados compõem o um modelo geral para construção de um sistema visual de monitoração. Tais módulos são implementados de acordo com a aplicação desejada. Este modelo representa uma seqüência completa desde a aquisição até a análise final problema em questão.

A seguir são descritos cada módulo de acordo com as suas características, de funcionamento, objetivos e formas de implementação.

3.1.Aquisição

O primeiro passo do processo de reconhecimento é a aquisição da imagem. Para tanto, é necessária a utilização de um sensor de luz. Um exemplo de dispositivo utilizado para este fim é a câmera CCD (*Charge Coupled Device*). Este tipo de dispositivo utiliza uma matriz de

células fotossensíveis que atuam como capacitores, armazenando carga elétrica na proporção da energia luminosa incidente. Um conjunto de prismas e filtros de cor decompõe a imagem em seus componentes de cor RGB (*Red*, *Green* e *Blue*), onde cada componente é capturado por um CCD independente. Como a imagem gerada por esses dispositivos é analógica, torna-se necessária a utilização de outro dispositivo para converter o sinal analógico em digital, chamado digitalizador.

Na aquisição deve ser observada a visão fornecida pela câmera, a qual deve ser apropriada ao objetivo do problema. Para melhorar a adequação podem ser usadas técnicas de calibração de contexto.

3.1.1. Visão de câmera

O fator que mais influencia no processamento de imagens de trânsito é a visão fornecida pela foto adquirida [MATSUSHITA, 2002], a qual depende do posicionamento da câmera segundo os ângulos horizontal θ e vertical ϕ (Figura 7).

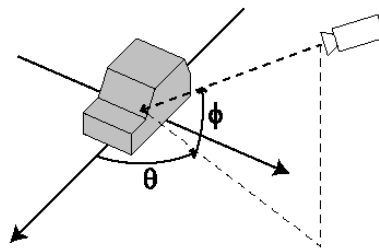


Figura 7: Ângulos de aquisição: horizontal θ e vertical ϕ .

Para aplicações de monitoração geralmente são utilizadas câmeras em posição fixa e com os ângulos de aquisição constantes. Estes podem ser ocasionalmente alterados quando existir a disponibilidade de equipamento *pan-tilt*. Porém essas alterações não são comuns porque podem implicar em reajustes em outros módulos. Porém em aplicações de apoio ao motorista as câmeras estão a bordo e os ângulos mudam dinamicamente e exigem adaptação automática de parâmetros do sistema.

Assim, a visão da câmera influencia no nível de detalhamento oferecido na imagem e nas aplicações possíveis como, por exemplo, a classificação de categoria de veículo. Assim, a visão frontal (Figura 8-d) oferece um conjunto limitado de detalhes do veículo, o qual pode ser utilizado para a classificação da categoria, mas oferece dificuldades para a diferenciação entre categorias similares. A visão lateral (Figura 8-a;c) oferece mais detalhes que facilitam

essa classificação. Porém, a visão diagonal, através da combinação das visões frontal e lateral (Figura 8-b), mostra-se como a de maior quantidade de características eficientes para a classificação da categoria.

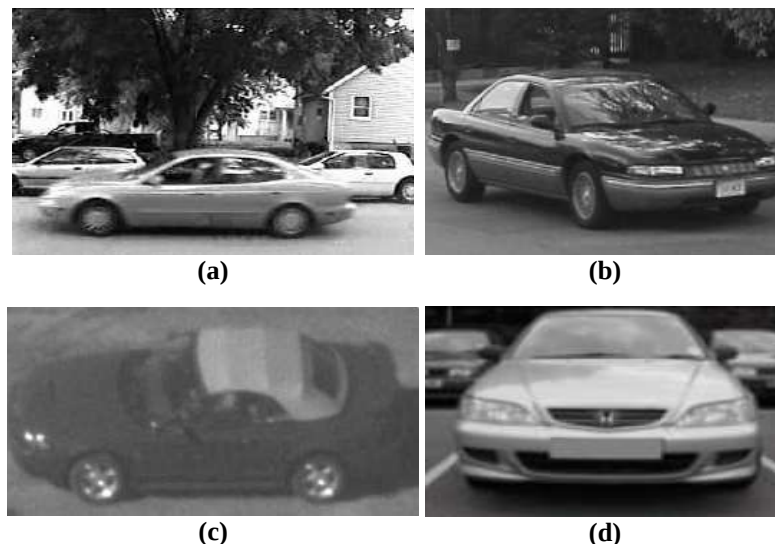


Figura 8: Veículos individuais e com imagem próxima em ângulos variados.
(a) Lateral [DORKO, 2003] (b) Frontal, diagonal e lateral [HAN, 2005] (c) Lateral e superior [GUO, 2005] (d) Frontal [PETROVIC, 2004].

Também exerce influência a distância do veículo em relação a câmera. Na (Figura 8) são mostradas visões que oferecem uma imagem muito próxima de veículos individuais, o que possibilita uma rica extração de características [DORKO, 2003; HAN, 2005; GUO, 2005] para a classificação de categorias e modelos. Porém, a distância pode ser parcialmente controlada de acordo com a capacidade de zoom da câmera, porém com o risco de perda de qualidade.

Entretanto, as câmeras de monitoração de trânsito geralmente utilizam uma distância maior para aumentar a abrangência e visualizar múltiplas pistas e veículos em apenas uma imagem (Figura 9; Figura 10). Deste modo é possível o acompanhamento de trajetória dos veículos e a classificação de eventos em seus movimentos. Porém as imagens de cruzamento de pistas exigem um processamento mais complexo por mostrarem trajetórias em várias direções e sentidos (Figura 9).

Porém, um problema encontrado nessas imagens múltiplas é a oclusão (sobreposição) entre veículos, a qual pode ser parcial ou total. A oclusão parcial ou a grande proximidade entre dois veículos causam a perda de regiões de veículo, o que pode prejudicar a extração de características para a classificação. Além disso, podem causar o equívoco do sistema em

fundir as regiões de dois veículos e considerá-las como sendo de apenas um veículo. Outro problema em imagens abrangentes é a mudança de da distância e perspectiva causada por movimentação de veículos em diferentes posições na imagem.



Figura 9: Imagens em cruzamento de pistas.

(a) Visão múltipla, superior (aérea) em cruzamento [HINZ, 2003] (b) Visão múltipla, superior, diagonal e variável em cruzamento [KOLLER, 1993].



Figura 10: Imagens em pistas múltiplas.

(a) Visão múltipla, superior, diagonal, traseira e lateral [KOLER, 1994] (b) Visão múltipla, superior e frontal em estrada [WORRALL, 1994].

A (Figura 11) mostra a visão de uma câmera a bordo de um veículo em movimento, o que adiciona a complexidade do fundo em movimento e totalmente variável [TZOMAKAS, 1998]. O ângulo de visão é variável devido ao movimento dinâmico e imprevisível dos veículos, o que dificulta a tarefa de detecção e segmentação. Outro problema é a trepidação freqüente do veículo que pode afetar a qualidade através de adição de ruído na imagem.



Figura 11: Visão múltipla, traseira e variável em câmera a bordo de veículo.

3.1.2. Calibração de contexto

No contexto de trânsito a monitoração de ambiente externo possui pouco controle sobre fatores críticos para o processamento das imagens como, por exemplo, iluminação e posição de objetos. Para atenuar esse descontrole a calibração é realizada a partir da utilização máxima de conhecimentos prévios sobre o ambiente que possam incrementar o sucesso das operações.

Na calibração de câmera fotográfica são determinados os parâmetros necessários para estabelecer equações de projeção entre medidas do mundo real e a geometria da imagem [BAS, 1997]. Uma vez que estas relações são encontradas, um sistema de vídeo pode estimar informações em coordenadas reais para a inspeção de cenas.

A calibração de contexto pode ser realizada através de algumas técnicas como: delimitação de regiões, marcação de trajetórias esperadas, caracterização geométrica, caracterização radiométrica, tratamento de perspectiva e projeção de ponto de fuga.

Para descobrir e classificar veículos são necessários alguns de seus atributos na imagem como, por exemplo, o local, o comprimento, a largura e a velocidade. Entretanto, os algoritmos de análise de tráfego geralmente usam câmeras fotográficas estáticas, calibradas com antecedência em relação ao plano espacial do chão, operando em cenas nas quais são esperados veículos seguindo trajetórias aproximadamente fixas. Estas condições são encontradas na maioria dos sistemas de monitoração de trânsito e facilitam a calibração de contexto.

A delimitação de regiões pode ser inclusiva ou exclusiva [SILVA, 2001]. Na inclusiva são delimitadas regiões contendo os caminhos possíveis de veículos a serem considerados. Na exclusiva são marcadas as áreas a serem ignoradas que não necessitam ser analisadas (Figura 12).

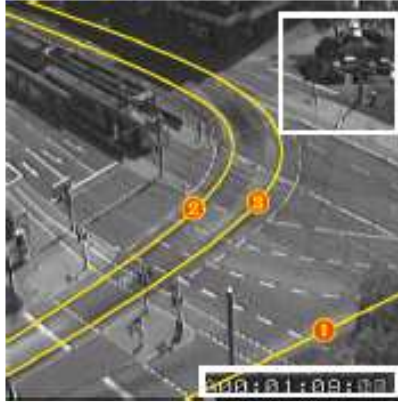


Figura 12: Calibração de regiões incluídas e excluídas em cruzamento de trânsito.
(as curvas traçam os caminhos de veículos e os retângulos delimitam as áreas a serem ignoradas)

Considerando que a maioria das estradas são superfícies planas, pode-se assumir que o movimento dos veículos é paralelo às linhas das pistas [MALIK, 1997]. Assim é possível especificar pontos ou retas correspondentes entre o plano da pista e o plano do sistema global de coordenadas (Figura 13).

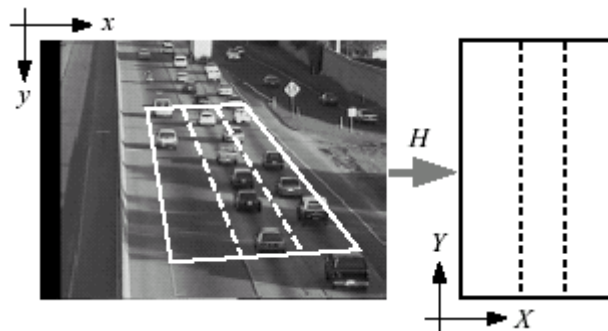


Figura 13: Correspondência entre o plano da pista e o sistema global.
A projeção geométrica H é usada para mapear as coordenadas da imagem (x,y) para o sistema global de coordenadas (X,Y) .

Foi desenvolvida em [GUPTE, 2002] uma ferramenta de calibração que permite marcar as pistas de tráfego e seus sentidos na imagem (Figura 14). Também são marcados os locais de detecção e cálculo da velocidade. A partir dessas marcas e do conhecimento de suas medidas, o sistema pode computar as dimensões aproximadas dos veículos.

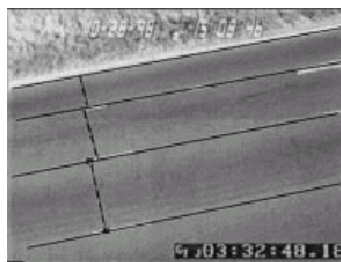


Figura 14: Ferramenta de calibração por linhas.
Ferramenta para marcação das linhas de divisão entre as pistas, e para designar a largura das pistas.

Foi modelado por [ZHANG, 2006] um caso em cruzamento múltiplo com variações de trajetória determinadas pelos semáforos (Figura 15.a). Neste ambiente complexo é realizada a calibração dos caminhos possíveis (direções e sentidos) (Figura 15.b). Também são calibrados os fatores de escala para tratar o problema de mudança de perspectiva devido ao movimento em diferentes distâncias da câmera (Figura 15.c).

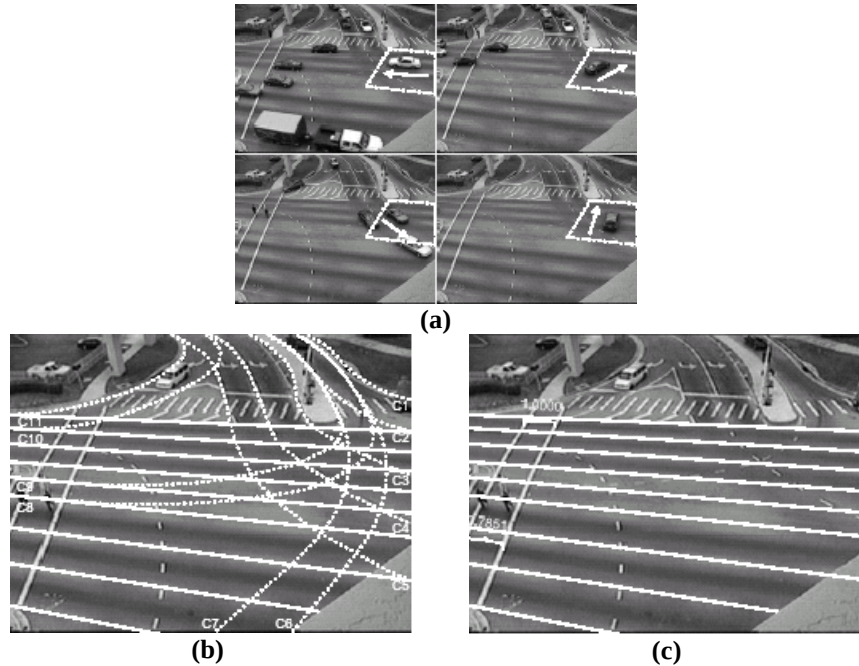


Figura 15: Calibração de cruzamento múltiplo de vias urbanas.

(a) Movimentação em diferentes sentidos na mesma área, porém em momentos diferentes determinados por semáforos; (b) Delimitação dos caminhos possíveis; (c) Delimitação de faixas de escala que representam as diferenças de perspectiva na profundidade da visão da câmera.

Na calibração realizada em [SULLIVAN, 1997] é esperado que o tráfego flua ao longo de direções fixas na imagem. Assim, é detectado o surgimento do veículo em duas regiões diferentes (retângulos sólidos) (Figura 16). A terceira região (pontilhada) é usada para confirmação de presença de veículo e pode se mover lateralmente.



Figura 16: Calibração de regiões esperadas para detecção.

Posições de geração de hipótese (retângulos sólidos) e verificação de veículo (retângulo pontilhado).

Em [HINZ, 2003] é usado um modelo que consiste em características geométricas e radiométricas. Geometricamente, o veículo é modelado como um objeto 3D que inclui informações relativas à projeção da sombra (Figura 17). Esta é aproximada a partir da data e da hora que determinam o ângulo de incidência solar, em conjunto com o ângulo e foco de

visão. Como característica radiométrica é incluída a uniformidade de cor entre o teto e o capô, considerando que esta é uma medida relativa e independente de mudanças de iluminação.

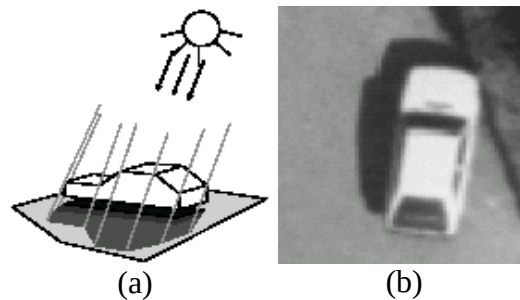


Figura 17: Modelo de veículo com projeção de sombra, pára-brisa, teto e capô. (a) projeção de sombra derivando da data, dia e parâmetros de orientação de imagem. (b) Visão superior do veículo mostra a uniformidade de cor no o teto e no capô.

O comprimento focal e o ângulo horizontal (*pan*) podem ser computados a partir da altura, do ângulo vertical (*tilt*) e de quatro pontos laterais da estrada selecionados na imagem. Assim, a partir dos pontos podem ser projetadas linhas paralelas na direção do infinito. Essa projeção é chamada de “ponto de fuga” e é representado pela interseção destas linhas (Figura 18).



Figura 18: Ponto de fuga representado pela interseção das linhas projetadas ao infinito.

No exemplo de [WORRALL, 1994] a convergência das linhas paralelas é bem nítida (Figura 19-a), e o ponto de fuga pode ser determinado com precisão. Assim, a linha central, que converge ao ponto de fuga, está aproximadamente na direção do ângulo vertical da câmera. Deste modo é usado um par de linhas paralelas para recuperar a escala e o comprimento focal (Figura 19-b).



Figura 19: Calibração a partir de um par de linhas paralelas. (a) Par de linhas paralelas; (b) Calibração da escala, comprimento focal e ponto de fuga.

3.2.Pré-processamento

Após a aquisição e digitalização da imagem, o próximo passo é o pré-processamento. A função chave do pré-processamento é melhorar a imagem, com o objetivo de aumentar as chances de sucesso dos processos seguintes [GONZALES, 1992]. Nesta etapa, são utilizadas técnicas para correção de imagem e também de redução de complexidade.

O pré-processamento é realizado como etapa preparatória anterior ao método principal, sendo que o seu objetivo é obter uma imagem mais apropriada para melhorar a qualidade do resultado ou aperfeiçoar seu desempenho. Geralmente essa etapa é importante, já que influencia no resultado de todas as etapas posteriores do processamento da imagem. Isso é importante em ambientes externos e com pouco controle, como é o caso da monitoração de trânsito.

3.2.1. Redução de complexidade

A redução usada para diminuir a complexidade da imagem em relação ao espaço de cores, resolução ou valores brutos. Porém, como em qualquer tipo de redução, isso implica em perda de informação e deve ser feito com o cuidado de se manter as características principais e necessárias para o posterior processamento.

Uma técnica básica para redução de complexidade é a conversão do espaço de cores da imagem. Normalmente uma imagem está representada no espaço RGB e alguns autores a convertem para outros espaços como HSV e HSI, pois nesses espaços as cores são mais facilmente separáveis [XIE, 2005].

Outra técnica muito comum é a conversão do espaço 3D do RGB para o espaço 1D de níveis de cinza, no qual um ponto é representado por apenas um valor de intensidade de cinza [THIANG, 2001].

A diminuição de complexidade também pode ser feita através de redução da resolução (escala), como por exemplo, [PETROVIC, 2004] reduz de 640x480 para 50x120 com o objetivo de descartar detalhes irrelevantes e destacar a forma principal dos veículos.

3.2.2. Correção de imagem

A correção visa atenuar a variação de luminosidade que é característica de imagens com iluminação natural, a qual varia de acordo com o horário e as condições climáticas (Figura 20). Uma técnica usada é a aplicação de uma função linear na intensidade e subtração do resultado na imagem original [MOUTINHO, 2005]. Porém a melhor forma é utilizar métodos tolerantes à iluminação [TSAI, 2005] ou que se adaptem automaticamente a sua variação [MESSELODI, 2005; BING-FEI WU, 2005].

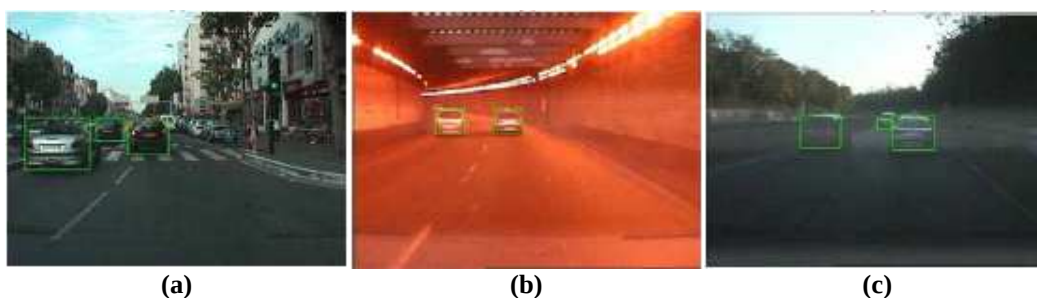


Figura 20: Ambientes variados com iluminação natural ou artificial.
(a) Urbano; (b) Túnel; (c) Estrada [KHAMMARI, 2005].

Outro problema tratado é o ruído na imagem, o qual pode ser atenuado através de filtros do tipo passa-baixa [JOLLY, 1996; LIM, 2002]. O ruído é gerado no sinal recebido pela câmera ou na transmissão de dados até o equipamento receptor de imagem.

Além disso, para realçar o contraste da imagem é usada a técnica de equalização de histograma, a qual incrementa o desempenho de métodos baseados na análise de bordas ou contornos [LIU, 2005; 73]. Também para a melhoria do contraste são usados filtros do tipo passa-alta [WU, 2005; PUMRIN, 2003].

A ocorrência de sombra na pista ou nos veículos é muito comum e causa interferências no processamento de imagens. A sombra pode incidir parcialmente ou totalmente sobre os veículos, deformando-os e alterando a sua tonalidade. Para tratamento desse problema são utilizadas técnicas de limiarização.

3.3.Segmentação

A segmentação de veículo é a etapa inicial e básica de qualquer aplicação de monitoração de trânsito, sendo em imagem estática ou em movimento (seqüência temporal). De acordo com a abrangência da visão, a segmentação pode ser única em uma região pequena ou múltipla com abrangência de uma ou mais pistas. Segmentar consiste em delimitar a região

onde cada veículo se encontra, a qual pode ser retangular ou aproximada sobre o contorno do veículo. Como resultado final, a segmentação destaca o veículo do fundo restante da imagem.

Geralmente a segmentação é dividida em dois passos, sendo o primeiro a geração de hipóteses de regiões candidatas, as quais no seguinte passo serão confirmadas ou não como sendo veículos. As técnicas usadas podem variar de acordo com a análise de imagem individual ou de sequência temporal, respectivamente nomeadas como estática e temporal.

De um modo geral, a segmentação divide uma imagem em suas partes constituintes ou objetos de interesse. Cada uma destas partes é uniforme com respeito a algumas características como, por exemplo, forma, cor ou textura. Para realizar essa segmentação são utilizadas duas propriedades básicas: descontinuidade e similaridade [GONZALES, 1992]. Através da descontinuidade, o particionamento da imagem é baseado no subconjunto de pontos de um objeto que o separa do restante da imagem. Essas técnicas buscam evidenciar os limites entre os objetos, através da detecção de pontos isolados, linhas e bordas. Porém, na similaridade, a segmentação é baseada na união de regiões que tenham características semelhantes. Isso pode ser feito através de técnicas de limiarização, crescimento por regiões, união e divisão de regiões.

3.3.1. Estática Individual

A segmentação estática utiliza apenas uma imagem para a análise, através de técnicas espaciais e inserindo conhecimento prévio de características que indiquem a localização de veículos ou outras regiões (calibração de contexto). Técnicas utilizadas incluem a análise de cor, bordas, contorno e forma de veículo, sendo estas analisadas através de ferramentas computacionais como SVM, RNA, *Wavelet*, PCA.

A partir de visão variável em estacionamento, [TSAI, 2005] realiza segmentação de regiões por cor. Para isso transforma RGB em 2D através de PCA, extraindo os autovetores e verificando-os por método estatístico para gerar regiões candidatas. Após valida os veículos por tamanho mínimo de 36x36, constrói contornos (*chain coding*) e mapa de distâncias, extrai os coeficientes de *Wavelet* (*Daubeches*) e normaliza-os para o conjunto [-1, 0, 1]. No final faz uma análise conjunta de contornos e *Wavelet*. O método utiliza dois grupos de treinamento, sendo um composto por imagens de veículos e outro por não-veículos (planos de fundo) (Figura 21).

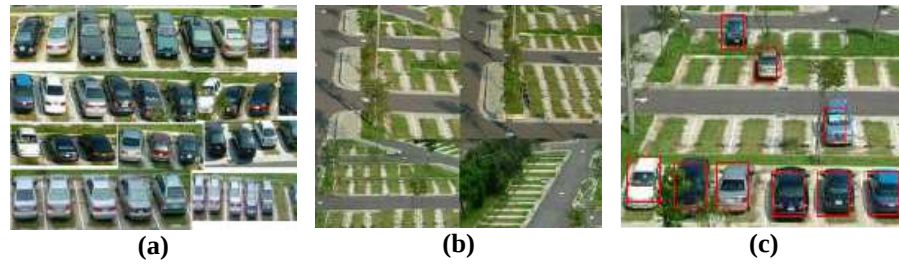


Figura 21: Amostras para segmentação por cor.

(a) Amostras de veículos; (b) Amostras de não-veículos; (c) Exemplos de segmentação.

A partir de visão traseira superior (256x240), [ROJAS, 1997] converte do RGB para no plano 2D de cromaticidade (HSV) e neste classifica os pontos em pista ou não-pista. Após realiza uma projeção *image-to-ground* e une as regiões sobrepostas ou muito próximas, encaixando-as em retângulos que são considerados como veículos (Figura 22). Testa apenas 88 imagens, porém inclui exemplos com variação de luminosidade de acordo com condições climáticas ensolarado, nublado, sombreado e chuvoso.

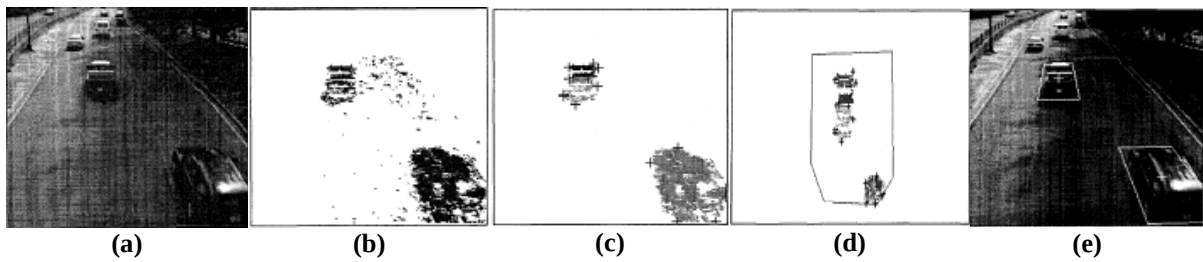


Figura 22: Segmentação de veículos em monitoração.

(a) Imagem original; (b) Pontos classificados como não-pista (c) Regiões conectadas; (d) Projeção *image-to-ground*; (e) Veículos detectados no plano original da imagem.

Em [WANG, 2005] a segmentação é realizada em visão traseira superior, em conjunto de 5120 imagens contendo 9200 veículos. Nestas as regiões de bordas são extraídas através da análise dos coeficientes *Wavelet* (*Daubechies*). Em seguida para a delimitação das regiões são aplicados os passos mostrados na (Figura 23). Após esses passos as regiões resultantes são consideradas como veículos.

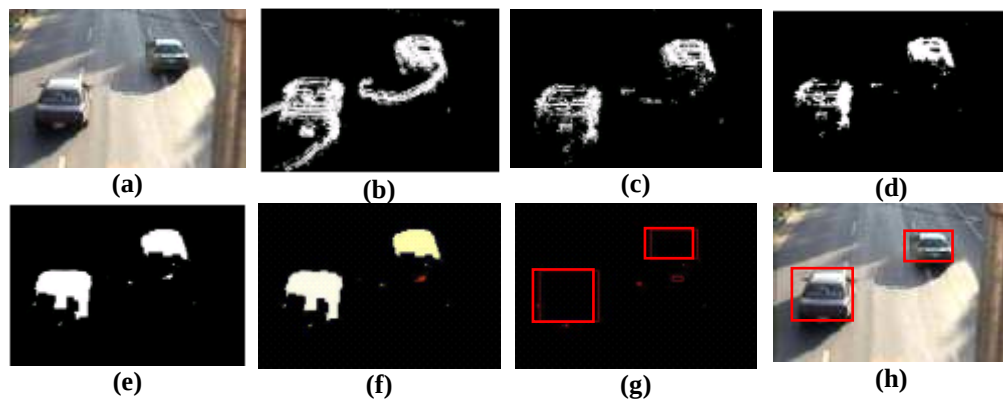


Figura 23: Passos para segmentação estática.

(a) Original; (b) Bordas; (c) Remoção de sombra; (d) Filtro de média 3x3; (e) Fechamento morfológico; (f) Componentes conexos; (g) Exclusão de pequenos retângulos; (h) Resultado final.

3.3.2. Estática Temporal

A segmentação temporal inclui informações sobre o movimento do veículo, utilizando técnicas espaço-temporais na análise de mais de um quadro de imagem em sequência ou operações com quadros de referência.

As imagens em sequência são representadas pelos dois ou três quadros mais recentes. O quadro ou fundo de referência pode ser estático, representado por uma imagem da pista sem veículos, ou adaptativo iniciado a partir de um quadro vazio (sem veículos) que em sequência é atualizado a cada próximo quadro analisado.

A partir de imagens extraídas de sequência de vídeo (15 fps) em visão lateral superior, [GUPTE, 2000] extrai os mapas de bordas dos dois últimos quadros e combina-os com a operação lógica XOR. Neste resultado remove o ruído por filtragem e dilatação morfológica. As regiões resultantes são agrupadas por altura e largura, e de acordo com essas dimensões identifica os veículos.

Analizando imagens de vídeo (5 fps, *jpeg* 320x240) em visão traseira superior, [PUMRIN, 2003] constrói fundo de referência adaptativo a partir da subtração entre o mapa de bordas (*Sobel*) de dois quadros em sequência. Para isolar os veículos realiza a diferença entre o quadro atual e o de referência, e deste resultado aplica a operação morfológica de fechamento (2x2).

Em visão frontal superior [XIE, 2005] calcula a diferença em tons de cinza entre um fundo adaptativo e os três últimos quadros. No resultado desta operação são analisadas as bordas por *Sobel* para segmentar os veículos. No final elimina a sombra através de limiarização no espaço HSV.

A subtração entre os três últimos quadros é realizada por [JOLLY, 1996] em visão lateral, a qual é seguida de binarização e fechamento morfológico. São extraídas as bordas pelo método de *canny*. Destas realiza somatórios das linhas e colunas e os analisa para encontrar a base e o comprimento da base do veículo.

Na visão lateral, superior e distante (320x240), [HUANG, 2004] realiza a subtração de fundo adaptativo, binarização, remoção de ruído, abertura morfológica, fechamento morfológico, e rotulação de componentes conexos. A validação é realizada através da análise de tamanho mínimo e separação de oclusão.

Em [ZHANG, 2000] na visão lateral calcula a diferença entre os quadros $t-1$ e $t-2$, após a diferença entre os quadros $t-2$ e $t-3$, e no final executa a operação lógica AND entre esses dois resultados. Realiza a segmentação por difusão e fusão (*quadtree*) em RGB, detecta bordas por *Sobel* e remove ruído com filtro de maioria 3x3, o que resulta no contorno final do veículo.

