PAMELA ROSA TANCREDI

**MONITORAMENTO DO ACESSO DE VEÍCULOS DE CARGA EM VIAS
URBANAS**

Dissertação apresentada à Escola de Engenharia de São Carlos, como parte integrante dos requisitos para a obtenção do título de Mestre em Ciências, Programa de Pós Graduação em Engenharia de Transportes. Área: Planejamento e Operação de Sistemas de Transportes.

Orientador: Prof. Titular D. João Alexandre Widmer

São Carlos, SP

2012

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

T156m Tancredi, Pamela Rosa
Monitoramento do acesso de veículos rodoviários de
carga em vias urbanas. / Pamela Rosa Tancredi ;
orientador João Alexandre Widmer. São Carlos, 2012.

Dissertação (Mestrado - Programa de Pós-Graduação em
Engenharia de Transportes e Área de Concentração em
Planejamento e Operação de Sistemas de Transportes)--
Escola de Engenharia de São Carlos da Universidade de São
Paulo, 2012.

1. Monitoramento de veículos de carga. 2.
Processamento de imagens. 3. *MATLAB*. I. Título.

FOLHA DE JULGAMENTO

Candidata: Engenheira **PÂMELA ROSA TANCREDI**.

Título da dissertação: "Monitoramento do acesso de veículos rodoviários de carga em vias urbanas".

Data da defesa: 04/06/2012

Comissão Julgadora:

Resultado:

Prof. Titular **João Alexandre Widmer (Orientador)**
(Escola de Engenharia de São Carlos/EESC)

Aprovada

Prof. Associado **José Reynaldo Anselmo Setti**
(Escola de Engenharia de São Carlos/EESC)

Aprovada

Prof^a. Dr^a **Ana Paula Camargo Larocca**
(Escola Politécnica/USP)

Aprovada

Coordenador do Programa de Pós-Graduação em Engenharia de Transportes e
Presidente da Comissão de Pós-Graduação
Prof. Associado **Paulo Cesar Lima Segantine**

AGRADECIMENTOS

Agradeço a Deus, pela sabedoria e persistência concedida.

A família, Cleunice, Percio, Nath e Thay, meus melhores amigos, pelas palavras de força e incentivo, mesmo nos momentos em que eu nem sabia o que precisava ouvir.

Ao Wilson Ricardo, noivo, amigo, companheiro e maior incentivador desse trabalho.

Aos amigos do Departamento de Engenharia de Transportes, pelas noites em claro, pelas gargalhadas, por todo o companheirismo. Em especial, aos amigos André Cunha e Glêvson Franco, pelos primeiros conselhos na área da pesquisa em processamento de imagens, pela assistência didática, pelos inúmeros artigos, pelas correções, enfim, por toda a ajuda essencial ao desenvolvimento do trabalho.

Ao professor João Alexandre Widmer, pela orientação, acompanhamento e aprendizado.

Aos membros da banca avaliadora, professora Ana Paula Camargo Larocca e professor José Reynaldo Setti, pelas críticas e elogios durante a defesa e especialmente pela contribuição no melhoramento da versão final deste trabalho. Agradeço ainda ao Professor Adilson Gonzaga, pelos valiosos conselhos no exame de qualificação.

Aos professores e funcionários do Departamento de Transportes da Escola de Engenharia de São Carlos, pelo apoio e dedicação ao ensino. Em especial a Heloísa e Beth, por não medirem esforços no auxílio aos alunos.

As amigas de infância, Élide e Carol, meus sinceros agradecimentos por todo o incentivo e pensamentos bons. Aos amigos da Universidade Federal de Mato Grosso, obrigada por fazerem os 5 anos da graduação em engenharia civil, os mais felizes que eu poderia ter. Em especial as minhas queridas Renata Conciani, Livia Fuji, Sara Brígida, Thais Cristina, Jéssica Gálio, Juliana Friedrich, vocês foram e ainda são as melhores amigas que alguém pode querer. Aos professores Luiz Miguel e Marilda Serra, que ainda na faculdade, acreditaram no meu projeto. Aos queridos amigos que tive o prazer de conhecer e conviver durante os 3 anos de

mestrado, em São Carlos e São Paulo, obrigada pelos inúmeros momentos de alegria, em especial a minha tia querida e madrinha Patrícia Tancredi, Eunice e Raquel que por inúmeras vezes dividiram suas casas comigo.

Aos amigos das empresas Arcadis Logos e Figueiredo Ferraz, obrigada por apreciarem e incentivarem o desenvolvimento acadêmico dentro do ambiente de trabalho, em especial a Daniela Pereira, Marta Duarte, Osvaldo Azevedo, Adriana Vivan e Ana Salles (*in memorian*) pelos inúmeros conselhos e palavras de incentivo. Aos amigos Clarissa Yebra, Sidnei Ono, Poliana Avelar e Adriano Bittar, vocês sabem que foram essenciais ao desenvolvimento desta pesquisa.

Aos meus novos companheiros de Escola Politécnica, Carla, Tabiani, Fabi, Alessandro e Henry, obrigada por toda a gentileza, a amizade, pela cobertura nos trabalhos, pelas listas de exercícios, por me ajudarem demais nos meses finais do mestrado.

Finalmente, obrigada a todos, que não citei o nome, mas que de alguma maneira contribuíram para a realização do projeto de mestrado, que comecei a construir ainda em Cuiabá, e hoje é realidade.

Ao CNPQ - Conselho Nacional de Desenvolvimento Científico e Tecnológico, pela bolsa concedida para o desenvolvimento desta pesquisa.

RESUMO

Tancredi, P. R. (2012). Monitoramento do acesso de veículos de carga em vias urbanas. São Carlos, 2012, 89 p. Dissertação (Mestrado) – Escola de Engenharia de São Carlos, Universidade de São Paulo.

O monitoramento do trânsito de veículos rodoviários de carga em vias urbanas é dependente do trabalho de profissionais que precisam estar integralmente atentos ao tráfego de uma determinada via, e assim, identificar a presença desses veículos em meio a cenas complexas. Sistemas automáticos de identificação de veículos em quadros de vídeos podem agregar valor na qualidade de vida das cidades, pois tendem a diminuir os congestionamentos e incrementam segurança ao trânsito. Nesta pesquisa, propõe-se um modelo capaz de identificar e segregar veículos rodoviários de carga em vias urbanas, por meio de técnicas de processamento e análise de vídeos, gravados por câmeras posicionadas em diversos ângulos em dispositivos de acessos a vias urbanas. O sistema foi desenvolvido em linguagem MATLAB® e foram abordadas técnicas referentes às etapas de detecção, segmentação e processamento de objetos em fundos complexos. Ao final do trabalho, o sistema é aplicado em cenários de filmagem na cidade de São Carlos - SP e os resultados obtidos são discutidos.

Palavras-chaves: Detecção de veículos, Monitoramento de veículos de carga, Processamento de imagens, MATLAB.

ABSTRACT

Tancredi, P. R. (2012). Monitoring of freight vehicles access on urban roads. São Carlos, 2012, 89 p. Dissertação (Mestrado) – Escola de Engenharia de São Carlos, Universidade de São Paulo.

The monitoring of the road freight vehicles traffic on urban roads relies on the work of professionals who need to be fully aware of the traffic on a specific road, and thereby identify the presence of these vehicles in complex scenes. Automatic vehicle identification in video shots may add value to the city life quality, due to congestion diminishment and traffic security improvement. In this research, a model able to identify and segregate road freight vehicles on urban roads is proposed, by means of video analysis and processing, that were recorded by cameras positioned at different angles on highways accesses on urban roads. The system was developed using MATLAB®. Techniques related to detecting, targeting and processing steps have been used to identify objects in complex backgrounds. At the end of the work, the system is applied to several scenarios that have been filmed in the city of São Carlos –SP and the results are discussed.

Keywords: Detection of vehicles, Vehicle tracking, Image processing, MATLAB.

SUMÁRIO

LISTA DE FIGURAS.....	iii
LISTA DE TABELAS	vi
Capítulo 1	1
INTRODUÇÃO	1
1.1 Introdução e Justificativa.....	1
1.2 Importância do Estudo	2
1.3 Objetivo.....	2
1.4 Organização do Trabalho.....	2
Capítulo 2	5
REVISÃO BIBLIOGRÁFICA	5
2.1 Transporte Urbano de Cargas.....	5
2.2 Restrições de Tráfego.....	7
2.3 Processamento de Imagens Digitais.....	7
2.3.1 Operações lógicas com imagens.....	11
2.3.2 Técnicas para segmentação baseadas em análise de imagens...	14
A) Limiarização (<i>Thresholding</i>)	14
B) Características Geométricas	16
2.3.3 Detecção de fundo baseada em detecção de movimento	17
a) Diferença entre quadros de plano de fundo (<i>Background frame</i> <i>differencing</i>)	17
B) Diferença entre quadros (<i>Inter-frames differencing</i>)	19
C) Estimativa de Plano de Fundo (<i>Background Estimation</i>)	20
D) Misturas Gaussianas.....	21

2.3.4 Reconhecimento e Interpretação de imagens	22
2.4 Detecção e Identificação de Veículos	23
2.4.1 Sistemas de Detecção e Identificação de Veículos Baseados em Processamento de Imagem	26
2.5 Softwares e Ferramentas	39
Capítulo 3	41
MÉTODO	41
3.1 Modelo de identificação de veículos utilizando estimativa de background	41
Capítulo 4	59
APLICAÇÃO DO MÉTODO	59
4.1 Modelo Implementado	59
4.2 Equipamentos de Hardware	59
4.3 Aplicação aos “Portões de entrada da cidade de São Carlos, SP”.	60
Capítulo 5	71
CONSIDERAÇÕES FINAIS	71
5.1 Conclusões	71
5.2 Sugestões para Trabalhos Futuros	72
REFERÊNCIAS	73
APÊNDICES	81

LISTA DE FIGURAS

Figura 2. 1 – Método para o processamento de imagens digitais.	9
Figura 2. 2 – Elementos de Análises de imagens	10
Figura 2. 3 - Exemplo de operação AND entre imagens monocromáticas: (a) X, (b) Y, (c) $X \wedge Y$	12
Figura 2. 4 - Exemplo de operação OR entre imagens monocromáticas: (a) X, (b) Y, (c) $X \vee Y$	13
Figura 2. 5 - Exemplo de operação XOR entre imagens monocromáticas: (a) X, (b) Y, (c) $X \oplus Y$	13
Figura 2. 6 - Exemplo de operação NOT sobre imagem monocromática: (a) X, (b) NOT X.....	13
Figura 2. 7 – Instalação do laço de indução	24
Figura 2. 8 – Sensores indutivos	25
Figura 2. 9 – Quadro típico utilizado – resolução de 640 x 480 <i>pixels</i>	27
Figura 2. 10 – Detecção e rastreamento de veículos – resultados.....	28
Figura 2. 11 – Modelos utilizados para a classificação	29
Figura 2. 12 – Modelo de imagem utilizada no trabalho e a detecção de veículos	29
Figura 2. 13 – Exemplo de classificação correta	30
Figura 2. 14 – Exemplo de classificação equivocada – o sistema identifica pedestres como veículos.....	30
Figura 2.15 – Modelo de classificação proposto	31
Figura 2. 16 – Exemplo de classificação correta	31
Figura 2. 17 – Exemplo de classificação equivocada – o sistema identifica três veículos como um único	31

Figura 2. 18 - Fixação da câmera para coleta da perspectiva superior do veículo.....	32
Figura 2. 19 - Modelo de imagem coletada para análise.....	33
Figura 2. 20 – Evolução da adaptação do plano de fundo.	34
Figura 2. 21 – a) quadro de entrada; b) plano de fundo estimado; c) resultado da diferença entre a e b; d) agrupamento pelas características do “vizinho mais próximo”.	36
Figura 2. 22 - (i) Quadro de entrada; (ii) Plano de fundo inicial, obtido com o primeiro valor de limiar T; (iii) Iteração n.75, (iv) iteração n.120, (v) estimativa do plano de fundo.....	37
Figura 2. 23 – Comparação entre (b) Método de Harasse <i>et. al.</i> , 2006 e (c) Misturas Gaussianas.....	38
Figura 3. 1 – Modelo para identificar veículos rodoviários de carga utilizando MATLAB®	42
Figura 3. 2 – Componente do bloco Estimativa de Plano de Fundo.....	44
Figura 3. 3 – Bloco Rastreamento do Veículo.....	46
Figura 3. 4 – Componente do bloco Rastreamento do Veículo.....	47
Figura 3. 5 – Parâmetros da função de análise de Blob.....	49
Figura 3. 6 – Parâmetros da função de análise de Blob.....	50
Figura 3. 7 – Subcomponente do componente <i>Region Filtering</i>	51
Figura 3. 8 – Componente do bloco Mostrando os Resultados.....	52
Figura 3. 9 – Janela para inserção dos parâmetros do programa.....	54
Figura 3. 10 – Sequência do trabalho de programação	55
Figura 3. 11- Croqui do método para a coleta de dados.	56
Figura 4. 1 – Diagrama da sequência da aplicação do sistema	59
Figura 4. 2 – Ilustração do teorema de Tales.....	61
Figura 4. 3 – Esquema de cálculo da relação <i>pixel</i> x milímetro do trabalho.....	62

Figura 4. 4 – Erro de identificação de veículo	63
Figura 4. 5 – Erro de identificação de objeto	63
Figura 4. 6 – Resultados de detecção.....	67
Figura 4. 7 - Resultados de detecção.....	68
Figura 4. 8 - Resultados de detecção.....	69

LISTA DE TABELAS

Tabela 3. 1 – Padrões do vídeo de entrada	43
Tabela 4. 1 – Resultados das análises.....	64
Tabela 4. 2 – Definição dos parâmetros do modelo	65

CAPÍTULO 1

INTRODUÇÃO

1.1 INTRODUÇÃO E JUSTIFICATIVA

A circulação inadequada de caminhões no perímetro urbano pode gerar um grande número de impactos negativos de ordem social, econômica e ambiental que vão além do aumento dos congestionamentos nas regiões mais centrais. Problemas como restrições de tráfego e o aumento do risco sobre os demais veículos e sobre os pedestres, conflitos entre o transporte urbano e o transporte de carga, obstruções ao trânsito com consequente redução da velocidade, maiores tempos de viagem, incremento na poluição sonora, atmosférica e visual, deterioração precoce dos pavimentos urbanos, afetam a qualidade de vida da população e de certo modo, acabam aumentando o custo do serviço de transporte de carga urbana e assim o custo final dos produtos. Tratando-se dos veículos de transporte de cargas, particularmente nas zonas urbanas das capitais e grandes cidades brasileiras, há alguns anos são adotadas medidas restritivas à circulação destes veículos.

A restrição do acesso de veículos rodoviários de carga pode ser considerada uma medida de rápida implantação, porém com difícil monitoramento, pois dependem da capacidade humana de percepção e atenção. Sistemas baseados na captação automática de imagens de veículos em meio ao tráfego, por exemplo, como os estudados em (Gupte *et. al.*, 2002; Toulminet, *et. al.*, 2002; Sullivan *et. al.*, 1997) tiveram como foco a identificação de veículos por meio da análise de imagens. Atualmente utilizando a capacidade humana de reconhecimento, uma simples câmera de vigilância colocada num local relevante pode ser a forma mais econômica para a fiscalização da entrada de veículos não autorizados em determinadas zonas urbanas, no entanto é necessário que existam operadores que constantemente visualizam e avaliam as imagens enviadas. Essa análise é dificultada pela presença de outras cenas que tiram a atenção até mesmo dos operadores com as visões melhor treinadas. Assim, o desafio do emprego da visão computacional aliada a técnicas de inteligência artificial é realizar tarefas semelhantes à visão humana, acrescentando a capacidade de analisar várias imagens em poucos segundos.