Em vídeo (25 fps, 352x288) com visão superior e variada de cruzamento, [MESSELODI, 2005] utiliza diferença de fundo adaptativo, o qual é atualizado pela média. Detecta bordas dos objetos em movimento (mapa de diferenças) através da média de convolução de dois passos de filtro Gaussiano 1D (Figura 24). Remove ruído por filtragem e executa operações morfológicas para unir os componentes próximos, excluindo os objetos muito pequenos. São analisados 70000 quadros contendo 1325 veículos.

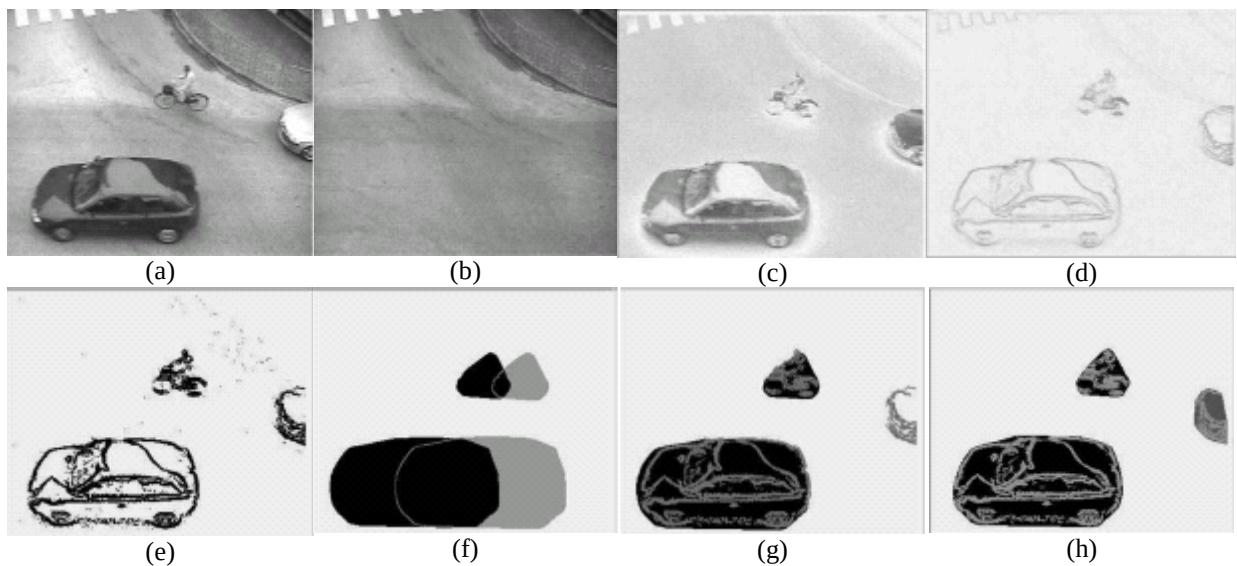


Figura 24: Passos de segmentação temporal em monitoração.

(a) Quadro atual. (b) Fundo de referência atualizado. (c) Mapa de diferenças. (d) Gradiente do mapa de diferenças. (e) Bordas do mapa de diferenças. (f) *Blobs* dos objetos ativos (em cinza) e suas posições esperadas (em preto). (g) *Blobs* esperados (em preto) e objetos após o movimento. (h) objetos detectados: dois objetos ativos e um novo objeto em entrada.

3.3.3. Dinâmica Individual

A imagem dinâmica é resultante de aquisição a partir de câmera em movimento. Esse caso é encontrado em câmera a bordo de veículo com o objetivo de apoio ao motorista. Nesse contexto o fundo é dinâmico e análise é realizada em apenas uma imagem, pois a análise em sequência temporal somente se justifica quando o fundo dos quadros adjacentes é muito similar (ou estático). Algumas técnicas utilizadas incluem a análise características como bordas, contorno, cor, geometria e ponto de fuga.

A partir de visão traseira e frontal, [THOMANEK, 1994] divide uma área de busca em 30 janelas verticais e detecta as bordas verticais e horizontais (Figura 25). A partir destas extrai e analisa os contornos extraídos, gerando assim as principais linhas horizontais do veículo.

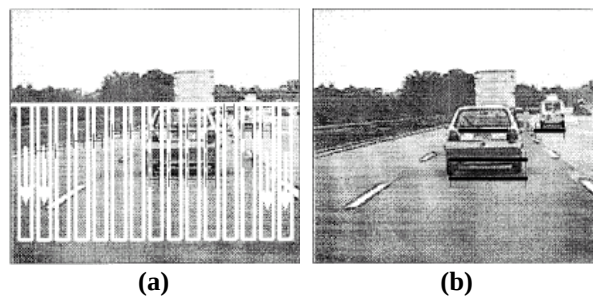


Figura 25: Detecção de linhas principais de veículo.

(a) Janelas de busca para extração das bordas horizontais; (b) Linhas principais traçadas após encadeamento de bordas e análise de contorno.

No trabalho de [BETKE, 2000] são usadas imagens extraídas de uma hora de vídeo. Nele os carros são localizados através de busca por retângulo analisando as bordas, validando por *aspect ratio* (valor próximo a um), e por correlação com modelos (Figura 26). Também detecta a região de pista através de análise de cor no espaço HSV e bordas (Figura 27).

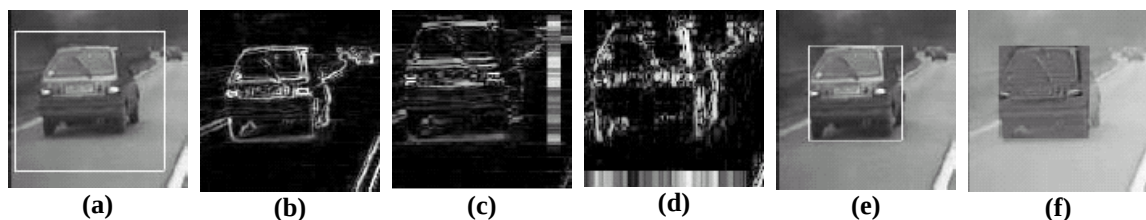


Figura 26: Passos da segmentação dinâmica.

(a) Imagem com região de busca; (b) Bordas; (c) Bordas e somas horizontais; (d) Bordas e somas verticais; (e) Esboço da região candidata após limiarização dos valores dos somatórios; (f) Um modelo de carro sobreposto sobre a região esboçada, sendo que a o coeficiente de correlação é de 0,74.

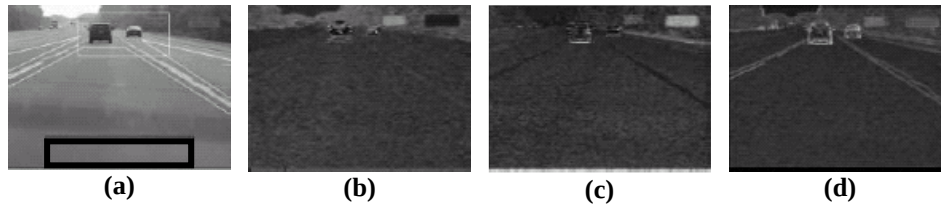


Figura 27: Passos da detecção de pista.

- (a) A região do retângulo preto é usada para estimar os valores de média e variância da cor de pista;
 (b) Componente *hue*; (c) Componente *saturation*; (d) Soma das bordas, *hue* e *saturation*.

A segmentação também é aplicada na reconstrução da estrutura 3-D de uma cena a partir de uma ou mais imagens. Isto é um problema clássico de Visão de Computador e tem aplicações importantes à orientação de veículo com piloto automático.

Analisando série de vídeo contendo 17 min e 25000 imagens, com conhecimento prévio do foco da câmera em 16 mm e da altura em 1.6 m, [CAPPELLO, 1994] segmenta veículos considerando-os como regiões com bordas verticais e horizontais dentro da pista, após a calibração automática de contexto através da detecção de pontos de fuga que resulta na delimitação de pista, construções, árvores, sinais, postes e na reconstrução 3D básica (Figura 28).

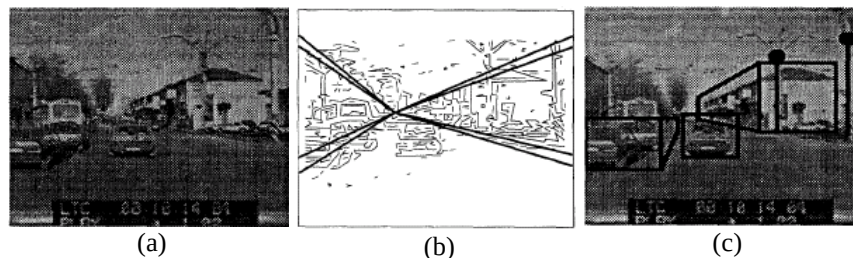


Figura 28: Segmentação de ambiente 3D.

- (a) Imagem original; (b) Mapa de bordas com os pontos de fuga e as linhas de perspectiva; (c) Representação de blocos identificando os veículos e as construções detectados.

3.4.Representação e Descrição

Geralmente, a saída do estágio de segmentação gera dados brutos representados em pontos (*pixel*). Neste caso pode ser necessário converter os dados para uma forma conveniente, possibilitando o processamento numérico de acordo com o método subsequente de reconhecimento. Dois tipos de representação podem ser utilizados: contorno ou regional.

A representação por contorno é apropriada quando o foco está em características da forma externa como, por exemplo, em cantos. Já a regional é apropriada quando o foco está em propriedades internas, tais como textura ou forma estrutural. No entanto, em algumas aplicações estas representações coexistem.

Escolher a representação é apenas parte da solução para a transformação de dados brutos em uma forma conveniente para o processamento subsequente. Um método para descrever os dados, tal que as características de interesse sejam realçadas, também deve ser utilizado.

A descrição lida com a extração de características que resultam em algumas informações qualitativas de interesse ou que são básicas para diferenciar uma classe de objetos de outra [GONZALES, 1992].

Várias técnicas são usadas para representação e descrição. Essas são citadas e descritas nas seções dos outros módulos de segmentação e reconhecimento.

3.5.Reconhecimento e interpretação

O último estágio no processo de análise da imagem envolve reconhecimento e interpretação. Reconhecimento é o processo que estabelece um rótulo a um objeto baseado na informação fornecida pelos seus descritores. Interpretação envolve o estabelecimento de significado a um grupo de objetos reconhecidos.

Para resolver problemas de reconhecimento pode-se partir de três abordagens conforme [GONZALES, 1992]:

- Estatística: conjuntos de medidas de características (na forma de n-tuplas ou vetores) são extraídos das imagens e métodos estatísticos são utilizados para separar as classes. Dentre os métodos utilizados cita-se os classificadores estatísticos, métodos probabilísticos, regras de decisão etc.;

- Estrutural: padrões são representados em uma forma simbólica (*strings* ou *árvores*), e os métodos de reconhecimento são baseados em casamento de símbolos ou em modelos que tratam padrões de símbolos como sentenças, a partir de uma linguagem artificial;

- Neural: na abordagem neural o reconhecimento é realizado utilizando-se Redes Neurais Artificiais.

Algumas das principais aplicações possíveis de reconhecimento e interpretação são: confirmação de presença, classificação de categoria, rastreamento de trajetória e interpretação de eventos.

3.5.1. Classificação de Categoria

Os veículos podem ser classificados quanto à categoria a partir de descritores que representam algumas das características dominantes de sua forma visual. Algumas das técnicas descritoras utilizadas são: medidas geométricas, bordas, contorno, modelo 3D, filtro de *Gabor*, retângulo, esqueleto, gradiente, pontos brutos e autovetores. Esses são analisados por métodos como: PCA, SVM, RNA, *Bayes*, RBF e CCN. As visões de câmera mais usadas são a superior, lateral, frontal e diagonal, de acordo com o contexto de monitoração de trânsito.

Este tipo de aplicação utiliza técnicas de classificação que tipicamente dividem o conjunto de imagens em dois conjuntos, de treinamento e teste. O conjunto de treinamento é utilizado pelo método classificador para realizar a aprendizagem das características principais e discriminantes entre as classes de objetos.

No caso de categoria de veículo, a aprendizagem pode ser supervisionada, na qual os exemplos de treinamento são previamente rotulados com a categoria desejada. A partir disso o classificador é treinado para separar corretamente o maior número possível de exemplos. Após o treino, o classificador é desafiado a responder de acordo com os exemplos não rotulados do conjunto de teste, o qual é desconhecido para o mesmo, utilizando apenas o conhecimento adquirido no treinamento. Após o teste as respostas do método são comparadas com um gabarito para determinar o nível acerto do método.

A partir de veículo segmentado em visão superior variável em cruzamento de trânsito (Figura 15), [ZHANG, 2006] inicia com rotação, transformação e escala para normalização com os modelos. Compara dois métodos de extração de autovetores, PCA e o proposto *Eigen-vehicles*, que são classificados por SVM (Tabela 4). A base de imagens é composta por 390 veículos divididos entre treino (90) e teste (300). A categoria carro é a mais numerosa, sendo que ela se apresenta na proporção de 4:1 em relação às restantes.

Tabela 4: Acerto de classificação em SVM por PCA e *Eigen-vehicles*.

Acerto Categoria	PCA (%)	<i>Eigen-vehicles</i> (%)
Carro	86	64
Camionete	53	65
Van	55	54

Em visão Lateral, [JOLLY, 1996] extrai o contorno (canny) poligonal (sete lados) e compara com modelos através de classificação estatística, em cinco categorias (Figura 29). O

teste de classificação é realizado em 393 imagens de veículos (Tabela 5). O acerto geral nos testes de classificação é 91%.

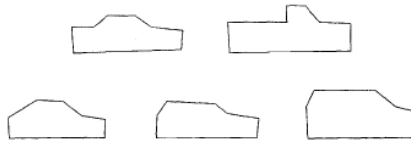


Figura 29: Modelos poligonais de sete lados.
(sedan, camionete, carro hatchback, station wagon, e van)

Tabela 5: Matriz de confusão para modelos poligonais e minimização metropoli.

Alvo \ Teste	Sedan	Camionete	Hatchback	Station	Van
Sedan	202	4	6	0	0
Camionete	2	32	0	0	0
Hatchback	1	0	38	3	0
Station	2	4	1	45	1
Van	2	2	1	3	44

Coletando imagens laterais em entrada de garagem, [LIM, 2002] extrai o *Gabor Jet* a partir de filtragem do veículo por *Gabor* (8 orientações e 4 escalas). Após esse descritor é comparado com cada um dos 12 modelos (4 por categoria) e escolhido o de maior similaridade para as seguintes categorias: carro *sedan*, *van* e camionete. Como resultados dos testes são classificados corretamente 93% dos veículos.

Analisando imagem lateral superior distante, [HUANG, 2004] extrai o contorno do veículo. A classificação é realizada através da análise das seguintes proporções entre as medidas do contorno: altura, comprimento, *aspect ratio*, *compact ratio*. As classes são: carro, van, camionete, caminhão, caminhão baú, ônibus e *trailer* (Figura 30). A taxa geral de classificações corretas é 91.35% e por categoria é mostrada na tabela (Tabela 6), na qual os números entre parênteses indicam as oclusões ocorridas entre veículos.



Figura 30: Exemplos de veículos em visão lateral.

Tabela 6: Matriz de confusão em proporções entre medidas de contorno.

Teste Alvo	Carro	Van	Camionete	Caminhão	Cam. baú	Ônibus	Trailer	Total	%
Carro	202 (11)	1	3	0	0	0	0	206 (11)	98
Van	4(3)	25(3)	5(2)	0	0	0	0	34(8)	66
Camionete	2	2	10	0	0	0	0	14	71
Caminhão	0	0	0	10(4)	0(1)	0	0	10(5)	93
Cam. baú	0	0	0	0	6	0(1)	0(1)	6(2)	75
Ônibus	0	0	0	(1)	0	9(1)	0	9(2)	91
Trailer	0	0	0	(1)	0	0	3(1)	3(2)	80
								282(30)	

Comparando o retângulo que contém o contorno lateral do veículo com os modelos, [OTA, 2005] classifica em: carro *sedan*, *wagon*, perua (van) e carro *hatch-back*. Neste processo, é calculada a diferença entre o retângulo do veículo candidato com cada um dos modelos armazenados (Figura 31). A partir disso é escolhido o modelo com a menor diferença com o candidato. Foram realizados apenas 34 testes, nos quais a taxa de acerto geral é 27/34 (80%) (Tabela 7).



Figura 31: Diferença entre retângulo de candidato e modelo.

Tabela 7: Matriz de confusão em retângulo de candidato e modelo.

Alvo	Sedan	Mini Van	Hatchback	Wagon
Teste				
Sedan	10/10	0/6	5/9	1/9
Mini Van	0/10	5/6	0/9	0/9
Hatchback	0/10	0/6	4/9	0/9
Wagon	0/10	1/6	0/9	8/9

A partir de visão superior traseira, [FORESTI, 1999] classifica através da extração do esqueleto morfológico estatístico (Figura 32), determinando a sua orientação e comparando com modelo 3D (36 variações possíveis de ângulos). Foram testados aproximadamente 3000 veículos em diferentes de iluminação e fundo, porém em conjunto limitado de ângulo horizontal de visão (0°, 90°, 180°, 270°)(Tabela 8).

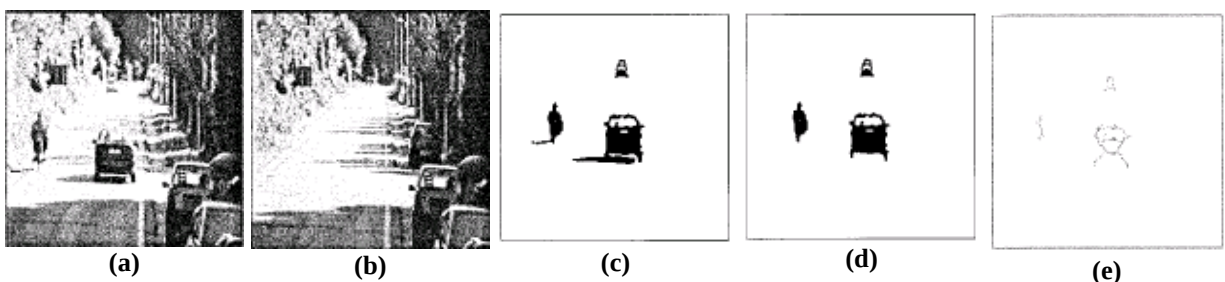


Figura 32: Passos da construção de esqueleto morfológico estatístico.

(a) Original; (b) Fundo de referência; (c) Limiarização; (d) Remoção de sombra; (e) Esqueleto m.e.

Tabela 8: Matriz de confusão em esqueleto morfológico estatístico e modelos 3D.
(valores expressos em porcentagem)

Teste \ Alvo	Carro	Ônibus	Moto	Van	Lorry
Carro	74	3	5	7	2
Ônibus	3	86	1	3	7
Moto	6	2	71	4	2
Van	6	2	3	79	5
Lorry	2	5	1	4	88

Em visão superior e variada em cruzamento (Figura 24, seção 3.3.2), [MESSELODI, 2005] projeta modelos 3D na imagem (*convex hull*) e procura correlação com modelos armazenados. Após usa SVM para classificar entre as oito categorias (Tabela 9), alcançando 85% de acerto .

Tabela 9: Matriz de confusão em *convex hull* e modelos 3D.

Teste \ Alvo	Bicicleta	Moto	Carro	Van	Ônibus Urbano	Ônibus Interurbano	Lorry	Outros
Bicicleta	37	8	3	1	0	0	0	2
Moto	5	138	7	2	0	0	1	1
Carro	0	0	884	24	0	0	14	0
Van	0	0	5	27	0	0	13	1
On. Urbano	0	0	0	0	10	0	0	0
On. Interurbano	0	0	0	0	0	4	0	0
Lorry	0	0	0	2	1	0	9	0
Outros	14	6	12	4	0	0	3	24
Total	56	154	915	67	11	6	47	76

Veículos em faixas múltiplas (Figura 33) são classificados por [HSIEH, 2006] com base no tamanho e linearidade, confirmando por análise do mesmo veículo em vários quadros em seqüência, em: carro, perua (van), caminhão e caminhão baú (inclui ônibus). No sistema são realizados tratamentos como eliminação de sombra, separação de oclusão e normalização de tamanho de veículo de acordo com a posição na pista. São testadas três seqüências de vídeo, incluindo comparação de desempenho para análise em apenas um ou múltiplos quadros para determinar a classificação (Tabela 10).



Figura 33: Pistas com faixas múltiplas.

Tabela 10: Desempenho por tamanho e linearidade, com 1 ou múltiplos quadros.**Legenda:** Vn (vídeo n), c (carro), v (van), cb (caminhão baú e ônibus), cm (caminhão), AC (acerto em %).

	Veículos reais				Classificação em 1 quadro					Classificação em múltiplos quadros				
	c	v	cb	cm	c	v	cb	cm	AC	c	v	cb	cm	AC
V1	310	142	32	22	290	113	25	16	87.7	301	129	29	20	94.8
V2	1994	162	143	146	1814	124	121	120	89.1	1860	139	131	135	92.6
V3	167	78	89	41	151	63	75	34	86.1	159	70	81	38	92.8

A partir de contorno de veículo em visão frontal, superior e distante (Figura 34), [RAD, 2005] analisa largura do retângulo e velocidade de veículo para classificar em: moto, carro, ônibus e micro ônibus, confirmando em vários quadros para o mesmo candidato. Os testes resultam em acerto de 96% em uma sequência de vídeo com 400 quadros (320x320), sendo que cada um destes possuem em média 27 veículos.

**Figura 34: Veículos em visão frontal, superior e distante.**

Utilizando 18 atributos de contorno lateral, disponíveis no banco de imagens da UCI [MLRepository, 1998], [PRAMPERO, 1999] classifica através de modelo híbrido com MLP, RBF (com *K-means*) e CCN (*Cascade Correlation Net*) (Tabela 11). Os resultado destes são analisados em conjunto de acordo com diferentes combinações (Tabela 12). As categorias são: ônibus de dois andares, *chevrolet van*, *saab 9000* e *opel manta 400*. O conjunto é formado por 960 imagens, sendo distribuído em 240 exemplos por categoria.

Tabela 11: Desempenho de testes por classificador individual.

Classificador	Acerto (%)
CCN	70,29
RBF	79,40
MLP	88,05

Tabela 12: Desempenho de testes por combinação dos classificadores.

Combinação	Acerto (%)
Maioria	71,38
Soma máxima	90,15
Média	69,08
Votação por MLP	89,15
Votação por RBF	86,48

Em visão frontal (640x480) de veículo estacionado, [PETROVIC, 2004] classifica entre 77 modelos distintos. O veículo é único e bem centrado na imagem, a qual é convertida para níveis de cinza e redimensionada para 50x120. Em seguida recorta a região de interesse a

partir da placa como referência, sendo que tal região irá aproximadamente incluir os faróis e o pára-choque. Desta região extrai *square mapped gradients* (Figura 35), classificando com *dot product* a partir de modelos de referência. Os testes são comparados com outros oito extratores, com inclusão ou não de redução PCA (Tabela 13). O conjunto de dados é dividido em 105 para o treino e 1027 para o teste.

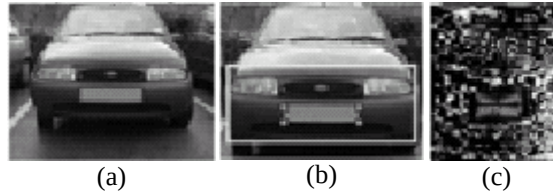


Figura 35: Passos de reconhecimento de modelo de veículo.
(a) Imagem original. (b) Região de interesse. (c) *Square mapped gradients*.

Tabela 13: Taxa de acerto em veículos estacionados.

Extrator	Taxa de acerto (%)		<i>n dims.</i> PCA (95%)
	Sem PCA	Com PCA	
<i>Locally normed grads</i>	47.3	79.3	91
<i>Edge orientation</i>	54.9	57.9	96
<i>Raw pixel values</i>	62.8	4.5	68
<i>Harris corner detector</i>	72.4	25.0	70
<i>Sobel edge responses</i>	75.0	74.7	85
<i>Spectrum phase</i>	82.5	48.4	97
<i>Direct normed grads</i>	84.5	83.9	93
<i>Square mapped grads</i>	97.7	95.1	95

Em sequência, [PETROVIC, 2004] realiza testes em outro conjunto de imagens capturadas em entrada de estacionamento (Figura 36). Essas imagens são caracterizadas por rotação e variação de iluminação. São comparados quatro extratores distintos (Tabela 14). É importante destacar a viabilidade do descritor *raw pixel values* que é extraído diretamente dos valores dos pontos da imagem em escala de cinza, e que obteve acertos de 62.8 e 91.2.



Figura 36: Veículos em entrada de estacionamento.

Tabela 14: Taxa de acerto em veículos em entrada de estacionamento.

Extrator	Taxa de acerto (%)	
	Sem PCA	Com PCA
<i>Raw pixel values</i>	91.2	6.4
<i>Harris corner detector</i>	96.6	40.7
<i>Direct normed grads</i>	98.0	93.1
<i>Square mapped grads</i>	98.0	93.6

Das bordas *Sobel* [TAN, 2000] extrai o histograma de orientação para determinar o ângulo de comparação com modelos 3-D em visão superior variável. A classificação é através

de correlação com todos os modelos (Figura 37). As classes são: carro, microônibus, van e van alta (Figura 38).

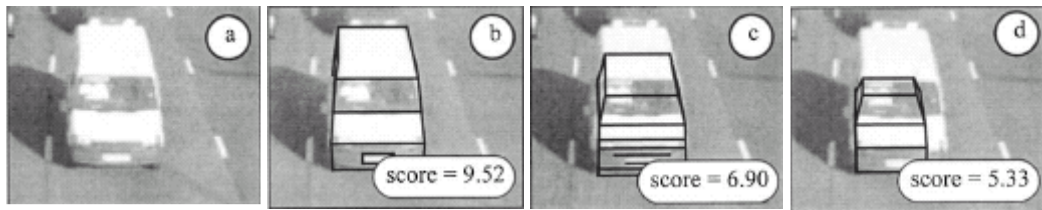


Figura 37: Passos de classificação por modelo 3D.

(a) Original; (b) Correlação com van alta; (c) Correlação com microônibus; (d) Correlação com carro.

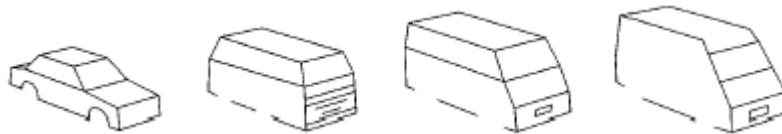


Figura 38: Categorias em modelos 3D: carro, microônibus, van e van alta.

Em [KOLLER, 1993] analisa visão superior variável de dois vídeos, um de 80 quadros em curva de estacionamento (Figura 39) e outro de 50 quadros em cruzamento (Figura 9.b, seção 3.1.1). São utilizados modelo poligonais 3D de 12 medidas (Figura 40), os quais possuem possibilidade de rotação para adequação à posição do veículo na imagem. A classificação é realizada através da distância de *Mahalanobis* entre os segmentos de bordas do veículo e os dos modelos 3D projetados em 2D (após a rotação).



Figura 39: Sequência em curva de estacionamento.



Figura 40: Categorias em modelos 3D de 12 medidas.

(carro sedan, carro hatchback, carro station wagon, perua (van) e camionete)

Em visão traseira superior (320x280), [SIDLA, 2004] gera milhares de janelas sobrepostas de áreas candidatas. As janelas são de tamanhos diferentes de acordo com a perspectiva e diferença de tamanho entre as classes. Dessas áreas o autor extrai os 20 primeiros autovetores (PCA) e os classifica via SVM em automóveis ou caminhões. As imagens de treinamento foram editadas centralizando o veículo em imagem de 128x128 e alinhando os faróis (Figura 41). O melhor desempenho foi com SVM, mas também realizou comparações com outros classificadores (Tabela 15). O conjunto de imagens possui 2697 automóveis, 577 caminhões e 2222 fundos, sendo estes divididos em conjuntos de 60% para treino e 40% para teste.

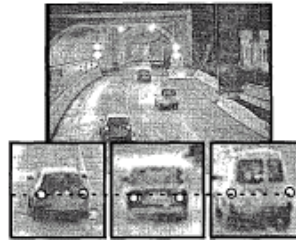


Figura 41: Veículos centralizados em imagem 128x128 e alinhamento de faróis.

Tabela 15: Comparação de desempenho entre classificadores.

Classificador	Acerto (%)
<i>Nearest Neighbour (NN)</i>	95.8
<i>k-Nearest Neighbour (k-NN)</i>	98.2
<i>SVM</i>	99.4
<i>Probabilist Neural Net (PNN)</i>	95.6
<i>Learning Vector Quantisatiort (LVQ)</i>	91.7

3.5.2. Confirmação de presença

Essa aplicação é utilizada em auxílio aos métodos segmentadores para confirmar a presença de veículo em regiões previamente segmentadas como “candidata a veículo”. Para isso as regiões são classificadas em apenas duas classes: veículos e não-veículos (fundo). Alguns descritores utilizados são: bordas, *Wavelet*, *Gabor* e autovetores. Esses são analisados por SVM, RNA, PCA e AG.

Nas visões traseira e frontal superiores (720x576), [BAI, 2006] calcula a diferença em níveis de cinza entre os dois últimos quadros e detecta bordas por *Canny*. Para validar o veículo dilata as bordas, extrai 12 atributos do tipo *haar-like* e os classifica em *Adaboost*. Sendo que no treinamento são usadas 3405 amostras positivas e 3386 negativas.

Zehang Sun realizou um estudo incremental a partir de veículos já localizados em área retangular de visão traseira (360x248) (Figura 42). Para isso escreveu uma seqüência de cinco artigos com as mesmas imagens e mesmo classificador SVM como confirmador, porém com técnicas distintas para extrair os descritores. O conjunto de dados possui 1051 amostras positivas e 1051 negativas, sendo estes divididos em subconjunto de 70% para o treinamento, 20% para a validação e 10% para teste.



Figura 42: Exemplos de imagens traseiras

No primeiro trabalho [SUN, 2002] redimensiona as imagens para o tamanho 64x64, divide-as em 16 partes de 16x16 agrupadas em 9 janelas sobrepostas de 32x32. Em cada uma destas executa filtragem de *Gabor* e no resultado classifica (SVM) para confirmar a presença veículos. Também compara os resultados com PCA (tratamento e redimensionamento) em conjunto com redes neurais.

No trabalho seguinte [SUN, 2002] realiza tratamento de contraste através de subtração de função linear na intensidade da imagem, e também por equalização de histograma. Após isso, os veículos são centralizados em regiões de tamanho 32x32, das quais são extraídos os coeficientes de *Wavelet (Haar)*. Esses são normalizados, sendo que os negativos extremos para -1, os positivos extremos para +1 e os restantes para 0 (zero). São realizados testes em três configurações distintas de SVM, e comparação de desempenho com redes neurais MLP.

Em seguida [SUN, 2002] analisa em conjunto três resoluções diferentes, sendo elas 360x248 (100% do original), 180x124 (50%) e 90x62 (25%). Nestas realiza as operações mostradas em (Figura 43). A partir da combinação dos resultados das três resoluções são determinadas as regiões candidatas a veículos. Cada uma destas é redimensionada em 32x32, da qual são extraídos os coeficientes *Wavelet (Haar)* de cinco níveis com exclusão da sub-banda HH, o que resulta em 768 coeficientes que são analisados por SVM para confirmação de veículos.

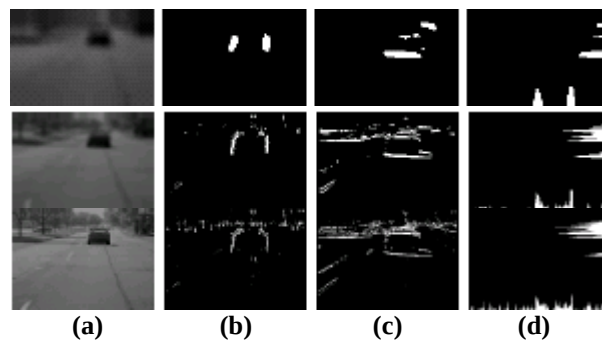


Figura 43: Geração multi-resolução de regiões candidatas.
Resoluções: 90x62 (1ª linha), 180x124 (2ª) e 360x248 (3ª)
(a) filtragem passa-baixa; (b) bordas verticais; (c) bordas horizontais;
(d) máximos e mínimos, verticais e horizontais.

Também realçando o contraste, com veículos recortados e centralizados em 32x32, em [SUN, 2004] extrai 200 autovetores (PCA), e a partir de algoritmos genéticos são escolhidos 46, os quais são classificados (SVM) para confirmar veículos. O autor também discute reconstruções adquiridas por métodos de seleção específica (Figura 44) e outros métodos automatizados de seleção de autovetores (Figura 45).

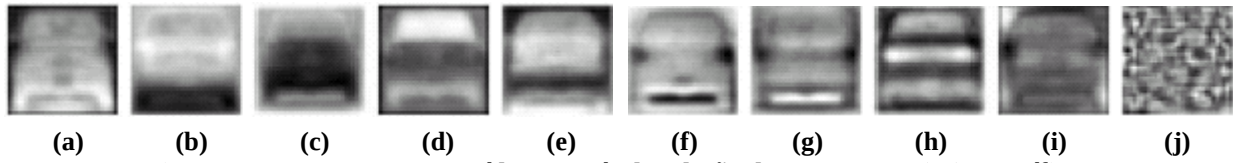


Figura 44: Imagens reconstruídas através da seleção de autovetores (av) específicos.
(a) av 1 (b) av 2 (c) av 3 (d) av 4 (e) av 5 (f) av 8 (g) av 10 (h) av 12 (i) av 15 (j) av 150

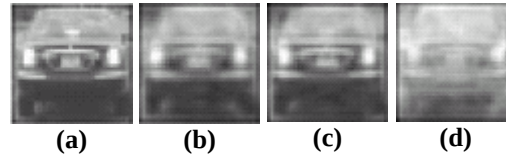


Figura 45: Imagens reconstruídas através da seleção automatizada de autovetores.
(a) original; (b) primeiros 50 av; (c) av por SFBS; (d) av por algoritmos genéticos.

No último artigo, [SUN, 2005] utiliza a mesma análise em três resoluções de [SUN, 2002], porém os candidatos são divididos em 16 regiões de tamanho 8x8. Estas são agrupadas em outras 16 regiões (com sobreposição parcial), as quais recebem a filtragem de *Gabor* para descrever os veículos candidatos. Os filtros deste são escolhidos por algoritmo genético, o qual tem como função de avaliação o SVM para aperfeiçoamento das gerações.

3.5.3. Rastreamento de trajetória

Acompanhar um veículo significa rastrear a sua trajetória em sequência temporal de vídeo. As técnicas como filtro de *Kalman* e distância entre centróides são usadas visando identificar a posição do veículo em quadros consecutivos, resultando no traçado de seu caminho.

O rastreamento é muito usado no contexto de monitoração de trânsito, no qual geralmente as imagens oferecem visão abrangente de grande área com múltiplos veículos. Exemplos disso são imagens de cruzamentos de pistas, avenidas com várias pistas e auto-estradas.

O filtro de *Kalman* é o mais utilizado para acompanhar o movimento de veículos. Através desse filtro é possível acompanhar movimentos em variadas visões, como por exemplo, visão variável [FORESTI, 1999; YANG, 2001], traseira [SOTELO, 2005; JUN-WEI HSIEH, 2006], frontal [HSIEH, 2006], frontal superior [VISSER, 2002].

Alguns autores adicionam à filtragem de *Kalman* outros parâmetros para melhorar o acompanhamento, como por exemplo, predição de posição em visão frontal superior (distante)

[RAD, 2005], distância mínima com semelhança em tamanho em lateral superior [GUPTE, 2000], coerência de caminho e forma em lateral superior (distante) [HUANG, 2004].

Outro método comum é a distância euclidiana entre os centróides de veículo em quadros consecutivos, como, por exemplo, é utilizada em visão traseira [PUMRIN, 2003; JUNXIAN WANG, 2005; ZHANG, 2003; SRINIVASA, 2002; HEISELE, 1995]. Em visão traseira e frontal superiores essa distância é usada em conjunto com análise de histograma HSV do veículo e QX-LMS [SADOOGHI-YAZDI, 2006]. Outra distância utilizada é a de *Hausdoff* que é usada em conjunto com entropia cruzada [HANDMANN, 2000].

Em [MESSELODI, 2005] é encontrado um método híbrido para visão superior variada (cruzamento), utilizando filtro de *Kalman* em imagem *convex hull*, e sendo selecionadas cinco áreas de 16x16 de cada veículo para acompanhá-lo e prever o seu próximo movimento.

3.5.4. Interpretação de eventos

Este tipo de aplicação adiciona uma análise de maior nível em semântica com o objetivo de interpretar o comportamento dos veículos no trânsito. Isso a partir de seus movimentos e interações com o meio ambiente.

Um exemplo abrangente com vários tipos de interpretações é o sistema *Autoscope* [MICHALOPOULOS, 1991], que há 17 anos desenvolve aplicativos em ITS, os quais se encontram instalados em 55 países e abrangendo um conjunto de 45.000 câmeras de vídeo instaladas. Esses aplicativos são divididos em cinco tipos, de acordo com o contexto de atuação: rodovia, tráfego urbano, túnel, coleta, análise, salvamento e segurança.

Em **rodovias** realiza monitoração de fluxo, detecção de congestionamento, detecção de incidente, contagem de fluxo em rampa de acesso, controle de limites variáveis de velocidade, segurança de regiões em obras, e estimação de tempo em viagem.

No **tráfego urbano** realiza a detecção da presença de veículo, detecção de filas, contagem de veículos, detecção de movimentos de retorno, ultrapassagem de semáforo vermelho, monitoração de fluxo, detecção de veículos parados, e detecção de veículo em sentido errado.

Em **túnel** detecta veículos parados, tráfego lento, veículos lentos, veículos em sentidos errados, detecção de fumaça, monitoração de fluxo, e controle de ventilação.

Na **coleta e análise de informação** é monitorado o volume de tráfego, velocidades de veículos, tempo de ocupação, intervalo de tempo entre veículos, classificação de veículo, estatísticas de incidente ou de evento.

E no **salvamento e segurança** são detectados veículos parados em áreas proibidas, veículos no sentido incorreto, detecção de ultrapassagem de semáforo vermelho, veículos de grande porte que excedem velocidades seguras, entrada sem autorização em áreas restritas.

3.6.Base de Conhecimento

Em um sistema de análise de imagens, o conhecimento sobre o domínio do problema pode ser codificado na forma de uma base de conhecimento. Este conhecimento pode ser representado tanto na forma de regiões simples, assim como regiões detalhadas de uma imagem, onde a informação de interesse a ser localizada é conhecida. Além disso, para guiar a operação e cada módulo de processamento, a base de conhecimento também controla a interação entre eles. Isto significa que a comunicação entre os módulos de processamento geralmente é baseada no conhecimento prévio de qual deve ser um resultado.

4. Proposta e avaliação de desempenho do modelo computacional

O objetivo principal deste trabalho é a classificação de categoria e cor de veículo a partir da análise de uma imagem digital. O modelo proposto para abordagem deste problema, conforme o diagrama (Figura 46), inclui sete módulos de processamento.

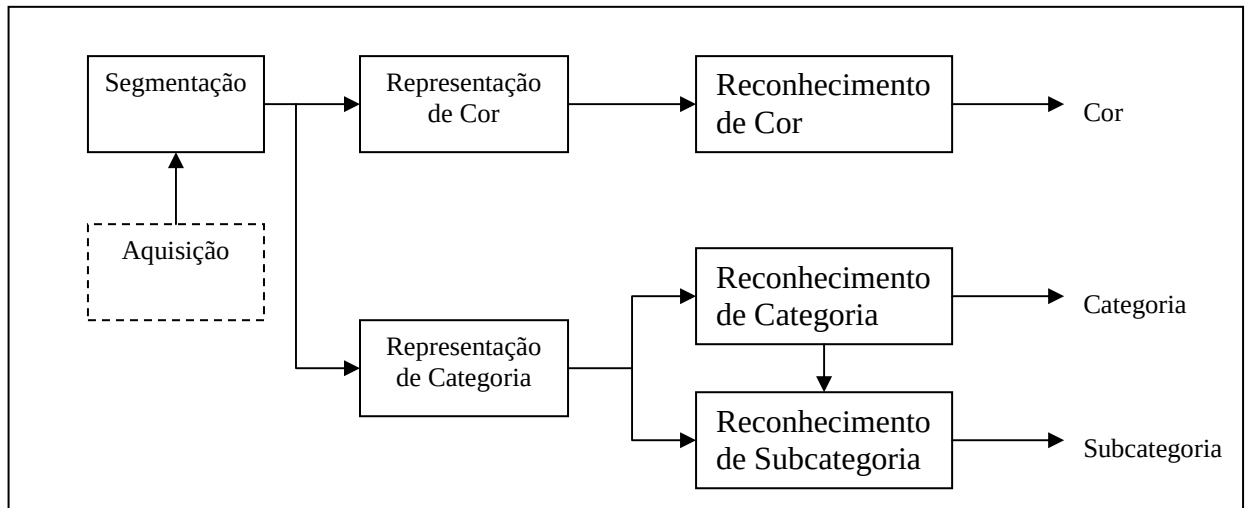


Figura 46: Diagrama de módulos funcionais.

O processamento de baixo nível, referente à aquisição das imagens, é realizado a partir da infra-estrutura e do sistema desenvolvido em [GUINGO, 2003]. As imagens foram tomadas em uma das vias de acesso do Campus Fundão da UFRJ e estão em visão diagonal, assim como é mostrado em (Figura 47).



Figura 47: Imagem de veículo em via de acesso à UFRJ.

No nível intermediário a segmentação é realizada através da marcação poligonal de veículos. Para isto foi construída uma ferramenta interativa que recebe como entrada as imagens adquiridas e permite segmentar manualmente o contorno do veículo e, em adição, escolher os rótulos de categoria e cor (Figura 48).



Figura 48: Segmentação interativa de veículo.
(a) Marcação poligonal de veículo; (b) Seleção de categoria e cor predominante de veículo.

A representação possui naturezas diferentes para as classes categoria e cor. Para a categoria o foco está em descrever a forma do veículo e para a cor está em marcar um ponto representativo da cor predominante. Sendo assim, a representação utiliza descritores regionais e geométricos para categorias de veículos e tonais para cores. Para a redução dimensional dos vetores descritores de categorias são utilizados o redimensionamento de imagem e a análise de componentes principais (PCA). Para os descritores tonais somente é necessária a redução por PCA. E na decorrelação dos dados também é utilizada a análise PCA para ambos os tipos de classes.

Enfim, no nível alto, o reconhecimento de categoria e cor de veículo é realizado através de classificação por Redes Neurais. O modelo utilizado é MLP com análise WTA na interpretação das suas respostas. São utilizados modelos diferentes para essas classes. A partir da categoria também é realizado o reconhecimento de subcategoria.

A plataforma *Matlab 7.1 R14 SP3* foi escolhida para desenvolver e testar os componentes funcionais deste estudo. É utilizada a biblioteca especializada *Image Processing Toolbox 5.1* [DEMUTH; BEALE, 2005] que contém um conjunto de funções para processamento de imagem. Também é utilizado o *Neural Network Toolbox 4.0.1* [DEMUTH; BEALE, 2005] para implementar e simular as redes neurais. O equipamento utilizado é um microcomputador com processador *Intel Pentium III* de 650 MHZ, memória RAM de 448 MB e disco rígido de 80 GB.

4.1.Aquisição e segmentação

A aquisição é originada por sinal de passagem em sensor de laço magnético, o qual ativa um aplicativo que realiza a coleta de três imagens seqüenciais em intervalo de 10 ms

(Figura 49). A imagem é salva em cores com profundidade de 24 bpp (*bits per pixel*) em BMP (*Bitmap*) e resolução 320x240. O modelo de câmera utilizado é a *Samsung Digital Color Camera SCC-C4203*.



Figura 49: Aquisição de seqüência de 3 imagens.
(a linha tracejada horizontal destaca o movimento dos veículos entre as imagens)

Para organizar a base de imagens é necessário um processo prévio de classificação e segmentação manuais dos veículos nas imagens. Isso posteriormente é usado para orientar os treinamentos e validar os resultados obtidos pelos classificadores. Para classificar a categoria de um veículo (Figura 50) é necessário segmentar a sua região e descartar as restantes contidas no ambiente ao seu redor (plano de fundo) (Figura 51). A partir desse segmento deve ser escolhido um ponto de coleta para a cor predominante do veículo. Esse ponto, de acordo com a categoria de veículo, pode ser escolhida em regiões como o capô ou teto.



Figura 50: Veículo em trânsito.



Figura 51: Recorte retangular colorido com fundo colorido.

Foi escolhido segmentar a região completa do veículo porque adiciona informações do comprimento e contorno do veículo que são relevantes para o reconhecimento da categoria. A partir disso o recorte completo do veículo pode ser retangular ou poligonal. É usado o polígono por ser o mais próximo ao contorno do veículo e, se necessário, o retângulo pode ser deduzido a partir do polígono.

Para a segmentação e classificação manuais foi construído o módulo Marcador para manipular cada imagem de veículo e definir as seguintes informações:

- Contorno: marcação do polígono aproximado ao contorno de veículo.
- Rótulo de categoria: seleção da categoria entre as disponíveis em lista pré-definida e fixa;
- Rótulo de modelo: seleção o modelo (ou subcategoria) em lista dinâmica;
- Rótulo de Cor: seleção da cor principal do veículo, em lista pré-definida e fixa;
- Ponto de Cor: marcação de um ponto no veículo, o qual representa a sua cor predominante;

Este módulo marcador oferece as seguintes facilidades:

- Seleção inicial dos arquivos de imagens é realizada por pasta;
- Navegação automática em seqüência, com opção de “pular” as já marcadas;
- Zoom automático para melhor visualização durante a marcação do contorno poligonal;
- Adição *on-line* de até 70 modelos;
- Seleção visual do modelo por similaridade de forma;
- A seleção de imagens é tripla, de acordo com a característica do contexto (Figura 52);
- Transposição automática das informações idênticas (rótulo do modelo, da categoria e da cor) da primeira imagem para as outras duas da mesma tripla;
- Transposição automática da marcação poligonal da primeira imagem para as outras duas da mesma tripla, com possibilidade de ajuste manual (Figura 52);
- Transposição automática das informações básicas (rótulo do modelo, da categoria e da cor) para as imagens da próxima tripla, facilitando a marcação de imagens divididas em pastas separadas de acordo com a categoria;
- Informações selecionadas são salvas no nome do próprio arquivo da imagem.

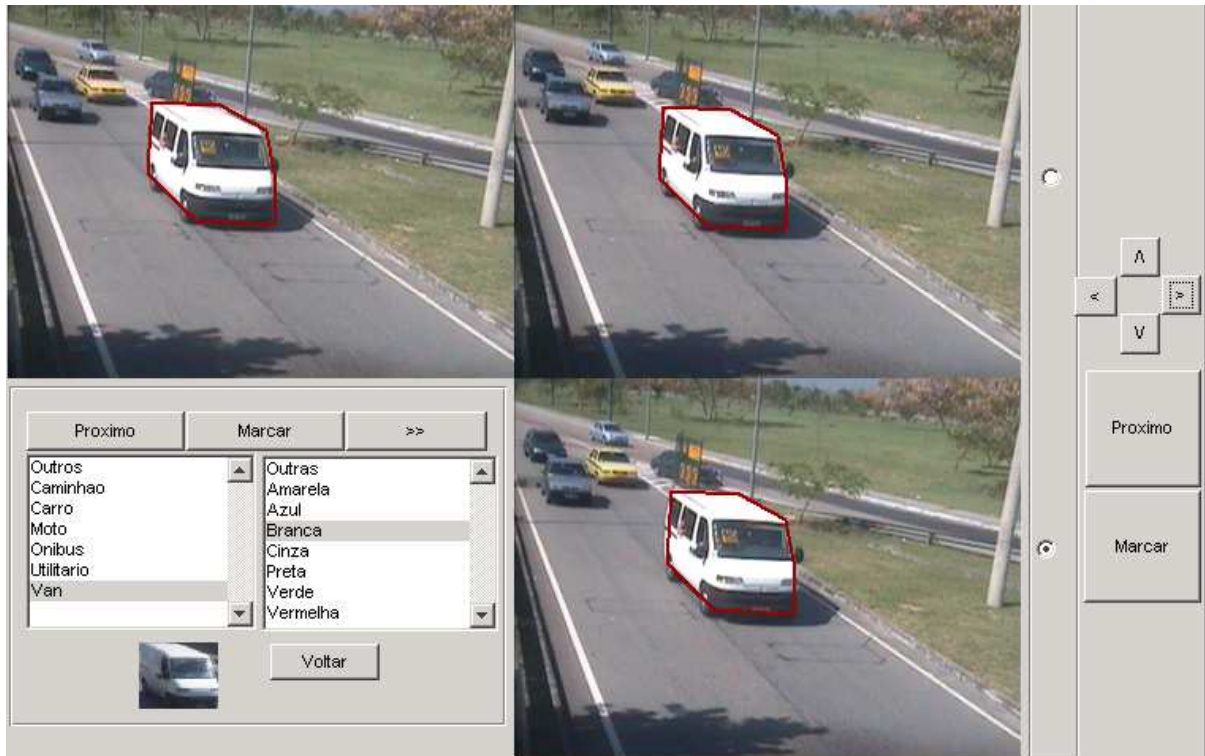


Figura 52: Tela de marcação manual de cor, categoria e contorno de veículo.

4.2.Descriptores de Categoria

Um veículo é um objeto que pode ser classificado de acordo com a sua forma visual. Esta descrição deve conter informações que, de algum modo, representem as formas e as diferenças entre categorias.

Assim, são utilizados dois grupos distintos de descritores, os geométricos e os regionais. Os geométricos extraem medidas numéricas que representam características de proporções do veículo. Os descritores regionais revelam características de aspecto visual do veículo. Esses descritores foram escolhidos de acordo com os estudos exemplificados na (Seção 3).

4.2.1. Geométricos

Os geométricos extraem medidas que representam informações de tamanhos e proporções do veículo, a partir da forma da sua silhueta. A silhueta é formada pela região delimitada pelo contorno externo do veículo, a qual é preenchida por pontos de cor branca (Figura 53.b).

Para a extração é utilizado o método *RegionProps* (Matlab), do qual foram selecionadas as medidas descritas logo a seguir. Deste modo resultam em poucos valores

numéricos que originam vetores de dimensões muito reduzidas em comparação aos regionais. Neste caso, o tamanho do vetor extraído já possui dimensão fixa, composto por um grupo de dez medidas, e não necessita de padronização de tamanho.

- *Area* (área): o número de pontos (*pixels*) da silhueta;
- *ConvexArea* (área convexa): o número de pontos na região da imagem convexa (*convex hull*) (Figura 53.c). A imagem convexa é o menor polígono convexo que contém a silhueta;

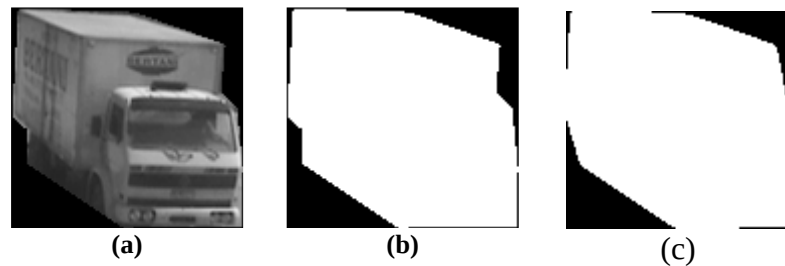


Figura 53: Imagem Convexa.
(a) Original; (b) Silhueta; (c) Imagem convexa;

- *Eccentricity* (excentricidade): a excentricidade da elipse que tem os mesmos momentos de 2ª ordem da silhueta. É a relação da distância entre o *foci* da elipse e seu comprimento de eixo principal. O valor está entre os valores extremos 0 (círculo) e 1 (linha);
- *EquivDiameter* (diâmetro equivalente): o diâmetro de um círculo com a mesma área da silhueta. Calculado como $\sqrt{4 * Area / \pi}$;
- *Extent* (extensão): a proporção entre os números de pontos que estão na silhueta e os que estão no menor retângulo que a contém. Calculado como a área da silhueta dividida pela área do retângulo;
- *FilledArea* (área preenchida): o número de pontos não nulos que correspondem à silhueta, com todos os buracos preenchidos;
- *MajorAxisLength* (comprimento do maior eixo): comprimento (em pontos) do maior eixo da elipse que tem os mesmos momentos centrais normalizados de 2ª ordem que a silhueta;
- *MinorAxisLength* (comprimento do menor eixo): comprimento (em pontos) do menor eixo da elipse que tem os mesmos momentos centrais normalizados de 2ª ordem que a silhueta;
- *Orientation* (orientação): o ângulo (em graus) entre o eixo x e o maior eixo da elipse que tem os mesmos momentos de 2ª ordem da silhueta;
- *Solidity* (solidez): é a proporção entre os números de pontos que estão na imagem convexa e estão na silhueta. Calculado como a área da silhueta dividida pela área convexa.

4.2.2. Regionais

Os descritores regionais preservam características de aspecto visual do veículo e, assim, possibilitam ao classificador uma análise que se aproxima da visão humana. Isso porque esses descritores são apresentados como imagens resultantes de transformações a partir dos recortes originais dos veículos. Ou seja, são descritos como o conjunto de valores brutos dos pontos do recorte resultantes de alguma transformação sobre o original. Deste modo são simples de extrair e representativos em informações visuais.

Os descritores regionais testados são: cinza, binário, silhueta, retângulo e bordas. O descritor regional cinza é o mais simples e corresponde aos pontos do recorte original do próprio veículo (Figura 54). Porém este descritor carrega informações muito redundantes e sensíveis a alterações de tonalidades resultantes da iluminação natural. Isso porque ele é simplesmente o conjunto de valores brutos dos pontos (*raw pixels*) da região do veículo. Então, a partir de transformações do recorte cinza, são derivados os descritores binários e de bordas com o objetivo de extrair as formas essenciais ou principais dos veículos. Com este mesmo objetivo são usados os descritores de retângulo e silhueta, porém originados do polígono da região do veículo marcado na segmentação manual. Sendo assim, alguns regionais refletem características da estrutura interna do contorno do veículo (binário, cinza, borda) e outras da sua forma externa (silhueta, retângulo).



Figura 54: Descritor cinza.

Para produzir o descritor cinza é usado o método *rgb2gray* (Matlab) que converte imagens de RGB para escala de cinza eliminando as informações de cor, porém mantendo a luminância. A viabilidade deste descritor é exemplificada no trabalho [PETROVIC, 2004], no qual é alcançado o acerto de até 91.2% na classificação entre 77 classes.

O descritor binário (Figura 55) é resultante da limiarização dos tons de cinza do intervalo [0,255] para [0,1]. Para isso é usado o método *im2bw* (Matlab) que converte a imagem de escala de cinza para a binária, baseando-se em limiar. A imagem binária produzida tem valores de 1 (branco) para todos os pontos da imagem com luminância maior

que um limiar e 0 (preto) para todo os outros pontos. O método *graythresh* (Matlab) é usado para calcular automaticamente o limiar global da imagem, usando a estratégia de [OTSU, 1979]. A vantagem deste descritor é a obtenção de um vetor com pouca variação de valores (0 ou 1), em comparação com o intervalo muito maior do cinza (256 valores), mas que mesmo assim represente e preserve a forma visual do veículo. Essa redução é importante para reduzir a complexidade dos cálculos executados para a classificação.



Figura 55: Descritor binário com fundo preto.

O próximo descritor é resultante da extração de bordas (Figura 56) que representam as formas essenciais do veículo. Para este descritor o método utilizado é o de [CANNY, 1986] que é considerado de ótimo desempenho. O *Canny* descobre bordas a partir de máximos locais do gradiente da imagem, o qual é calculado usando a derivada de um filtro Gaussiano. Este método usa dois limiares para descobrir bordas fortes e fracas, e só inclui as bordas fracas na produção se elas são conectadas a bordas fortes. Esse descritor, por também ser binário, reduz a complexidade dos cálculos e, em adição, revela as principais formas estruturais do contorno e do esqueleto do veículo. O método usado é *edge* (Matlab) que produz uma imagem binária com um (1) onde a função encontra bordas e zero (0) nas restantes.



Figura 56: Descritor de Bordas com fundo preto.

Em todos os tipos acima (cinza, binário e bordas), antes da extração o fundo contido no recorte pode ser anulado para destacar o veículo, ou seja, o veículo ser exatamente segmentado do fundo (Figura 57). Deste modo, somente a informação do veículo é preservada, descartando os valores que constam no fundo do recorte, os quais podem ser considerados como sendo “ruídos” e influenciam no resultado destes descritores. Para isto, a partir do polígono do contorno do veículo, os pontos contidos na região do veículo são mantidos e os da região externa são alterados para o valor zero (cor preta).



Figura 57: Anulação do fundo cinza do recorte.

Os descritores restantes, silhueta e retângulo, são retirados da segmentação manual, em forma de polígono, realizada sobre a imagem do veículo. A silhueta (Figura 58) é resultante do preenchimento em branco neste polígono e mostra o contorno aproximado do veículo. O outro descritor é extraído do menor retângulo que contém o polígono da região do veículo, e mostra a sua forma de modo mais simples. O retângulo carrega a informação mais básica das dimensões do veículo, ou seja, a altura e o comprimento. Também a silhueta, aparentemente simples, tem informação sobre a área sobre o contorno, o qual é forma externa principal do veículo.



Figura 58: Descritor Silhueta.

O tratamento final desses descritores é realizado através da vetorização da imagem resultante de cada um deles. Assim os valores do vetor descritor são os mesmos valores de tonalidade da imagem resultante, a qual pode ser cinza (cinza) ou binária (binário, borda, silhueta, retângulo). Nesses casos o recorte de um veículo está disposto em formato matricial de tamanho $m \times n$ que deve ser vetorizado para o tamanho $1 \times p$, no qual $p = m \times n$, e p deve ser constante para todos os exemplos. Porém, neste estudo os recortes de veículos possuem tamanhos diferentes (ex.: carro e ônibus) devem ser redimensionados para um tamanho fixo antes da vetorização.

Redimensionamento

Além da necessidade de padronização de tamanho, uma desvantagem dos descritores regionais é que o tamanho grande do vetor extraído causa uma alta complexidade no treinamento do classificador. Isso porque um descritor regional resulta em uma imagem e o seu tamanho total (altura $\times$ largura) representa a dimensão do espaço numérico que o classificador deve analisar. Deste modo, por exemplo, um recorte de um caminhão pequeno

(Figura 54) de tamanho 110x116 é representado por um vetor de dimensão 1x12760 para esses descritores.

Para compensar esse problema o recorte do veículo é redimensionado para um tamanho menor, desde que se mantenha a informação necessária à classificação, e somente após isso extrair os descritores. Isto oferece a vantagem de diminuir a complexidade e o custo operacional do processo posterior de extração, pois primeiramente é executado o redimensionamento que é mais simples se comparado aos métodos de extração e de PCA. Além disso, os métodos de redimensionamento têm naturalmente o objetivo de reduzir as dimensões mantendo o aspecto visual do original.

O redimensionamento também é simples de ser executado porque é aplicado sobre imagem individual, independente de outras do conjunto. Isso é uma vantagem em relação a outros métodos como o PCA, que também pode ser usado para redimensionamento. Porém, o PCA depende da análise global de todas as imagens do conjunto e de operações de muita complexidade em matrizes numéricas muito grandes.

A partir destas necessidades são realizados dois tipos de redução de tamanho e padronização de dimensões, mantendo ou não a relação de aspecto do veículo (*aspect ratio*). Neste tipo, o recorte é dimensionado sem relação de aspecto para a forma quadrada, ou seja, altura e largura de mesmo valor, independente dos valores originais (Figura 59). Isso resulta na perda da relação de aspecto e a conseqüente deformação visual de veículos que possuam a altura muito diferente do comprimento, assim como são exemplos um ônibus ou caminhão (Figura 60). Outro efeito indesejado é que os veículos originalmente muito distintos passam a ter tamanhos muito próximo como, por exemplo, carro e ônibus.

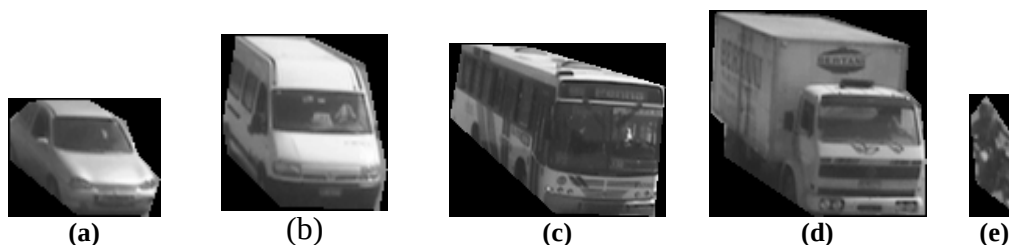


Figura 59: Exemplos de recortes originais.
(a) Carro; (b) Van; (c) Ônibus; (d) Caminhão; (e) Moto.