O desenvolvimento desta pesquisa de mestrado propõe um sistema que seja capaz de segregar veículos rodoviários de carga, por meio da análise de imagens de vídeos gravados em vias urbanas. A ferramenta computacional visa apoiar o trabalho de controladores de trânsito que, ao observarem simultaneamente vídeos de tráfego urbano de vários locais, devem voltar sua atenção para situações em que o veículo de interesse esteja presente na imagem captada pela câmera.

1.2 IMPORTÂNCIA DO ESTUDO

Atualmente, todo o monitoramento da entrada dos veículos rodoviários de carga nas vias urbanas das cidades brasileiras, quando existente, é realizado de forma manual, ou seja, agentes fiscalizadores ficam locados em pontos estratégicos das cidades, autuando os veículos infratores. Porém, esse método de fiscalização demanda a intensa utilização de recursos humanos e ainda está sujeito a falhas, pois, mesmo os fiscais com maior experiência estão sujeitos a instante de cansaço e desatenção.

Assim, um sistema que automatize o processo de segregação dos veículos rodoviários de carga entre os outros veículos em um local de interesse, como por exemplo, os principais acessos de uma cidade, tem grande importância e potencial como facilitador do trabalho de controle e restrição do acesso de veículos rodoviários de carga.

1.3 OBJETIVO

O objetivo desta pesquisa é propor e avaliar um sistema de processamento e interpretação de imagens de vídeos que permita a pré-seleção de imagens, onde estejam presentes veículos rodoviários de carga com dimensões e características tais, que mereçam a atenção dos controladores de trânsito de modo a visualizar com maior facilidade a cena que contém tal veículo em vias de acesso a áreas urbanas.

1.4 ORGANIZAÇÃO DO TRABALHO

Além deste capítulo introdutório, este trabalho está estruturado em cinco capítulos assim distribuídos:

No Capítulo 2 – Revisão Bibliográfica – é apresentada a revisão da literatura sobre transporte urbano de cargas no Brasil, métodos de processamento e interpretação de imagens, métodos de detecção de veículos em meio a cenas de tráfego, softwares e ferramentas para análise e processamento de imagens.

No Capítulo 3 – Método – descreve o método empregado na realização desse trabalho, apresentando o diagrama de blocos utilizado para a identificação e segregação dos veículos rodoviários de carga nas cenas, assim como, as funcionalidades dos blocos implementados e suas respectivas variáveis adotadas.

No Capítulo 4 – Aplicação do Sistema - são apresentados os resultados obtidos na aplicação do sistema desenvolvido a uma série de vídeos de vias urbanas de acesso à cidade de São Carlos, SP.

Capítulo 5 – Considerações finais - são apresentadas as conclusões sobre a aplicação do sistema desenvolvido, assim como sugestões e recomendações para trabalhos futuros de desenvolvimento do conhecimento nessa área.

CAPÍTULO 2

REVISÃO BIBLIOGRÁFICA

A revisão bibliográfica apresentada situa o trabalho desenvolvido no contexto da Engenharia de Transportes e da Logística Urbana. O capítulo tem início com o estudo sobre transporte urbano de cargas. Logo após, destaca os aspectos importantes e necessários para o desenvolvimento prático do trabalho, que é o estudo sobre a detecção de objetos e o processamento de imagens digitais.

2.1 TRANSPORTE URBANO DE CARGAS

A movimentação urbana de carga está diretamente relacionada com a vida da população nas cidades e cresce na medida em que a necessidade da entrega “porta a porta” cresce também. Os estabelecimentos comerciais assim como os consumidores precisam que seu produto seja entregue com eficiência e qualidade. No entanto, a circulação inadequada de caminhões no perímetro urbano gera um grande número de impactos negativos de ordem social, econômica e ambiental, que vão além do aumento dos congestionamentos nas regiões mais centrais. Problemas como restrições de tráfego e o aumento do risco sobre os demais veículos e sobre os pedestres, conflitos entre o transporte urbano e o transporte de carga, obstruções ao trânsito com consequente redução da velocidade, maiores tempos de viagem, incremento na poluição sonora, atmosférica e visual, deterioração precoce dos pavimentos urbanos, afetam a qualidade de vida da população e de certo modo acabam aumentando o custo do serviço de transporte de carga urbana e assim o custo final dos produtos (Caixeta e Martins, 2001). De fato, observa-se a queda da qualidade de vida em diversos aspectos nos centros urbanos e o problema do transporte de carga urbana só aumenta essa queda, pois cresce muito além da capacidade da infraestrutura das cidades.

Lima Jr. (2005) enfatiza bem que apesar dos inúmeros impactos negativos, a movimentação de cargas urbanas é fundamental para a economia tanto global quanto local. A modernização do ambiente industrial, o aumento significativo da competitividade e o crescimento dos mercados implicam que todos os diferenciais

logísticos passam a ser importantes, com destaque para os movimentos das cargas urbanas. Assim, o operador passa a buscar incessantemente a redução do custo total da cadeia de suprimentos, onde os custos de distribuição e coleta urbana são parcela importante. Transportar o maior volume de carga, com o menor número de veículos passa a ser uma prioridade de economia.

Sendo inegável a importância dos sistemas de transportes de cargas urbanas, a eficiência e a produtividade desses sistemas dependem de esforços conjuntos envolvendo o setor público e privado. O objetivo do sistema de transporte de carga é o deslocamento de bens de um ponto para outro da rede logística, respeitando as restrições de integridade de carga e confiança (Novaes, 1986¹ *apud* Faruolo, 2007). O setor público deve proporcionar as condições necessárias de infraestrutura e regulamentação e ao setor privado cabe o desenvolvimento de sistema logístico do transporte de cargas urbanas da melhor maneira possível para racionalizar suas atividades.

As movimentações urbanas de carga têm como principais dificuldades no sistema viário das cidades os raios de curvatura, largura de vias e deficiências ou inexistência de pontos de carga e descarga (Chiesa, 1999; Lima Jr., 2005). Por outro lado, veículos rodoviários de carga, trafegando muitas vezes com excesso de peso, provocam a deterioração dos pavimentos urbanos, que não foram projetados para o trânsito de veículos pesados.

O problema gerado pelo trânsito de veículos rodoviários de cargas em vias urbanas é tema de diversos estudos pelo mundo, visando resolver ou minimizar os impactos gerados por essa movimentação. Nesse contexto, está inserido o conceito de *City Logistics*, ou Logística Urbana, que busca otimizar a movimentação de cargas e veículos dentro das cidades da melhor maneira possível de modo a proporcionar a combinação entre o bem estar social, a economia, e o meio ambiente (Taniguchi *et. al.*, 2001; BESTUFS, 2008). Thompson, (2003²) *apud* Dutra, 2004, definem *city logistics* como um processo de planejamento, baseado em um sistema de

¹ **NOVAES, A.G.N.**; Sistemas de transportes, 3 volumes, Editora Edgard Blucher Ltda, São Paulo, 1986.

² **THOMPSON, R. G.** *AusLink green paper submission, Freight and Logistics Group*, Department of Civil and Environmental Engineering, The University of Melbourne, February, 2003.

integração, que promove inovações e reduz os custos econômico, social e ambiental. Ainda segundo Dutra (2004) a caracterização do fluxo de veículos e, mais especificamente, de caminhões (quantidade, tipo, número de eixos, dimensões e pesos etc.) é importante para o devido planejamento viário (dimensionamento de pavimentos e vias, dimensões de caixas de vagas de estacionamento etc.).

2.2 RESTRIÇÕES DE TRÁFEGO

Medidas de restrição ao trânsito de veículos são utilizadas por diversos países com êxito há pelo menos 30 anos. Em se tratando de veículos em geral, principalmente automóveis, as principais medidas de restrição veicular são o rodízio de veículos e a cobrança de pedágios urbanos.

Para os veículos rodoviários de carga, além de estarem sujeitos às mesmas restrições dos automóveis, que tem como principal objetivo monitorar o trânsito e a emissão de poluentes (Jones e Hervik, 1992), também tem, como principal medida de restrição de tráfego, a implantação das zonas de restrição de circulação, onde os veículos rodoviários de carga são proibidos de circular em determinados horários. Segundo Miles *et. al.* (1998) o objetivo principal dos sistemas de monitoramento de acesso é monitorar a entrada de veículos pré- determinados por características que possuem (placa, tipo, etc.) em áreas definidas, durante determinados períodos de tempo, reduzindo o impacto do trânsito na área ou ligação com restrição, além dos objetivos supracitados. Essas medidas são adotadas em cidades de grande e médio porte, pois usualmente tem baixo custo de implantação, porém o custo de operação pode ser alto de acordo com a fiscalização adotada.

2.3 PROCESSAMENTO DE IMAGENS DIGITAIS

Grande parte dos sistemas de detecção e análise de imagens em vídeos, inicialmente detecta o movimento com o objetivo de segmentar a imagem, separando o objeto em movimento do fundo. Posteriormente, podem ser implementadas técnicas para o acompanhamento da trajetória do objeto e a gravação das imagens para análise futura. O processamento de detecção do movimento normalmente envolve tanto a modelagem do ambiente, que consiste na representação do fundo da imagem – *background* - como a segmentação do objeto

em movimento e a sua posterior classificação em uma categoria pré-determinada (Wang *et. al.*, 2004). As variantes para a correta detecção do movimento na cena são as condições da aquisição da imagem, tais como: resolução da câmera, taxa de quadros por minuto, ruído de aquisição e condições do ambiente externo, como iluminação, sombras, ventos, reflexos e etc. (Toyama *et. al.*, 1999).

Assim, as imagens adquiridas devem ser manipuladas de modo a segmentar o objeto desejado, do restante da cena e reconstruí-lo no “mundo virtual”. À técnica de manipular imagens com o objetivo de obter dados de entrada e saída em forma de imagens dá-se o nome de Processamento de Imagens. Segundo Gonzalez e Woods (2007), é dividida nas seguintes etapas: aquisição da imagem, pré-processamento, segmentação, representação e/ou descrição e reconhecimento e/ou interpretação.

A aquisição da imagem é feita por meio de um dispositivo de entrada – *hardware* - como câmeras digitais, por exemplo. O pré-processamento visa deixar a imagem mais “limpa”, corrigir possíveis falhas na aquisição e ainda realçar características importantes, como bordas e vértices. A etapa de segmentação é responsável por agrupar os *pixels* pertencentes a um mesmo objeto ou região de interesse. A representação e descrição busca extrair características de interesse que possibilitem a descrição e agrupamento em classes de objetos. Por fim, a etapa de reconhecimento procura atribuir um rótulo ao objeto, baseado nas informações de uma base pré-determinada (Gonzalez e Woods, 2007).

As etapas fundamentais para o processamento de imagens digitais, segundo Gonzalez e Woods (2007) resultam na Figura 2. 1.

Figura 2. 1 – Método para o processamento de imagens digitais.

Fonte: Adaptado de Gonzalez e Woods (2007).

Para o estudo das técnicas de análise e processamento de imagens, Gonzalez e Woods (2007), as classificam em três áreas: processamento de baixo nível, processamento de nível intermediário e processamento de alto nível. Embora essas subdivisões não possuam fronteiras definidas, elas fornecem um esquema para a categorização dos vários processos que são componentes dos métodos de análise digital de imagens. A Figura 2. 2 mostra a divisão adotada pelos autores. As linhas tracejadas sobrepostas indicam a inexistência de limites definidos entre os processos. Por exemplo, a limiarização pode ser vista como um melhoramento (pré-processamento) ou como uma ferramenta de segmentação, dependendo da aplicação.

Figura 2. 2 – Elementos de Análises de imagens

Fonte: Adaptado de Gonzalez e Woods (2007).

Processamentos de baixo nível envolvem operações primitivas tais como redução de ruído, realce de contrastes e nitidez. Um processo de baixo nível é caracterizado pelo fato de que as entradas e as saídas são imagens.

O nível médio de processamento de imagens envolve tarefas como segmentação, descrição desses objetos para reduzi-los a uma forma adequada para o processamento computacional e classificação (reconhecimento) de objetos individuais. Um processo de nível médio é caracterizado pelo fato das entradas serem imagens, mas seus resultados são atributos extraídos das imagens, por exemplo, bordas, contornos, e a identidade dos objetos individuais.

Finalmente, o processamento de alto nível tem como característica de entrada, geralmente, um conjunto pequeno de dados e envolvem os processos de reconhecimento e interpretação. Nesta etapa estão envolvidos métodos de reconhecimento por decisão teórica, métodos estruturais para o reconhecimento e métodos para a interpretação de imagens.

Pode-se observar que a maioria das técnicas utilizadas pelo processamento de baixo nível e de nível intermediário são baseadas em teorias bem definidas e formuladas. O mesmo não ocorre na área de reconhecimento e especialmente na área de interpretação, pois o conhecimento e compreensão dos princípios

fundamentais são muito mais especulativos do que quantitativos. Essa falta relativa de entendimento resulta, em último caso, em uma formulação de restrições e idealizações com a intenção de reduzir a complexidade da tarefa a um nível fácil de compreensão. Assim, o produto gerado é um sistema com capacidades operacionais especializadas.

Após as etapas de aquisição e pré-processamento das imagens, uma das etapas mais importantes na detecção de veículos é a segmentação da imagem para separar os veículos do fundo. Existem várias abordagens que tratam do assunto, com diferentes graus de eficácia. O método de segmentação, para ser considerado eficaz, precisa separar com precisão veículos do fundo da imagem, ser rápido o suficiente para operar em tempo real, ser insensível às condições meteorológicas e de luminosidade, e exigir uma quantidade mínima de inicialização, ou seja, atualização de fundo (Gupte *et. al.*, 2002; Toyama *et. al.*, 1999).

2.3.1 OPERAÇÕES LÓGICAS COM IMAGENS

Após uma imagem ter sido adquirida e digitalizada, a mesma pode ser representada como uma matriz de inteiros e, portanto ser manipulada numericamente utilizando operações lógicas ou aritméticas, sendo que para imagens binárias, ou seja, com valores de pixel de 0 e 1, são aplicadas operações lógicas. Estas operações podem ser efetuadas *pixel a pixel* ou orientadas à vizinhança. No primeiro caso, elas podem ser descritas pela seguinte notação:

$$X \text{ opn } Y = Z \quad (1)$$

Onde X e Y são imagens (matrizes) ou escalares, Z é obrigatoriamente uma matriz e opn é um operador aritmético (+, -, x e /) ou lógico (AND, OR, XOR).

Considerando duas imagens X e Y de igual tamanho, estas imagens podem ser processadas *pixel a pixel* utilizando um operador aritmético ou lógico, produzindo uma terceira imagem Z , cujos pixels correspondem ao resultado de $X \text{ opn } Y$ para cada elemento de X e Y .

Todas as operações lógicas (ou booleanas) conhecidas podem ser aplicadas entre imagens, com qualquer número de nível de cinza, mas são mais bem compreendidas quando vistas em imagens binárias.