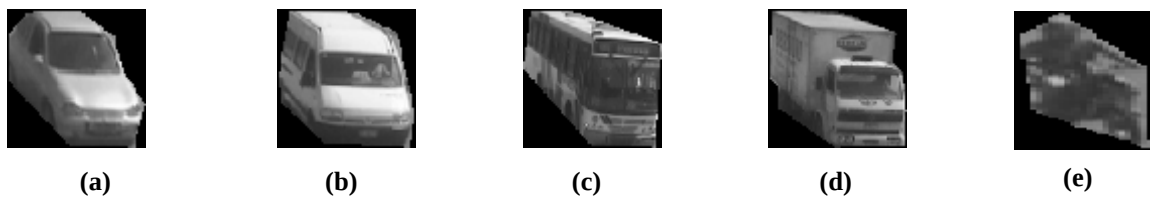


Figura 60: Exemplos de redimensionamento sem relação de aspecto.
(a) Carro; (b) Van; (c) Ônibus; (d) Caminhão; (e) Moto.

No outro tipo, a relação de aspecto é mantida através do redimensionamento do veículo, reduzindo na mesma proporção à altura e largura. Porém, para padronização de tamanho é necessária a sobreposição do recorte resultante num fundo preto de tamanho fixo, o qual deve ser grande o suficiente para encaixar todos os tamanhos de veículos. Porém isto causa um aumento da área de fundo, já que esta é baseada no tamanho do maior veículo. Neste caso pode ser importante centralizar os veículos em relação ao centro de massa (centróide) para padronizar a posição dos veículos em relação ao fundo preto e evitar grandes translações nos vetores descritores (Figura 62).

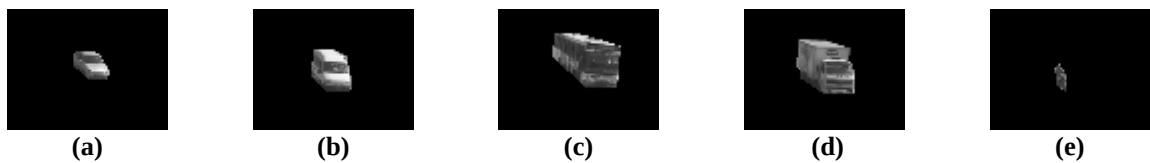


Figura 61: Exemplos de redimensionamento com relação de aspecto.
(a) Carro; (b) Van; (c) Ônibus; (d) Caminhão; (e) Moto.

Sendo assim, para a escolha das dimensões ideais, o recorte cinza original (Figura 54.b) foi redimensionado para vários tamanhos relativos ao original da imagem completa (320x240), com manutenção da relação de aspecto (Figura 62) e sem manutenção da mesma (Figura 63). O método de interpolação é o *nearest*, por ser o menos complexo.

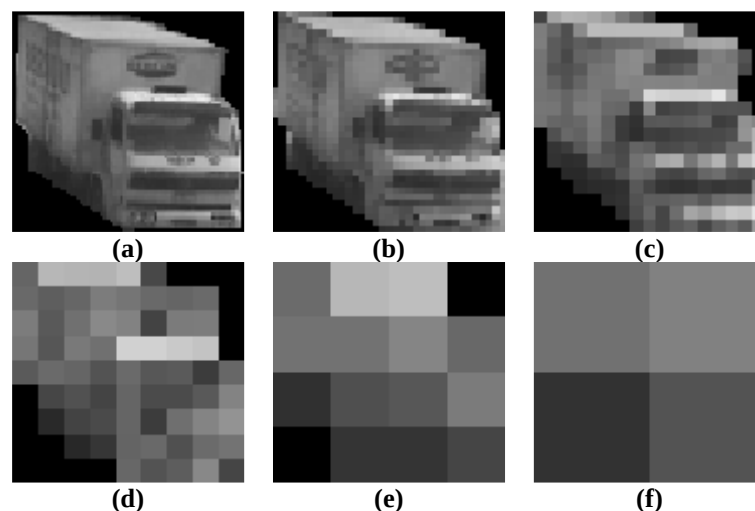


Figura 62: Recortes cinza redimensionados sem relação de aspecto.
(Porc. – Escala Imagem – Tamanho Vetor). Para facilitar a visualização da diferença de resolução e detalhes, após a redução o recorte é dimensionado (*zoom*) para uma escala visível.
(a) 25% - 69x69 – 4761; (b) 12,5% - 35x35 – 1225; (c) 6,25% - 17x17 – 289;
(d) 3,13% - 9x9 – 81; (e) 1,56% - 4x4 – 16; (f) 0,78% - 2x2 – 4.

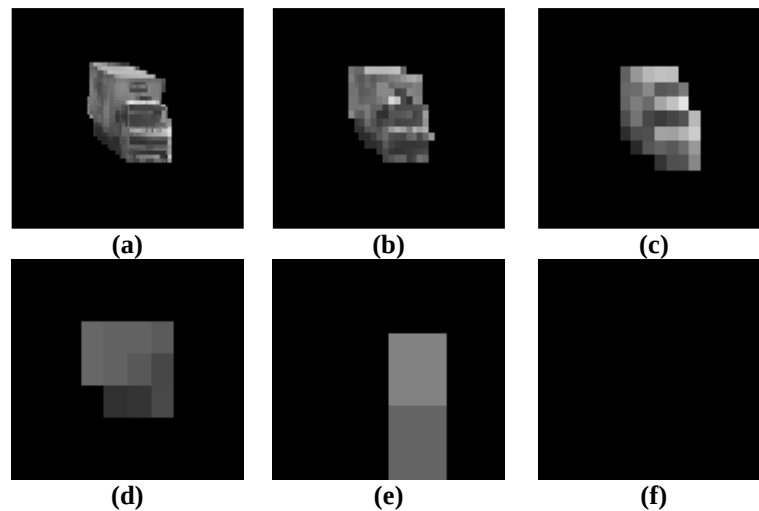


Figura 63: Recortes cinza redimensionados com relação de aspecto.
(Porc. – Escala Imagem – Tamanho Vetor). Para facilitar a visualização da diferença de resolução e detalhes, após a redução o recorte é dimensionado (*zoom*) para uma escala visível.

(a) 25% - 60x80 – 4800; (b) 12,5% - 30x40 – 1200; (c) 6,25% - 15x20 – 300;
(d) 3,13% - 8x10 – 80; (e) 1,56% - 4x5 – 20; (f) 0,78% - 2x2 – 4.

Pode-se observar que até a escala 6,25% (Figura 62.c, Figura 63.c) a forma do objeto ainda se mantém, porém a partir desta para menor ela começa a ser distorcida. É importante destacar que tal escala tem dimensão de aproximadamente 1/40 da original, reduzindo assim a complexidade do vetor de 12760 para apenas 289 ou 300. Nota-se o mesmo em outros descritores regionais reduzidos (Figura 64).

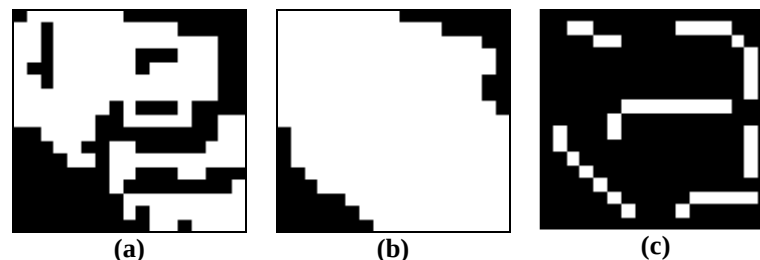


Figura 64: Recortes de outros descritores redimensionados sem proporção em 6,25%
(a) Binário; (b) Silhueta; (c) Bordas.

4.3.Descritores de Cor

A cor de um veículo pode ser classificada de acordo com o seu valor tonal. Basicamente pode ser usado um ponto que represente a cor predominante do veículo. Deste ponto são extraídos os valores de acordo com o modelo de cores escolhido para a representação.

Neste estudo as imagens adquiridas são representadas em uma matriz formada pelos três planos RGB, da qual são extraídos os valores referentes ao ponto marcado como sendo o representante da cor predominante do veículo.

Deste modo, é verificada a classificação a partir dos valores tonais da sua cor em RGB e também os valores resultantes de transformações para outros espaços de cores. As transformações adicionais utilizadas são HSI, Lab e Ycbr. Em todos estes espaços de cores a dimensão do vetor extraído é fixa em três, pois é representado pelos valores tonais do ponto marcado como cor predominante durante a marcação manual. Apesar do tamanho reduzido deste vetor, são utilizadas as técnicas *Z-score* e PCA para descorrelacionar os dados.

4.3.1. Modelo RGB

O modelo RGB (*Red, Green, Blue*) é composto pela mistura das cores primárias vermelho, verde e azul. Os três parâmetros deste modelo definem um espaço tridimensional com eixos ortonormais, no qual as coordenadas são definidas na faixa $[0,1]$. Sendo assim, cada uma das cores primárias corresponde a um dos vértices do cubo localizado sobre os eixos do espaço (Figura 65). Neste modelo as diferentes cores são pontos dentro deste cubo, os quais são definidos como vetores se estendendo a partir da origem. A diagonal do cubo, que possui igual quantidade de cada cor primária, representa os níveis de cinza que são distribuídos desde o ponto de origem $(0,0,0)$ (preto) ao ponto extremo $(1,1,1)$ (branco) (Gonzalez & Woods, 1992).

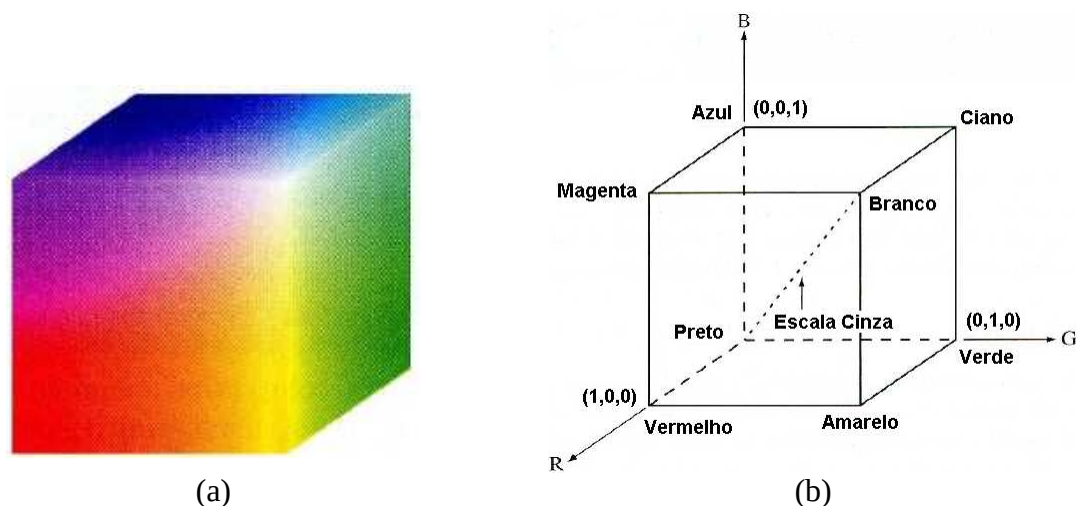


Figura 65: Modelo RGB.
(a) Cubo de cores; (b) Planos de cores.

O modelo RGB é ideal para a geração de imagem colorida, seja na aquisição de imagem, assim como na visualização em monitor de vídeo. Uma imagem RGB pode ser adquirida usando três filtros respectivamente sensíveis às cores, vermelha, verde e azul. Quando um destes filtros é exposto a uma cena, o resultado é uma imagem monocromática cuja intensidade é proporcional à resposta do filtro em relação à luz recebida. Repetindo este

processo para cada filtro são produzidas três imagens monocromáticas correspondente aos componentes RGB da cena, os quais combinados resultam na imagem colorida.

4.3.2. Modelo HSI

O modelo RGB não descreve as cores em termos práticos para a interpretação humana. Por exemplo, ninguém se refere à cor de um automóvel fornecendo a porcentagem das cores primárias que a compõem. Quando um ser humano vê um objeto colorido, ele o descreve através de sua tonalidade (*hue*), saturação e brilho. A tonalidade é o atributo de cor que descreve uma cor pura (ex.: vermelho), na qual a saturação fornece uma medida do grau no qual a cor pura está diluída na luz branca. Brilho é um descritor subjetivo que é complexo de se medir. Este descritor incorpora a noção acromática de intensidade (níveis de cinza), o qual é um dos termos utilizados na descrição da sensação de cor. Esta intensidade é mensurável e de fácil interpretação. O modelo HSI (*Hue, Saturation, Intensity*) desacopla o componente de intensidade das informações que carregam a descrição de cor (tonalidade e saturação) de uma imagem. Desde modo, o HSI é a ferramenta ideal para desenvolver algoritmos baseados nas descrições de cores que são naturais e intuitivas aos seres humanos. Este modelo pode ser construído de acordo com planos triangulares (Figura 66).

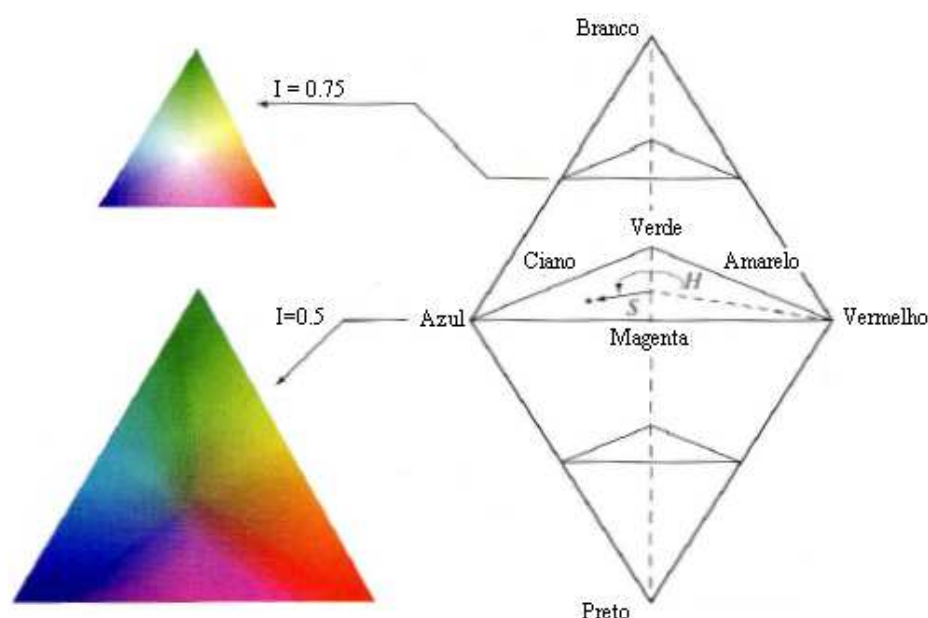
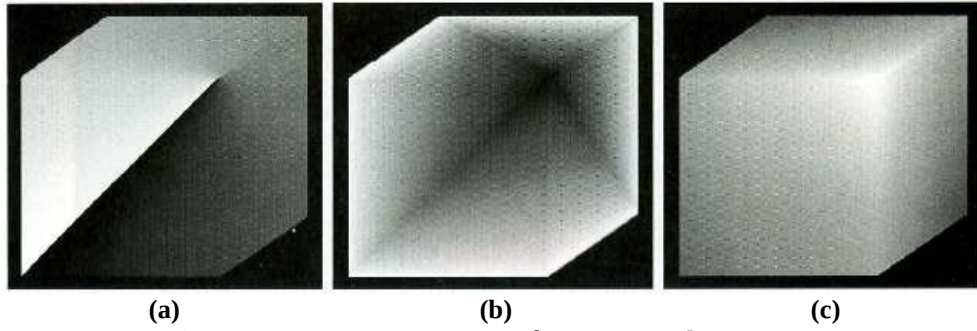


Figura 66: Modelo HSI baseado em planos triangulares.

Comparando os planos ao modelo, podem ser extraídas as imagens que representam os componentes HSI (Figura 67) referentes à representação do cubo RGB (Figura 65).



(a) (b) (c)
Figura 67: Componentes HSI referentes ao cubo RGB.
 (a) H (tonalidade); (b) S (satura  o); (c) I (intensidade).

A partir de uma imagem no formato RGB, o componente H de cada ponto   obtido usando a equa  o (Equa  o 1).

$$H = \begin{cases} \theta & \text{if } B \leq G \\ 360 - \theta & \text{if } B > G \end{cases}, \text{ onde: } \theta = \cos^{-1} \left\{ \frac{\frac{1}{2}[(R-G) + (R-B)]}{[(R-G)^2 + (R-B)(G-B)]^{1/2}} \right\}$$

Equa  o 1: Componente H do modelo HSI.

O componente S   fornecido por (Equa  o 2).

$$S = 1 - \frac{3}{(R+G+B)} [\min(R, G, B)]$$

Equa  o 2: Componente S do modelo HSI.

Finalmente o componente I obtido por (Equa  o 3).

$$I = \frac{1}{3}(R+G+B)$$

Equa  o 3: Componente I do modelo HSI.

Na implementa  o destas equa  es   utilizada a fun  o *rgb2hsi* descrita em [MOUTINHO, 2005].

4.3.3. Modelo YCbCr

A transmiss  o de televis  o utiliza a codifica  o YCbCr como m  todo de representa  o de cores. Neste sistema, a cor de um ponto   representada por uma componente de lumin  ncia, representada por Y, e por duas componentes de cor Cb e Cr.

A compatibilidade com os sistemas preto e branco   uma das vantagens da utiliza  o deste sistema de cores, pois na componente Y h  a informa  o da figura em tons de cinza. Informa  es sobre a sensibilidade do sistema visual humano s o usadas como par  metros para a obten  o da lumin  ncia Y, que se refletem nas constantes “0.59” para verde, “0.30”

para vermelho e “0.11” para azul. As equações mostradas logo abaixo definem este espaço de cores, a partir do espaço RGB.

$$Y = 0,59G + 0,30R + 0,11B$$

Equação 4: Componente Y do modelo YCbCr.

$$Cb = B - Y$$

Equação 5: Componente Cb do modelo YCbCr.

$$Cr = R - Y$$

Equação 6: Componente Cr do modelo YCbCr.

Na implementação destas equações é utilizada a função *rgb2ycbcr*.

4.3.4. Modelo Lab

O espaço de cores LAB separa uma imagem em três componentes, onde L representa a luminosidade ou brilho, e as componentes A e B representam parcelas de cores. O espaço LAB busca linearizar a perceptividade de cores, de forma a fazer as variáveis L, A e B reproduzirem a escala logarítmica do olho humano. A vantagem do sistema LAB é a que a representação de brilho da variável L está mais adaptada à forma com que o olho humano percebe a luminosidade.

A conversão de RGB para LAB necessita de uma conversão prévia para o espaço de cores XYZ como

$$\text{mostrado na equação } \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0,49018626 & 0,30987954 & 0,19993420 \\ 0,17701522 & 0,81234418 & 0,01066060 \\ 0 & 0,01007720 & 0,98992280 \end{bmatrix} * \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

Equação 7). As equações (Equação 8), (Equação 9) e (Equação 10) mostram a conversão de XYZ para o formato LAB, a qual inclui uma transformação cúbica representada pela função $f(w)$ (Equação 11).

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0,49018626 & 0,30987954 & 0,19993420 \\ 0,17701522 & 0,81234418 & 0,01066060 \\ 0 & 0,01007720 & 0,98992280 \end{bmatrix} * \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

Equação 7: Conversão de RGB para XYZ.

$$L = \begin{cases} 116 \left(\frac{Y}{Y_0} \right)^{1/3} & \text{para } \left(\frac{Y}{Y_0} \right) > 0,008856 \\ 903,3 \left(\frac{Y}{Y_0} \right) & \text{para } \left(\frac{Y}{Y_0} \right) \leq 0,008856 \end{cases}$$

Equação 8: Componente L do modelo Lab.

$$A = 500 \left(f \left[\frac{X}{X_0} \right] - f \left[\frac{Y}{Y_0} \right] \right)$$

Equação 9: Componentes A do modelo Lab.

$$B = 200 \left(f \left[\frac{X}{X_0} \right] - f \left[\frac{Y}{Y_0} \right] \right)$$

Equação 10: Componentes B do modelo Lab.

$$f(w) = \begin{cases} w^{1/3} & \text{para } w > 0,008856 \\ 7,787w + 0,1379 & \text{para } 0 \leq w \leq 0,008856 \end{cases}$$

Equação 11: Transformação $f(w)$ para componentes A e B do modelo Lab.

Nas equações logo acima, os valores de X_0 , Y_0 e Z_0 representam o branco padrão, escolhido como o D65 neste trabalho, ou seja, uma temperatura de cor de 6500K, muito usado em monitores e câmeras. Os valores para este branco são $X_0=0,950456$, $Y_0=1$ e $Z_0=1,088754$.

4.4.Classificação Neural

O modelo neural foi escolhido para abordar a classificação de categoria, assim como da cor predominante de veículo. Esta escolha foi feita com base na complexidade do problema e as características das redes neurais.

Um exemplo de aplicação de RNA é [PRAMPERO, 1999] que utiliza MLP para classificar categorias através de descritores geométricos, obtendo 88% de acerto.

Outro motivo para escolha deste classificador é que este trabalho faz parte do grupo de pesquisa LABIC, o qual aplica RNA em aplicações de classificação como, por exemplo, reconhecimento de caracteres e detecção de faces. Especificamente na área de ITS, este trabalho complementa o anterior de que realiza leitura e reconhecimento de placas automotivas.

Em relação às características de arquitetura da RNA, foi identificada necessidade de uso da análise PCA para redução de dimensões e descorrelação de dados nos vetores descritores. Isso é realizado para todos os descritores de categoria e cor.

Descorrelação

Apesar das reduções, a dimensão do vetor de descritores regionais em torno de 300 ainda é grande e, além disso, os componentes dos vetores podem estar altamente correlacionados (redundantes). Isso é devido à natureza dos descritores regionais que resultam em imagens a partir de transformações das originais.

Um procedimento efetivo para executar esta operação é a Análise de Componente Principal (PCA) [JOLLIFFE, 1986]. Esta técnica tem três efeitos: ortogonaliza os componentes dos vetores (de forma que eles fiquem descorrelacionados); ordena os componentes ortogonais resultantes (componentes principais) de forma que os de maior variação venham em primeiro lugar; e elimina os componentes que contribuem menos para essa variação no conjunto dos dados.

Na ortogonalização desses vetores foi adotada a técnica de normalização conhecida como *Z-score* [JOHNSON; WICHERN, 1992]. Esta técnica apresenta como resultado o número de desvios-padrão que cada elemento do conjunto de dados está distante da média do conjunto. Para isso os dados são normalizados para que suas médias sejam zero e os desvios padrões (e variância) unitários. A equação (Equação 12) mostra como é feito este cálculo, o qual é implementado no Matlab pelo procedimento *prestd*.

$$pn = \frac{(p - meanp)}{stdp}$$

Equação 12: Normalização Z-score em *prestd*.

Onde,

pn = vetor de entrada normalizado;

p = vetor de entrada;

meanp = média dos vetores de entrada;

stdp = desvio-padrão dos vetores de entrada.

Deste modo, primeiramente os vetores descritores são normalizados com *Z-score*, de forma que eles tenham média zero e variância unitária. Num segundo passo, é utilizada a análise PCA (*prepca*, Matlab) com fator de corte 0.005. Isto significa que serão eliminados os componentes principais que contribuem menos que 0,5% à variação total no conjunto de dados. Com isto as informações com pouca significância são eliminadas do conjunto e as restantes, as mais significativas são selecionadas para a classificação. Reduções significativas

são alcançadas após essa análise. Por exemplo, o vetor de descritor geométrico reduz de 10 para aproximadamente 5 (veja seção 4.6). E o descritor cinza reduz de 300 para aproximadamente 25 (veja seção 4.6), o qual se comparado com um veículo de tamanho original 12760 (Figura 54), após redução para 6.25% e tratamento por PCA, representa uma redução de 1/510 no tamanho do vetor descritor.

O modelo de classificação adotado visa classificar a categoria e, a partir desta, classificar a subcategoria. O objetivo é classificar as categorias mais gerais e após identificar as subcategorias que possam ser especializadas das primeiras. Os descritores utilizados são divididos nos tipos distintos regionais e geométricos com o objetivo de determinar o melhor desempenho e razão entre custo e benefício de processamento e acerto.

Para a cor a classificação é realizada diretamente sobre valores de pontos de imagem e busca avaliar o acerto possível de se alcançar, pois esses valores são muito influenciados negativamente por fatores ambientais.

4.5. Estudo de Caso

O estudo de caso é composto por imagens de trânsito adquiridas nas quatro pistas (Figura 68) que compõem a via de acesso ao Posto 3 da UFRJ. As pistas 1 e 2 são de entrada no Campus e as restantes são de saída. As câmeras instaladas fazem parte do projeto de segurança da Prefeitura da UFRJ em convênio com o Instituto de Matemática (IM-/UFRJ). Deste projeto faz parte o Sistema Kapta Acesso que utiliza as fotos para localizar automaticamente os veículos e identificá-los através de leitura da placa.

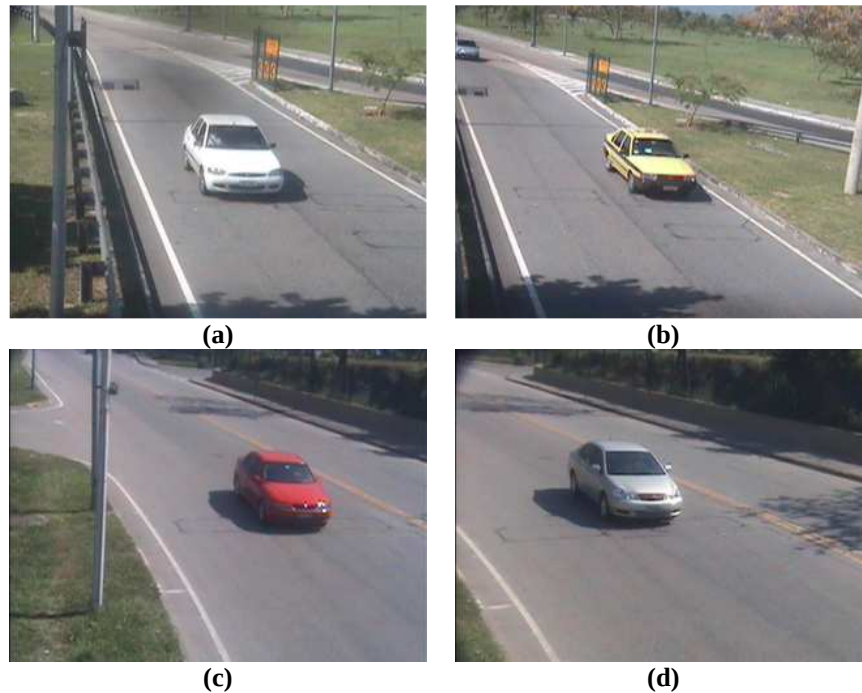


Figura 68: Imagens de veículos nas pistas do Posto 3 de acesso à UFRJ.
(a) Pista 1; (b) Pista 2; (c) Pista 3; (d) Pista 4.

A visão da câmera é diagonal e abrange parcialmente os ângulos frontal, lateral e superior. Esta visão abrange vários veículos em uma imagem, porém o objetivo é focado apenas no veículo que está próximo ao laço magnético e originou a aquisição (Figura 69). A iluminação é natural, nos períodos da manhã e tarde. As imagens foram adquiridas no dia 15 de setembro de 2006 no horário entre 09h e 17h.

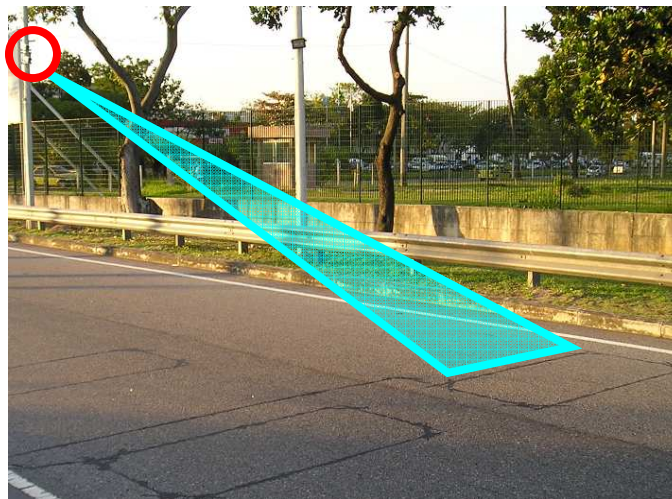


Figura 69: Posição da câmera 1 (círculo) relativa ao laço magnético (retângulo na pista).

Quanto ao ambiente é considerado parcialmente controlado, pois a posição do veículo é esperada próxima ao laço magnético. Porém, alguns fatores afetam o desempenho do classificador. Um desses é fundo complexo das imagens, ou seja, que inclui pavimento,

vegetação, postes, grades e outros veículos. Também a via possui duas pistas e tráfego livre e, conseqüentemente, múltiplos veículos na mesma imagem e oclusão entre eles (Figura 70).



Figura 70: Pistas com múltiplos veículos e oclusão entre eles.

Outro fator é observado em categorias com diferença de escala no mesmo conjunto como, por exemplo, o caso do ônibus ou *van* que variam de acordo com os modelos. Assim como a pequena mudança de escala nos mesmos modelos de veículos que são fotografados em distâncias diferentes devido à variação de velocidade e posição (Figura 71).



Figura 71: Mudança de escala devido a variação de distância da câmera.
(a) Ônibus em escala 91x83; (b) Ônibus em escala 105x93;

Além disso, o ângulo de aquisição que é composto da visão lateral e frontal é afetado por pequenas variações na do veículo que resultam em rotação de sua forma e conseqüente deformação. Isso é característico da via que possui duas pistas e não possui separação física entre elas, assim um veículo na trajetória correta e outro deslocado podem ser fotografados pela mesma câmera (Figura 72).



Figura 72: Deformação de veículos devido ao deslocamento de trajetória.
(a) Trajetória correta; (b) Trajetória deslocada.
(a linha pontilhada central é referência para mostrar o deslocamento de trajetória)

Outro fator é a cor, a qual varia em veículos da mesma categoria (Figura 73). Além disso, a iluminação natural varia de acordo com o ângulo de incidência solar, o que resulta em

efeitos indesejados como alteração de tonalidades, sombra e reflexo. Esses fatores podem influenciar nos valores extraídos por descritores dependentes da tonalidade.



Figura 73: Variações de cor na mesma categoria.

A forma dos veículos entre os exemplos da mesma categoria também é um fator determinante. Analisando de modo geral as categorias, o ônibus é bem definido na forma e possui poucas variações como, por exemplo, ônibus e micro ônibus. A moto também tem uma forma bem definida. Já a *van* possui algumas variações, mas também é definida em sua forma média. Porém o carro, apesar de ter dimensões aproximadamente uniformes, é heterogêneo quanto à forma dos modelos. E a categoria caminhão apresenta formas muito heterogêneas e possui pouca uniformidade quanto à aparência e ao tamanho.