As operações lógicas constituem a base de uma série de ferramentas para processamento de imagens, dentre as quais está a linguagem do MATLAB®.

Onde:

AND – E, representada pelo operador, $\wedge$.

OR – OU, representada pelo operador, $\vee$.

XOR - OU Exclusivo, representada pelo operador, $\oplus$.

NOT – Negação

SUB - Subtração de imagens

AND, OR e XOR, são operações binárias, enquanto NOT é uma operação unitária.

A operação Subtração é particularmente útil para rastrear objetos em movimento em uma sequência de imagens com fundo constante.

As Figuras 2.3 a 2.6 demonstram exemplos de operações lógicas com imagens.

Figura 2. 3 - Exemplo de operação AND entre imagens monocromáticas: (a) X, (b) Y, (c) $X \wedge Y$.

Fonte: Marques Filho e Vieira Neto, 1999.

Figura 2. 4 - Exemplo de operação OR entre imagens monocromáticas: (a) X, (b) Y, (c) $X \vee Y$.

Fonte: Marques Filho e Vieira Neto, 1999.

Figura 2. 5 - Exemplo de operação XOR entre imagens monocromáticas: (a) X, (b) Y, (c) $X \oplus Y$.

Fonte: Marques Filho e Vieira Neto, 1999.

Figura 2. 6 - Exemplo de operação NOT sobre imagem monocromática: (a) X, (b) NOT X.

Fonte: Marques Filho e Vieira Neto, 1999.

Um composto das operações binárias NOT e AND é o filtro de diferença onde, a função $NOT(AND(a, b))$, sendo a e b, valores do *pixel*, compara duas imagens *pixel* a *pixel* e como resposta, tem-se uma imagem com apenas os *pixels* que não se repetem nas duas imagens.

2.3.2 TÉCNICAS PARA SEGMENTAÇÃO BASEADAS EM ANÁLISE DE IMAGENS

Algumas das principais questões para detecção de objetos são apresentadas e discutidas nesse item.

A) LIMIAÇÃO (THRESHOLDING)

Kastrinaki *et. al.* (2003) consideram essa técnica a mais simples para segmentação de imagens. Baseia-se no princípio de que os veículos são objetos compactos com intensidade luminosa diferente do restante da cena. Assim, para pequenas regiões, é possível separar o veículo do fundo utilizando níveis de limiarização (Kastrinaki *et. al.*, 2003). Simplificadamente, a segmentação divide uma imagem de entrada em suas partes constituintes ou objetos. Cada uma dessas partes é uniforme e homogênea no que diz respeito a algumas propriedades da imagem, como por exemplo, cor e textura. Os algoritmos de segmentação geralmente são baseados em duas propriedades básicas: descontinuidade e similaridade. Na primeira categoria, a divisão da imagem é baseada no subconjunto de pontos de um objeto que o separa do restante da imagem. As técnicas de segmentação nesta categoria buscam evidenciar os limites entre os objetos por meio da detecção de pontos isolados e da detecção de linhas e bordas na imagem (Rosin, 1998). Na segunda categoria, a segmentação é baseada nas técnicas de limiarização, crescimento por regiões, união e divisão de regiões (Gonzalez e Woods, 2007). Na limiarização é analisada a similaridade dos níveis de cinza da imagem extraíndo os objetos de interesse por meio da seleção de um valor de limiar T que separa os agrupamentos de níveis de cinza.

Uma das dificuldades da limiarização de uma imagem é a melhor determinação do valor de limiar, ou seja, o ponto de separação dos *pixels*. A escolha deste limiar pode assumir um critério subjetivo de um operador humano, com a inspeção visual do histograma ou tentativa e erro. Contudo, já foram propostos diversos métodos

para que esta escolha se dê automaticamente baseados em critérios entrópicos ou estatísticos da imagem [(Beghdadi *et. al.*, 1995³; Pun, 1980⁴) *apud* Lopes, 2003; Otsu, 1979]. Os resultados destes métodos levam a uma boa qualidade de segmentação, tendo como critério de qualidade uma opinião subjetiva e que, algumas vezes, é adotada como parâmetro para avaliação do método.

As técnicas de limiarização automática (*autothreshold*) convertem automaticamente uma imagem definida em níveis de cinza em uma imagem definida em preto e branco (binária). O método de Otsu, por exemplo, determina o limiar através da divisão do histograma da imagem de entrada, de tal modo que, a variância para cada um dos grupos de *pixels* é minimizada, ou seja, limiarizada dividindo os *pixels* de uma imagem de n níveis de cinza em duas classes, C_0 e C_f , que representam o objeto e o fundo respectivamente. Assim, $C_0 = \{0, 1, \dots, T\}$ e $C_f = \{T + 1, T + 2 \dots G - 1\}$, sendo T o valor do limiar escolhido e G o número de níveis de intensidade da imagem.

Seja ∂_w^2 a variância dentro da classe, ∂_B^2 a variância entre as classes e ∂_T^2 a variância total. Um limiar ótimo pode ser obtido pela seguinte função:

$$\varphi = \frac{\partial_B^2}{\partial_T^2} \quad (2)$$

Onde φ é o nível de cinza que pode ser utilizado para diferenciar C_0 de C_f , determinado no intervalo $0 \leq \varphi \leq 1$.

Métodos de escolha de limiar autoadaptativos podem ser utilizados para as mudanças de iluminação, mas não podem evitar a falsa detecção de sombras ou erros de detecção das partes do veículo com intensidades similares ao fundo (Park, 2001⁵, *apud* Kastrinaki *et. al.*, 2003). Operadores morfológicos de escala de cinza

³ **BEGHDADI, A.; NÉGRATE, A.L.; DE LESEGNO, P.V.** *Entropic Thresholding Using a Block Source Model*. CVGIP: Graphical Models and Image Processing, (57), p.197-205, 1995.

⁴ **PUN, T.** *A New Method for Gray-Level Picture Thresholding Using the Entropy of the Histogram*. Signal Processing, p. 223-237, 1980.

⁵ **Park, Y.**, *Shape-resolving local thresholding for object detection*, *Pattern Recognition Letters* 22, 883–890, 2001.

tem sido propostos para detecção de objetos e identificação, pois são insensíveis a variação de iluminação (Won *et. al.*, 2001⁶ *apud* Kastinaki *et. al.*, 2003).

B) CARACTERÍSTICAS GEOMÉTRICAS

Esse método considera o conhecimento *a priori* das características dos objetos a serem segmentados, a partir de informações de simetria, cantos e bordas (Bertozzi *et. al.*, 1997; Matthews *et. al.*, 1996; Goerick *et. al.*, 1996). Os resultados para essa técnica são geralmente bons e sua qualidade varia de acordo com a adaptação das características escolhidas para carregar a base de informações com as condições do meio (Kastrinaki *et. al.*, 2003).

Assim, os objetos são descritos por suas características (formas, dimensões, luminosidade), que permitam a identificação em meio ao seu ambiente (Klausmann *et. al.*, 1999⁷; Dickinson e Wan, 1989⁸, *apud* Kastrinaki *et. al.*, 2003). Klausmann *et. al.* (1999) empregam uma abordagem de regressão lógica utilizando as próprias características pré-extraídas dos veículos, a fim de detectá-lo em seu plano de fundo. Deve ser criado um modelo específico para cada tipo de veículo, a fim de formar o banco de dados para ser consultado pelo sistema e assim reconhecer o veículo. Isso cria um problema, uma vez que existem muitas formas geométricas para veículos contidos na mesma classe de veículo. Além disso, o modelo pressupõe que haja pouca mudança nas características de intensidade dos veículos. Na prática, no entanto, alterações na iluminação ambiente, sombras, oclusões geram variações nas imagens dos veículos do mesmo tipo. Para superar esses problemas, sistemas como, Dickinson e Wan, (1989⁸); Wan e Dickinson, (1989⁹) *apud* Kastrinaki *et. al.*, (2003); Demuth e Beale, (2006); Nunes e Prado, (2002); Ribeiro e Lima, (1999); Piccioli *et. al.*, (1996) entre outros, empregam redes neurais,

⁶ Won, Y.; Nam, J.; Lee, B. H.; *Image pattern recognition in natural environment using morphological feature extraction*, in: F.J. Ferri (Ed.), pp. 806–815, SSPR&SPR 2000, Springer, Berlin, 2001.

⁷ Klausmann, P., Kroschel, K., Willersinn, D., *Performance prediction of vehicle detection algorithms*, *Pattern Recognition* 32, 2063–2065, 1999.

⁸ Dickinson, K.W. ; Wan, C.L., *Road traffic monitoring using the TRIP II system*, *IEE Second International Conference on Road Traffic Monitoring*, Conference Publication Number 299, 56–60. February , 1989.

⁹ Wan, C.L.; Dickinson, K.W., *Road traffic monitoring using image processing—a survey of systems, techniques and applications*, *IFAC Control Computers, Communications in Transportation*, Paris, France 1989.

onde um banco de dados pré-definido pelo usuário com as imagens de parâmetro serve como comparação e treinamento para a rede neural, e assim, o sistema de reconhecimento é codificado. Além disso, o sistema explora a capacidade para interpolar entre diferentes formas conhecidas (Mantri e Bullock, 1995¹⁰ *apud* Kastrinaki *et. al.*, 2003).

2.3.3 DETECÇÃO DE FUNDO BASEADA EM DETECÇÃO DE MOVIMENTO

Este tipo de técnicas visa responder ao problema: como obter automaticamente a imagem estática na cena, ou seja, como segregar o que é plano de fundo dos objetos em movimento?

Piccardi (2004) faz uma revisão sobre as principais técnicas para a segmentação entre o plano de fundo e o primeiro plano. A seguir, são apresentadas as principais abordagens.

A) DIFERENÇA ENTRE QUADROS DE PLANO DE FUNDO (*BACKGROUND FRAME DIFFERENCING*)

O método de subtração de fundo é um dos mais utilizados para a extração do plano de fundo de cenas em movimento (Wang, *et. al.*, 2004; Cheung e Kamath, 2004, 2005). A técnica consiste em detectar a região de movimento (veículos, pessoas, animais, balanço de árvores e etc.) por meio da diferença de *pixels* entre a imagem de análise e a referência de fundo pré-carregada, ou seja, identifica objetos em movimento a partir da parte de um quadro de vídeo, que difere significativamente de um modelo de plano de fundo. Essa é uma técnica relativamente simples, porém apresenta diversos pequenos erros, que podem comprometer o resultado final do processamento, como a sensibilidade a mudança de iluminação, detecção de objetos de fundo não-estacionários, tais como folhas em movimento, chuva e sombras de objetos em movimento. Assim, o modelo é dependente da atualização da imagem do plano de fundo (Cheung e Kamath, 2004, 2005; Friedman e Russell, 1997).

¹⁰ Mantri, S., Bullock, D., *Analysis of feedforward-backpropagation neural networks used in vehicle detection, Transportation Research Part C 3*, 161–174, 1995.

Os modelos de atualização comumente partem de um modelo estatístico para avaliar cada *pixel*, e então, classificá-lo como parte da imagem de primeiro plano ou de plano de fundo. No trabalho de Cheung e Kamath (2004) verificou-se que boas técnicas de pré e pós-processamento do vídeo podem ser determinantes para melhorar a detecção de objetos em movimento. Por exemplo, pequenos objetos em movimento, como folhas de uma árvore, podem ser removidos por transformação morfológica dos quadros após a sua correta identificação. A velocidade com que ocorrem essas atualizações do plano de fundo é extremamente importante para o resultado. Modelos de adaptação lenta do plano de fundo não apresentam bons resultados para mudanças rápidas e passageiras na cena de fundo (como uma nuvem que passa sobre a cena, fazendo sombra) (Oliver *et. al.*, 2000¹¹; Haritaoglu *et. al.*, 2000¹², *apud* Cheung e Kamath, 2004). Nessa situação, por exemplo, os *pixels* do fundo podem ser classificados incorretamente como *pixels* de primeiro plano. Por outro lado, modelos rápidos de atualização de fundo podem lidar com esse tipo de mudança, mas falhar em baixas taxas de quadros além de serem muito suscetíveis a ruídos (McFarlane e Scho, 1995¹³; Remagnino *et. al.*, 1997¹⁴ *apud* Cheung e Kamath, 2004).

Uma abordagem probabilística para a segmentação é descrito em Friedman e Russell (1997). Os autores usam um algoritmo de maximização da confiança para classificar cada *pixel* como objetos que “se movem em sombra”, ou se fazem parte do plano de fundo. Karmann e Brandt (1990¹⁵) *apud* Gupte *et. al.* (2002), fazem uma abordagem de segmentação usando a subtração de plano de fundo adaptativo. O filtro de Kalman é utilizado para prever a imagem de plano de fundo durante o

¹¹ **OLIVER, N., ROSARIO, B., AND PENTLAND, A.** *A Bayesian computer vision system for modeling human interactions.* IEEE Transactions on Pattern Analysis and Machine Intelligence 22, p. 831-843, 2000.

¹² **HARITAOGU, D. HARWOOD, AND L. DAVIS.** *A real-time surveillance of people and their activities.* IEEE Transactions on Pattern Analysis and Machine Intelligence 22, p. 809-830, 2000.

¹³ **MCFARLANE N. AND SCHO C.** *Segmentation and tracking of piglets in images.* Machine Vision and Applications 8, p. 187-193, 1995.

¹⁴ **REMAGNINO P. ET AL.** *An integrated and pedestrian model-based vision system.* In Proceedings of the Eighth British Machine Vision Conference, p. 380-389, 1997.

¹⁵ **KARMANN, K.P. AND BRANDT, A. VON.** *Moving object recognition using an adaptive background memory.* In Time-Varying Image Processing and Moving Object Recognition, 2 – edited by V. Capellini, 1990.

intervalo até a próxima atualização. O erro entre a previsão e a imagem de fundo real é usado para atualizar as variáveis de estado do filtro de Kalman. Este método tem a vantagem de adaptar-se automaticamente às mudanças das condições de clima e de luminosidade. No entanto, ele precisa ser inicializado e treinado com uma imagem do fundo, sem quaisquer veículos ou outra interferência presentes.

B) DIFERENÇA ENTRE QUADROS (*INTER-FRAMES DIFFERENCING*)

Outra abordagem é a diferença entre quadros, utilizada por Lipton, (1998); Desa e Salih, (2004), entre outros, que consiste na subtração de quadros sucessivos. Este método é insensível às condições de iluminação e tem a vantagem de não exigir inicialização com uma imagem de fundo, porém, o resultado pode gerar muitos ruídos de difícil eliminação.

Diferença entre quadros é uma das formas mais simples de subtração de fundo. O quadro atual é simplesmente subtraído a partir do quadro anterior e, se a diferença nos valores de *pixel* para um dado *pixel* é maior do que um limiar T , o *pixel* é considerado como parte do primeiro plano.

$$|frame_i - frame_{i-1}| \geq T \quad (3)$$

Um problema presente nesse método é que para objetos com valores de intensidade uniformemente distribuídos (tais como a lateral dos veículos), os *pixels* interiores a borda do objeto podem ser interpretados como parte do fundo. Outro problema é que os objetos devem estar continuamente em movimento. Se um objeto permanece por mais de um período de quadro, torna-se parte do fundo.