4.5.1. Categorias e Subcategorias

Existem várias categorias de veículos trafegando nas vias brasileiras. Porém não são encontrados exemplares de todas as categorias neste estudo de caso. Sendo assim, neste contexto devem ser escolhidas e reconhecidas somente as categorias entre as encontradas no conjunto de imagens. Também são usadas como referências as categorias citadas em trabalhos e nas leis de trânsito para determinar as categorias deste estudo de caso.

Deste modo, de acordo com o [DENATRAN, 2006], as categorias da frota brasileira de veículos são as seguintes: automóvel, bonde, caminhão, caminhão-trator, caminhonete, camioneta, chassi-plataforma, ciclo motor, ônibus, micro-ônibus, motocicleta, motoneta, triciclo, quadriciclo, reboque, semi-reboque, side-car, trator de esteira, trator de rodas, utilitário e outros.

Por sua vez, as principais categorias utilizadas nos trabalhos citados na seção (3.5.1) são: carro (*sedan*, *hatchback*, *station wagon*, *sallon*), camionete, *van*, *van* alta, caminhão, caminhão pequeno, *trailer*, jipe, moto, *lorry*, bicicleta, moto e ônibus (urbano, interurbano, micro, de dois andares).

Então foi realizada a inspeção visual no conjunto de imagens das quatro pistas. Durante esse processo os veículos foram divididos nas categorias dominantes e ao final da inspeção foram identificadas as seguintes: ônibus, *van*, carro, caminhão e moto (Figura 74).



Figura 74: Exemplos de categorias de veículos.

No agrupamento geral de imagens, a categoria carro (27947 exemplos) é predominante em relação às outras (4920 exemplos), representando 85% do total. Sendo assim, o número de exemplos selecionados foi reduzido (Tabela 16) para se manter uma proporção aproximada em relação às outras categorias (Tabela 17). Porém a categoria moto foi encontrada somente na pista três, devido à configuração do módulo de aquisição (sensibilidade muito alta de captura do sinal do laço magnético). Nas seguintes tabelas o grupo Global é composto pela união das imagens das quatro pistas.

**Tabela 16: Redução de quantidade de exemplos de carros.
(grupo Global é composto pela união das imagens das quatro pistas)**

Grupo	Carros Originais	Carros Selecionados	Porc. dos selecionados em relação aos originais
Pista 1	7346	570	7,76
Pista 2	6516	537	8,24
Pista 3	7695	810	10,53
Pista 4	6390	540	8,45
Global	27947	2457	8,79

Tabela 17: Total de exemplos por pista e categoria (número de exemplos).

Categ. Grupo	Caminhão	Carro	Ônibus	Van	Moto	Total
Pista 1	210 (12 %)	570 (34 %)	351 (21 %)	555 (33 %)	0 (%)	1686
Pista 2	120 (8 %)	537 (35 %)	627 (40 %)	261 (17 %)	0 (%)	1545
Pista 3	486 (18 %)	810 (30 %)	525 (19 %)	684 (25 %)	213 (8 %)	2718
Pista 4	174 (12 %)	540 (38 %)	396 (28 %)	318 (22 %)	0 (%)	1428
Global	990 (13 %)	2457 (33 %)	1899 (26 %)	1818 (25 %)	213 (3 %)	7377

Os grupos de modelos de veículo são muito diversificados e não foram identificados por análise visual superficial. O agrupamento aproximado dos modelos foi realizado durante o processo manual de classificação, no qual foram adicionados na medida em que surgiam.

Nesta análise foram agrupados os modelos dominantes em 26 variações de caminhões, 7 vans, 4 ônibus, 1 moto e 30 carros (Figura 75).

A partir desses modelos foram definidas algumas subcategorias agrupadas de acordo com a semelhança aproximada de forma. Todas as categorias foram subdivididas, com exceção do carro que possui modelos muito similares entre si que tornam complexa a tarefa de agrupamento. Assim, os modelos foram agrupados a partir do grupo global de imagens em 11 subcategorias de acordo com a (Tabela 18) (Figura 75).

Tabela 18: Totais de subcategorias no grupo global

Código	Subcategoria	Total
1	Caminhão de Carroceria Média	219
2	Caminhão de Carroceria Grande	84
3	Caminhão de Baú Médio	381
4	Caminhão de Baú Grande	54
5	Caminhão de outro tipo	234
6	Van Pequena	726
7	Van Média	1089
8	Ônibus Médio	414
9	Ônibus Grande	1482
	Total	4683



Figura 75: Subcategorias de veículos, de acordo com os seus respectivos códigos.
(as dimensões dos modelos foram normalizadas por motivo de visualização) (os carros foram omitidos)

Durante o processo manual de marcação algumas imagens foram descartadas por motivo de pouca qualidade e consideradas inapropriadas para o estudo (Figura 76).



Figura 76: Imagens descartadas por falta de qualidade.

4.5.2. Cores

Do mesmo modo em que foram definidas as categorias, a escolha das cores deve ser de acordo com o contexto. Assim, foi selecionada o grupo da pista 1 como referência e a partir da sua inspeção visual os veículos foram separados através das similaridades tonais.

Foram observados superficialmente 8533 veículos e divididos em sete grupos dominantes de cores (Figura 77.a).

Para comparar e validar a escolha das cores, foi analisado o BD do sistema SaproS que contém um cadastro de 11808 veículos pertencentes a usuários da UFRJ. Neste cadastro os proprietários classificaram os seus veículos de acordo com 17 cores. Este grupo de cores inicialmente não era fixo e foi determinado pelos próprios usuários na medida em que realizavam os seus cadastros. Porém de acordo com as maiores distribuições foram agrupadas em apenas oito (Figura 77.b), sendo que as restantes de menor influência foram unidas no grupo “outras” (vinho, amarela, dourada, creme, marrom, roxo, grená, laranja, rosa, fantasia).

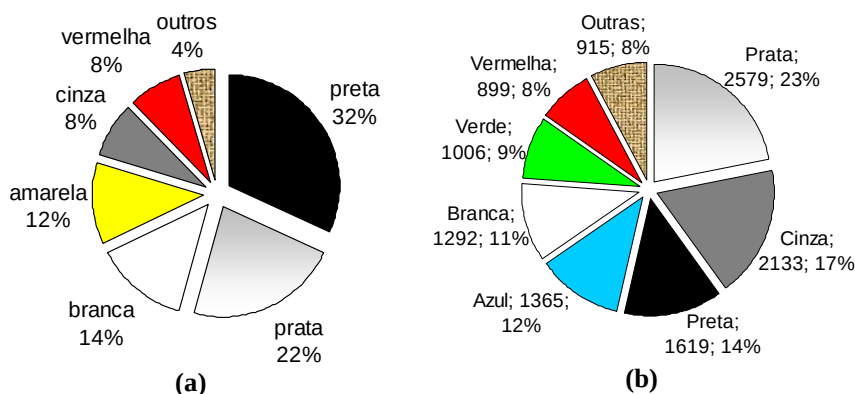


Figura 77: Distribuição de referência para cores.
(a) Inspeção visual no grupo Pista 1: 8.533 veículos;
(b) Cadastro UFRJ: 11.808 veículos (situação em agosto/2006) ;

A partir dessas duas referências foram analisadas as distribuições e escolhidas as 8 cores (Figura 78) a serem utilizadas no estudo de caso para reconhecimento da cor predominante.

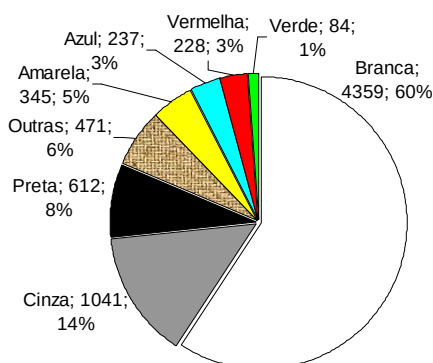


Figura 78: Distribuição de referência para cores no grupo Global.

4.6. Configuração e Treinamento dos Classificadores

Nos testes são utilizadas algumas estratégias gerais e outras específicas para categoria ou cor. Os testes são aplicados individualmente em cada grupo de imagens, inclusive no global. Nesses testes é verificada a capacidade das estratégias em tratar veículos com pequenas variações de escala, tonalidade, posição e ângulo. Aqui é descrita uma configuração básica a ser utilizada e também as possíveis variações que possam influenciar no desempenho do classificador.

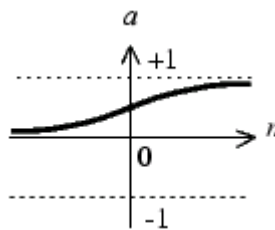
A partir disso um grupo de imagens foi selecionada para testar variações na extração de descritores e no treinamento do classificador. A pista 3 foi escolhida porque contém o maior número de exemplos e também por ser a única que contém a categoria de motos.

O classificador usado neste trabalho é uma Rede Neural do tipo MLP com três camadas e a função de transferência *logsig* (Matlab). Também foram realizados testes com outras funções de transferência (*tansig*, *hardlin*, ...). Os pesos (e *biases*) de cada camada são inicializados com *initnw* (Matlab). O treinamento é realizado com a função de retro-propagação *traingdx* (Matlab). O erro deste é medido de acordo com a função de erro MSE.

A função de transferência *logsig* (sigmóide logarítmica) [VOGL, et al, 1988] calcula a saída de uma camada a partir sua entrada através da fórmula (Equação 13) que gera o gráfico da (Figura 79), com valores entre zero e um (excluindo estes).

$$f(n) = \frac{1}{1 + e^{-n}}$$

Equação 13: Função de transferência logsig.



$$a = f(n)$$

Figura 79: Gráfico da função de transferência logsig.

Traingdx é uma função de treinamento do tipo retro-propagação (*backpropagation*) que atualiza valores de peso (e *bias*) de acordo com gradiente descendente, constante de momento e taxa de aprendizagem adaptativa. Os seguintes parâmetros são utilizados para esta função:

- Número de máximo de ciclos para treinar (*epochs*): 1000 (padrão sugerido);
- Meta de erro (*goal*): 0 (para possibilitar o alcance do menor erro possível);
- Taxa de aprendizagem (*lr*): 0.01 (padrão sugerido);
- Proporção para aumentar a taxa de aprendizagem (*lr_inc*): 1.05 (padrão sugerido);
- Proporção para diminuir a taxa de aprendizagem (*lr_dec*): 0.7 (padrão sugerido);
- Aumento máximo de erro (*max_perf_inc*): 1.04 (padrão sugerido);
- Constante de momento (*mc*): 0.9 (padrão sugerido);
- Gradiente de erro mínimo (*mingrad*): $1e^{-10}$ (padrão sugerido);
- Tempo de máximo para treinar (*time*): infinito (sem limite de tempo)

Para cada ciclo do treinamento, se o erro diminuir em direção à meta, então a taxa de aprendizagem é incrementada pelo fator *lr_inc*. Senão, se o erro aumentar por mais do que um fator *max_perf_inc*, a taxa de aprendizagem é decrementada pelo fator *lr_dec* e a mudança que aumentaria o erro não é realizada. O treinamento pára quando qualquer uma das seguintes condições acontece:

- O número máximo de ciclos (*epochs*) é alcançado;
- O tempo máximo (*time*) foi excedido;
- Erro foi minimizado à meta (*goal*);
- O gradiente de erro cai abaixo do mínimo (*mingrad*);

A função *initnw* implementa o método de [NGUYEN; WIDROW, 1990] que gera valores iniciais de peso (e *bias*) para cada camada, de forma que as regiões ativas dos neurônios serão distribuídas de forma aproximadamente uniforme no espaço da entrada. Desde modo poucos neurônios são desperdiçados e o treinamento acontece mais rapidamente, pois cada área do espaço de entrada possui neurônios.

É utilizada uma única camada oculta de neurônios. Como o problema em questão é de classificação de padrões, segundo [MAREN, 1999] a utilização de apenas uma camada oculta é considerada suficiente. Quanto ao número de neurônios na camada escondida, foi adotada a média geométrica entre o número de entradas e de saídas da rede [MAREN, 1999]. Serão realizados outros testes com duas e três camadas ocultas.

Também é utilizada a separação dos dados em dois grupos para treinamento e teste (simulação). A separação inicial é de 70% para treino e 30% para teste, os quais são agrupados separadamente em cada classe. Isso é realizado para garantir a mesma representatividade de cada categoria compondo o grupo de treinamento pela mesma porcentagem de cada categoria. Com isso é evitada a ocasional predominância de alguma classe caso os dados fossem separados a partir do grupo global e na mesma ordem em que as imagens foram capturadas pela câmera. Também serão realizados testes com a redução no

número de exemplos de treino em 45% e 20%, conseqüentemente aumentando a quantidade de teste para 55% e 80%.

Além disso, no grupo global de treinamento é aplicada ordenação aleatória para misturar os exemplos de várias classes. Isso é realizado porque este conjunto global inicialmente está em ordem de classe, ou seja, primeiramente todos os exemplos de uma determinada classe, após todos de outra classe e, assim por diante. Esta ordenação pode influenciar no desempenho do treinamento.

Na camada de saída da RNA é colocada a quantia n de neurônios igual ao de classes e os valores alvos em formato binário (Tabela 19). Isso melhora a capacidade de separação entre as classes, já que são utilizados vetores ortonormais como alvos no treinamento.

Tabela 19: Valores de alvos binários.

Classe	Valores alvos
1	1 0 0 0 0
2	0 1 0 0 0
3	0 0 1 0 0
:	:
N	0 . . . 1

Esses valores de saída da simulação de rede neural são interpretados de acordo com a estratégia *Winner Takes All* (WTA), onde o maior valor entre todas as saídas é alterado para um (1) e o restante é alterado para zero (0). Deste modo os resultados do teste são analisados para revelar o valor de maior peso na resposta, melhorando a descoberta e interpretação de resultados muito próximos ao alvo. Dessa forma também se tem índice zero de rejeição. É calculado o erro MSE entre as respostas da simulação e as novas respostas geradas pelo WTA para estimar a magnitude das alterações que foram necessárias para normalizar os valores para 0 e 1. Este erro, em conjunto com os erros de treino e simulação, são utilizados para medir a qualidade da rede neural. Também serão realizados testes com interpretação de códigos alvos naturais e sem WTA.

Em relação ao modelo proposto de RNA, são também realizados testes com variação de em alguns parâmetros que influenciam no desempenho do treinamento. Os parâmetros são: algoritmo de treinamento, função de transferência, número de camadas ocultas, interpretação de saída.

A redução e descorrelação por análise PCA é utilizada com fator de corte 0,5% conforme exemplos encontrados na literatura, porém são testados alguns fatores como 10%, 5%, 1%, 0% e 0,25%.

Dentre as diferentes configurações avaliadas os melhores resultados foram obtidos com os seguintes parâmetros: fator PCA é “0.005” para os descritores regionais de categoria e para os tons de cores. Algoritmo de treinamento *trainidx*, função de transferência *logsig* e o número de camadas ocultas igual a um (1) para os classificadores. E interpolação *nearest neighbor* (vizinho mais próximo) para o algoritmo de redimensionamento.

Os resultados são comentados neste capítulo, porém as tabelas completas são mostradas nos apêndices deste trabalho.

4.6.1. Classificação de Categoria

Para realizar os testes foram selecionadas duas vertentes principais de acordo com o grupo de descritores. Também foram analisadas algumas variações de configuração no método para verificação de suas influências e validação das escolhas iniciais.

Para os descritores regionais a estratégia básica constitui em:

- não equalizar;
- manter o fundo original do recorte;
- não centralizar;
- redimensionar sem relação de aspecto.

Em sequência é utilizada outra estratégia, porém a proporção é mantida através da sobreposição do recorte em um fundo preto de tamanho fixo. Essas estratégias são aplicadas sobre os descritores cinza, bordas, silhueta, retângulo e binário.

Já para os descritores geométricos a estratégia adotada constitui em extrair as dez medidas da silhueta do veículo em tamanho original.

Os dados dos descritores são divididos em 70% para treino (Tabela 20) e 30% para teste (Tabela 21).

Tabela 20: Conjunto de treinamento para categoria em 70% do total (número de exemplos).

Grupo \ Categ.	Caminhão	Carro	Ônibus	Van	Moto	Total
Pista 1	147	399	246	389	0	1181
Pista 2	84	376	439	183	0	1082
Pista 3	340	567	368	479	149	1903
Pista 4	122	378	277	223	0	1000
Global	693	1720	1329	1273	149	5164

Tabela 21: Conjunto de teste para categoria em 30% do total (número de exemplos).

Grupo \ Categ.	Caminhão	Carro	Ônibus	Van	Moto	Total
Pista 1	63	171	105	167	0	505
Pista 2	36	161	188	78	0	463
Pista 3	146	243	158	205	64	815
Pista 4	52	162	119	95	0	428
Global	297	737	570	545	64	2213

A variação dos testes é realizada através da execução ou não dos seguintes itens (quando aplicáveis): equalização cinza, centralização no fundo padrão, presença de fundo cinza fixo, tratamento dos dados por PCA, quantidade de imagens para treino/teste e aplicação de ordem aleatória na separação dos dados. Aos geométricos se aplica redimensionamento apenas por PCA, mas não se aplica fundo cinza. O item centralizado se aplica apenas ao regional com relação de aspecto.

A escala selecionada é 6,25% da imagem original que determina as seguintes dimensões para os vetores extraídos dos recortes dos veículos, de acordo com o tipo de redimensionamento (Tabela 22). Também são testadas outras escalas em 12,5% e 3,13%. O método para redimensionamento é o *nearest*, por ser o mais simples, porém também é testado o *bicubic*.

Tabela 22: Dimensões de teste

Redimensionamento	Resolução	Dimensão
Regional Deformado	15x20	300
Regional Não deformado	17x17	289
Geométrico	-	10
Momentos	-	7

Foram inicialmente testados os descritores cinza e os geométricos para cada uma das pistas, inclusive para o grupo global. Isso demonstrou a viabilidade e uniformidade de resultados entre as diferentes pistas.

Tabela 23: Número de neurônios na saída de acordo com a classe.

Classe	Grupo	Saídas
1-Categoria	Pista 1	4
1-Categoria	Pista 2	4
1-Categoria	Pista 3	5
1-Categoria	Pista 4	4
1-Categoria	Global	5
2-Subcategoria	Pista 3	11
2-Subcategoria	Global	11

Uma característica particular do tipo de captura utilizado neste estudo é que cada veículo possui 3 fotos em sequência. Estes veículos possuem uma aparência similar, já que o intervalo entre as suas imagens é de apenas 10ms. Isso poderia ser identificado como um vício no treino, já que poderiam aparecer imagens similares entre os exemplos de treino e os de teste. Sendo assim, para verificar a influência deste fato são realizados treinos com a seleção de apenas uma imagem de cada trio, o que é denominado de “seleção de captura”.

4.6.2. Classificação de Cor Predominante

Para os descritores tonais a estratégia básica consiste em classificar o valor do ponto de cor selecionado na segmentação. Em sequência é utilizada outra estratégia, porém também são selecionados e classificados os pontos da vizinhança de 8 do ponto selecionado e é realizada uma análise de maioria para determinar a cor predominante do conjunto.

Os dados dos descritores são divididos em 70% para treino (Tabela 24) e 30% para teste (Tabela 25).

Tabela 24: Conjunto de treinamento para cores em 70% do total (número de exemplos).

Cor \ Grupo	Branca	Cinza	Preta	Outras	Amarela	Azul	Vermelha	Verde	Total
Pista 1	691	208	113	36	38	34	46	15	1181
Pista 2	668	145	130	34	19	48	25	13	1082
Pista 3	1094	206	97	82	137	63	61	15	1755
Pista 4	596	172	88	29	48	21	27	17	998
Global	3049	731	428	181	242	166	159	60	5016

Tabela 25: Conjunto de teste para cores em 30% do total (número de exemplos).

Cor \ Grupo	Branca	Cinza	Preta	Outras	Amarela	Azul	Vermelha	Verde	Total
Pista 1	296	89	49	15	16	14	20	6	505
Pista 2	286	62	56	14	8	21	11	5	463
Pista 3	469	88	41	35	58	27	26	6	750
Pista 4	256	74	38	13	21	9	12	7	430
Global	1307	313	184	77	103	71	69	24	2148

Os vetores tonais possuem dimensão 3 em qualquer um dos modelos de cor. Foram inicialmente testados os descritores tonais para cada uma das pistas, inclusive para o grupo

global. Isso demonstrou a viabilidade e uniformidade de resultados entre as diferentes pistas. O número de neurônios na camada de saída é constante em 8 para qualquer modelo de cor.

4.7. Testes e Resultados

Nesta seção são apresentados apenas resultados consolidados, sendo que os detalhados estão listados no apêndice deste trabalho.

4.7.1. Classificação de Categoria

Foram realizados basicamente três tipos de experimentos:

- Um modelo neural treinado e testado por pista;
- Um modelo neural treinado com uma pista e testado com as demais;
- Um modelo neural treinado e testado com todas as pistas.

Na série de experimentos envolvendo a construção de um classificador específico para cada pista os resultados foram muito bons (Tabela 26, Tabela 28). Observa-se que os melhores resultados foram alcançados com o uso dos descritores regionais e redimensionamento sem relação de aspecto e fundo preto. Com um índice de acerto aproximadamente 25 pontos percentuais abaixo vê-se aqueles obtidos com os descritores de momento.

Tabela 26: Desempenho para categorias sem relação de aspecto e com fundo preto.

Descritor Grupo	Níveis de Cinza	Binário	Bordas	Momentos
Pista 1	100	98.42	97.43	76.44
Pista 2	100	100	99.35	83.59
Pista 3	99.63	99.51	99.14	69.82
Pista 4	100	99.30	99.53	69.63
Global	99.68	96.02	98.78	64.53
Média	99,86	98,65	98,85	72,80
Desvio padrão	0,19	1,58	0,84	7,37

Tabela 27: Matriz de confusão para descritor cinza deformado para categoria na pista 3.

Alvo Teste	Caminhão	Carro	Ônibus	Van	Moto
Caminhão	145	0	0	1	0
Carro	0	243	0	0	0
Ônibus	0	0	157	1	0
Van	0	0	0	203	0
Moto	1	0	0	0	64

Na (Tabela 27) vê-se a matriz de confusão para a pista 3 em descritores regionais cinzas. As poucas confusões existentes se restringem a falsas classificações de uma van como ônibus e uma van como caminhão. Uma falsa classificação que chama atenção é a de um

caminhão classificado como moto. Isto pode ser parcialmente justificado pelo redimensionamento sem manutenção da relação de aspecto que, como pode ser visto na (Figura 60), pode fazer com que uma moto adquira características de um veículo bem maior.

Tabela 28: Desempenho para categorias sem relação de aspecto e com fundo original.

Grupo	Descritor	Níveis de Cinza	Binário	Bordas	Momentos
Pista 1		100	98.81	95.45	72.48
Pista 2		99.57	99.57	98.49	90.50
Pista 3		99.14	99.26	96.93	68.22
Pista 4		99.77	100	97.43	72.90
Global		98.19	98.42	96.61	56.17
Média		99,33	99,21	96,98	72,05
Desvio padrão		0,71	0,62	1,11	12,33

Tabela 29: Desempenho para categorias sem relação de aspecto.

Grupo	Descritor	Silhueta	Retângulo
Pista 1		93.12	92.28
Pista 2		92	92.22
Pista 3		90	90.55
Pista 4		93	92.76
Global		91	90

Tabela 30: Desempenho para categorias com relação de aspecto e com fundo preto.

Grupo	Descritor	Níveis de Cinza	Binário	Bordas	Momentos
Pista 1		98.81	98.02	94.85	61.58
Pista 2		98.92	98.27	95.25	74.73
Pista 3		98.65	94.85	90.31	60
Pista 4		99.53	98.60	96.73	60.05
Global		84.64	90.47	91.73	59
Média		96,11	96,04	93,77	63,07
Desvio padrão		6,42	3,46	2,66	6,58

Tabela 31: Desempenho para categorias com relação de aspecto e com fundo original.

Grupo	Descritor	Cinza	Binário	Bordas	Momentos
Pista 1		97.82	98.02	96.24	60.00
Pista 2		98.27	98.27	96.11	73.87
Pista 3		97.67	95.83	91.29	61
Pista 4		98.83	98.83	98.13	56.54
Global		95.80	92.36	89.83	60
Média		97,68	96,66	94,32	62,28
Desvio padrão		1,14	2,66	3,56	6,70

Tabela 32: Desempenho para categorias com relação de aspecto.

Grupo \ Descritor	Silhueta	Retângulo
Pista 1	97.23	93.07
Pista 2	98.27	93.95
Pista 3	95.58	90.55
Pista 4	98.13	95.33
Global	97.10	91.15

O descritor geométrico também apresentou resultados similares ao descritor regional cinza para classificação de categorias (Tabela 33). Na (Tabela 35) vê-se a matriz de confusão para a pista 3 com descritores geométricos. Também ocorrem algumas confusões de classificações de caminhão e ônibus como van.

Tabela 33: Desempenho para categorias para descritor Geométrico.

Grupo \ Descritor	Geométrico
Pista 1	98,22
Pista 2	100
Pista 3	98,90
Pista 4	100
Global	82,06
Média	95,84
Desvio padrão	7,74

Tabela 34: Matriz de confusão para descritor geométrico para categorias na pista 3.

Alvo \ Teste	Caminhão	Carro	Ônibus	Van	Moto
Caminhão	143	0	0	0	0
Carro	0	243	0	0	0
Ônibus	0	0	154	0	0
Van	3	0	3	205	0
Moto	0	0	0	0	64

Para avaliar a robustez dos classificadores testou-se o uso do classificador específico de uma pista para a classificação de categorias nas demais pistas (Tabela 35, Tabela 36). Os Melhores resultados são obtidos com descritores regionais cinza com fundo preto. Observa-se que os acertos mais altos estão na diagonal, pois se referem à classificação de cada pista por sua própria rede. Em alguns casos, principalmente com descritores com fundo original (Tabela 36) observa-se que existe similaridade de acerto entre as pistas 1 e 2, e entre as pistas 3 e 4. Isto é parcialmente explicado pela similaridade visual entre as respectivas pistas.

Tabela 35: Teste de Robustez da rede de cada Pista com dados das outras pistas, com fundo preto.

	Pista 1	Pista 2	Pista 3	Pista 4	Média	Desvio
Pista 1	100	92,01	58,32	53,50	75,96	23,46
Pista 2	60,20	99,78	65,51	63,08	72,14	18,55
Pista 3	69,31	86,83	99,63	97,90	88,42	13,94
Pista 4	75,64	86,39	93,08	100	88,78	10,37

Tabela 36: Teste de Robustez da rede de cada Pista com dados das outras pistas, com fundo original.

	Pista 1	Pista 2	Pista 3	Pista 4	Média	Desvio
Pista 1	99,80	89,85	52,73	51,87	73,56	24,89
Pista 2	90,10	100	57,39	51,17	74,67	24,02
Pista 3	36,63	28,08	99,39	93,22	64,33	37,17
Pista 4	66,34	56,37	93,34	98,83	78,72	20,58

Tabela 37: Teste de Robustez da rede de cada Pista com dados das outras pistas, geométricos.

	Pista 1	Pista 2	Pista 3	Pista 4	Média	Desvio
Pista 1	98.81	71.49	71.64	71.50	78,36	13,63
Pista 2	57.23	100	97.20	90.89	86,33	19,77
Pista 3	68.91	94.17	99,26	96.50	89,71	14,02
Pista 4	64.36	96.11	94.14	100	88,65	16,38

O terceiro tipo de experimento visa avaliar o desempenho de um classificador global treinado com exemplos extraídos do conjunto de todas as pistas. Quando o treinamento é realizado com o agrupamento global das pistas, o acerto se mantém na faixa de 99% para o descritor regional cinza e em 97% para os geométricos (Tabela 38). Apesar das diferenças entre os veículos, para os descritores regionais cinzas ocorrem poucas confusões referentes a alguns carros classificados com caminhão e ônibus, e algumas vans como caminhão (Tabela 39). Porém, os geométricos resultam em várias confusões (Tabela 40), as quais também são justificadas pela influência sofrida em relação ao tamanho dos veículos.

Tabela 38: Desempenho comparativo entre descritores para categorias no grupo global.

Descritor	Cinza	Binário	Bordas	Silhueta	Retângulo	Geométrico	Momentos
Acerto	99,68	98,19	96,88		85,59	97,7	61,95

Tabela 39: Matriz de confusão para descritor cinza sem fundo no grupo global.

Alvo \ Teste	Caminhão	Carro	Ônibus	Van	Moto
Caminhão	297	3	0	2	0
Carro	0	732	0	0	0
Ônibus	0	0	570	0	0
Van	0	2	0	543	0
Moto	0	0	0	0	64

Tabela 40: Matriz de confusão para descritor geométrico no grupo global.

Alvo \ Teste	Caminhão	Carro	Ônibus	Van	Moto
Caminhão	279	0	0	3	0
Carro	3	737	0	21	0
Ônibus	12	0	561	0	0
Van	3	0	9	521	0
Moto	0	0	0	0	64

4.7.2. Classificação de Subcategoria

Os testes com subcategorias estão divididos em duas estratégias principais, sendo elas a “direta” e a “cascata”. A estratégia direta utiliza como classes o agrupamento de todas as categorias e subcategorias no mesmo treinamento e teste. E a outra em cascata utiliza dois

níveis de classificação, sendo que o primeiro determina a categoria principal (carro, moto, caminhão, ônibus ou van) e o nível final especializa as subcategorias das classes que possuem subdivisões (caminhão, ônibus e van).

Na estratégia direta, o acerto se mantém na faixa de 98% para o descritor regional cinza (Tabela 41). Apesar da especialização em subcategorias, as confusões são poucas e esperadas (Tabela 42).

Tabela 41: Desempenho comparativo entre descritores para subcategorias no grupo pista 3.

Descritor	Cinza	Binário	Bordas	Silhueta	Retângulo	Geométrico	Momentos
Acerto	98,89	96,44	92,38		85,01	86,36	65,11

Tabela 42: Matriz de confusão para descritor cinza em subcategorias da pista 3.

(1) Caminhão Carroceria M; (2) Caminhão Carroceria G; (3) Caminhão Baú M; (4) Caminhão Baú G; (5) Van P; (6) Van M; (7) Ônibus M; (8) Ônibus G; (9) Carro; (10) Moto; (11) Caminhão Outro.

Teste \ Alvo	1	2	3	4	5	6	7	8	9	10	11
1	35	1	0	0	0	0	0	0	0	0	0
2	0	15	0	1	0	0	0	0	0	0	0
3	1	0	55	0	0	0	0	0	0	0	0
4	0	0	0	8	0	0	0	0	0	0	0
5	0	0	0	0	58	1	0	0	0	0	0
6	0	0	2	0	0	144	0	0	0	0	0
7	0	0	0	0	0	0	49	0	0	0	0
8	0	0	0	0	0	0	0	108	0	0	0
9	0	0	0	0	0	0	0	0	243	0	0
10	0	0	0	0	0	0	0	0	0	64	0
11	0	0	1	0	0	2	0	0	0	0	26

O descritor geométrico encontra dificuldades maiores para classificar as subcategorias e cai para a faixa de 85% (Tabela 44). Ainda na estratégia direta, porém no classificador global (Tabela 43), os acertos se mantêm nas mesmas faixas obtidas com os classificadores específicos.

Tabela 43: Desempenho comparativo entre descritores para subcategorias na grupo global.

Descritor	Cinza	Binário	Bordas	Silhueta	Retângulo	Geométrico	Momentos
Acerto	98,05	92,16	92,97		82,64	85,27	54,03

Tabela 44: Matriz de confusão para descritor geométrico em subcategorias do grupo global.
(1) Caminhão Carroceria M; (2) Caminhão Carroceria G; (3) Caminhão Baú M; (4) Caminhão Baú G;
(5) Van P; (6) Van M; (7) Ônibus M; (8) Ônibus G; (9) Carro; (10) Moto; (11) Caminhão Outro.

Teste \ Alvo	1	2	3	4	5	6	7	8	9	10	11
1	27	4	0	0	0	0	0	0	0	3	3
2	6	9	4	0	0	0	0	0	0	0	0
3	0	0	48	6	0	0	0	0	0	0	0
4	0	0	0	3	0	0	0	0	0	0	0
5	0	0	3	0	52	0	3	0	0	12	0
6	3	0	0	0	0	147	0	0	0	0	0
7	0	0	0	0	6	0	43	0	0	0	0
8	0	0	0	0	0	0	3	108	0	0	0
9	0	0	0	0	0	0	0	0	243	49	0
10	0	0	0	0	0	0	0	0	0	0	0
11	0	3	3	0	0	0	0	0	0	0	23

Na estratégia em cascata, primeiramente os exemplos da pista 3 são classificados nas principais categorias no primeiro nível da cascata. A partir disso os exemplos classificados como participantes de algum grupo especializado são separados para serem utilizados no segundo nível.

Os exemplos separados no primeiro nível da cascata são encaminhados para os respectivos classificadores especializados em subcategorias. Os acertos do último nível se mostram na faixa acima de 99% e são melhores que a estratégia direta. Assim, a estratégia em cascata obteve aumento do acerto para todas as classes de subcategorias (Tabela 45).

Tabela 45: Desempenho comparativo entre classificação direta e em cascata de subcategoria.

	Estratégia Direta	Estratégia Cascata
Caminhão	95,86	99,32
Ônibus	100	100
Van	98,53	100

4.7.3. Classificação de Cor

A classificação de cor é realizada diretamente a partir dos valores do ponto principal marcado para cada veículo.

Os classificadores especializados por pista apresentam acertos em torno de 80% (Tabela 46). No grupo global de exemplos ocorre uma pequena queda de desempenho, o que pode ser parcialmente justificada pela pequena variação de iluminação entre as pistas. Testes sem PCA causam uma perda considerável de desempenho para (Tabela 47). O espaço Ycbcr se mostrou melhor quando não é usado o PCA.

Tabela 46: Resultados para os descritores de cores com PCA.

Descritor Grupo	RGB	HSI	Lab	Ycbr
Pista 1	92.48	88.91	92.28	92.08
Pista 2	89.20	80.13	87.90	89.85
Pista 3	78.93	81.73	81.73	83.20
Pista 4	78.37	77.21	80	76.51
Global	76.78	72.31	77.85	74.36
Média	83,15	80,06	83,95	83,2
Desvio padrão	7,16	6,11	5,97	7,84

Tabela 47: Resultados para os descritores de cores sem PCA.

Descritor Grupo	RGB	HSI	Lab	Ycbr
Pista 1	58.61	92.28	58.61	58.61
Pista 2	63.50	89.42	12.10	61.77
Pista 3	62.53	77.73	62.53	62.53
Pista 4	59.53	71.16	59.53	59.53
Global	60.82	76.64	67.66	60.82
Média	61,00	81,45	52,09	60,65
Desvio padrão	2,03	9,00	22,63	1,60

5. Análise e Conclusões

5.1. Análise

Conforme apresentado no capítulo (3), os resultados divulgados em diferentes abordagens estão, em média, na faixa de 91% de acerto, um pouco abaixo dos aqui alcançados. A comparação entre os resultados encontrados na literatura e os alcançados nesta pesquisa não podem ser tomados como absolutos face às diferenças em termos de quantidade, qualidade e contexto das imagens utilizadas, assim como nas categorias consideradas. Porém servem como parâmetros para uma avaliação superficial de desempenho.

Comparando com resultados encontrados por outros pesquisadores (Seção 3.5.1), este trabalho alcançou melhores resultados para a classificação de categorias e subcategorias, sendo que alguns descritores similares são utilizados. Por exemplo, em (Tabela 48) observa-se que outros autores alcançaram no máximo 97%, enquanto que neste trabalho foram alcançados acertos de até 100%.

Tabela 48: Acertos na classificação de categorias por outros autores.

Autor	Descritor	Acerto
[PETROVIC, 2004]	Gradient	97 %
[LIM, 2002]	Gabor	97 %
[JOLLY, 1996]	Bordas	91 %
[HUANG, 2004]	Geométrico	91 %
[PRAMPERO, 1999]	Geométrico	90 %
[HSIEH, 2006]	Geométrico	89 %
[MESSELODI, 2005]	Modelo 3D	85 %

Nos experimentos realizados neste trabalho de pesquisa, destaca-se a utilização de um redimensionamento prévio nas imagens que, além de reduzir a dimensão e a complexidade do problema, permitiu um ganho em termos do nível de acerto.

No estudo de caso realizado observa-se uma certa robustez e similaridade de resultados entre as diferentes abordagens testadas. Observa-se que o classificador global está no mesmo nível de acerto alcançado pelos classificadores específicos.

No caso da classificação em subcategorias observou-se que a melhor abordagem é a do cascadeamento de classificadores, que implementam a quebra do problema original em problemas mais simples.

Os descritores regionais, principalmente o regional em níveis de cinza com a imagem reduzida a “6.25%”, foram os que apresentaram melhor resultado e robustez. Os descritores geométricos, embora de menor dimensão e com resultados próximos ao regional cinza, são

mais dependentes da etapa de segmentação que precisa identificar com maior precisão a silhueta do veículo.

Mesmo visualmente pode-se observar que os descritores regionais cinza geram padrões que facilitam o trabalho de discriminação e de classificação das redes neurais. Para efeito de visualização deste fato, em (Figura 80) o vetor descritor de cada veículo é disposto no sentido horizontal (linha) e agrupado de acordo com a categoria. Deste modo, cada linha dos grupos representa um veículo e cada coluna corresponde a um dos atributos do descritor. Assim, é possível confirmar visualmente que cada uma das categorias possui forma similar entre os seus veículos e distinta em relação às outras. Isso também se mantém após a aplicação da normalização *Z-score* (Figura 81) e a análise PCA (Figura 82), inclusive para outros descritores regionais (Apêndice).

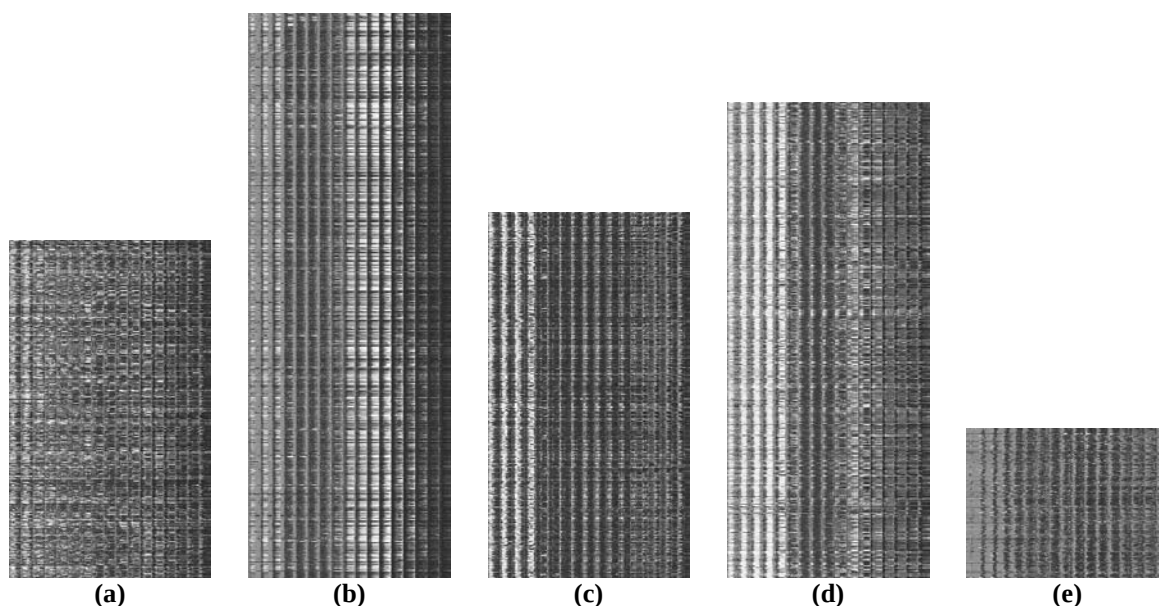


Figura 80: Representação visual do conjunto da Pista 3, descritor cinza com fundo. Redimensionamento em 17x17 (289) sem relação de aspecto.
(a) Caminhão-481 exemplos; (b) Carro-802; (c) Ônibus-520; (d) Van-677; (e) Moto-211.

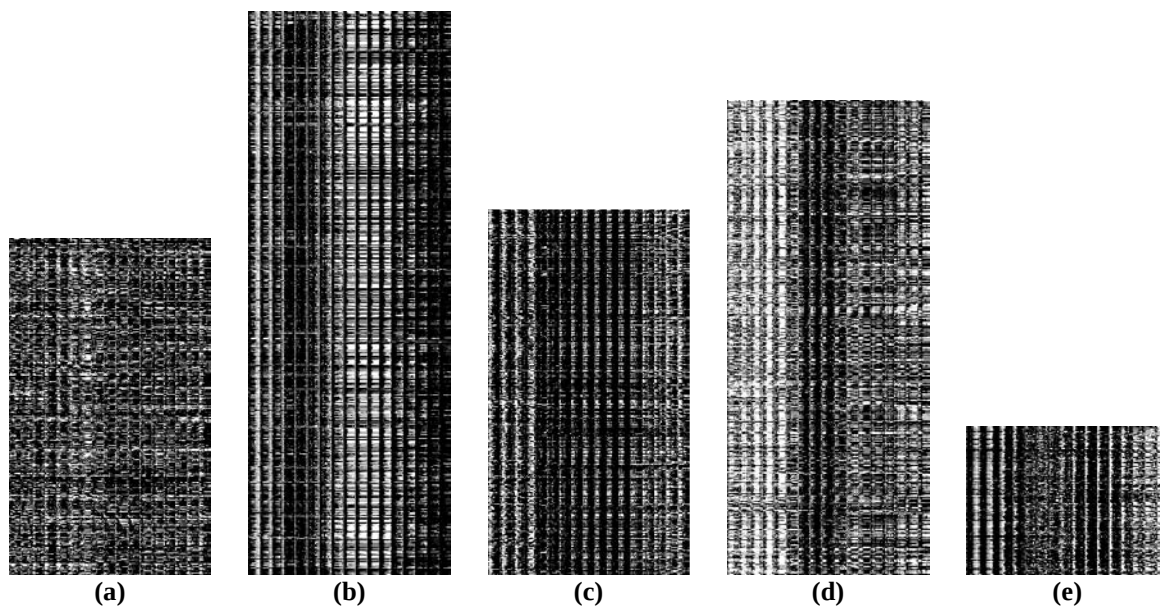


Figura 81: Representação visual do conjunto da Pista 3, descritor cinza com fundo e normalização.
 (a) Caminhão; (b) Carro; (c) Ônibus; (d) Van; (e) Moto.

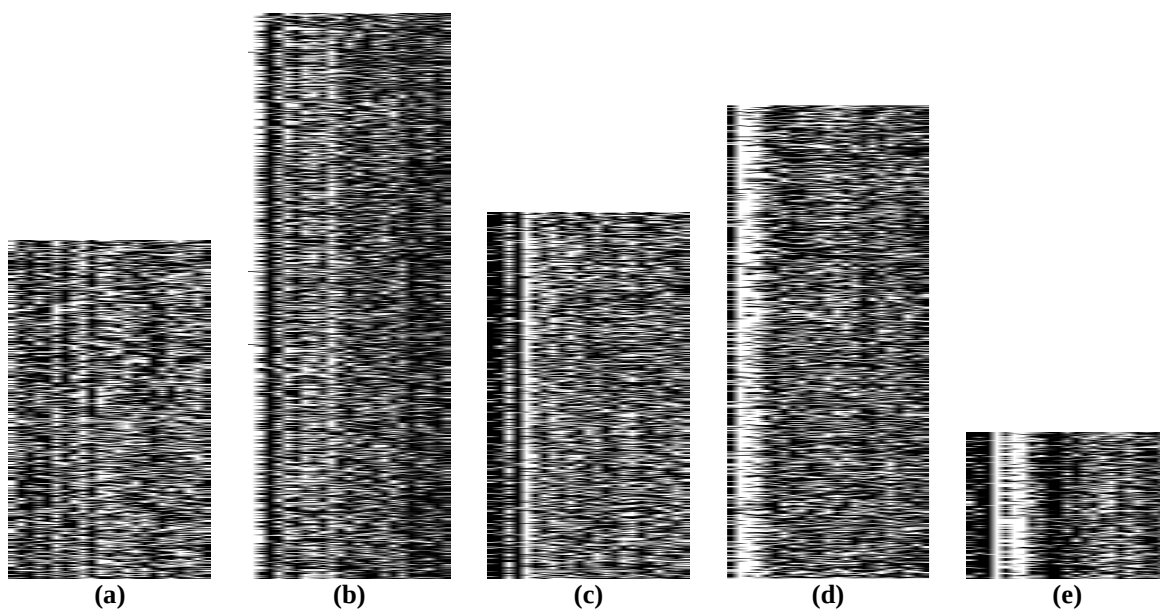


Figura 82: Representação visual do conjunto da Pista 3, descritor cinza com fundo e PCA.
 (dimensão 23)
 (a) Caminhão; (b) Carro; (c) Ônibus; (d) Van; (e) Moto.

5.2. Conclusões

Este trabalho atingiu plenamente os objetivos quanto ao aprendizado e aos resultados alcançados no estudo de caso real que foi realizado.

Uma contribuição importante foi a identificação da viabilidade de utilização de descritores simples, como o regional cinza, para classificação da categoria de veículos.

Outra contribuição importante foi a revisão abrangente do estado da arte em trabalhos relativos a ITS. Esta revisão servirá de base para trabalhos futuros.

5.3. Trabalhos futuros

Expandir o estudo para imagens obtidas por infravermelho em situações de baixa luminosidade.

Considerar a uso de imagens adquiridas com outros ângulos de câmera.

Considerar a captura com mais de uma câmera para formar e trabalhar com imagens em 3D.

6. Apêndice 1 - Rotinas

- São listados neste apêndice as principais rotinas dos módulos de Extrator e Treinador e Testador.

Rotina 1: Módulo Extrator

Na implementação da representação foi desenvolvido um único algoritmo capaz de extrair em conjunto os descritores das classes de categorias e cores.

Todos esses descritores são salvos em arquivo no formato de matriz, na qual cada linha contém a descrição de uma imagem. O algoritmo de extração contempla a variação de estratégias, comentada acima, a qual é escolhida de acordo com os seguintes parâmetros:

- Descritores: quais descritores serão extraídos;
- Fundo Cinza: indica se o fundo cinza será mantido ou não no recorte do veículo;
- Deformação: indica o tipo de redimensionamento, com deformação ou não;
- Centralizar: indica se o veículo será centralizado ou não no recorte;
- Equalizar: indica se haverá pré-processamento com equalização de histograma;
- Escala: indica qual o fator de redução de escala;
- Tipo de Classe: indica para qual classe (categoria, subcategoria ou cor) serão extraídos os descritores.

A partir destes parâmetros o algoritmo de Extração pode ser resumido de acordo com os seguintes passos:

```

Selecionar descritores
Ler arquivos de imagens marcadas de acordo com os índices
Para cada imagem do conjunto,
    Ler marcação da imagem (marcas, categoria, modelo, cor, ponto de cor)
    Concatenar rótulos de categoria, modelo, cor em matriz de valores alvos
    Se descritor é geométrico,
        Gerar imagem com silhueta a partir do polígono da região do veículo
        Extrair medidas geométricas a partir da silhueta
        Concatenar medidas em vetor descritor
        Agrupar vetor em matriz
    Se descritor é de cor,
        Extrair valores de RGB do ponto de cor
        Converter valores para o espaço de cor selecionado
        Concatenar valores em vetor
        Agrupar vetor em matriz
    Se descritor é regional,
        Se necessário remover fundo cinza e adicionar recorte do veículo em fundo preto
        Se necessário aplicar equalização de histograma na região do recorte do veículo
        Se tipo de redimensionamento é sem deformação,
            Se necessário centralizar o recorte do veículo em fundo preto
        Realizar redimensionamento de acordo com a escala e tipo de redimensionamento
        Realizar a transformação necessária do recorte do veículo para o descritor selecionado
        Linearizar o recorte transformado para o formato de vetor
        Agrupar vetor em matriz
    Salvar em arquivo as matrizes de descritores e a matriz de valores alvo
  
```

O algoritmo possibilita extrair, ao mesmo tempo, vários descritores regionais em diferentes escalas, descritores geométricos e tonais, e salvá-los em arquivo único. O formato de nomeação dos arquivos de descritores é o seguinte:

dadosPistan-Eqp1-Centp2 –Deforp3 –FdCzap4--d--e.mat, onde:

- *n*: número da pista entre 1 e 4;
- *p1*: valor lógico (1=V, 0=F) indicando a presença ou não de equalização;
- *p2*: valor lógico indicando a presença ou não de centralização;
- *p3*: valor lógico indicando a presença ou não de deformação;
- *p4*: valor lógico indicando a presença ou não de fundo cinza;
- *d*: lista de descritores que foram extraídos
- *e*: escala de redução.

Exemplos de nomes de arquivos:

- *dadosPista2-Eq0-Cent0-Defor1-FdCza0---ExCinza-ExBin-ExCanny-ExSilh-ExRetang-ExMoment--0.0625.mat*
- *dadosPista2-Eq0-Cent0-Defor1-FdCza1---ExRGB-ExHSV-ExHSI-ExYcbcr-ExLab--0.0625.mat*

Rotina 2: Módulo Treinador e Testador

Algoritmo de classificação

Na implementação da classificação para categoria e cor, o algoritmo de Treinamento e Teste pode ser resumido de acordo com os seguintes passos:

Ler arquivos de descritores selecionados
 Ler matriz de descritores e códigos alvos de acordo com a pista. Se a pista for global, ler e agrupar as 4 pistas
 Montar rótulos e códigos alvos das classes de acordo com o seu tipo (categoria, subcategoria, cor)
 Selecionar escalas de descritores
 Converter a matriz de códigos alvo para os códigos alvos das classes
 Converter a matriz de códigos alvo para o formato binário
 Se necessário, selecionar imagem de captura
 Separar matriz de descritores de acordo as classes
 Para cada matriz de classe
 Separar matriz de classe nos conjuntos de treinamento e teste
 Se necessário, ordenar aleatoriamente a matriz de códigos alvo e os conjuntos de treino e teste
 Agrupar todos as matrizes de códigos alvo e os conjuntos de treinamento e teste de cada classe
 Se necessário, ordenar aleatoriamente a matriz de códigos alvo global e os conjuntos globais de treino e teste
 Se necessário, aplicar normalização e PCA nos conjuntos globais de treino e teste
 Calcular número de neurônios na camada oculta e na camada de saída
 Criar Rede Neural
 Treinar Rede Neural
 Calcular o erro MSE do treino
 Simular Rede Neural
 Calcular o erro MSE da simulação
 Aplicar na resposta da simulação a análise *Winner Takes All* (WTA)
 Calcular o erro MSE da aplicação de WTA
 Calcular a quantidade de respostas corretas e incorretas a partir das respostas WTA
 Construir matriz de confusão da simulação
 Salvar em arquivo os resultados de treino e simulação

O algoritmo possibilita treinar em lote vários descritores em várias rodadas e salvar seus resultados em arquivo único. As informações do arquivo de resultados são as seguintes:

Pista, Classe, Equalizado, Centralizado, Deformado, Fundo Cinza, Descritor, Dimensão, Quantia/Porc de treinamento, Quantia de treinamento, Quantia de teste, Características antes do PCA, Características depois do PCA, Número de Saídas da Rede, Neurônios Ocultos, Porc. de Corretos, Porc. de Incorretos, Quantia de Corretos, Quantia de Incorretos, Fator PCA, Ordem Aleatória, Maximo de Ciclos, Final de Ciclos, Erro de treino (MSE), Erro de teste (MSE), Erro de WTA (MSE), Captura selecionada, Tempo de PCA, Tempo de Treino. O formato de nomeação dos arquivos de resultados é o seguinte:

Classec-dadosPistan-Eqp1-Centp2 –Deforp3 –FdCzap4--d--e-resultados-bPcaP-PorcQ - PcaOutroT.txt , onde:

- *c*: tipo de classe (categoria, subcategoria ou cor);
- *n*: número da pista entre 1 e 4;
- *p1*: valor lógico (1=V, 0=F) indicando a presença ou não de equalização;
- *p2*: valor lógico indicando a presença ou não de centralização;
- *p3*: valor lógico indicando a presença ou não de deformação;
- *p4*: valor lógico indicando a presença ou não de fundo cinza;
- *d*: lista de descritores que foram extraídos;
- *e*: escala de redução;
- *P*: valor lógico indicando a presença ou não de PCA;
- *Q*: valor indicando o limiar de corte de PCA;
- *T*: valor lógico indicando a presença ou não parâmetros de PCA armazenados anteriormente.

Exemplos de nomes de arquivos:

- Classe1-Pistas1234--Eq0-Cent0-Defor0-FdCza0---ExRProps--0.0625-resultados-bPca1-Porc0.70 - PcaOutro0.txt
- Classe3-Pistas1234--Eq0-Cent0-Defor1-FdCza1---ExRGB-ExHSV-ExHSI-ExYcbcr-ExLab--0.0625-resultados-bPca1-Porc0.70 -PcaOutro0.txt

Rotina 3: Módulos Marcador e indexador

O Marcador recebe as imagens originais adquiridas, as quais apresentam o seguinte padrão na nomeação dos arquivos:

Px_aaaammdd_hhmmsszzz_w.bmp (ex.: P2_20060915_092628468_1.bmp), onde:

- *x*: número de um dígito representando a pista entre 1 e 4;
- *aaaammdd*: representa a data de aquisição composta de ano, mês e dia;
- *hhmmsszzz*: representa a hora de aquisição composta de hora, minuto, segundo e milissegundo;
- *w*: número de um dígito representando a ordem entre 1 e 3 entre o conjunto de imagens coletadas do mesmo veículo.

A marcação gerada pelo módulo é incluída entre colchetes no nome da própria imagem, a qual passa a ter o seguinte padrão de nomeação:

$Px_aaaammdd_hhmmsszzz_w [p_{x1_p_{y1}} \dots p_{xn_p_{yn}} c_l_r_t_x_t_y].bmp$, onde

- $p_{x1_p_{y1}} \dots p_{xn_p_{yn}}$: n coordenadas (x,y) representando os vértices do polígono do contorno;
- c : rótulo da categoria, representado por um número entre 1 e 5;
- l : rótulo da subcategoria, representado por um número entre 1 e 70;
- r : rótulo da cor, representado por um número entre 1 e 8;
- $t_x_t_y$: coordenadas (x,y) do ponto de cor.

Neste padrão de marcação os números são separados pelo caractere ‘_’, como por exemplo:

$P2_20060915_092628468_1[160_145_115_145_93_121_96_92_148_92_161_129_160_145_3_1_5_128_119].bmp$

O módulo Indexador é utilizado logo após o Marcador, percorrendo as pastas selecionadas e listando as imagens que foram marcadas. Esses índices de arquivos são organizados por pista em arquivos nomeados como “pistan.txt”, onde n é a pista. O conteúdo do arquivo contém, para cada imagem de veículo, as seguintes informações:

pista;hora;minuto;pasta;nome;categoria;modelo;cor;marcas;pontocor

Estas informações são separadas por “;”, assim como no seguinte exemplo:

- $4;9;15;E:\img\P4\1\caminhao;P4_20060915_093430546_1[187_128_122_128_82_98_84_26_145_30_17_8_42_188_94_187_128_2_2_4_153_67].bmp;2;2;4;187_128_122_128_82_98_84_26_145_30_178_42_18_8_94_187_128;153_67$

7. Apêndice 2 - Resultados de Classificação de Categoria

A seguir são listados alguns resultados, sendo que as seguintes definições são seguidas:

- Proporção para redimensionamento: 6,25 %
- Fator de corte PCA: 0,005
- Máximo de ciclos para treino: 1000, com exceção de alguns testes com 2000 para cores
- P.Treino: porcentagem de exemplos a serem usados no treino, padrão em 70%
- Deformação: indica o tipo de redimensionamento
- Dimensão de vetor: quando não informada é 289 para deformado e 300 para não deformado
- F.Cinza: indica se o fundo cinza está presente (1) ou não (0) no retângulo que contém o polígono do veículo; quando não informado é 1;
- Centr.: indica se o recorte foi centralizado (1) ou não (0); aplicado somente em descritor sem deformação; quando não informado é 1;
- Ciclos: indica o número final de ciclos alcançados após o treino; quando não informado é 1000;
- Grupo: é o agrupamento que está sendo testado; Quando não informado é P.3 (pista 3);
- Saídas: o número de neurônios na saída da rede para a categoria depende da presença de moto no conjunto, já para subcategoria é sempre 11;
- Acerto: porcentagem de acerto do teste (simulação) da RNA;
- D.PCA: número de dimensões de vetor após processamento PCA;
- E.Treino: erro MSE alcançado no final de treinamento de RNA;
- E.Teste: erro MSE alcançado no final de teste (simulação) de RNA;
- E.WTA: erro MSE do processamento WTA;
- T.PCA: tempo em segundos de processamento PCA;
- T.Treino: tempo em segundos de treinamento de RNA;

Descritor Cinza

O descritor cinza deformado foi testado em todas as pistas e apresentou acertos em torno de 99% para categoria de veículo (Tabela 49). Porém, quando o acerto chega a 100% os erros MSE são maiores, o que indica um esforço para ajuste através do WTA. Para este descritor a presença de fundo cinza possui não tem grande influência nos resultados.

Tabela 49: Resultados para o descritor cinza deformado para categoria em todas as pistas.

Grupo	F.Cinza	D.PCA	Acerto	Ciclos	E.Treino	E.Teste	E.WTA	T.PCA	T.Treino
P.1	0	26	100	344	0,1590	0,0002	0,0002	16,62	30,79
P.1	1	24	99,21	574	0,0004	0,0056	0,0020	19,49	49,28
P.2	0	32	100	319	0,0113	0,7210	0,7210	14,15	29,76
P.2	1	26	100	328	0,1800	1,9300	1,9300	15,11	28,08
P.3	1	23	99,63	1000	0,0001	0,0013	0,0009	38,77	127,1
P.4	0	27	100	323	0,1270	2,1600	2,1600	12,78	27,97
P.4	1	24	99,77	325	0,1440	0,0006	0,0002	13,37	24,98
Global	0	26	99,23	1000	0,0009	0,0023	0,0018	280,37	414,06
Global	1	26	99,68	1000	0,0005	0,0014	0,0005	276,51	414,6

Nos resultado acima foi utilizada permutação na ordenação dos vetores na matriz de treinamento. Entretanto a ausência de permutação resulta em perda considerável no desempenho (Tabela 50).

Tabela 50: Resultados para o descritor cinza deformado sem permutação na pista 3.

D.PCA	Acerto	Ciclos	E.treino	E.teste	E.WTA	T.PCA	T.Treino
25	88,71	337	0,2660	0,0405	0,0116	38,5	48,18

O processamento PCA é essencial para o descritor cinza em categorias, pois a sua ausência causa grande perda de desempenho para a faixa de 46%. Isso é esperado porque esse descritor é representado por valores brutos do recorte em tons de cinza.

Tabela 51: Resultados para o descritor cinza deformado sem PCA na pista 3.

Acerto	Ciclos	E.treino	E.teste	E.WTA	T.PCA	T.Treino
46,26	1000	0,1368	0,1354	0,0893	0,03	2018,95

O descritor cinza é robusto quanto a redução na quantia de exemplos de treinamento e aumento na quantia de teste, assim como é demonstrado nas porcentagens de 45% e 20% para treino da RNA, seja para a pista 3 (Tabela 52) ou global (Tabela 53).

Tabela 52: Resultados para o descritor cinza deformado com redução do conjunto de treino na pista 3.

P.Treino	Q.Treino	Q.Teste	D.PCA	Acerto	Ciclos	E.treino	E.teste	E.WTA	T.PCA	T.Treino
0,45	1224	1494	23	98,66	813	0,0007	0,0057	0,0019	18,61	69,05
0,20	544	2174	23	97,1	332	0,0198	0,0100	0,0020	7,33	16,13

Tabela 53: Resultados do descritor cinza deformado com redução do conjunto de treino na pista global.

P.Treino	Q.Treino	Q.Teste	D.PCA	Acerto	Ciclos	E.treino	E.teste	E.WTA	T.PCA	T.Treino
0,45	3321	4056	23	98,1	1000	0,0033	0,0070	0,0040	135,57	272,94
0,20	1476	5901	22	96,97	1000	0,0019	0,0110	0,0051	31,4	145,33

Quanto à seleção de captura das imagens o descritor cinza manteve o desempenho em torno de 98% (Tabela 54), no qual foram selecionadas apenas as imagens da captura dois do conjunto de três disponíveis por veículo. Essa seleção, além de diminuir para 1/3 das quantias de treino e teste, também diminui o número de exemplos similares.

Tabela 54: Resultados do descritor cinza deformado com seleção de captura na pista 3.