A principal dificuldade deste é a determinação do valor de limiar. Isso também é um problema para os outros métodos. O limite é normalmente encontrado empiricamente, o que pode ser complicado, ou pode ser encontrado por meio da análise de histogramas ou ainda com o auxílio das ferramentas de visualização do MATLAB® (Piccardi, 2004).

No trabalho de Gupte *et. al.* (2002) os autores utilizam um método de subtração autoadaptativo de plano de fundo para a segmentação, que é similar em princípio ao método descrito por Friedman e Russell (1997). No entanto, o método utilizado pode

ser considerado mais simples e mais robusto para a atualização do plano de fundo. O método extrai automaticamente o plano de fundo de uma sequência de vídeo e assim, não necessita de uma inicialização manual. Kim *et. al.* (2001) utilizam a técnica de diferença entre quadros de plano de fundo para complementar o método de diferença entre quadros, melhorando assim os resultados da imagem do plano de fundo.

C) ESTIMATIVA DE PLANO DE FUNDO (*BACKGROUND ESTIMATION*)

A segmentação, rastreamento e análise de objetos de primeiro plano de interesse (*foreground*) são tarefas fundamentais dentro da área de processamento de imagens e detecção de objetos. Muitas abordagens (Friedman e Russell, 1997; Kim *et. al.*, 2001; Gupte *et. al.*, 2002; Wang, *et. al.*, 2004; Cheung e Kamath, 2004, 2005) para a detecção e rastreamento de objetos são baseadas em subtração de plano de fundo, onde cada quadro é comparado com um modelo de plano de fundo, sendo este, atualizado para cada novo quadro de entrada. No entanto, a maioria deles presume que a sequência de imagens de treinamento utilizada para obter o modelo do fundo estará livre de objetos em primeiro plano. Esse pressuposto não se aplica a ambientes não controlados, tais como ruas, rodovias, terminais de transporte e qualquer ambiente ao ar livre, onde uma mudança de iluminação pode tornar o modelo de plano de fundo existente ineficaz (Reddy *et. al.*, 2009). Assim, é necessária a utilização de uma sequência de imagens para estimar o plano de fundo da cena, onde partes desse plano estão ocluídas.

O método de Estimativa de Plano de Fundo nada mais é que, a filtragem mediana, dos planos de fundo, ou seja, tal como com a diferença entre quadros, o fundo é subtraído do quadro atual e limiarizado para determinar os *pixels* de primeiro plano.

Em 1995, McFarlane e Schofield, realizaram uma pesquisa financiada pelo poder público, para o controle dos leitões em grandes fazendas comerciais e introduziram em uma aplicação prática o conceito de Estimativa de Plano de Fundo, com movimentos lentos. Desde então, pode-se perceber uma ampla aplicação na literatura deste método para a segregação de cenários de fundo.

O método funciona da seguinte maneira: se um *pixel* no quadro atual tem um valor maior do que o *pixel* de fundo correspondente, o *pixel* de fundo é incrementado por 1 (um). Da mesma forma, se o elemento de imagem atual é menor do que o *pixel* de fundo, o fundo é diminuído de 1 (um). Deste modo, o fundo eventualmente converge para uma estimativa onde a metade dos *pixels* de entrada são maiores do que o plano de fundo, e a outra metade, tem valores menores do que os do plano de fundo.

Um problema desse método é que se a taxa de atualização do plano de fundo for alta (30 quadros/s), a possibilidade da ocorrência de “sombras” nos objetos detectados como primeiro plano é maior. Esse problema é minimizado na taxa usual para atualização em problemas reais, 15 quadros/s (Piccardi, 2004).

D) MISTURAS GAUSSIANAS

Entre os métodos de alta complexidade a filtragem Kalman e as mistura Gaussianas (*Gaussian mixture models* – GMMs), tem ampla abordagem na literatura. Ambos tem suas vantagens, porém, o principal diferencial entre as misturas Gaussianas e a filtragem de Kalman é que a última técnica não tem resultados satisfatórios para grande parte das aplicações, no que diz respeito à eliminação das sombras dos objetos detectados, tais como no método da Estimativa de Plano de Fundo (Piccardi, 2004).

No GMM, o plano de fundo não é um quadro de valores e sim, um modelo paramétrico. Cada localização de pixel é representada por um número (ou mistura) de Gaussian, conforme equação.

$$F_{(i_t=\mu)} = \sum_{i=1}^k \omega_{i,t} \cdot \eta_{(\mu,0)} \quad (4)$$

Onde:

K - número de distribuições gaussianas

T - soma de limiar de pesos para teste de fundo

ω - peso inicial.

η - Número de observações da ocorrência da mistura

O μ é a média de cada função gaussiana ou componente. Pode ser considerado como um palpite do valor de pixel no próximo quadro, assumindo que os pixels são em sua maioria componentes do plano de fundo. No GMM, cada pixel é classificado baseado na distribuição de Gaussianas que o representa mais eficazmente como fundo. Além disso, ele realiza aproximações quadro a quadro para atualizar o modelo de plano de fundo. Assim, uma distribuição de Gaussianas que represente fundo terá grande peso e baixa variância, ou seja, ocorra frequentemente e varie pouco no tempo.

Piccardi (2004) conclui que é impossível fazer uma comparação entre as técnicas sem uma referência significativa, pois cada uma das abordagens apresenta vantagens e desvantagens em variados aspectos. Assim, alguns métodos parecem ser melhores princípios, como por exemplo, Misturas Gaussianas, fornecendo um bom grau de acurácia para grande parte de suas aplicações, porém, se comparados sob o ponto de vista de velocidade de processamento e requisitos de memória, métodos simples como média padrão, diferença entre quadros e Estimativa de Plano de Fundo, pode fornecer uma precisão aceitável em aplicações específicas.

2.3.4 RECONHECIMENTO E INTERPRETAÇÃO DE IMAGENS

O reconhecimento de imagens é o processo de atribuir rótulos a um objeto, baseado na informação fornecida previamente pelo seu descritor. A interpretação envolve a atribuição de um significado ao conjunto de objetos reconhecidos (Gonzalez e Woods, 2007). O reconhecimento de padrões em uma imagem depende muito da qualidade da própria imagem, portanto, para que não ocorram erros no reconhecimento, é necessário que a imagem passe por um tratamento anterior para eliminar possíveis interferências que poderão causar erros no resultado final (Gonzalez e Woods, 2007).

A área de reconhecimento de padrões está presente nas pesquisas há vários anos, antes mesmo do desenvolvimento computacional, e antes da década de 1960, era principalmente constituída da pesquisa teórica na área de estatística. No entanto, com o desenvolvimento da computação houve um aumento da procura por aplicações práticas de reconhecimento padrão. Dessa forma, a maioria dos sistemas

inteligentes construídos para tomada de decisões utiliza-se dessas técnicas (Theodoridis e Koutroumbas, 2006).

A análise de imagens está dividida nos processos de detecção, identificação e de reconhecimento de padrões relevantes, tendo como objetivo principal o aumento do desempenho e eficiência de uma tarefa que tem como etapa a análise de imagens. Assim, um dos principais objetivos dos métodos de análise automática de imagens é dotar um computador com a capacidade similar dos seres humanos de reconhecer e compreender uma imagem a partir de um padrão visto anteriormente (Gonzalez e Woods, 2007).

Um padrão é uma descrição quantitativa de um objeto ou algum outro ponto de interesse e é formado por um ou mais descritores, ou características. Ao ato de gerar descritores dá-se o nome de extração de características. Os padrões utilizados para descrever um objeto podem ser quantitativos (comprimento, área, textura etc.), ou qualitativos (Gonzalez e Woods, 2007). A área de reconhecimento de padrões trata da classificação de elementos extraídos e da sua análise computadorizada com a menor interferência humana tornando possível a detecção de falhas.

Não existe uma técnica que seja de consenso a resolver todos os problemas, principalmente porque essa técnica deve estar associada ao reconhecimento e à análise das informações (Young, 2004).

2.4 DETECÇÃO E IDENTIFICAÇÃO DE VEÍCULOS

São diversas as técnicas para identificar veículos em meio ao tráfego, e podem ser classificadas como invasivas e não invasivas, ou seja, quanto à interferência na via ou local que será monitorado. As técnicas mais utilizadas no Brasil serão brevemente descritas neste item.

Laço de indução ou loop indutivo é um método invasivo de detecção, pois os sensores são instalados abaixo do pavimento. Seu funcionamento é como o de um sensor de massa metálica, ou seja, é capaz de detectar o início e o final da passagem de um veículo utilizando como premissa a massa metálica que interfere no campo eletromagnético. A contagem dos eixos de um veículo pode ser efetuada

por meio da contagem de passagem sobre o laço e a operação conjunta de dois laços permite medir a velocidade dos veículos.

Figura 2. 7 – Instalação do laço de indução

Fonte: Catálogo Comercial

A identificação por rádio frequência - RFID é um método de identificação automático, baseado no armazenamento e posterior recuperação de informação através da utilização de dispositivos conhecidos por RFID *tags* ou *transponders*. Trata-se de micro-dispositivos que se comunicam por meio de ondas de rádio, trocando qualquer tipo de informação, como identidade, temperatura e localização. Similar ao código de barras, que consiste em um leitor óptico para os códigos impressos nos itens, o RFID utiliza um leitor de rádio frequência dos *tags*, que são colocados nos itens a serem controlados. Os dados dentro do *tag* identificam os veículos e tem como outras aplicações a identificação de um item em uma linha de fabricação, por exemplo, ou mercadorias em trânsito, localização, identificação de animais ou indivíduos.

No Brasil, o uso mais popular do RFID é o sistema de cobrança automática de pedágios do tipo “Sem Parar”. Instalado no vidro dianteiro do carro, o chip é alimentado pela bateria autônoma, ativado pelo campo magnético. Seus sinais são captados por leitores que registram data, horário, praça e o custo do pedágio, liberando a cancela para a passagem do veículo. Depois, a cobrança da tarifa é enviada para a casa do cliente a cada mês, junto com a fatura do cartão de crédito.

Outro método de detecção de veículos são os sensores indutivos, ou sensores de proximidade. Os sensores indutivos geram um campo eletromagnético sobre a área inspecionada. Assim, quando um objeto metálico passa na região monitorada, parte do campo é absorvido, provocando um sinal de saída do sensor. As principais vantagens na utilização desses sensores são: a grande durabilidade, uma vez que não há contato físico com o objeto detectado, é insensível a trepidações e as condições climáticas. Como desvantagem, não permite a classificação de veículos.

Figura 2. 8 – Sensores indutivos

Fonte: Catálogo Comercial

Trabalhos relacionados à retirada de características veiculares utilizando as técnicas acima são descritos por. Sun Gu Sun *et. al.* (2003), que utilizaram raios infravermelhos para retirar parâmetros de identificação de veículos. Similarmente, Urazghildiiev *et. al.* (2007), mostram que a altura de um veículo pode ser obtida através da análise de dados provenientes de um sensor de radar.

Técnicas não invasivas, como a detecção de veículos utilizando a análises de imagens estão amplamente descritas na literatura. Conforme visto nos itens anteriores propõe diversas técnicas para a detecção de veículos em meio a cenas reais.

Em uma fase inicial os pesquisadores abordaram a análise de imagens estáticas obtidas por câmeras fotográficas. Essas técnicas apresentam como maior dificuldade a possível confusão entre os veículos e outros objetos presentes na imagem (Crisman e Thorpe, 1988; Kuehnle, 1991; Handmann *et. al.*, 2000).

Nos últimos anos, várias pesquisas foram desenvolvidas com o objetivo de identificar veículos em meio à análise de imagens capturadas em vídeos, utilizando ou não o movimento como premissa de detecção (Cunha, 2011¹⁶; Silva e Gonzaga, 2007; Toulminet *et. al.*, 2002; Gupte *et. al.*, 2002; Cucchiara e Piccardi, 1999; Giachetti, Campani e Torre, 1998; Sullivan *et. al.*, 1997; Franke e Kutzbach, 1996). Essas se estendem desde a utilização em sistemas de vigilância de vias e cruzamentos, até a aplicação em veículos independentes.

No trabalho de Sun *et. al.* (2006), os autores realizam uma revisão sobre a evolução dos sistemas de detecção de veículos, que abrange os principais métodos presentes na literatura. Os autores argumentam ainda que os principais motivos para a continuidade da pesquisa em detecção automática de veículos são: o prejuízo financeiro e o alto número de vítimas nos acidentes com veículos, o avanço de tecnologias e pesquisas na área de visão computacional nas últimas décadas e o avanço da tecnologia de processamento computacional, acessível até mesmo em computadores pessoais.

2.4.1 SISTEMAS DE DETECÇÃO E IDENTIFICAÇÃO DE VEÍCULOS BASEADOS EM PROCESSAMENTO DE IMAGEM

Os sistemas de processamento de imagem desenvolvem-se como uma importante ferramenta para detecção automática de veículos, como apresentado em trabalhos de (Jazayeri *et. al.*, 2009; Buch *et. al.*, 2008; Morris e Trivedi, 2006; Mohottala *et. al.*, 2003; Gupte e Papanikolopoulos, 2000; Zhang e Siyal, 2000). Esta tecnologia já viabiliza sistemas de reconhecimento de padrões faciais, padrões de assinaturas entre outras aplicações. No monitoramento de tráfego, são aplicados os sistemas de reconhecimento de caracteres para o reconhecimento automático de placas veiculares. Esses sistemas para identificação de veículos propõem distinguir as categorias dos veículos pelos padrões coletados por câmeras de vídeo e/ou fotografia digitais, efetuando análises das características dessas imagens coletadas.

¹⁶ CUNHA, A. L. B.N. Publicação eletrônica [mensagem pessoal]. Mensagem pessoal recebida por <pam.tancredi@usp.br>, em 01/11/2011. Tese de doutorado em desenvolvimento no Departamento de Engenharia de Transportes da Escola de Engenharia de São Carlos – EESC/USP.

Jazayeri *et. al.* (2009) propõem uma análise em tempo real de vídeos capturados a partir de uma câmera instalada em um veículo em movimento que segue no tráfego, e tem como objetivo obter parâmetros de segurança, direção e trajetória (rastreamento) dos veículos seguidos. Os autores analisaram os vídeos em formato AVI (*Audio video interleave*) para os períodos noturnos e diurnos. O método foi implementado na linguagem C++ utilizando a biblioteca do OpenCV (*Open Source Computer Vision Library*), e os cenários de “vizinhança” foram calibrados em MATLAB[®]. Na Figura 2. 9 é apresentado um perfil típico analisado.

Figura 2. 9 – Quadro típico utilizado – resolução de 640 x 480 *pixels*

Fonte: Jazayeri *et. al.* (2009).

Para a análise e processamento dos vídeos, os autores utilizaram técnicas de probabilidade baseadas em HMM – *Hidden Markov Model*¹⁷. Os principais parâmetros a serem calibrados foram as distâncias máximas em relação ao veículo filmado. O resultado do trabalho foi a detecção e segmentação dos veículos, possibilitando a contagem automática de fluxo, conforme pode ser visto na Figura 2. 10.

¹⁷ Métodos baseados em *Hidden Markov Models* (HMMs), ou Modelos Ocultos de Markov, são métodos estocásticos, que consistem na geração de modelos de classes acústicas, que fazem uso de informações temporais, o que pode ser bastante útil para algumas aplicações, como é o caso dos sistemas de reconhecimento dependentes de atualizações.

Figura 2. 10 – Detecção e rastreamento de veículos – resultados.