Q.Treino	Q.Teste	D.PCA	Acerto	Ciclos	E.treino	E.teste	E.WTA	T.PCA	T.Treino
634	272	22	98,9	519	0,2380	0,0041	0,0010	7,62	26,32

Também para o descritor cinza não deformado o ótimo desempenho se mantém, porém a melhor configuração mostra leve melhoria quando é utilizado a centralização e o fundo cinza é excluído do recorte (Tabela 55).

Tabela 55: Resultados do descritor cinza não deformado na pista 3.

Centr.	F.Cinza	D.PCA	Acerto	Ciclos	E.treino	E.teste	E.WTA	T.PCA	T.Treino
1	0	31	99,14	1000	0,0007	0,0034	0,0006	14,83	155,86
1	1	23	98,53	1000	0,0029	0,0042	0,0020	16,48	135,69
0	1	41	97,91	1000	0,0019	0,0058	0,0029	15,06	208,55
0	0	41	97,42	1000	0,0017	0,0094	0,0037	15,2	211,21

Para a subcategoria o descritor cinza também mostrou ótimo desempenho, isso considerando que o número de classes passa para onze e a separação continua sendo feita com sucesso na faixa de 98% (Tabela 56). Isso também alcançado para o descritor cinza deformado na pista global (Tabela 57).

Tabela 56: Resultados para o descritor cinza deformado em subcategorias na pista 3.

D.PCA	Acerto	Ciclos	E.treino	E.teste	E.WTA	T.PCA	T.Treino
24	98,53	1000	0,0009	0,0024	0,0016	45,3	190,31

Tabela 57: Resultados para o descritor cinza deformado em subcategorias na pista global.

D.PCA	Acerto	Ciclos	E.treino	E.teste	E.WTA	T.PCA	T.Treino
26	98,05	1000	0,0020	0,0029	0,0018	284,7	591,48

Descritor Binário

O descritor binário deformado apresentou resultados similares satisfatórios em torno de 99% e com melhor desempenho na presença de fundo cinza no recorte (Tabela 58). O mesmo desempenho pode ser observado na pista global (Tabela 59).

Tabela 58: Resultados para o descritor binário deformado para categoria na pista 3.

F.Cinza	D.PCA	Acerto	Ciclos	E.treino	E.teste	E.WTA	T.PCA	T.Treino
1	30	99,26	567	0,3230	0,0030	0,0014	40,03	92,15
0	27	98,77	1000	0,0006	0,0039	0,0025	37,61	142,39

Tabela 59: Resultados para o descritor binário deformado para categoria na pista global.

D.PCA	Acerto	Ciclos	E.treino	E.teste	E.WTA	T.PCA	T.Treino
29	98,19	1000	0,0038	0,0060	0,0029	296,49	459,41

Também para o descritor binário não deformado o ótimo desempenho se mantém, porém a melhor configuração mostra leve melhoria quando é utilizado a centralização e o fundo cinza é excluído do recorte (Tabela 60).

Tabela 60: Resultados do descritor binário não deformado na pista 3.

Centr.	F.Cinza	D.PCA	Acerto	Ciclos	E.treino	E.teste	E.WTA	T.PCA	T.Treino
1	0	32	98,90	1000	0,0058	0,0049	0,0027	15,49	164,52
0	1	37	96,44	1000	0,0070	0,0136	0,0061	15,35	197,25
0	0	37	95,95	1000	0,0070	0,0135	0,0059	13,78	192,27
1	1	27	94,97	1000	0,0188	0,0161	0,0067	19,60	142,88

Para a subcategoria o descritor binário também mostrou ótimo desempenho com sucesso na faixa de 96% (Tabela 61). Isso também alcançado para o descritor binário deformado na pista global (Tabela 62).

Tabela 61: Resultados para o descritor binário deformado em subcategorias na pista 3.

D.PCA	Acerto	Ciclos	E.treino	E.teste	E.WTA	T.PCA	T.Treino
30	96,44	1000	0,0016	0,0043	0,0026	38,96	225,19

Tabela 62: Resultados para o descritor binário deformado em subcategorias na pista global.

D.PCA	Acerto	Ciclos	E.treino	E.teste	E.WTA	T.PCA	T.Treino
28	92,16	1000	0,0079	0,0095	0,0059	325,59	611,94

Descritor de Bordas

O descritor de bordas deformado foi testado em todas as pistas e apresentou acertos em torno de 99% para categoria de veículo (Tabela 63). Para este descritor a presença de fundo cinza possui não tem grande influência nos resultados.

Tabela 63: Resultados para o descritor de bordas deformado para categoria em todas na pista 3.

F.Cinza	D.PCA	Acerto	Ciclos	E.treino	E.teste	E.WTA	T.PCA	T.Treino
0	50	99,51	466	0,0001	0,0020	0,0012	31,08	117,74
1	49	97,18	450	0,0006	0,0107	0,0039	32,39	114,73

Também para o descritor de bordas não deformado o melhor desempenho cai para a faixa de 96%, a qual é alcançada quando é utilizado a centralização e o fundo cinza é excluído do recorte (Tabela 64). Este acerto se mantém quando no teste da mesma configuração na pista global (Tabela 65).

Tabela 64: Resultados do descritor de bordas não deformado na pista 3.

Centr.	F.Cinza	D.PCA	Acerto	Ciclos	E.treino	E.teste	E.WTA	T.PCA	T.Treino
1	0	43	96,44	1000	0,0159	0,0128	0,0035	16,7	222,19
0	1	57	92,15	1000	0,0121	0,0269	0,0081	14,4	280,9
0	0	57	91,41	1000	0,0114	0,0278	0,0084	15,29	287,55
1	1	43	89,45	1000	0,0287	0,0228	0,0115	18,93	220,6

Tabela 65: Resultados do descritor de bordas não deformado no grupo global.

D.PCA	Acerto	Ciclos	E.treino	E.teste	E.WTA	T.PCA	T.Treino
49	96,88	1000	0,0024	0,0111	0,0060	238,22	674,28

Para a subcategoria o descritor de bordas também mostrou ótimo desempenho, porém o acerto cai para a faixa de 92% (Tabela 66). Isso também é alcançado para esse descritor na pista global (Tabela 67).

Tabela 66: Resultados para o descritor de bordas deformado em subcategorias na pista 3.

D.PCA	Acerto	Ciclos	E.treino	E.teste	E.WTA	T.PCA	T.Treino
49	92,38	513	0,0015	0,0105	0,0051	31,35	149,15

Tabela 67: Resultados para o descritor de bordas deformado em subcategorias na pista global.

D.PCA	Acerto	Ciclos	E.treino	E.teste	E.WTA	T.PCA	T.Treino
49	92,97	1000	0,0050	0,0098	0,0060	232,04	835,94

Descritor Silhueta

Para o descritor de silhueta não deformado o melhor desempenho é obtido quando é utilizado a centralização e o fundo cinza é excluído do recorte (Tabela 68).

Tabela 68: Resultados do descritor silhueta não deformado na pista 3.

Centr.	F.Cinza	D.PCA	Acerto	E.treino	E.teste	E.WTA	T.PCA	T.Treino
1	0	25	97,42	0,0121	0,0092	0,0032	15,34	138,26
1	1	23	97,06	0,0131	0,0108	0,0061	15,31	135,59
0	1	33	96,44	0,0119	0,0144	0,0064	14,64	163,91
0	0	33	96,2	0,0147	0,0148	0,0062	15,37	169,01

Descrito Retângulo

No descritor de retângulo não deformado o melhor está na faixa de 90%, principalmente quando é utilizado a centralização e o fundo cinza é excluído do recorte (Tabela 69).

Tabela 69: Resultados do descritor de retângulo não deformado na pista 3.

Centr.	F.Cinza	D.PCA	Acerto	E.treino	E.teste	E.WTA	T.PCA	T.Treino
1	0	29	90,94	0,0278	0,0286	0,0109	16,89	154,41
0	0	29	90,92	0,0283	0,0280	0,0112	16,76	152,35
1	1	29	89,85	0,0282	0,0298	0,0118	17,49	148,52
0	1	29	89,82	0,0289	0,0283	0,0123	14,42	151,86

O descritor de retângulo deformado apresentou resultados satisfatórios em torno 85% na pista global (Tabela 70).

Tabela 70: Resultados para o descritor de retângulo deformado para categoria na pista global.

D.PCA	Acerto	E.treino	E.teste	E.WTA	T.PCA	T.Treino
31	85,59	0,0327	0,0363	0,0150	175,65	620,93

Para a subcategoria o descritor de retângulo também mostrou desempenho na faixa de 85% (Tabela 71). Isso também é alcançado na pista global (Tabela 72).

Tabela 71: Resultados para o descritor de retângulo deformado em subcategorias na pista 3.

D.PCA	Acerto	E.treino	E.teste	E.WTA	T.PCA	T.Treino
28	85,01	0,0180	0,0181	0,0087	14,69	206,78

Tabela 72: Resultados para o descritor de retângulo deformado em subcategorias na pista global.

D.PCA	Acerto	E.treino	E.teste	E.WTA	T.PCA	T.Treino
31	82,64	0,0218	0,0232	0,0101	166,89	647,87

Descritor Geométrico

O descritor geométrico apresentou resultados similares satisfatórios em torno entre 97% e 100% (Tabela 73) em todas as pistas, inclusive na global. A única exceção é o acerto de 91% na pista 3.

Tabela 73: Resultados para o descritor geométrico para categorias na pista 3.

Grupo	D.PCA	Acerto	E.treino	E.teste	E.WTA	T.PCA	T.Treino
P.1	5	98,61	0,0059	0,0078	0,0045	1,54	44,85
P.2	5	100	0,0001	0,0003	0,0003	1,35	42,34

P.3	5	91,41	0,0199	0,0189	0,0145	3,82	68,99
P.4	5	100	0,9170	0,0001	0,0001	1,21	37,34
Global	5	97,7	0,0155	0,0089	0,0038	30,56	182,37

O processamento PCA é essencial para o descritor geométrico, pois a sua ausência causa grande perda de desempenho para a faixa entre 33% e 40% (Tabela 74).

Tabela 74: Resultados para o descritor geométrico sem PCA em todas as pistas.

Grupo	Acerto	Ciclos	E.treino	E.teste	E.WTA	T.Treino
P.1	33,86	1000	0,1796	0,1796	0,1516	51,24
P.2	40,6	816	0,1700	0,1699	0,1271	39,16
P.3	29,82	376	0,1545	0,1545	0,1263	32,86
P.4	37,85	1000	0,1981	0,1982	0,1129	41,8

Para a subcategoria o descritor geométrico também mostrou sucesso na faixa de 86% (Tabela 75).

Tabela 75: Resultados para o descritor geométrico em subcategorias na pista 3 e global.

Grupo	D.PCA	Acerto	E.treino	E.teste	E.WTA	T.PCA	T.Treino
P.3	5	86,36	0,0150	0,0148	0,0107	4,24	121,33
Global	5	85,27	0,0196	0,0179	0,0091	30,77	321,02

Descritor Momentos

O descritor de momentos deformado apresentou resultados médios de até 72% com ausência de fundo cinza no recorte (Tabela 76).

Tabela 76: Resultados para o descritor momentos deformado para categoria na pista 3.

F.Cinza	D.PCA	Acerto	Ciclos	E.treino	E.teste	E.WTA	T.PCA	T.Treino
0	6	72,64	1000	0,0790	0,0854	0,0406	3,31	77,7
1	7	68,71	1000	0,0874	0,0920	0,0370	2,97	83,09

Para o descritor de momentos não deformado o desempenho cai para 61% na pista global (Tabela 77).

Tabela 77: Resultados para o descritor momentos não deformado para categoria na pista global.

D.PCA	Acerto	Ciclos	E.treino	E.teste	E.WTA	T.PCA	T.Treino
6	61,95	1000	0,1117	0,1084	0,0533	34,08	289,52

Para a subcategoria o descritor de momentos também mostra acerto de 65% para deformado e 54% para não deformado (Tabela 78, Tabela 79).

Tabela 78: Resultados para o descritor momentos deformado para subcategoria na pista 3.

D.PCA	Acerto	Ciclos	E.treino	E.teste	E.WTA	T.PCA	T.Treino
6	65,11	1000	0,0403	0,0447	0,0224	3,76	142,71

Tabela 79: Resultados para o descritor momentos não deformado para subcategoria na pista global.

D.PCA	Acerto	Ciclos	E.treino	E.teste	E.WTA	T.PCA	T.Treino
6	54,03	1000	0,0566	0,0581	0,0305	25,09	359,84

8. Apêndice 3 - Resultados de Classificação de Cor

A classificação de cor é realizada diretamente a partir dos valores do ponto principal marcado para cada veículo. Nenhum processamento adicional é realizado sobre esses dados.

Devida a pequena dimensão do vetor extraído de cada ponto, também são realizados testes com incremento de ciclos para 2000 na tentativa de melhorar o desempenho do acerto que resultou em valores em torno de 80%.

Descritor RGB

O descritor RGB apresenta acertos de até 80% na pista 3 (Tabela 80). Nesse caso o aumento de ciclos para 2000 incrementa o desempenho. Testes sem PCA causam uma perda considerável de desempenho para 62% (Tabela 81).

Tabela 80: Resultados para o descritor RGB para cores na pista 3 e global.

Grupo	D.PCA	Acerto	Ciclos	E.treino	E.teste	E.WTA	T.PCA	T.Treino
P.3	3	79,07	1000	0,0398	0,0400	0,0156	1,46	78,63
P.3	3	80,4	2000	0,0318	0,0362	0,0100	1,6	171,06
Global	3	70,64	2000	0,0394	0,0506	0,0173	26,54	605,2

Tabela 81: Resultados para o descritor RGB para cores sem PCA na pista 3.

Acerto	Ciclos	E.treino	E.teste	E.WTA	T.Treino
62,53	1000	0,0750	0,0747	0,0178	77,73

Descritor HSI

O descritor HSI apresenta acertos de até 80% na pista 3 (Tabela 82). O aumento de ciclos para 2000 decrementa o desempenho. Testes sem PCA causam um leve perda de desempenho para 76% (Tabela 83)

Tabela 82: Resultados para o descritor HSI para cores na pista 3 e global.

Grupo	D.PCA	Acerto	Ciclos	E.treino	E.teste	E.WTA	T.PCA	T.Treino
P.3	3	80	1000	0,0347	0,0383	0,0138	1,51	85,67
P.3	3	78,53	2000	0,0365	0,0403	0,0163	1,66	173,14
Global	3	76,17	2000	0,0342	0,0453	0,0114	18,78	557,28

Tabela 83: Resultados para o descritor HSI para cores sem PCA na pista 3.

Acerto	Ciclos	E.treino	E.teste	E.WTA	T.Treino
76,4	1000	0,0381	0,0397	0,0145	76,64

Descritor Lab

O descritor Lab também apresenta acertos de até 80% na pista 3 (Tabela 84), porém o número maior de ciclos decrementa o desempenho. Testes sem PCA causam uma perda considerável de desempenho para 62% (Tabela 85).

Tabela 84: Resultados para o descritor Lab para cores na pista 3 e global.

Grupo	D.PCA	Acerto	Ciclos	E.treino	E.teste	E.WTA	T.PCA	T.Treino
P.3	3	80,53	1000	0,0366	0,0394	0,0130	1,49	82,02
P.3	3	72	2000	0,0406	0,0464	0,0204	1,65	179,12
Global	3	74,73	2000	0,0362	0,0458	0,0148	43,74	556,04

Tabela 85: Resultados para o descritor Lab para cores sem PCA na pista 3.

Acerto	Ciclos	E.treino	E.teste	E.WTA	T.Treino
62,67	1000	0,1095	0,1088	0,0678	81,13

Descritor Ycbr

O descritor Ycbr apresenta acertos de até 82% na pista 3 (Tabela 84), o qual é alcançado com a ajuda do número maior de ciclos. Testes sem PCA causam uma grande perda de desempenho para 4% (Tabela 85).

Tabela 86: Resultados para o descritor Ycbr para cores na pista 3 e global.

Grupo	D.PCA	Acerto	Ciclos	E.treino	E.teste	E.WTA	T.PCA	T.Treino
P.3	3	70	1000	0,0449	0,0508	0,0211	1,49	82,01
P.3	3	82,67	2000	0,0307	0,0353	0,0092	1,63	180,89
Global	3	78,73	2000	0,0307	0,0403	0,0113	15,87	508,68

Tabela 87: Resultados para o descritor Ycbr para cores sem PCA na pista 3.

Acerto	Ciclos	E.treino	E.teste	E.WTA	T.Treino
4,4	1000	0,1614	0,1599	0,0549	82,19

9. Apêndice 4 – Redes Neurais

Redes Neurais Artificiais (RNA) são técnicas computacionais que propõem um modelo matemático baseado na estrutura neural de organismos inteligentes, mais especificamente o cérebro humano. Segundo [HAYKIN, 1999], uma rede neural é um processador maciçamente e paralelamente distribuído, constituído de unidades de processamento simples que têm a propensão natural para armazenar conhecimento experimental e torná-lo disponível para o uso. A principal característica das Redes neurais é a capacidade de aprender a partir de exemplos e, assim, classificar novos padrões.

Desde o começo, os trabalhos em RNA têm sido motivados pelo reconhecimento em que o cérebro humano processa informações de forma mais eficiente, se comparado com os computadores convencionais. O cérebro é um sistema de processamento de informação altamente complexo, não-linear e paralelo que tem a capacidade de organizar seus componentes estruturais (neurônios) de forma a realizar certos tipos de processamento muito mais rapidamente que o mais rápido computador digital hoje existente [HAYKIN, 1999]. Uma rede neural artificial se assemelha ao cérebro em dois aspectos [HAYKIN, 1999]:

- O conhecimento é adquirido pela rede a partir de seu ambiente através de um processo de aprendizagem;
- Forças de conexão entre neurônios, conhecidas como pesos sinápticos, são utilizadas para armazenar o conhecimento adquirido.

O primeiro modelo de um neurônio artificial foi proposto pelo neurofisiologista McCulloch e pelo matemático Pitts [MCCULLOCH; PITTS, 1943]. A Figura (Figura 83) mostra um esquema do neurônio artificial proposto por eles. A operação dessa unidade de processamento pode ser resumida da seguinte forma:

- Sinais são apresentados às entradas ($p_1, p_2, \dots, p_n$);
- Cada sinal é multiplicado por um número, ou peso, que indica sua influência na saída da unidade ($W_1, W_2, \dots, W_p$);
- É realizada uma soma ponderada dos sinais, produzindo um nível de atividade (Σ);
- Se este nível de atividade exceder um certo limite (*threshold*) a unidade produz uma determinada resposta de saída (Y), senão o neurônio permanece inativo.

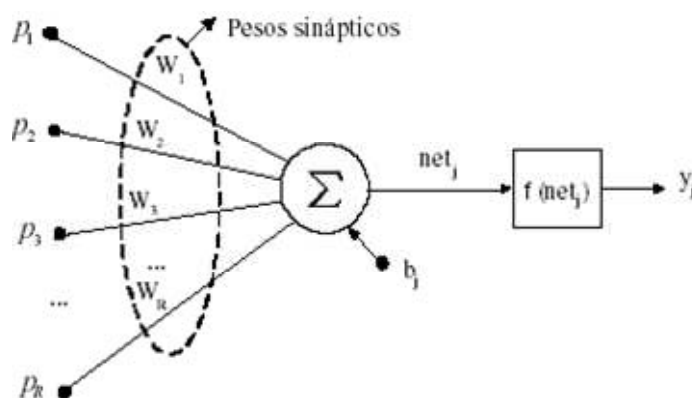


Figura 83: Neurônio artificial de McCulloch / Pitts.

As entradas do neurônio artificial podem ser comparadas com os estímulos do neurônio natural. Os pesos são valores que representam o grau de importância que cada entrada possui em relação àquele determinado neurônio. A função de ativação é uma função de ordem interna, cuja atribuição é decidir o que fazer com o valor resultante da soma das entradas ponderadas.

A característica mais importante das RNA é a capacidade de aprender a partir de exemplos. A aprendizagem ocorre quando a rede neural atinge uma solução generalizada para uma classe de problemas. Segundo [HAYKIN, 1999], a aprendizagem é um processo pelo qual os parâmetros livres de uma RNA são adaptados através de um processo de estimulação pelo ambiente no qual a rede está inserida. O conjunto de regras bem definidas para a solução de um problema de aprendizagem é denominado algoritmo de aprendizagem. Na **aprendizagem supervisionada** um agente externo é utilizado para indicar à rede qual a resposta desejada para o padrão de entrada.

A rede *Perceptron* Multicamadas (*Multi Layer Perceptron* - MLP) [BISHOP, 1995; HAYKIN, 1999] é formada por neurônios do tipo *Perceptron* derivados do modelo de McCulloch e Pitts. A MLP é utilizada no modo de alimentação progressiva (*feed-forward*). Neste modo o sinal de saída de um neurônio não é utilizado como entrada para os neurônios das camadas anteriores, ou para os neurônios da própria camada, mas somente para os da camada posterior (Figura 84).

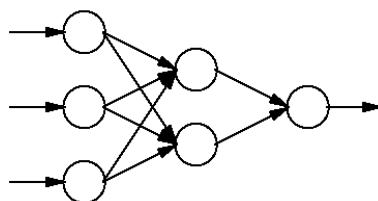


Figura 84: Rede em modo de alimentação progressiva (feed-forward).

Para o treinamento dessas redes utiliza-se o algoritmo *Backpropagation* [RUMELHART; WILLIAMS, 1986]. Durante o treinamento, a rede opera em uma sequência de dois passos. Primeiro, um padrão é apresentado à camada de entrada da rede. A atividade resultante flui através da rede, camada por camada, até que a resposta seja produzida pela camada de saída.

No segundo passo, a saída obtida é comparada à saída desejada para esse padrão particular. Se esta não estiver correta, o erro é calculado e propagado a partir da camada de saída até a camada de entrada. Os pesos das conexões entre as unidades das camadas internas vão sendo modificados conforme o erro é retro-propagado. O ciclo é repetido até que o erro esteja dentro de um limite aceitável. O número de vezes que o conjunto de treinamento completo é apresentado à rede é chamado de ciclos do treinamento, ou seja, cada vez que o conjunto de treinamento completo é apresentado conta-se uma ciclo.

Depois que a rede estiver treinada e o erro estiver em um nível satisfatório, ela poderá ser utilizada como uma ferramenta para classificação de novos dados. Para isto, novos dados são apresentados à camada de entrada, são processados nas camadas intermediárias e os resultados são apresentados na camada de saída, como no treinamento, mas sem a retro-propagação do erro. A saída corresponde à interpretação da rede para a nova entrada apresentada.

10. Apêndice 5 – Visualizações em conjunto

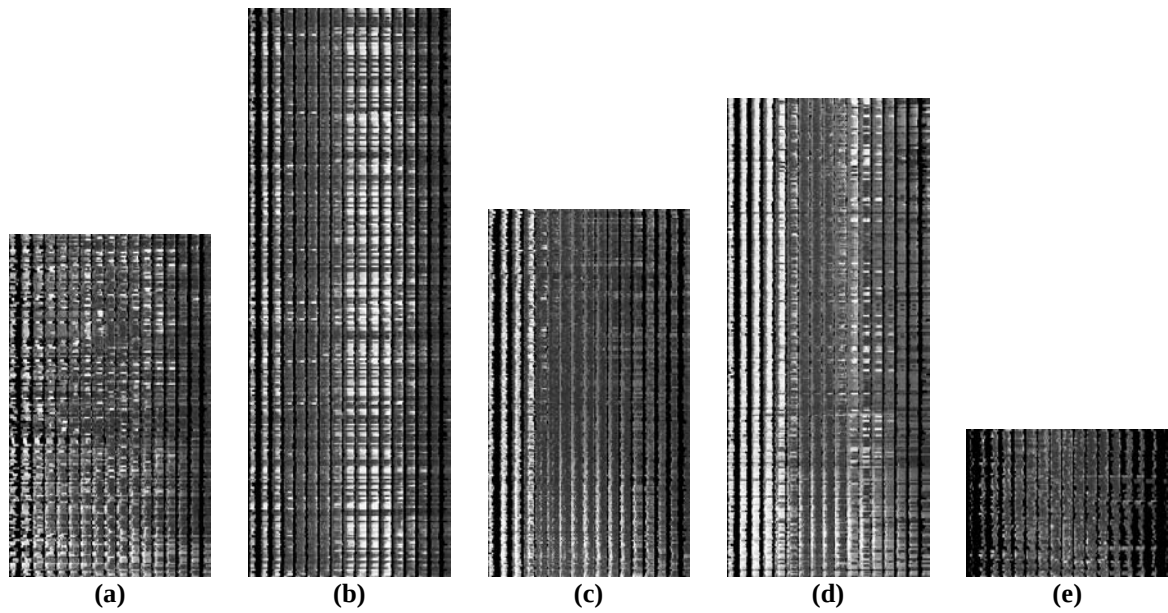


Figura 85: Representação visual do conjunto da Pista 3, descritor cinza sem fundo.

(a) Caminhão; (b) Carro; (c) Ônibus; (d) Van; (e) Moto.

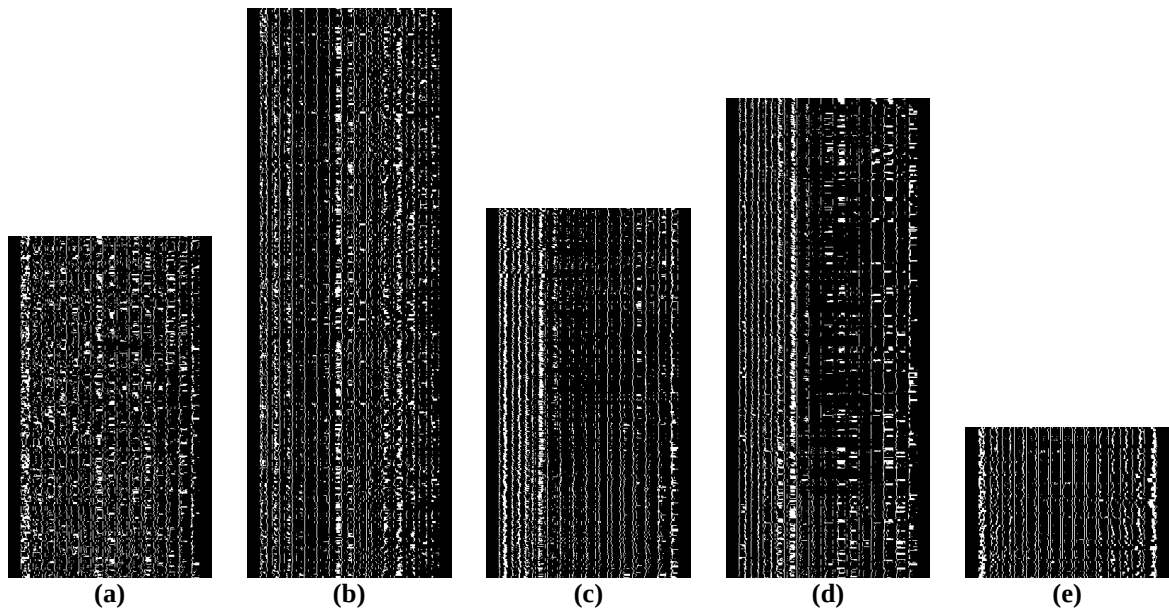


Figura 86: Representação visual do conjunto da Pista 3, descritor de bordas sem fundo.

(a) Caminhão; (b) Carro; (c) Ônibus; (d) Van; (e) Moto.

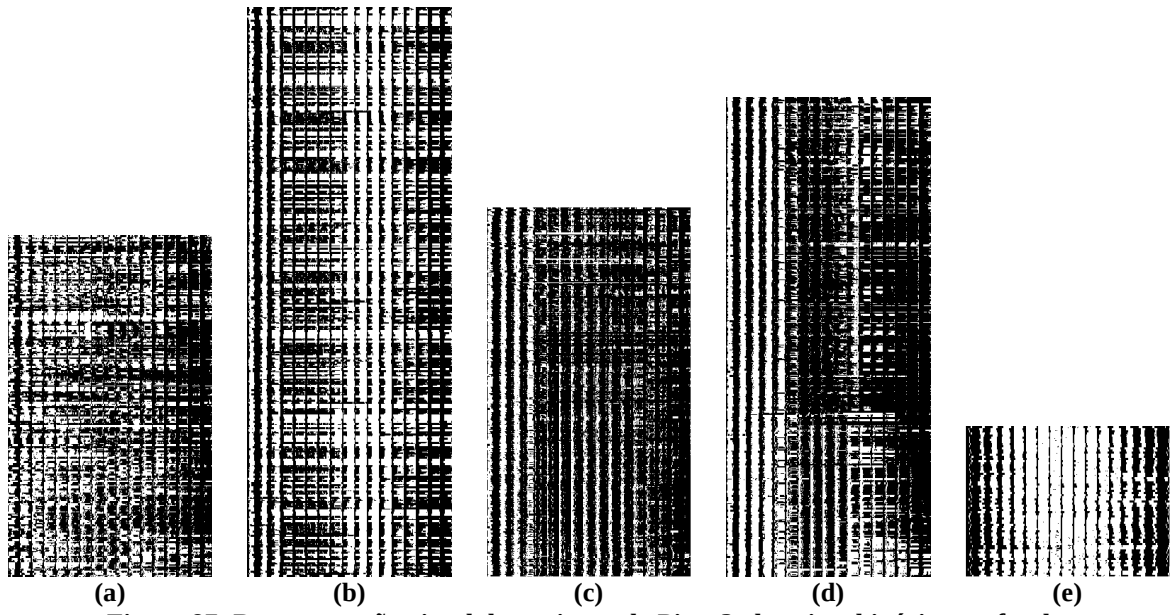


Figura 87: Representação visual do conjunto da Pista 3, descritor binário sem fundo.

(a) Caminhão; (b) Carro; (c) Ônibus; (d) Van; (e) Moto.