Fonte: Jazayeri *et. al.*, 2009.

As principais dificuldades para o desenvolvimento do método foram variabilidades do ambiente, como: cores, formas e tipos de veículos, a detecção do veículo realizada em movimento, as variações do ambiente e iluminação; e características do processamento dos dados para identificação das imagens em contrapartida do balanço da câmera/veículo, ruídos, e qualquer característica que possa intervir na análise (Jazayeri *et al.*, 2009).

No trabalho de Buch *et. al.* (2008), os autores apresentam um sistema de detecção e classificação de veículos em meio ao cenário de tráfego urbano, com o intuito de auxiliar na vigilância de locais públicos sujeitos a tráfego intenso. A detecção é realizada pelo formato geométrico do veículo e a classificação é obtida por meio de imagens que captam medidas de comprimento e largura dos veículos, que são comparadas com modelos pré-carregados (Figura 2. 11). O modelo utiliza a Estimativa de Plano de Fundo para detectar movimento nas cenas e misturas gaussianas para gerar as máscaras do plano inicial de aprendizagem.

Figura 2. 11 – Modelos utilizados para a classificação

Fonte: Buch *et. al.*, 2008.

O trabalho classifica os veículos em quatro categorias: ônibus/caminhão, van, carros/táxi e motocicletas/bicicletas. O índice de acerto para a classificação foi de cerca de 90% sob as variadas condições de ambiente. As figuras abaixo demonstram algumas análises do trabalho.

Figura 2. 12 – Modelo de imagem utilizada no trabalho e a detecção de veículos

Fonte: Buch *et. al.* (2008).

Figura 2. 13 – Exemplo de classificação correta

Fonte: Buch *et. al.* (2008).

Figura 2. 14 – Exemplo de classificação equivocada – o sistema identifica pedestres como veículos

Fonte: Buch *et. al.* (2008).

Morris e Trivedi (2006) classificam os veículos em três classes: sedan, caminhões e caminhonetes+SUV+van (Figura 2.15). O sistema foi desenvolvido para tráfego em rodovias por meio de comparações de informações pré-carregadas e armazenadas em um banco de dados que caracterizavam os tipos de veículos, em função das medidas do comprimento e altura, com os valores obtidos nas imagens. O desempenho apresentado é de aproximadamente 65%. As Figuras 2.16 e 2.17 demonstram algumas análises do trabalho.

Figura 2.15 – Modelo de classificação proposto

Fonte: Morris e Trivedi (2006).

Figura 2. 16 – Exemplo de classificação correta

Fonte: Morris e Trivedi (2006).

Figura 2. 17 – Exemplo de classificação equivocada – o sistema identifica três veículos como um único

Fonte: Morris e Trivedi (2006).

Similarmente, os mesmos autores (Morris e Trivedi⁽²⁾, 2006) apresentam um estudo para a abordagem de classificação para a visão lateral de rodovias. Um único modelo Gaussiano de fundo é utilizado para a segmentação. Para cada modelo de primeiro plano segmentado, um vetor característico com base nesses modelos é calculado. A avaliação sobre uma sequência de 24 horas de vídeos, mostra o aproveitamento de cerca de 75% para a classificação. A precisão pode ser

aumentada através da combinação rastreamento e classificação. Assim, o esse trabalho foca na segmentação auto-adaptativa, ao contrário do trabalho anterior dos mesmos autores, que tinha como foco a classificação dos veículos, assumindo conhecida a segmentação dos dados de entrada de vídeo. A comparação entre as características baseadas em imagem e em medidas da imagem, como entre características de *pixels* e tamanho de regiões, é abordada no trabalho, de modo que, as características do objeto são classificadas de acordo com um “vizinho” mais próximo. O Filtro de Kalman é aplicado para rastrear a região do primeiro plano das imagens.

Mohottala *et. al.* (2003) propõem a classificação de quatro categorias de veículos, por meio da análise de imagens capturadas por uma câmera fixada acima do fluxo dos veículos, conforme Figura 2. 18. Modelos das imagens coletadas são mostrados na Figura 2. 19.

Figura 2. 18 - Fixação da câmera para coleta da perspectiva superior do veículo
Fonte: Mohottala *et. al.* (2003).

Figura 2. 19 - Modelo de imagem coletada para análise

Fonte: Mohottala *et. al.* (2003).

O trabalho consiste de um algoritmo de reconhecimento das imagens, baseado no método *eigen-window*, onde as imagens obtidas são comparadas com imagens de um banco de dados gerado para o treinamento do algoritmo. O método *eigen-window* propõe a divisão da imagem em pequenas janelas para classificação dos veículos, selecionando as que efetivamente são relevantes ao estudo (Ohba e Ikeuchi, 1997). Os mesmos autores afirmam que o critério para escolha de janelas consiste em medir os “cantos” com características parecidas facilmente identificadas (detecção), a eliminação de janelas similares para delimitar o objeto (singularidade) e a definição do quão estável são esses “cantos”, através da análise do alcance da rotação dos ângulos entre os eixos que os formam (confiabilidade). Ohba e Ikeuchi (1997), afirmam ainda que o número de janelas normalmente encontradas pode ser aumentado, ocupando uma quantidade excessiva de espaço, e devido à similaridade entre estas, com tendência ao erro do método de identificação.

O sistema de Mohottala *et. al.* (2003), combinado com inteligência artificial, é capaz de “aprender” com experiências passadas. Assim, a imagem obtida é classificada, depois de analisadas as características dos padrões e de matrizes geradas de acordo com o padrão de maior similaridade.

Uma dificuldade dessa forma de filmagem esta na superfície visível em função da altura do pórtico limitando, em alguns casos, a identificação de veículos mais longos.

Gupte *et. al.* (2002) propuseram a detecção e classificação de veículos baseado em análise de regiões. O trabalho desenvolve um método de estimativa rápida de plano de fundo, utilizando em conjunto as técnicas de diferença entre quadros de plano de fundo e plano de fundo autoadaptativo, onde a todo novo quadro os valores dos

pixels são testados se pertencem ou não ao plano de fundo, além de promoverem a atualização constante do valor de limiar T , a cada novo plano de fundo. O classificador proposto faz a distinção entre caminhões e não caminhões, baseados em características de dimensões. Técnicas de calibração de câmera são necessárias para a retirada dessas características. Como resultado, em uma sequência de 20 minutos, 70% dos veículos foram corretamente classificados. A sequência da Figura 2.20 mostra o resultado da técnica adaptativa para a segmentação do plano de fundo.

(a) Initial background provided to the algorithm.

(b) Image of the scene after camera's orientation has changed.

(c) Current background after 4 seconds

(d) Current background after 6 seconds.

Figura 2. 20 – Evolução da adaptação do plano de fundo.

Fonte: Gupte *et. al.* (2002).

Trabalhos como o de Bertozzi *et. al.* (1997); Matthews *et. al.* (1996) e Goerick *et. al.*, (1996) exploram as características geométricas dos veículos como forma de detectá-los em meio a cenas de tráfego em vias. Bertozzi *et. al.*, 1997, propuseram um método baseado na identificação dos vértices do veículo para localizá-lo nas

imagens. Matthews *et. al.* (1996) exploraram em seu trabalho um sistema para detectar bordas verticais, e assim localizar a lateral esquerda e direita do veículo. O método consiste em registrar o perfil vertical da extremidade da imagem, aplicando logo em seguida um filtro suavizador. Goerick *et. al.* (1996) utilizam um método para retirar as informações de bordas que consiste em sequências de um código binário representando a direção das variações de níveis de cinza vizinhos a cada *pixel* da imagem.

Os trabalhos de Mazaheri e Mozaffari (2011); Reddy *et. al.* (2009); Harasse *et. al.* (2006); tratam da aplicação de técnicas de estimativa de fundo, para a resolução de problemas relacionados a fundos complexos, aplicados principalmente em sistemas de vigilância de ambientes e detecção de objetos. As técnicas de Estimativa de Plano de Fundo mostram ótimos resultados se comparados com outras técnicas, pois sua característica dinâmica permite a atualização do plano de fundo nas cenas a cada mudança de iluminação, cor, novo objeto estático na cena, e assim sucessivamente.

Mazaheri e Mozaffari (2011) trazem em seu trabalho a programação de um algoritmo para detecção e identificação de veículos em meio a cenas urbanas de tráfego, baseado em duas fases de detecção dos veículos: a primeira utiliza transformada de Hough¹⁸ para detectar as bordas do veículo e a segunda fase utiliza detecção de Kernel, que é uma técnica de análise de agrupamento (*Cluster*), que avalia a significância dos resultados da primeira fase de detecção e os agrupa para a fase de reconhecimento. Os objetos detectados são subtraídos do plano de fundo que foi previamente estimado e guardado como em um “livro de códigos”, que posteriormente são comparados com os resultados da fase de detecção dos objetos em movimento para garantir que esses não fazem parte do plano de fundo. Na etapa de reconhecimento e classificação dos veículos, os autores utilizam a técnica de

¹⁸ A transformada de Hough é uma técnica de extração de características utilizada na análise de imagens, visão computacional e processamento digital de imagens. O objetivo da técnica é o agrupamento de pontos num conjunto de imagens que tendem a ter características em comum necessárias para a classificação dentro de um grupo. Por exemplo, ao encontrar formas simples, tais como retas, círculos ou elipses, utilizando um detector de bordas, devido a imperfeições nos dados de imagem ou no próprio detector, podem surgir pontos faltosos ou desvios entre a linha/círculo/elipse ideal e os pontos obtidos. Shapiro e Stockman. Computer Vision, Prentice-Hall, 2001.

redes neurais. Os resultados finais para a segregação dos veículos geraram muitos erros provocados pela sobreposição dos objetos. Na etapa de reconhecimento, esse erro pode ser corrigido, pois algumas características como distância entre os faróis, altura e comprimento máximo dos veículos, já estavam previamente armazenados na rede neural. A Figura 2.21 mostra o resultados da primeira fase do trabalho, a segregação do plano de fundo e o agrupamento para a avaliação dos objetos identificados como elementos do primeiro plano.

Figura 2. 21 – a) quadro de entrada; b) plano de fundo estimado; c) resultado da diferença entre a e b; d) agrupamento pelas características do “vizinho mais próximo”.

Fonte: Mazaheri e Mozaffari (2011).

Reddy et. al. (2009), propõem um método de Estimativa de Plano de Fundo, que analisa sequências de imagens, bloco por bloco. Os autores definem bloco como uma região da imagem que tem $N \times M$ *pixels* de tamanho. Cada bloco é definido automaticamente por conceitos de vizinhanças de *pixels*, onde o agrupamento ocorre de acordo com suas características em comum. Como diferencial, o método proposto por Reddy et. al. (2009) trata da imagem do plano de fundo iterativamente, ou seja, ao contrário das técnicas de diferença entre quadros e diferença entre

quadros de plano de fundo, em que todos os quadros são armazenados na memória até formar o quadro final de plano de fundo. O método processa os quadros sequencialmente. e a cada novo processamento da cena, o bloco candidato a ser elemento do plano de fundo é substituído pelo bloco da iteração anterior. Os algoritmos foram programados em linguagem MATLAB®. A Figura 2.22 demonstra o resultado da estimativa do plano de fundo após o término das iterações.

Figura 2. 22 - (i) Quadro de entrada; (ii) Plano de fundo inicial, obtido com o primeiro valor de limiar T ; (iii) Iteração n.75, (iv) iteração n.120, (v) estimativa do plano de fundo.

Fonte: Reddy et. al. (2009).

Harasse *et. al.* (2006), propuseram um algoritmo para a estimativa do fundo de cenas em ambientes complexos. O método proposto segue as etapas de cálculo do vetor de características para cada *pixel*, a estimativa do momento ideal para a mudança do plano de fundo entre dois quadros e a atribuição de *pixels* observados nessa hipótese, finalizando com a estimativa bem próxima do real para o plano de fundo de cada *pixel*. Os próprios autores definem as principais contribuições desse trabalho, são elas: a introdução de uma estimativa robusta de plano de fundo, um método ideal para o equilíbrio entre o brilho das imagens, corrigindo a variação das

cores, e a rapidez no cálculo dos *pixels* acumulados independente do tamanho dos quadros.

O método foi desenvolvido para a aplicação na contagem de passageiros em veículos de transporte público, onde o objetivo é reconhecer a presença de passageiros parados no ponto de ônibus. A imagem é adquirida por meio de uma câmera fixa instalada no interior do ônibus que tem a visão externa através da porta. Apesar de desenvolvido para um contexto particular de veículos de transporte, o método desenvolvido por Harasse *et. al.* (2006), é apresentado de modo a ser usado em qualquer outro contexto, com restrições semelhantes, por exemplo, quando a câmera passa por vibrações e cenas com plano de fundo complexo. O método agrega o conceito de estimativa de *pixels* de vizinhança, pois, em todas as imagens, o plano de fundo nunca está completamente aparente. Os autores utilizam métodos estatísticos para a formulação do algoritmo e fazem uma comparação do seu método com a utilização de misturas gaussianas, que são amplamente utilizadas para a resolução de problemas referentes a estimativas de dados ocultos. A Figura 2. 23 mostra o resultado dessa comparação, que resultou em pequeno número de falsos componentes de primeiro plano quando comparado com o modelo que utiliza misturas gaussianas.

Figura 2. 23 – Comparação entre (b) Método de Harasse *et. al.*, 2006 e (c) Misturas Gaussianas.

Fonte: Harasse *et. al.*, 2006

Da revisão bibliográfica sobre detecção e identificação de veículos foi possível perceber que, embora algumas técnicas baseadas em probabilidade e estatística, com complexas implementações, para a segmentação entre o plano de fundo da filmagem e o primeiro plano, tais como apresentadas em Jazayeri *et. al.* (2009), Buch *et. al.* (2008), entre outros, possam produzir um desempenho superior,

técnicas simples, como a estimativa da média no tempo dos quadros de fundo podem produzir bons resultados com baixa complexidade computacional, como nos trabalhos de Mazaheri e Mozaffari (2011); Reddy *et. al.* (2009); Harasse *et. al.* (2006).

2.5 SOFTWARES E FERRAMENTAS

No mercado existem vários softwares que fazem os trabalhos de aquisição, detecção, processamento e extração de características de imagem. Cada um deles apresenta diferenças essencialmente na linguagem de inserção das funções. A linguagem de programação mais utilizada em trabalhos de pesquisadores do meio da computação são as linguagens C e C++, por meio da biblioteca *Open CV*, que fornece uma infinidade de possibilidades na programação e combinação de algoritmos. Entretanto, o MATLAB[®] (*Matrix Laboratory*), tem em uma linguagem simplificada para a modelagem de operações com imagens e possibilita a implementação de operações simples, mas também permite a inserção de funções e algoritmos. Outros programas semelhantes em linguagem Fortran, Basic ou C exigem um maior conhecimento em programação de sistemas (Sun *et. al.*, 2006; Gonzalez, *et. al.*, 2009). Na apresentação das referências bibliográficas de publicações em importantes jornais e revistas na área de computação e processamento de sinais, a linguagem C é a mais empregada, pela maior capacidade de processamento, ficando a cargo do MATLAB[®] as operações lógicas entre imagens para a calibração dos cenários de vizinhança, como em Jazayeri (2009).

O MATLAB[®] é uma linguagem de programação, bastante utilizada em ambiente de pesquisa técnica na engenharia, dado que é apropriada ao desenvolvimento de aplicativos de natureza técnica, possuindo várias coleções de bibliotecas, as *toolboxes* – caixas de ferramentas, para as mais diversas áreas, como equações diferenciais, estatística, processamento de sinais, e especialmente processamento de imagens. Possui um ambiente genérico de computação científica, apresentando uma linguagem matricial simples, compacta e próxima da matemática (Math Woks, 2011).

Além disso, a ferramenta SIMULINK® - Simulação e projetos baseados em modelos - permite a criação e edição de algoritmos de forma simplificada. O SIMULINK® é um ambiente de simulação multi-domínios e projetos baseados em modelos para sistemas dinâmicos. Ele fornece um ambiente gráfico interativo e um conjunto personalizável de bibliotecas de blocos que permitem projetar, simular, programar e testar uma variedade de sistemas de variáveis no tempo, incluindo processamento de vídeo e de imagens.

O SIMULINK® suporta sistemas lineares e não lineares modelados em tempo contínuo, discreto ou a combinação dos dois. Em sua modelagem, o SIMULINK® apresenta uma interface simplificada que possibilita a construção de modelos através de diagrama de blocos utilizando as operações “clique e arrastar” do mouse, por meio da sua vasta biblioteca com diversos blocos prontos que podem ser personalizados e permite a criação de seus próprios blocos.

Após definido o sistema, pode-se simular, permitindo assim a exibição de resultados. A alteração de parâmetros é permitida a qualquer momento e a verificação é imediata. Resultados podem ser enviados ao *workspace* do MATLAB®, assim como objetos do *workspace* são inseridos através dos blocos de entrada de dados. Além disso, depois de finalizado o sistema, é possível a geração de código fonte em linguagem C a partir de sua linguagem de programação.

CAPÍTULO 3

MÉTODO

Um sistema para a detecção de veículos rodoviários de carga em meio a vias urbanas deve ser capaz de identificar esses veículos, independente das condições do vídeo e ser insensível às condições ambientes. Assim, o método desenvolvido nesse trabalho, buscou atender a essas condições, identificando e caracterizando os fatores que interferem na qualidade dos resultados.

A seguir apresenta-se o modelo de detecção de objetos em movimento aplicado nesta pesquisa.

3.1 MODELO DE IDENTIFICAÇÃO DE VEÍCULOS UTILIZANDO ESTIMATIVA DE BACKGROUND

O sistema de detecção e segmentação de veículos rodoviários de carga em vias urbanas apresentado neste trabalho foi desenvolvido utilizando o SIMULINK® e segue a estrutura do modelo disponível em MATLAB® - *Tracking Cars Using Background Estimation*, que foi desenvolvido pela *Mathwork* para rotular objetos em movimento em um vídeo de entrada.

O sistema de programação acontece por meio de blocos, que são resultado das funções existentes nas *toolboxes* do MATLAB. Assim, esses blocos, após dispostos em forma de fluxograma de tarefas, podem ser “abertos”, permitindo sua modificação, exclusão ou inserção de novas funções. Os parâmetros podem ser especificados e editados dentro de cada bloco, como mostra a Figura 3. 1.

Figura 3. 1 – Modelo para identificar veículos rodoviários de carga utilizando MATLAB®

Para iniciar o sistema, o usuário deve ter instalado em seu computador o MATLAB® 2010 ou versão anterior. Um duplo *click* no arquivo “*viptraffic_win.mdl*” inicia o sistema. Na primeira janela, o bloco Seleção de Vídeo, um duplo *click* direciona para a pasta onde estão armazenados os arquivos de vídeos, que pode estar no próprio computador ou em um dispositivo de armazenamento móvel.

O arquivo de vídeo deve seguir os padrões apresentados na Tabela 3.1:

Tabela 3. 1 – Padrões do vídeo de entrada

Padrões do Vídeo de Entrada	
Extensão	.avi
Tamanho	160 x 120 <i>pixels</i>
Velocidade de Projeção do Vídeo	15 quadros/s

O padrão 160 x 120 *pixels*, também conhecido como padrão QSIF, é relativamente restritivo em termos qualidade e de uso em tempo real, se comparado com as imagens geradas por câmeras comerciais, as quais já utilizam padrões de alta definição de imagem *high definition* - HD ou *full HD* e 30 quadros por segundo. Esse padrão foi adotado para este trabalho em função das restrições de capacidade de processamento do MATLAB®. Não faz parte do objeto desta pesquisa a proposta ou codificação de um método computacional mais “poderoso”. Possibilidades com o uso de linguagem C ou C++ estão discutidas no item 2.5 da revisão bibliográfica.

Após a seleção do vídeo, o sistema converte automaticamente o vídeo de entrada padrão RGB para escala de cinza, por meio da função *RGB to intensity* do bloco Conversão da Imagem. No terceiro bloco, Estimativa de Plano de Fundo, inicia-se a etapa de obtenção da imagem do plano de fundo, utilizando a técnica de mesmo nome do bloco. O sistema interno do bloco faz com que o vídeo de entrada seja reaproveitado nas funções de seus componentes, *Temporal Median Estimator*, *Temporal Median* e *Motion Based Background Estimator*. Nesses componentes do bloco Estimativa do Plano de Fundo, é processado o fundo padrão, conforme mostra a Figura 3. 2.

Figura 3. 2 – Componente do bloco Estimativa de Plano de Fundo

Cada um dos três componentes do bloco Estimativa de Plano de Fundo é responsável por gerar dois tipos de dados de saída, o fundo estático (*background - BG*) e tudo que não é estático na filmagem, ou seja, a imagem válida (objetos em movimento - *Valid*). Com este procedimento são geradas 6 (seis) saídas, três BG e três *Valid*. Esses dois tipos de dados de saída são as entradas da função *merge* que é responsável por transformar múltiplos sinais de entrada em um único sinal de saída.

O componente *Temporal Median Estimator*, estima a ocorrência dos valores médios dos *pixels* de valor zero em relação aos pixels em movimento, ou seja, dentre os 19.200 *pixels* provenientes da resolução 120x160, o BG é formado pelos *pixels* que mantém suas propriedades por no mínimo dois quadros, que assim, assumirão o valor 0 (zero). Os *pixels* que sofrem modificações, assumem o valor 1 (um), caracterizando os elementos *Valid*, que correspondem aos os objetos que serão utilizados para a identificação dos veículos.

O estudo sobre a importância do posicionamento da câmera, quanto ao ângulo de filmagem e fixação do tripé, de modo que a câmera esteja suficientemente fixa, foi possível com a análise dos resultados da aplicação do segundo componente, *Temporal Median*, que segrega os *pixels* estáticos a cada 15 quadros e demonstra que, quanto mais estática a câmera estiver, melhores serão os resultados da estimativa do plano de fundo da imagem.

O terceiro e último componente, *Motion Based Background Estimator*, cria uma cópia continua dos *pixels* estáticos a partir dos primeiros quadros. O componente segrega os *pixels* estáticos dos *pixels* em movimento. Nesse componente não são necessários configurações de parâmetros, ele opera automaticamente após a entrada e conversão do vídeo para o padrão cinza de cores.

Uma vez estimado o plano de fundo, o sistema o subtrai de cada quadro para produzir imagens do primeiro plano - *foreground*.

Os dados de saída do bloco Estimativa de Plano de Fundo são as entradas para as linhas *Valid* e BG (Ver o detalhe da Figura 3. 3). A entrada *Vídeo in* é o vídeo original em níveis de cinza. Os dados de entrada *Valid* podem habilitar ou não a utilização do bloco Rastreamento do Veículo, ou seja, utilizando os operadores apresentados no item 2.3.1, a equação (7) é testada.

$$valid = NOT(AND(video\ in, BG)) \quad (7)$$

Assim, todos os dados em que a equação (7) é válida serão processados pelas funções do bloco Rastreamento do Veículo (Figura 3. 4) e o que não atende a equação *Valid*, não é processado por este bloco.

Figura 3. 3 – Bloco Rastreamento do Veículo

O quarto bloco, Rastreamento do Veículo, é responsável pela identificação, visualização e contabilização do número de ocorrências do objeto procurado, neste caso, os veículos rodoviários de carga (Figura 3. 4).

Figura 3. 4 – Componente do bloco Rastreamento do Veículo

A primeira operação realizada nos primeiros componentes do bloco da Figura 3.4, é a diferença entre o plano de fundo estimado no bloco Estimativa de Plano de Fundo e o vídeo de entrada, já em escala de cinza. Essa diferença subtrai da imagem os *pixels* e os *pixels* em movimento

O componente *Autothreshold BW* produz a partir da imagem, uma matriz de valores 0 – 1 de cada pixel, ou seja, este componente do bloco Rastreamento do Veículo aplica o método de limiarização de Otsu apresentado no item 2.3.2, onde, o *level* = 1, é o nível preto, o fundo, e *level* = 0, todos os outros objetos, nível branco.

O componente, *Close*, após a limiarização da imagem, opera por meio da função *strel – structuring element*, do MATLAB®, que é responsável por definir a vizinhança de cada *pixel*. A função *strel('square'w)*, foi aplicada em todos os vídeos de entrada, em que “w” é o número de ocorrência de 1 vizinho em forma de um quadrado, ou seja, uma matriz N x N, que é o caso em vídeos obtidos sob uma base estática, um tripé, por exemplo.

No componente *Region Filtering*, a imagem binarizada, ou seja, possui *pixels* com valores 0 e 1, é filtrada de acordo com os parâmetros escolhidos. Assim, o veículo é detectado de acordo com a região de interesse de análise (Figura 3. 4) e os limites definidos de valor de área (mínimo e máximo) selecionados nos parâmetros macros quando os veículos passam por baixo da linha branca. Esse é o sinal para que os que atendem às condições sejam detectados. A região padrão utilizada na detecção é formada pelos elementos que diferem em relação ao plano de fundo médio.

Os veículos são localizados em cada quadro do vídeo utilizando a função análise de *Blob*¹⁹ - BBOX. Esta função gera os valores de áreas pré-definidas, ou seja, cria objetos a partir das vizinhanças definidas pelos objetos formados pelos “*pixels* brancos” da imagem segmentada. Os parâmetros deste bloco são configurados nas janelas, ilustradas nas Figura 3. 5 e 3.6, que contem três abas: *Main*, *Blob Properties* e *Data Types*.

¹⁹ Na área de visão computacional, detecção blob refere-se a módulos visuais que são destinados a detectar pontos e/ou regiões na imagem que diferem em propriedades como brilho ou cor, quando comparados com a envoltória. \http://en.wikipedia.org/wiki/Blob_detection.

Figura 3. 5 – Parâmetros da função de análise de Blob.

Na aba *main* são definidos os parâmetros utilizados para a filtragem dos objetos desejados. A marcação dos itens *Área* e *Bounding Box* define uma saída dupla (*double*). A *Área* é definida por duas variáveis de espaço, uma mínima e outra máxima, para a definição de um objeto. *Bounding box* é responsável por delimitar o retângulo que marca os objetos detectados. A rotulação utilizada é a “8 conectados”, ou seja, juntamente com o 4 conectados, são os elementos de estrutura mais comuns para um plano cartesiano (Young , 2004), e definem o quanto e como os objetos podem encolher ou expandir.

Figura 3. 6 – Parâmetros da função de análise de Blob.

A aba *Blob Properties* define as propriedades dos objetos segmentados. Os objetos “brancos” segmentados correspondem a conjuntos de *pixels* com valor binário “1”. O parâmetro número máximo de *blobs*, foi definido como 100, e corresponde ao número máximo de segmentações por vídeo e depende do tempo de filmagem. Nos vídeos utilizados para o teste dos parâmetros, esse número foi suficiente, porém, se o número de detecções for ultrapassado, uma mensagem de aviso é enviada. A terceira aba *Data Types* não é aplicada a saídas simples ou duplas.

O contador no canto superior esquerdo da janela de resultados conta o número de detecções na região de interesse, a cada quadro (Vide detalhes da programação na Figura 3. 7). Ainda na Figura 3. 7 – Subcomponente do *componente Region Filtering* são determinados quais os parâmetros e funções que serão utilizados na função de detecção dos veículos, os parâmetros de área mínima e máxima. A condicionante de delimitação da região de análise (abaixo da linha branca) também é definida nesse bloco.

Figura 3. 7 – Subcomponente do componente *Region Filtering*

Finalmente, a Figura 3. 8 apresenta a programação para a visualização dos resultados, que são quatro janelas visualizadoras do processamento, contendo respectivamente o vídeo de entrada, o plano de fundo estimado, o processo de limiarização e os objetos identificados.

Figura 3. 8 – Componente do bloco Mostrando os Resultados

Os parâmetros de características dos objetos a serem detectados, necessários para a identificação dos veículos rodoviários de carga em meio à cena, são:

- Área mínima e área máxima dos veículos de interesse em relação à área útil de análise: esse parâmetro deve ser ajustado toda vez que o ângulo de filmagem da cena é trocado, ou toda vez que a distância entre a câmera estática e o ponto de passagem dos veículos é alterado. Ou seja, existem parâmetros a ser ajustados toda vez que o vídeo de entrada sofre uma mudança bruta nas suas características de coleta, ou seja, quando é modificado substancialmente o *setup* de padrão de imagens a serem identificadas pelo programa.
- Altura da linha branca - linha a partir da qual os veículos serão detectados: o parâmetro de altura da linha branca deve ser ajustado, assim como os parâmetros de área, toda vez que houver uma mudança significativa, principalmente na distância da filmagem. Existe a possibilidade de inserir duas linhas brancas estabelecendo um limite superior e inferior da faixa útil de detecção dos objetos. Essas linhas permitem reduzir o problema de falsas detecções em função de variações de pixels no plano de fundo da imagem.

Os valores adotados para cada vídeo de entrada são inseridos por meio da janela *block parameters* (Figura 3. 9), acessada por meio de duplo *click* na função parâmetros, que pode ser visualizada na Figura 3. 1. Nesta pesquisa adotou-se uma única linha delimitando a parte superior da imagem dado que o ângulo de filmagem selecionado é pequeno.

Figura 3. 9 – Janela para inserção dos parâmetros do programa

Assim, o algoritmo do modelo desenvolvido para identificar veículos rodoviários de carga em meio a vídeos de tráfego urbano, pode ser representado pela Figura 3. 10.

Figura 3. 10 – Sequência do trabalho de programação

O método utilizado para a coleta dos filmes de entrada neste trabalho é apresentado na Figura 3. 11.

Figura 3. 11- Croqui do método para a coleta de dados.

Uma câmera fixa foi disposta sobre um tripé, com altura H e distância D em relação à faixa de tráfego, de tal modo que a relação entre as distâncias produza um ângulo vertical de filmagem de aproximadamente 0° . A adoção dessa premissa de aquisição de filmagens facilita na hora da identificação visual, pelo controlador de tráfego, das características do veículo e na definição dos valores de áreas mínimas e máximas dos veículos rodoviários de carga em relação à cena filmada, pois minimiza as distorções nas dimensões obtidas pela relação ângulo de abertura e distância focal da câmera. Porém, se o vídeo for adquirido em outra posição ou outro ângulo de filmagem, isso não inviabiliza o funcionamento do sistema de detecção de veículos, pois se trata de um problema simples de ajuste da geometria para efeito da definição das áreas mínima e máxima a serem detectadas.