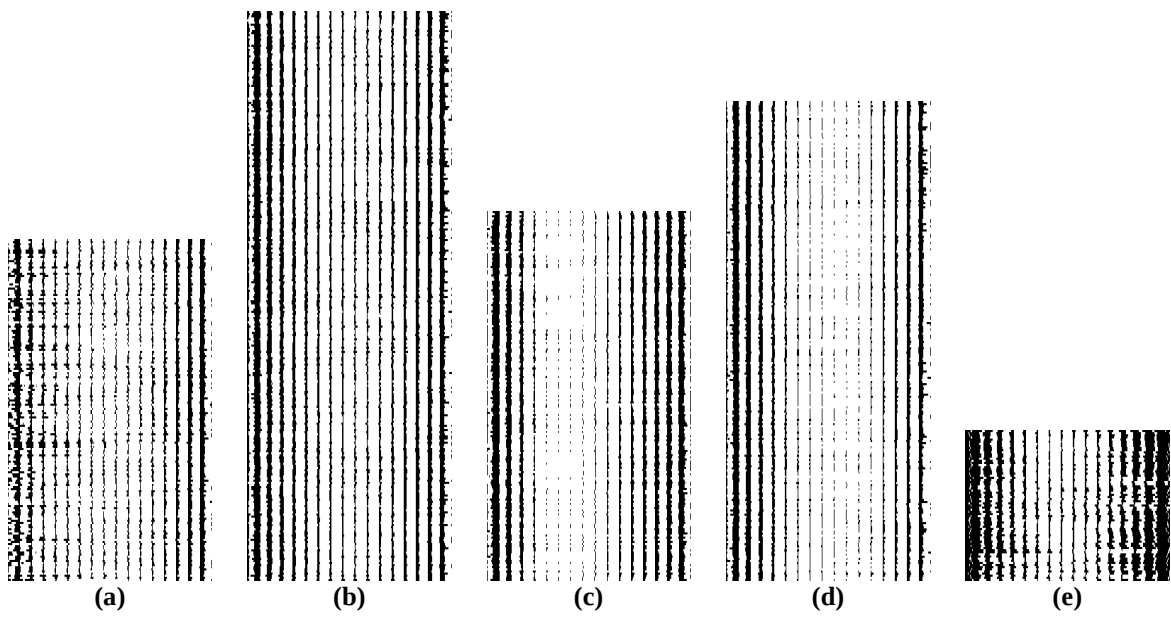


Figura 88 – Representação visual do conjunto da Pista 3, descritor silhueta sem fundo.

(a) Caminhão; (b) Carro; (c) Ônibus; (d) Van; (e) Moto.

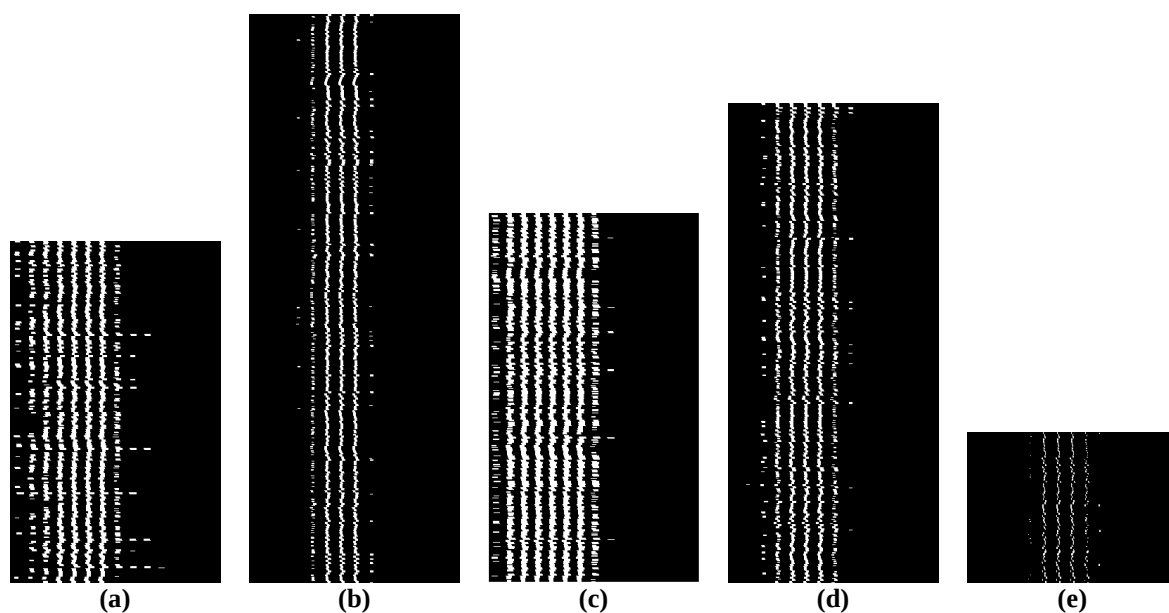


Figura 89: Representação visual do conjunto da Pista 3, descritor retângulo proporcional.
(a) Caminhão; (b) Carro; (c) Ônibus; (d) Van; (e) Moto.

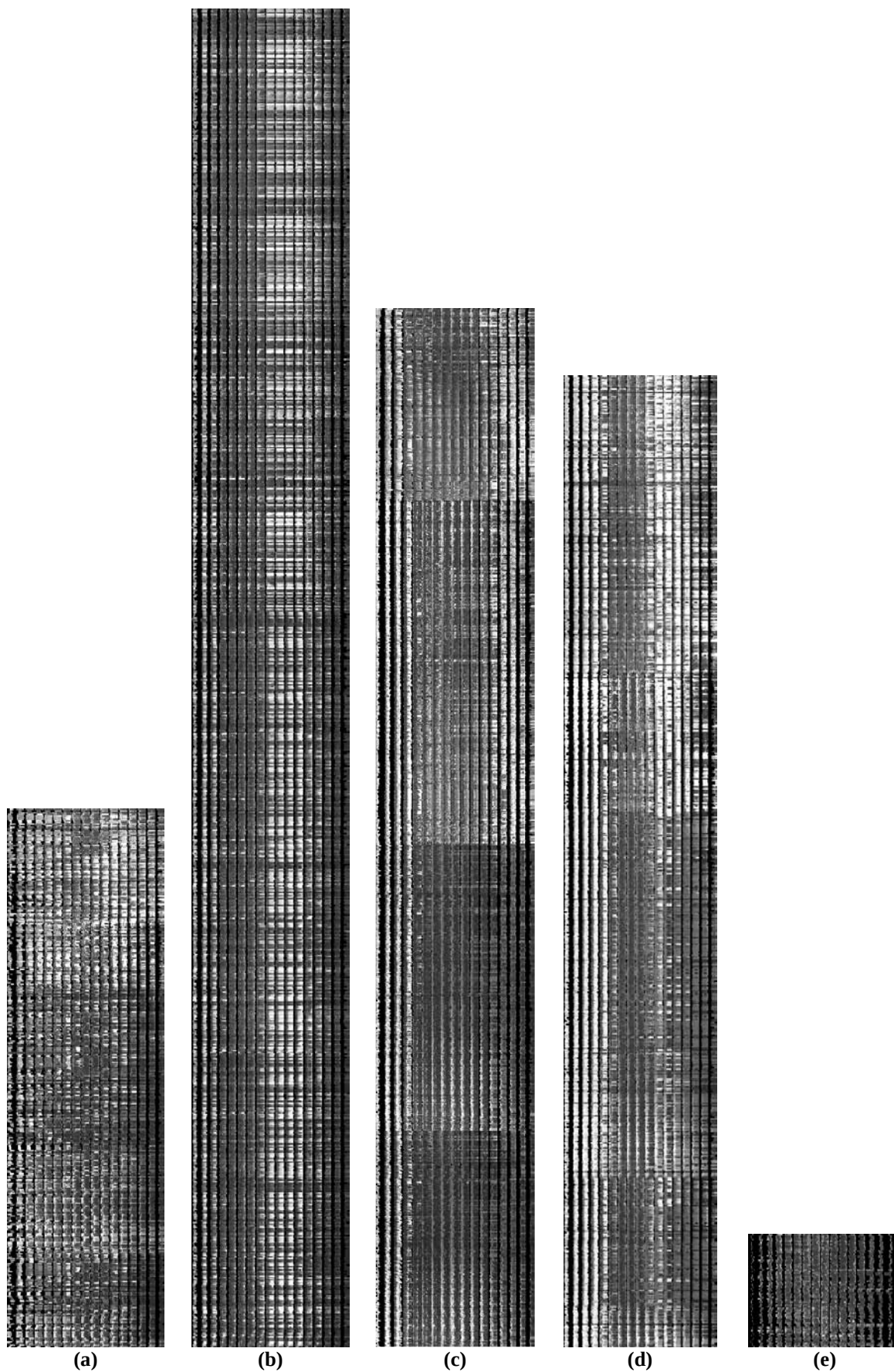


Figura 90- Representação visual do conjunto da Pista Global, descritor cinza sem fundo.
(a) Caminhão; (b) Carro; (c) Ônibus; (d) Van; (e) Moto.

11. Referências

- BAI, Hongliang ; WU, Jianping ; LIU, Changpin. Motion and haar-like features based vehicle detection. In: INTERNATIONAL MULTI-MEDIA MODELLING CONFERENCE, 12., 2006, Beijing. **Proceedings ...**, Beijing: IEEE, 2006
- BAS, E.K. ; CRISMAN, J.D. An easy to install camera calibration for traffic monitoring. Intelligent In: IEEE TRANSPORTATION SYSTEM CONFERENCE, 1997, Boston. **Proceedings ...** Boston, 1997. p. 362-366. ITSC 97
- BETKE, M. ; HARITAOGLU, E. ; DAVIS, L.S. Real-time multiple vehicle detection and tracking from a moving vehicle. **Machine Vision and Applications**, Berlin, v. 2, n. 2, p. 69–83, Aug. 2000.
- BISHOP, C. M. **Neural Networks for Pattern Recognition**. Oxford: Oxford University Press, 1995.
- CNT, **Confederação Nacional de Transporte**; Disponível em: <http://www.cnt.org.br/>. Acesso em: out. 2006.
- CANNY, J. A computational approach to edge detection. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, Piscataway, v.8, n. 6, p. 679-698. Nov. 1986.
- CAPPELLO, M. ; CAMPANI, M. ; SUCCI, A. Detection of lane boundaries, intersections and obstacles. In: INTELLIGENT VEHICLES '94 SYMPOSIUM, 1994, Paris. **Proceedings ...**, Tokyo, 2000.v. 1, p. 323-326.. VTC 2000.
- COLLINS, R. T. et al; **A system for video surveillance and monitoring**. , Pittsburgh PA: Robotics Institute, Carnegie Mellon University , May 2000. (CMU-RI-TR-00-12).
- DEMUTH, H. ; BEALE, M. **Neural network toolbox: user's guide version 4.0.1**, Natick, MA: MathWorks, 2005.
- _____ ; _____. **Image processing toolbox: user's guide version 5.1**. Natick, MA: MathWorks, 2005
- DENATRAN. Disponível em: http://www.denatran.gov.br/Frota_03.htm. Acesso em: out. 2006.
- DIEHL, C. P. **Toward efficient collaborative classification for distributed video surveillance**. 2000. Thesis (Doctoral) - Carnegie Institute of Technology, Carnegie Mellon University, Pittsburgh, PA, 2000.
- DORKO, G. ; SCHMID, C. Selection of scale-invariant parts for object class recognition. In: IEEE INTERNATIONAL CONFERENCE ON COMPUTER VISION, 9., 2003, Nice. **Proceedings ...** Nice, 2003. v. 2, p. 634-639.
- FORESTI, G. L. Object recognition and tracking for remote video surveillance. **IEEE Transactions on Circuits and Systems for Video Technology**, Piscataway, v. 9,n. 7, p. 1045-1062, Oct. 1999.

GONZALES, R. C. ; WOODS, R. E. **Processamento digital de imagens**, São Paulo: Edgard Blucher, 2000.

GUINGO, B. C. **Reconhecimento automático de placas de veículos automotores**. 2003. Dissertação (Mestrado em Informática), Instituto de Matemática, Núcleo de Computação Eletrônica, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2003.

GUO, Yanlin et al. Robust object matching for persistent tracking with heterogeneous features. In: JOINT IEEE INTERNATIONAL WORKSHOP ON VISUAL SURVEILLANCE AND PERFORMANCE EVALUATION OF TRACKING AND SURVEILLANCE, 2., 2005, Beijing. **Proceedings ...** Beijing, 2005. p. 81-88.

GUPTE, S. et al. Detection and classification of vehicles. **IEEE Transactions on Intelligent Transportation Systems**, Piscataway, v. 3, n. 1, p. 37-47, Mar. 2002.

_____ et al. Vision-based vehicle classification. In: 2000 IEEE INTELLIGENT TRANSPORTATION SYSTEMS CONFERENCE, 2000, Dearborn, MI. **Proceedings ...** Dearborn MI, 2000. p.46-51.

HAN, Dongjin et al. Vehicle class recognition from video-based on 3D curve probes. visual surveillance and performance evaluation of tracking and surveillance. In: JOINT IEEE INTERNATIONAL WORKSHOP ON, VISUAL SURVEILLANCE AND PERFORMANCE EVALUATION OF TRACKING AND SURVEILLANCE, 2., 2005, Beijing. **Proceedings ...** Beijing, 2005. p. 285-292.

HANDMANN, U. et al. An image processing system for driver assistance. **Image and Vision Computing** , Amsterdam, v. 18, n. 5, p. 367-376, Apr. 2000.

HAYKIN, S. **Redes neurais: princípios e prática**. 2. ed., Porto Alegre: Bookman, 2001

HEISELE, B. ; RITTER, W. Obstacle detection based on color blob flow. In: INTELLIGENT VEHICLES '95 SYMPOSIUM, 1995, Detroit. **Proceedings ...** Detroit: IEEE, 1995. p. 282-286.

HINZ, S. ; SCHLOSSER, C. ; REITBERGER, J. Automatic car detection in high resolution urban scenes based on an adaptive 3D-model. In: GRSS/ISPRS JOINT WORKSHOP ON REMOTE SENSING AND DATA FUSION OVER URBAN AREAS, 2., 2003. Berlin. **Proceedings ...** Berlin: IEEE, 2003 p.167-171

HSIEH, Jun-Wei et al. Automatic traffic surveillance system for vehicle tracking and classification, **IEEE Transactions on Intelligent Transportation Systems**, Piscataway, v. 7, n. 2, p. 175-187, Jun. 2006

HU, WEIMING et al.; A survey on visual surveillance of object motion and behaviors. **IEEE Transactions on Systems, Man and Cybernetics**, Piscataway, v. 34, Part C, n. 3, p. 334-352, Aug. 2004.

HUANG, Chung-Lin ; LIAO, Wen-Chieh. Vision-Based Vehicle Identification System, In: 17th INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION, 17.,2004, Trinity Lane. **Proceedings ...** Trinity Lane: IEEE, 2004. v. 4. p.364-367.

- ITSA, Intelligent Transportation Society of America. Disponível em: <http://www.itsa.org/>. Acesso em: out. 2006.
- ITS-AUSTRALIA. Disponível em: <http://www.its-australia.com.au/>. Acesso em: out. 2006.
- ITS-CANADA. Disponível em: <http://www.itscanada.ca/>. Acesso em: out. 2006.
- ITS-CHILE. Disponível em: <http://www.itschile.cl/>. Acesso em: out. 2006.
- ITS-DOT. Disponível em: <http://www.its.dot.gov/about.htm> . Acesso em: out. 2006.
- ITS-DOT-OVERVIEW. Disponível em: <http://www.itsoverview.its.dot.gov/>. Acesso em: out. 2006.
- ITS-KOREA. Disponível em: <http://www.itskorea.or.kr/>. Acesso em: out. 2006.
- ITSS. Disponível em: <http://www.ieee.org/itss>. Acesso em: out. 2006.
- ITS-SAS. Disponível em: <http://www.sasits.com/>. Acesso em: out. 2006.
- ITS-SINGAPORE. Disponível em: <http://www.itssingapore.org.sg>. Acesso em: out. 2006.
- JOHNSON, R. A. ; WICHERN, D. W. **Applied multivariate statistical analysis**. 3 ed. Appler Saddle River, NJ: Prentice-Hall, 1992. 642 p.
- JOLLIFFE, I. T. **Principal component analysis**. New York: Springer-Verlag, 1986
- JOLLY, M-P Dubuisson ; LAKSHMANAN, S. ; JAIN, A. K. Vehicle segmentation and classification using deformable templates. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, Piscataway, v. 18, n. 3, p. 293-308, Mar.1996.
- KHAMMARI, A. et al. Vehicle detection combining gradient analysis and AdaBoost classification. In: IEEE INTERNATIONAL CONFERENCE ON INTELLIGENT TRANSPORTATION SYSTEMS, 8., 2005, Vienna. **Proceedings ...** Vienna, 2005. p. 66-71.
- KOLLER, D. Moving object recognition and classification based on recursive shape parameter estimation. In: . ISRAELI CONFERENCE ON ARTIFICIAL INTELLIGENCE, 12., 1993. Tel-Aviv. **Proceedings ...** Tel-Aviv: IEEE, 1993. p. 27-28.
- KOLLER, D ; WEBER, J. ; MALIK, J. **Robust multiple car tracking with occlusion reasoning**. Berkeley: California Partners for Advanced Transit and Highways (PATH) - University of California, Berkeley. 2004. Working Papers. (UCB-ITS-PWP-94-1).
- LIM, T.R. ; GUNTORO, A.T. Car recognition using Gabor filter feature extraction. In: ASIA-PACIFIC CONFERENCE ON CIRCUITS AND SYSTEMS, 2002, Singapore. **Proceedings ...** Singapore: IEEE, 2002. v. 2, p. 451-455. APCCAS '02.
- LIU, T.; et al. Learning based symmetric features selection for vehicle detection. In: IEEE INTELLIGENT VEHICLES SYMPOSIUM, 2005. Las Vegas. **Proceedings ...** Las Vegas, 2005. p. 124-129.
- MCCULLOCH, W. S. ; PITTS, W. A logical calculus of ideas immanent in nervous activity. **Bulletin of Mathematical Biophysics**, New York, v. 5, n. 2, p. 115–133, 1943.

MALIK, J. ; RUSSELL, S. **Traffic surveillance and detection technology development: new traffic sensor technology final report**. Berkeley: California Partners for Advanced Transit and Highways (PATH) - University of California, Berkeley, 1997. Technical Report (UCB-ITS-PRR-97-6,).

MAREN, J. Configuring and optimizing the back-propagation network. In : MAREN, Aliana ; HARSTON, Craig ; PAP, Robert M. **Handbook of neural computing applications**, New York: Academic Press, 1999.

MATSUSHITA, T. et al. A feasibility study of vehicle-type classification based on symbolic representation. In: WORLD CONGRESS ON ITS, 9., 2002, Chicago. **Proceedings ...** Chicago: ITSSC, 2002.

MESSELODI, S. ; MODENA, C. M. ; ZANIN, M. A Computer vision system for the selection and classification of vehicles at urban road intersections. **Pattern Analysis and Applications**, London, v.8, n. 1-2, p. 17-31. Sept. 2005

MICHALOPOULOS, P.G. Vehicle detection video through image processing: the Autoscope system. **IEEE Transactions on Vehicular Technology**, Piscataway, v.40, n 1, Part 2, p. 21-29, Feb. 1991.

MLREPOSITORY, **Repository of machine learning data sets**. Irvine: University of California, 1998. Disponível em: <http://www.ics.uci.edu/~mlearn/MLRepository.html>. Acesso em: out. 2006.

MN/DOT, Minnesota Department of Transportation Office of Traffic Engineering/ ITS Section; **Evaluation of non intrusive technologies for traffic detection**; United States Department of Transportation Federal Highway Administration, 2001.

MOUTINHO, A. M. **Identificação de padrões faciais usando redes neurais**; 2005. Dissertação (Mestrado em Informática) - Instituto de Matemática, Núcleo de Computação Eletrônica, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2005.

NGUYEN, D. ; WIDROW, B. Improving the learning speed of 2-layer neural networks by choosing initial values of the adaptive weights. In: INTERNATIONAL JOINT CONFERENCE ON NEURAL NETWORKS, 1990, San Diego. **Proceedings ...** San Diego: IEEE, 1990. v. 3. p. 21-26, 1990.

OTA, R. et al. **Vehicle class recognition of street-parking vehicles from side-view range images**, In: WORLD CONGRESS ON INTELLIGENT TRANSPORT SYSTEMS AND SERVICES, 12., 2005, San Francisco. **Proceedings ...** San Francisco: IEEE, 2005.

OTSU, N. A threshold selection method from gray-level histograms; **IEEE Transactions on Systems, Man, and Cybernetics**, Piscataway, v. 9, n. 1, Jan. 1979. p. 62-66.

PETROVIC, V. S. , COOTES, T. F. Analysis of features for rigid structure vehicle type recognition. In: BRITISH MACHINE VISION CONFERENCE, 2004, London. **Proceedings ...** London: Kindon University, 2004.

_____ ; _____. Vehicle type recognition with match refinement. In: INTERNATIONAL CONFERENCE ON, PATTERN RECOGNITION, 17.,

2004. Cambridge, UK. **Proceedings ...** Cambridge, UK: IAPR, 2004. v. 3, p.95-98 ICPR 2004

PRAMPERO, P. ; CARVALHO, A. .P. L. Classifier combination for vehicle silhouettes recognition. In: INTERNATIONAL CONFERENCE ON IMAGE PROCESSING AND ITS APPLICATIONS, 7., 1999. Manchester UK. **Proceedings ...** Manchester, UK: IEEE, 1999. v.1, p.67-71.

PUMRIN, S. ; DAILEY, D. J.; Vehicle image classification via expectation-maximization algorithm. In: INTERNATIONAL SYMPOSIUM ON CIRCUITS AND SYSTEMS, 2003. Bangkok. **Proceedings ...** Bangkok: IEEE, 2003. v. 2. p. 468-471. ISCAS '03

RAD, Roya ; ZAMZAD, Mansour. Real time classification and tracking of multiple vehicles in highways. **Pattern Recognition Letters**, New York, v. 26 n. 10, p. 1597–1607, Jul. 2005

RITCHIE, S. G. et al **Anonymous vehicle tracking for real time freeway and arterial street performance measurement**; Berkeley: California Partners for Advanced Transit and Highways (PATH) - University of California, Berkeley. 2005. Research Reports (2005-9).

RODRIGUES, F. A. **Localização e reconhecimento de placas de sinalização utilizando um mecanismo de atenção visual e redes neurais**. 2002. Dissertação (Mestrado em Informática) - Coordenação do Curso de Pós-Graduação em Informática; Universidade Federal de Campina Grande, Campina Grande, PB. 2002.

ROJAS, J. C. ; CRISMAN, J. D. Vehicle detection in color images. In: IEEE CONFERENCE ON INTELLIGENT TRANSPORTATION SYSTEM, 1997. Boston. **Proceedings ...** Boston, 1997. p. 403-408.

RUMELHART, D. E. ; HINTON, G. E. ; WILLIAMS, R. J. Learning internal representations by error propagation. In: _____ : MCCLELLAND, J.L. **Parallel Distributed Processing: explorations in the microstructure of cognition**. Cambridge: MIT Press, 1986. v. 1. Foundations, p. 318-362.

SADOGHI-YAZDI, H. ; LOTFIZAD, M. ; FATHY, M. Car tracking by quantised input LMS, QX-LMS algorithm in traffic scenes. **IEE Proceedings Vision, Image and Signal Processing**, Stevenage, UK, v. 153, n. 1, p. 37-45, Feb. 2006.

SEDKY, M.H. ; MONIRI, M. ; CHIBELUSHI, C. C. Classification of smart video surveillance systems for commercial applications. In: IEEE CONFERENCE ON ADVANCED VIDEO AND SIGNAL BASED SURVEILLANCE, 2005, Como, Italy.. **Proceedings ...** Como, Italy, 2005. p. 638-643.

SIDLA, O. ; PALETTA, L. ; LYPETSKYY, Y. J. C. Vehicle recognition for highway lane survey. In: INTERNATIONAL IEEE CONFERENCE ON INTELLIGENT TRANSPORTATION SYSTEMS, 7., 2004. Washington, D.C. **Proceedings ...** Washington, D.C., 2004. p. 531-536.

SILVA, G. C.; **Automation of traffic flow measurement using video images**. 2001.; Thesis (Degree of M. Eng.) - Department of Computer Science and Engineering, University of Moratuwa, Sri Lanka, 2001.

SOTELO, M. A. et al. Road vehicle recognition in monocular images. In: IEEE INTERNATIONAL SYMPOSIUM ON INDUSTRIAL ELECTRONICS, 2005, Dubrovnik, Croatia. **Proceedings ...** Dubrovnik, Croatia, 2005. v. 4. p.1471-1476. ISIE 2005.

SRINIVASA, N.; Vision-based vehicle detection and tracking method for forward collision warning in automobiles. In: IEEE INTELLIGENT VEHICLE SYMPOSIUM, 2002. Versailles. **Proceedings ...** Versailles, 2002. v. 2. p. 626- 631

SULLIVAN, G. D. et al. Model-based vehicle detection and classification using orthographic approximations, **Image and Vision Computing**, Amsterdam, v. 15 n. 8, p. 649-654, 1997.

SUN, Zehang ; BEBIS, G. ; MILLER, R. On-road vehicle detection using evolutionary gabor filter optimization, **IEEE Transactions on Intelligent Transportation Systems**, Piscataway, v. 6, n. 2, p. 125-137, Jun. 2005.

_____ : _____ ; _____. Object detection using feature subset selection. **Pattern Recognition**, New York, v. 37, n. 11, p. 2165-2176, Nov. 2004.

_____ : _____ ; _____. On-road vehicle detection using Gabor filters and support vector machines. In: INTERNATIONAL CONFERENCE ON DIGITAL SIGNAL PROCESSING, 14., 2002. Santorini. **Proceedings ...** Santorini: IEEE, 2002. v. 2. p. 1019-1022.

_____ : _____ ; _____. Quantized wavelet features and support vector machines for on-road vehicle detection. In: INTERNATIONAL CONFERENCE ON CONTROL, AUTOMATION, ROBOTICS AND VISION, 7., 2002. Singapore. **Proceedings ...** Singapore, 2002. v. 3. p.1641-1646.

SUN, Zehang et al. A real-time precrash vehicle detection system. In: IEEE WORKSHOP ON APPLICATIONS OF COMPUTER VISION, 6., 2002. Orlando. **Proceedings ...** Orlando, 2002. p. 171-176.

TAN, T.N. ; BAKER, K.D. Efficient image gradient based vehicle localization. **IEEE Transactions on Image Processing**. Piscataway, v. 9, n. 8, p. 1343-1356, Aug. 2000

TARTARI, A. L. **Implementação de medidas tecnológicas para suporte estratégico na gerência de garagens subterrâneas**. 2002. ;Dissertação (Mestrado Profissionalizante em Engenharia) - Escola de Engenharia, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2002.

THIANG ; GUNTORO, A. T. ; LIM, R. Type of vehicle recognition using template matching method. In: INTERNATIONAL CONFERENCE ON ELECTRICAL, ELECTRONICS, COMMUNICATION AND INFORMATION, 2001, Jakarta. **Proceedings ...** Jakarta: IECI, 2001. CECT' 2001.

THOMANEK, F. ; DICKMANNS, E. D. ; DICKMANNS, D. Multiple object recognition and scene interpretation for autonomous road vehicle guidance. In: INTELLIGENT VEHICLES '94 SYMPOSIUM, 1994, Paris. **Proceedings ...** Paris: IEEE, 1994. p. 231-236.

TRANSITO_ONLINE. Disponível em:

<http://www2.rio.rj.gov.br/transitopub/entrada.cfm?pagina=7&menusel=6>. Acesso em: out. 2006.

TRANSITORIO. Disponível em: <http://transito.rio.rj.gov.br/>. Acesso em: out. 2006.

TSAI, Luo-Wei, et al. Vehicle detection using normalized color and edge map. **IEEE Transaction on Image Processing**, Piscataway, v. 16, n. 3, p. 850-864, Mar. 2007.

TZOMAKAS, C. ; SEELEN, W. **Vehicle detection in traffic scenes using shadows** Bochum: IR-IN, Institut für Neuroinformatik, Ruhr-Universität, Aug. 1998. (Technical Report 98-06).

VISSER, R. ; SEBE, N. ; LEW, M. S. Detecting automobiles and people for semantic video retrieval. In: INTERNATIONAL CONFERENCE ON, PATTERN RECOGNITION, 16., 2002. Québec City. **Proceedings ...** Los Amigos: IEEE, 2002. v. 2. ,p.733-736.

VOGL, T. P. et al. Accelerating the convergence of the backpropagation method. **Biological Cybernetics**, Heidelberg, v. 59, n. 4-5, p. 257-263, Sept. 1988.

WANG, Junxian ; BEBIS, G.; MILLER, R. Overtaking vehicle detection using dynamic and quasi-static background modeling. In: 2005 IEEE COMPUTER SOCIETY CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION, 2005, San Diego. **Proceedings ...** San Diego, 2005. p.:64-64.

WANG, Yuan-Kai ; CHEN, Shao-Hua **A robust vehicle detection approach..** In: IEEE CONFERENCE ON ADVANCED VIDEO AND SIGNAL BASED SURVEILLANCE, 2005, Como, Italy. **Proceedings ...** Como, Italy, 2005. p. 117-122.

WORRALL, A. D. , SULLIVAN, G. D. ; BAKER, K. D. **A simple, intuitive camera calibration tool for natural images**, In: CONFERENCE ON BRITISH MACHINE VISION, 5., 1994. York, UK. **Proceedings ...** York, UK: University of York, 1994. v. 2. p. 781-790.

WU, Bing-Fei ; LIN, Chuan-Tsai. A fuzzy vehicle detection based on contour size similarity. In: IEEE INTELLIGENT VEHICLES SYMPOSIUM, 2005, Las Vegas. **Proceedings...** Las Vegas, 2005. p.496 – 501.

XIE, Lei. et al. Robust vehicles extraction in a video-based intelligent transportation systems, In: INTERNATIONAL CONFERENCE ON COMMUNICATIONS, CIRCUITS AND SYSTEMS, 2005. Hong-Kong. **Proceedings. ...** Hong-Kong: IEEE, 2005. v. 2, p. 27-30.

YANG, Hao, et al. Efficient and robust vehicle localization. In: 2001 INTERNATIONAL CONFERENCE ON IMAGE PROCESSING, 2001, Thessalonik, **Proceedings ...** Thessaloniki: IEEE, 2001. v.2, p. 355-358.

ZHANG, Chengcui ; CHEN, Xin ; CHEN, Wei-Bang. A PCA-based vehicle classification framework. In: INTERNATIONAL CONFERENCE ON DATA ENGINEERING WORKSHOPS, 2006, 22., 2006, Atlanta, GA. **Proceedings ...** Atlanta, GA: IEEE, 2006, p.17-17.

_____. et al. Adaptive background learning for vehicle detection and spatio-temporal tracking. In: 2003 JOINT CONFERENCE OF THE INTERNATIONAL CONFERENCE ON INFORMATION, COMMUNICATIONS AND SIGNAL PROCESSING, PACIFIC RIM CONFERENCE ON MULTIMEDIA., 4., 2003, Singapore. **Proceedings ...** Singapore: IEEE, 2003. v. 2. p. 797-801.

_____ ; SIYAL, M.Y.; A new segmentation technique for classification of moving vehicles. In: 2000 IEEE VEHICULAR TECHNOLOGY CONFERENCE, 51., 2000, Tokyo. **Proceedings ...** , Tokyo, 2000.v. 1, p. 323-326.. VTC 2000.