Após a aquisição dos vídeos, os dados de entrada são carregados para o processamento computacional no sistema, lembrando que, para os objetivos de desenvolvimento do sistema desta pesquisa, foi adotado o pós-processamento, em função da limitação dos recursos computacionais do MATLAB®.

Em uma aplicação prática do método de auxílio no controle do acesso de veículos de carga a vias urbanas restritas os veículos de carga detectados no vídeo do tráfego, seriam retirados para outro dispositivo de visualização rápida, por exemplo, uma tela da sala de controle, contendo a imagem do veículo de interesse. Logo, o controlador de tráfego, atuando em uma central de controle de tráfego, poderia ocupar sua atenção para a imagem que contém o potencial veículo infrator, e assim, facilitar sua tomada de decisão em interceptar ou não tal veículo para averiguação.

CAPÍTULO 4

APLICAÇÃO DO MÉTODO

4.1 MODELO IMPLEMENTADO

O capítulo trata da descrição de um caso de aplicação, abordando os tópicos da sequência de execução assim como a descrição dos equipamentos utilizados.. A Figura 4. 1 – Diagrama da sequência da aplicação do sistema, apresenta a sequência utilizada.

Figura 4. 1 – Diagrama da sequência da aplicação do sistema

4.2 EQUIPAMENTOS DE HARDWARE

A câmera utilizada para a filmagem dos vídeos de tráfego foi uma câmera digital Sony Cyber-shot DSC-W320, colorida, com definição de 14.1 Mega *Pixels* para fotografias e 240x320 *pixels* de resolução e uma velocidade de 30 fps para gravação dos vídeos. Os padrões técnicos de distância focal e ângulo máximo de cobertura são respectivamente 26mm e 84°. A resolução, com valor baixo, mostrou-se suficiente, visto que os vídeos são processados na resolução 120x160 *pixels* devido às limitações de desempenho de processamento do software MATLAB®.

A entrada do vídeo no computador é através de uma porta de entrada USB, conectada manualmente. O arquivo de vídeo **.avi** fica armazenado na memória do próprio computador.

O microcomputador utilizado no processamento foi um notebook DELL INSPIRON 1545, com processador Intel Corel 2Duo T6600 de 2,2 GHz, memória RAM de 4GB e sistema operacional Windows 7 da Microsoft.

4.3 APLICAÇÃO AOS “PORTÕES DE ENTRADA DA CIDADE DE SÃO CARLOS, SP”.

Este item apresenta a aplicação do método em condições ambientais avulsas de filmagem com locação nas principais vias de acesso de uma cidade de médio porte. Como exemplo, foi tomada a cidade de São Carlos – SP. O mapa de localização do trabalho de campo está disponível no Apêndice A1.

O estudo teve início com a escolha dos locais para a filmagem. Como podemos observar no mapa do Apêndice A2 – Mapa de Localização das filmagens, foram identificados 3 (três) acessos principais à cidade de São Carlos, com potencial de entrada de veículos de carga rodoviário advindo das regiões lindeiras à cidade, carregados com os mais diversos tipos de carga, e que tem como destino, a descarga dentro do perímetro urbano. Foram tomadas como “portões de entrada” da cidade:

1. Avenida Getúlio Vargas – Rotatória de acesso à cidade proveniente da Rodovia SP-310;
2. Avenida Getúlio Vargas – Trevo com a Avenida Perimetral - recebe o fluxo proveniente do distrito industrial da cidade;
3. Rodovia Engenheiro Talles de Lorena Peixoto Júnior – Rotatória localizada próximo ao SAMU – Serviço de Atendimento Móvel de Urgência e Hospital Escola Municipal de São Carlos;
4. Rodovia SP 310 – Vista lateral da rodovia (vídeo utilizado para testar a maior distância de filmagem, 40,00 metros. O local era o único que fornecia essa condição de filmagem).

As filmagens utilizadas para os testes do sistema ocorreram durante 3 dias e a distância da câmera em relação a faixa de tráfego variou entre 20,00 e 40,00 metros em todas as filmagens. Foram realizadas filmagens sob variadas condições de iluminação, como pode ser consultado no Apêndice A3 – Arquivo de mídia DVD - Banco de dados de filmagens.

Foram escolhidas algumas amostras de vídeos para testar o sistema e outras amostras dos mesmos locais para confirmar a eficácia dos valores adotados para um determinado *setup*. Esta é a técnica utilizada para testar sistemas de identificação (controle) e de aprendizado em inteligência artificial.

Os veículos rodoviários de carga selecionados para serem rastreados e segmentados tem como característica as dimensões altura x comprimento, como o menor valor (4,00 x 14,00 metros) e como maior valor (4,00 x 30,00 metros), que são respectivamente, os valores utilizados para encontrar a relação *pixel* x milímetro, que preenche os parâmetros do modelo, área mínima e área máxima. Ou seja, qualquer veículo que esteja dentro dessa faixa de variação de medidas, deve ser segmentado pelo sistema. Esses valores podem ser obtidos por meio da resolução do problema de geometria enunciado por Tales ilustrado na Figura 4.2. Assim, conhecidos a distância focal da câmera e o ângulo de cobertura (26 mm e 84°, respectivamente), podem-se resolver as relações geométricas e assim obter o tamanho, em metros, do quadro de foco da filmagem. A distância focal, medida em milímetros, é a distância entre o centro ótico da lente e o sensor da câmera, é ela que define o quanto você consegue ver a partir de uma lente. Quanto maior o valor, mais “fechado” será o ângulo de visão de uma lente, (Catteli *et. al.*, 2009). Assim, para cada valor de distância de filmagem, nesse caso, 20, 30 e 40 metros, há uma relação de *pixel* x metro, o que nos leva a estimar os valores dos parâmetros a serem preenchidos na janela – parâmetros - da Figura 3.1.

Figura 4. 2 – Ilustração do teorema de Tales.

No esquema acima, as retas a , b e c são paralelas e as retas r e r' são transversais. De acordo com o Teorema de Tales, temos as seguintes proporcionalidades:

$$\frac{AB}{BC} = \frac{A'B'}{B'C'} \quad e \quad \frac{AB}{A'B'} = \frac{BC}{B'C'}$$

Assim, o esquema de cálculo para a obtenção do valor da relação *pixel* x milímetro para cada tomada de vídeo, por valor de distância de aquisição, pode ser representado na Figura 4.3.

Figura 4. 3 – Esquema de cálculo da relação *pixel* x milímetro do trabalho.

Os valores obtidos por essas relações geométricas são estimados em virtude de possíveis distorções da lente da câmera filmadora, porém tem acurácia suficiente para os objetivos deste projeto. Para a retirada de valores exatos de comprimento e altura dos veículos, técnicas de calibração de câmeras são necessárias, como no trabalho de Gupte *et. al*, (2002).

Assim, aplicando o método nos vídeos obtidos por meio das filmagens em campo, foram obtidos os resultados resumidos na Tabela 4.1, onde, todos os veículos rodoviários de carga foram identificados corretamente, e os erros, são relativos a eventos de falsa identificação.

Sendo o principal erro o excesso de identificação, podem-se considerar os resultados do modelo satisfatórios, pois, para o objetivo deste trabalho, que é de verificar a presença de um ou mais veículos rodoviário de carga nas cenas, a falsa identificação é mais bem aceita do que a falta de identificação. Os dois principais

erros do modelo foram: a sobreposição de veículos e a presença de ruídos nas imagens. No primeiro caso, dois veículos de passeio localizados muito próximos são identificados como um veículo rodoviário de carga (Figura 4.4).

Figura 4. 4 – Erro de identificação de veículo

O segundo tipo de erro é relativo ao evento ocorrido em algumas cenas, onde, o modelo identifica separadamente a região inferior e superior do caminhão (Figura 4. 5). Isso se deve ao fato que, por alguma descontinuidade na área, presença da sombra do próprio veículo, ou interferência do fundo da cena, o sistema compreende que se trata de 2 (dois) objetos. A correção depende do correto tratamento, a fim de eliminar as interferências do fundo complexo na etapa de detecção do plano de fundo.

Figura 4. 5 – Erro de identificação de objeto

Tabela 4. 1 – Resultados das análises

Vídeo	Duração (min)	Condições de Clima	Identificação visual		Identificação pelo modelo		
			Caminhão Pequeno	Caminhão Grande	Caminhão Pequeno	Caminhão Grande	Falsas detecções
<u>Cruzamento Getulio VargasxPerimetral.avi</u>	10:33	Ensolarado 10hs	11	8	11	8	6
<u>Cruzamento Getulio VargasxPerimetral Noite.avi</u>	01:32	Nublado	-	1	-	1	-
<u>Cruzamento Getulio VargasxPerimetral Noite 02.avi</u>	00:21	Nublado	1	-	1	-	-
<u>Rotatoria Getulio Vargas SP310 01.avi</u>	08:01	Ensolarado 09hs	3	5	3	5	-
<u>Rotatoria Getulio Vargas SP310 02.avi</u>	02:45	Ensolarado 13hs	1	-	1	-	-
<u>Rotatoria Getulio Vargas SP310 03 noite.avi</u>	10:22	Nublado	10	5	10	5	-
<u>Rotatoria Samu_xvid.avi</u>	17:11	Ensolarado 09hs	6	15	6	15	15
<u>Vista lateral Getulio Vargas 01.avi</u>	02:58	Ensolarado 09hs	2	2	2	2	-
<u>Vista lateral Getulio Vargas 02.avi</u>	01:25	Ensolarado 09hs	-	2	-	2	-
<u>Vista lateral Getulio Vargas 03.avi</u>	01:35	Ensolarado 13hs	-	4	-	4	-
<u>Vista lateral Getulio Vargas 04.avi</u>	08:59	Ensolarado 13hs	2	5	2	5	-
<u>Vista lateral Getulio I01.avi</u>	11:39	Ensolarado 09hs	8	14	8	14	-
<u>Vista lateral SP310_xvid.avi</u>	09:26	Ensolarado 11hs	16	22	16	22	-
TOTAL DAS ANÁLISES	92:08		60	83	60	83	21
TOTAL DE EVENTOS			143		143		21

A tabela 4.1 apresenta os números das análises discutidas no item anterior. Está apresentado o nome dos vídeos, a duração em minutos, as condições climáticas de obtenção das filmagens, o número de fenômenos de ocorrência de veículos rodoviários de carga em cada vídeo, obtidos primeiramente em uma inspeção visual e depois pela aplicação do modelo. Na última coluna, está apresentado o número de detecções falsas, ocasionadas pelos motivos expostos nos parágrafos anteriores.

Nesta pesquisa, são considerados veículos de grande porte todos os veículos de carga, excluindo-se os VUCs – Veículos urbanos de carga, vans e camionetes, que são veículos de pequeno porte.

Em cada análise foram modificados os parâmetros referentes ao valor do limiar “T”, utilizado na função *autothreshold*, a altura da linha branca e valores de área mínima e máxima de detecção. A escolha dos parâmetros deu-se em função da distância de filmagem. Assim, para cada vídeo, foram determinados os valores dos parâmetros, definidos na Tabela 4.2.

Tabela 4. 2 – Definição dos parâmetros do modelo

PARÂMETROS DE ANÁLISE					
Vídeo	Distância (m)	Área mín.	Área máx.	Limiar (T)	Altura da linha branca
Cruzamento Getulio VargasxPerimetral.avi	20,00	1100	2400	0,20	50
Cruzamento Getulio VargasxPerimetral_Noite.avi	20,00	1100	2400	0,20	50
Cruzamento Getulio VargasxPerimetral_Noite_02.avi	20,00	1100	2400	0,20	50
Rotatoria Samu_xvid.avi	20,00	1100	2400	0,30	50
Vista lateral_Getulio Vargas 01.avi	20,00	1100	2400	0,30	50
Vista lateral_Getulio Vargas 02.avi	20,00	1100	2400	0,30	50
Vista lateral_Getulio Vargas 03.avi	20,00	1100	2400	0,30	50
Vista lateral_Getulio Vargas 04.avi	20,00	1100	2400	0,30	50
Vista lateral_Getulio_I01.avi	20,00	1100	2400	0,30	50
Rotatoria Getulio Vargas_SP310_01.avi	30,00	490	1100	0,20	30
Rotatoria Getulio Vargas_SP310_02.avi	30,00	490	1100	0,20	30
Rotatoria Getulio Vargas_SP310_03_noite.avi	30,00	490	1100	0,15	60*
Vista lateral_SP310_xvid.avi	40,00	275	600	0,30	25

* O vídeo rotatória Getulio Vargas_SP310_03_noite.avi, teve o valor padrão da altura da linha branca alterado devido a presença de nuvens que causavam grande interferência na

segmentação do plano de fundo. Assim, a área de análise foi reduzida ao limite da altura máxima dos veículos.

A principal dificuldade encontrada nas análises em vídeos de tráfego de veículos é a presença de fundo complexo. No caso em análise, o plano de fundo é caracterizado pela presença de elementos comuns a cenários urbanos (árvores, placas, veículos passando na outra pista, pedestres, etc). Essa constatação também está presente nos trabalhos de Jazayeri *et. al.*, (2009) e Buch *et. al.*, (2008); e é minimizado nos trabalhos de Morris e Trivedi, (2006); Mohottala *et. al.*, (2003), pois os vídeos de entrada desses autores são tomados de um pórtico elevado acima do fluxo de veículos e assim, paralelo à faixa de tráfego, o que faz o plano de fundo tornar-se mais simples.

As sequências de Figura 4. 6 a 4.8 mostram alguns resultados de detecção dos veículos rodoviários de carga. O detalhe da contagem dos veículos por quadro no canto esquerdo das imagens pode ser visualizado.

Em uma extensão desta pesquisa os valores de área dos retângulos poderiam ser utilizados para a classificação dos veículos em categorias geométricas pré-definidas. Esta funcionalidade não foi implementada no sistema, pois são necessárias adaptações no modelo para a manipulação destes objetos, que fogem ao escopo desta dissertação.

Figura 4. 6 – Resultados de detecção

Figura 4. 7 - Resultados de detecção

Figura 4. 8 - Resultados de detecção

CAPÍTULO 5

CONSIDERAÇÕES FINAIS

Neste capítulo é apresentada uma análise sobre o desempenho do método aplicado e as conclusões sobre a aplicação do programa nos vídeos resultantes do trabalho de campo realizado em vias urbanas da cidade de São Carlos - SP, assim como, são sugeridas propostas para trabalhos futuros.

5.1 CONCLUSÕES

Este trabalho cumpre com o objetivo de contribuir para o estudo sobre a utilização de técnicas de processamento de imagens na detecção de veículos em vídeos de tráfego urbano.

O sistema apresentado permite um elevado índice de acerto na detecção de veículos rodoviários de cargas, em meio a cenas de tráfego urbano, por meio de características geométricas. O sistema é robusto e dentro do espectro de variação de iluminação investigado, não sofre influência da modificação de iluminação e posicionamento da câmera.

As variações de iluminação são absorvidas com a calibração dos parâmetros de limiar.

As variações de posicionamento da câmera são absorvidas pelo ajuste dos valores de área mínima e máxima, assim como pelo ajuste do valor da posição da linha branca, utilizada como apoio para a detecção.

A maior limitação encontrada foi referente ao problema do fundo complexo. Nesse tipo de filmagem, além de objetos de diversas formas presentes no fundo da cena, a sobreposição entre os veículos que passam nas faixas de tráfego ao mesmo tempo, algumas vezes, podem gerar falsas detecções.

Apesar das limitações encontradas pela falta de domínio em técnicas de programação, por se tratar de um trabalho em nível de mestrado realizado por

pesquisadora que não é do meio técnico da computação, foi possível, através da revisão e aprendizado das técnicas estudadas por autores do meio da computação e relatados em periódicos, modelar o problema utilizando o MATLAB®, uma plataforma popular e acessível para investigações técnico-científicas. As limitações do MATLAB em termos de capacidade de processamento, requerem o pós processamento das imagens, não permitindo, portanto, soluções em tempo real.

5.2 SUGESTÕES PARA TRABALHOS FUTUROS

A seguir, são apresentadas possíveis melhorias que podem ser aplicadas ao algoritmo e ao método de pesquisa para trabalhos futuros.

Para tornar o método aplicável para solução em tempo real, sugere-se o desenvolvimento de um modelo semelhante em linguagem C ou C++, por exemplo, aliada a técnicas de transmissão de dados.

O sistema desenvolvido pode ser complementado por meio da inserção de novas técnicas de filtragem dos ruídos das imagens com fundo complexo. Assim, as características geométricas dos veículos poderiam ser obtidas por meio de técnicas de calibração de câmeras (Marques, 2007; Neto, s.d.), uma vez que, conhecido o tamanho real de algum objeto na janela de filmagem, podem-se obter automaticamente as características de dimensões reais dos veículos, como por exemplo, altura, comprimento, largura, distância entre eixos e posteriormente realizar a classificação de acordo com o parâmetro escolhido.

Outra linha de pesquisa a ser considerada é a utilização de técnicas de Inteligência Artificial, como Redes Neurais Artificiais – RNA, para a partir de um banco de imagens pré armazenados, proceder-se a classificação dos veículos rodoviários de carga de acordo com alguma característica escolhida, do tipo dimensão ou cor (Dickinson e Wan, 1989; Wan e Dickinson, 1989; *apud* Kastrinaki *et. al.*, 2003; Demuth e Beale, 2006; Nunes e Prado, 2002; Ribeiro e Lima, 1999; Piccioli *et. al.*, 1996).

REFERÊNCIAS

- BERTOZZI, M., BROGGI, A. CASTELLUCCIO, S. *A Real-Time Oriented System for vehicle Detection*. J. Systems Architecture, p. 317-325, 1997.
- BESTUFS – Best Urban Freight Solutions. *Best Practice Update – E-Commerce and urban freight distribution*. Disponível em:
http://www.bestufs.net/cgi-bin/projectdb/project_db.pl. Acesso em: janeiro/2011.
- BUCH, N.; ORWELL, J.; VELASTIN, S.A. *Detection and Classification of Vehicles for Urban Traffic Scenes*. The Institution of Engineering and Technology. P. 182-187, 2008.
- CAIXETA, J.V.F e MARTINS, R.S. *Gestão Logística do Transporte de Cargas*. Atlas. São Paulo – SP, 2001.
- CATELLI, F., OURIQUE, P., da SILVA, F. *Qual é o campo de visão de uma objetiva de uma câmera fotográfica?*. In: Cad. Bras. Ens. Fis., v. 26, n. 2: p. 314-327, ago. 2009.
- CHEUNG, S. C.; C. KAMATH. *Robust techniques for background subtraction in urban traffic video*. Video Communications and Image Processing, SPIE Electronic Imaging, 2004.
- CHEUNG, S. C.; C. KAMATH. *Robust Background Subtraction with Foreground Validation for Urban Traffic Video*. EURASIP Journal on Applied Signal Processing, Vol. 14, p. 1-11, 2005.
- CHIESA, G. A. E. *Cargas em São Paulo*. Revista Transporte Moderno Nº 393, São Paulo - SP, 1999.
- CRISMAN, J.; THORPE, C. *Color vision for road following*. Proc. SPIE Conf. Mobile Robots, p. 246-249, 1988.
- COSTA, A. Notas de Aula. Mestrado ISEP/IST em Eng. Electrotécnica e Computadores. Lisboa, 2004.

CUCCHIARA, R.; PICCARDI, M. *Vehicle Detection under Day and Night Illumination*. Proc. Int'l ICSC Symp. Intelligent Industrial Automation, 1999.

DEMUTH, H., BEALE, M. *Neural Network Toolbox User's Guide*. The Mathworks Inc, 2006.

DESA, S. M.; SALIH, Q. A. *Image subtraction for real time moving object extraction*. Proceedings of the International Conference on Computer Graphics, Imaging and Visualization. 2004.

DUTRA, N. G. S. *O enfoque de "city logistics" na distribuição urbana de encomendas*. Tese doutorado apresentada ao Programa de Pós Graduação em Engenharia de Produção da Universidade Federal de Santa Catarina - PPGE/UFSC. Cap. 4. Florianópolis – SC, 2004.

FARUOLO L.B. *Análise de confiabilidade da pesagem dinâmica de veículos rodoviários*. Dissertação de mestrado em tecnologia, Centro Federal de Educação Tecnológica – CEFET, Rio de Janeiro - RJ, 2007.

FRANKE, U.; KUTZBACH, I. *Fast Stereo Based Object Detection for Stop and Go Traffic*. Intelligent Vehicles, p. 339-344, 1996.

FRIEDMAN, N. AND RUSSELL, S. *Image segmentation in video sequences*. In Proc. of the Thirteenth Conference on Uncertainty in Artificial Intelligence. Providence, Rhode Island, 1997.

GIACHETTI, A.; CAMPANI, M.; TORRE, V. *The Use of Optical Flow for Road Navigation*. IEEE Transactions Robotics and Automation, vol. 14, n. 1, 1998.

GOERICK, C., NOLL, D., WERNER, M. *Artificial Neural Networks in Real-time Car Detection and Tracking Applications*. Pattern Recognition Letters, vol.17, p. 335-343, 1996.

GONZALEZ, R.C.; WOODS R.E. *Processamento Digital de Imagens*, 3nd, Editora Blucher. São Paulo – SP, 2007.

GONZALEZ, R.C.; WOODS R.E AND EDDINS, S.E. *Digital Image Processing using MATLAB®*. Gatesmark Publishing. Second Edition, 2009.

GUPTE, S., PAPANIKOLOPOULOS, N. *Algorithms for vehicle classification*. Final Report 2000. University of Minnesota, July, 2000.

GUPTE, S.; MASOUD, O.; MARTIN, R. F. K.; PAPANIKOLOPOULOS, N. P. *Detection and Classification of Vehicles*. IEEE Transactions on Intelligent Transportation systems, Vol 3, n. 1, p. 37 - 47, 2002.

HANDMANN, U; KALINKE, T; TZOMAKAS, C; WERNER, M; SEELEN, W. V. *An Image Processing System for Driver Assistant*. Image and Computing, vol. 18, n. 5, 2000.

HARASSE, S., BONNAUD, L., DESVIGNES, M. *Backgorund Estimation for dynamic video scenes*. IEEE Industrial Electronics, IECON - 32nd Annual Conference on, p. 3519-3524, 2006.

Image Processing Toolboxes. User's Guide. Math Works, INC, 2011.

JAZAYERI, A.; CAI, H.; ZHENG, J.Y. *Vehicle Detection and Tracking in Car Video Based on Motion Model*. IEEE Transactions on Image Processing, 2009.

JONES, P.; HERVIK, A. *Restraining car traffic in European Cities: emerging role for road pricing*. Transportation Research, part A, vol. 26 A, n.2, p. 133-145, 1992.

KASTRINAKI, V.; ZERVAKIS, M.; KALAITZAKIS. *A survey of video processing techniques for traffic applications*. Image and Vision Computing, n. 21, p. 359-381, 2003.

KIM , J.B.; PARK , H.S.; PARK , M.H.; KIM, H.J., *A real-time region-based motion segmentation using adaptive thresholding and K-means clustering*, in: M. Brooks, D. Corbett, M. Stumptner (Eds.), pp. 213–224, Springer, Berlin, 2001.

KUEHNLE, A. Symmetry – *Based Recognition for Vehicle Rears*. Pattern Recognition Letters, vol. 12, p. 249-258, 1991.

LIMA JR. O.F. *A carga na cidade hoje e amanhã*. LALT - Laboratório de Aprendizagem em Logística e Transportes. UNICAMP, Campinas - SP, 2005.

LIPTON, A.J.; FUJIYOSHI, H. AND PATIL, R.S. *Moving target classification and tracking from real-time video*. In Proc. of the IEEE Workshop on Applications of Computer Vision, p. 8-14, 1998.

LOPES, F.M. *Um Modelo Perceptivo de Limiarização de Imagens Digitais*. Dissertação de Mestrado em Informática, Universidade Federal do Paraná – UFPR, Curitiba – PR, 2003.

MARQUES, C. C. S. C. *Um sistema de calibração de câmera*. Instituto Nacional de Matemática Pura e Aplicada – IMPA. Rio de Janeiro, 2007.

MARQUES FILHO, Ogê; VIEIRA NETO, Hugo. *Processamento Digital de Imagens*, Rio de Janeiro: Brasport, 1999.

MAZAHARI, M., MOZAFFARI, S. *Real Time Adaptive Background Estimation and Road Segmentation for Vehicle Classification*. Electrical Engineering (ICEE), 19th Iranian Conference on, 2011.

MILES, J.; WALKER, J.; MACMILLIAN, A.; ROUTLEDGE, I. *Access control in city centers: objectives, methods and examples*. Traffic Engineering and Control, v. 39, n.2, p. 648-654, 1998.

MOHOTTALA, S., KAGESAWA, M., IKEUCHI, K. *Vehicle class recognition using 3D CG models*. Proceedings of 2003 ITS World Congress, November, 2003.

MORRIS, B.; TRIVEDI, M. *Robust Classification and Tracking of Vehicles in Traffic Video Streams*. Proceedings of the IEEE ITSC – Intelligent Transportation Systems Conference, p. 1078-1083, 2006.

MORRIS, B. TRIVEDI, M. *Improved vehicle classification in long traffic video by cooperating tracker and classifier modules*. In Proceedings of AVSS 06, p.09. Washington D.C., USA, 2006. IEEE Computer Society.

NETO, A. A. S. *Calibração de Câmeras*, Pontifícia Universidade Católica do Rio de Janeiro – PUC-RIO, S.D.

NUNES, L.E.N.P., PRADO, P.P.L. Reconhecimento de objetos contidos em imagens através de redes neurais. *Revista Ciencias Exatas*, p. 77-88. Taubaté - SP, 2002.

OHBA, K. AND IKEUCHI, K. Detectability, Uniqueness, and Reliability of Eigen Windows for Stable Verification of Partially Occluded Objects. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 19, N.9, 1997.

OTSU, N. *A threshold selection method from gray-level histogram*. *IEEE Transactions on Systems, Man, and Cybernetics*, p.62-66, 1979.

PNEVMATIKAKIS, A. AND POLYMENAKOS, L. *Robust Estimation of Background for Fixed Cameras*. *IEEE, Proceedings of the 15th International Conference on Computing*, 2006.

PICCARDI, M. *Background Subtraction Techniques: A review*. *IEEE International Conference on Systems, Man and Cybernetics*. P. 2099-3004, 2004.

PICCIOLI, G., MICHELI, E., PARODI, P., AND CAMPANI, M. *Robust method for road sign detection and recognition*, 1996.

RIBEIRO, F. F. S., LIMA, A. C. C. *Proceedings of the IV Brazilian Conference on Neural Networks - IV Congresso Brasileiro de Redes Neurais - ITA - São José dos Campos - SP*, 1999.

REDDY, V., SANDERSON, C., LOVELL, B. C. An Efficient and Robust Sequential Algorithm for Background Estimation in Video Surveillance. *Image Processing (ICIP) - 16th IEEE International Conference on*, p. 1109-1112, 2009.

ROSIN, P. *Thresholding for change detection*. In Proceedings of IEEE Int'l Conference on Computer Vision, pgs 274–279, 1998

SILVA, E. A.; GONZAGA, A. *Controle de Fluxo de Veículos através da Segmentação de Movimentos usando Modelo de Misturas de Gaussianas*. In: 5º Congresso Brasileiro de Rodovias e Concessões, 2007, CBR&C, São Paulo - SP.

SULLIVAN, G.D.; BAKER, K.D.; WORRALL, A.D.; ATTWOOD, C.I.; REMAGNINO, P.M. *Model-based vehicle detection and classification using orthographic approximations*. Image and Vision Computing 15, pp 649-654, 1997.

SUN, Sun-Gu; PARK, J.; PARK, Y. W. *Identification of Military Ground Vehicles by Feature Information Fusion in FLIR Images*. Proceedings of the 3rd International Symposium on Image and Signal Processing and Analysis, 2003.

SUN, Z.; BEBIS, G; MILLER, R. *On-Road Vehicle Detection: A Review*. IEEE Transaction on Pattern Analyses and Machine Intelligence, vol. 28, n.5, 2006.

TANIGUCHI, R. G. THOMPSON, T. Y., R. VAN DUIN. *City Logistics: Network Modelling and Intelligent Transport Systems*. Pergamon, 2001.

THEODORIDIS, S.; KOUTROUMBAS, K. *Pattern Recognition*. 3. ed. San Diego: Elsevier, p.837, 2006.

TOYAMA, K.; KRUMM J.; BRUMITT, B.; MEYERS, B. *Wallflower: principles and practice of background maintenance*. In: Proceedings of 7th IEEE Conference on Computer Vision, Kerkyra, 1999.

TOULMINET, G.; BERTOZZI, M.; MOUSSET, S.; BENSRIHIR, A.; BROGGI, A. *Vehicle Detection by Means of Stereo Vision-Based Obstacles Features Extraction and Monocular Pattern Analysis*. IEEE Transactions on Image Processing, Vol 15, n.8, August, 2002.

URAZGHILDIIEV, I.; ROLF, R.; RIDDERSTRÖM, P.; RYDBERG, A.;
 ÖJEFORS, E.; WALLIN, K.; ENOCHSSON, P.; ERICSON, M.; LÖFQVIST, G.
Vehicle Classification Based on the Radar Measurement of Height Profiles.
 IEEE Transactions on Intelligent Transportation systems, Vol. 8, n. 2, June,
 2007.

ZHANG, C., SIYAL, M. Y. *A new segmentation technique for classification of moving vehicles.* Vehicular Technology Conference Proceedings, IEEE 51st, p.323-326 vol.1, Tokyo, 2000.

YOUNG, I.T.; Gerbrands, J.J.; Van Vilet, L.J. *Fundamentals of Image processing.* Quantitative Imaging Group - Department of Imaging Science and Technology, Faculty of Applied Sciences. Delft University of Technology – Delft The Netherlands. Disponível em:
<http://www.ph.tn.tudelft.nl/Courses/FIP/noframes/fip.html>. Acesso em novembro/2011

WANG, L.; HU, W.; TAN, T. e MAYBANKS, S. *A survey on visual surveillance of object motion and behaviors.* IEEE Transactions on Systems, Man and Cybernetics, vol.3, p.334-352, 2004.

APÊNDICES

APÊNDICE A2
LOCALIZAÇÃO DO TRABALHO DE CAMPO
Fonte: Google, 2012

Legenda:
Região de Filmagem - "Portões de Entrada" da Cidade de São Carlos
Região de Interesse - Contagem insignificante para a pesquisa
Localização da câmera de filmagem