

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
ESCOLA DE ENGENHARIA
DEPARTAMENTO DE ENGENHARIA ELÉTRICA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

CHRISTIANO BOUVIÉ DA SILVA

**RASTREAMENTO VEICULAR EM VÍDEOS DE TRÁFEGO
COM APLICAÇÃO EM CONTAGEM DE VEÍCULOS**

Porto Alegre

2012

CHRISTIANO BOUVIÉ DA SILVA

**RASTREAMENTO VEICULAR EM VÍDEOS DE TRÁFEGO
COM APLICAÇÃO EM CONTAGEM DE VEÍCULOS**

Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Engenharia Elétrica, da Universidade Federal do Rio Grande do Sul, como parte dos requisitos para a obtenção do título de Mestre em Engenharia Elétrica.

Área de concentração: Engenharia de Computação, Processamento de Sinais.

ORIENTADOR: Jacob Scharcanski

Porto Alegre

2012

CHRISTIANO BOUVIÉ DA SILVA

RASTREAMENTO VEICULAR EM VÍDEOS DE TRÁFEGO COM APLICAÇÃO EM CONTAGEM DE VEÍCULOS

Esta dissertação foi julgada adequada para a obtenção do título de Mestre em Engenharia Elétrica e aprovada em sua forma final pelo Orientador e pela Banca Examinadora.

Orientador: _____

Prof. Dr. Jacob Scharcanski, UFRGS

Doutor pela University of Waterloo – Waterloo, Canadá

Banca Examinadora:

Prof. Dr. Adilson Gonzaga, USP – São Carlos

Doutor pela Universidade de São Paulo – São Paulo, Brasil

Prof. Dr. Alexandre Balbinot, UFRGS

Doutor pela Universidade Federal do Rio Grande do Sul – Porto Alegre, Brasil

Prof. Dr. André Carlos Ponce de Leon Ferreira de Carvalho, USP – São Carlos

Doutor pela University of Kent – Canterbury, Reino Unido

Prof. Dr. Valner João Brusamarello, UFRGS

Doutor Universidade Federal de Santa Catarina – Florianópolis, Brasil

Coordenador do PPGEE: _____

Prof. Dr. João Manoel Gomes da Silva Jr.

Porto Alegre, novembro de 2012.

DEDICATÓRIA

Dedico este trabalho a minha mãe, Jaqueline, e a minha mulher, Rosângela, em especial pela dedicação e apoio em todos os momentos.

AGRADECIMENTOS

Ao Programa de Pós-Graduação em Engenharia Elétrica, PPGEE, pela oportunidade de realização de trabalhos em minha área de pesquisa.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, CAPES, pela provisão da bolsa de mestrado.

À Universidade Federal do Rio Grande do Sul, UFRGS, em conjunto com o Grupo Digicon S.A. pela oportunidade de participação no desenvolvimento do projeto coordenado pelo Professor Jacob Scharcanski.

Aos Professores do PPGEE e do Programa de Pós-Graduação em Computação, PPGC, pela sua orientação no desenvolvimento deste trabalho.

Aos colegas do PPGEE e PPGC, pelo seu auxílio nas tarefas desenvolvidas durante o curso.

RESUMO

Esta dissertação apresenta o desenvolvimento de um método, baseado em agrupamento de partículas, para realizar a contagem de veículos em vídeos de tráfego. Tal procedimento é importante em sistemas de tráfego inteligentes, ou como uma ferramenta auxiliar no planejamento de vias urbanas. Utilizando técnicas de processamento de imagens e agrupamento de partículas, o método proposto utiliza-se da coerência de movimento e posição espacial existente entre partículas extraídas das imagens de vídeo para agrupá-las, formando figuras convexas que são analisadas em busca de possíveis veículos. Essa análise leva em consideração a morfologia das figuras convexas e a informação de fundo da imagem, para unir ou dividir os agrupamentos. Após a identificação de um veículo, o mesmo é rastreado utilizando-se similaridade de histograma de cores, aplicado em janelas centradas nas partículas. A contagem dos veículos ocorre em laços virtuais definidos pelo usuário nas pistas desejadas, através da intersecção das figuras convexas rastreadas com estes laços virtuais. Testes foram realizados utilizando-se vídeos de seis cenas diferentes, totalizando 81.000 quadros. Os resultados das contagens de veículos obtidos foram comparados a dois métodos atuais. Um método possui abordagem similar ao método proposto (KIM, 2008), que tenta fixar agrupamentos de partículas em formas elipsoidais. O outro método (SÁNCHEZ *et al.*, 2011) rastreia objetos conectados, quando estes são diferentes do fundo, através da intersecção destes objetos entre quadros adjacentes. Considerando-se o universo total de veículos analisados, 1085 veículos, os resultados obtidos pelo método proposto apresentaram uma diferença absoluta na contagem dos veículos intermediária aos métodos comparativos, 53 veículos contra 66 e 27 para (KIM, 2008) e (SÁNCHEZ *et al.*, 2011) respectivamente, sendo o único método que contou menos veículos que o valor real, enquanto os métodos comparativos contaram veículos além do valor real. O método proposto perde 102 veículos, valor inferior ao método de (SÁNCHEZ *et al.*, 2011), 181, e praticamente o mesmo número que o método de (KIM, 2008), 101. Já os veículos detectados mais de uma vez apresentam valores inferiores para o método proposto, 49, em relação aos métodos comparativos, 167 para (KIM, 2008) e 208 para (SÁNCHEZ *et al.*, 2011).

Palavras-chaves: Agrupamento de Partículas, Contagem de Veículos, Processamento de Imagens, Processamento de Vídeo, Rastreamento de Veículos, Visão Computacional, Processamento de Sinais, Instrumentação.

ABSTRACT

This dissertation presents the developed of a method based on particle group, to conduct the count of vehicles in traffic videos. This procedure is important in intelligent traffic systems, or as an auxiliary tool in the planning of urban streets. Using image processing techniques and grouping of particles, the proposed method uses the coherence of spatial position and movement between particles extracted from the video footage to assemble them into convex figures that are parsed in search of possible vehicles. This analysis takes into account the morphology of convex figures and background information of the image, to merge or split the groups. After the identification of a vehicle, it is tracked using color histogram similarity, applied in Windows centered on particles. The count of vehicles occurs in user-defined virtual loops on the tracks desired, through the intersection of convex figures traced with these virtual loops. Tests were performed using videos of six different scenes, totaling 81,000 frames. The results of vehicle counts obtained were compared to two current methods. A method has similar approach to proposed method (KIM, 2008), which attempts to establish groups of particles in ellipsoidal shapes. The other method (SÁNCHEZ *et al.*, 2011) tracks connected objects, when these are different from the background, through the intersection of these objects between adjacent frames. Considering the total universe of vehicles examined, 1085, the results obtained by the proposed method showed an intermediate absolute difference in counting of vehicles to comparative methods, 53 against 66 and 27 vehicles for (KIM, 2008) and (SÁNCHEZ *et al.*, 2011) respectively. The proposed method is the only one that counted vehicles less than the real value, while comparative methods counted vehicles beyond the real value. The proposed method loses 102 vehicles, lower than value to (SÁNCHEZ *et al.*, 2011), 181, and roughly the same number as the method of (KIM, 2008), 101. The number of vehicles detected more than once are lower for the proposed method in relation to the comparative methods, 49 vehicles against 167 to (Kim, 2008) and 208 to (SANCHEZ *et al.*, 2011).

Keywords: Computer Vision, Image Processing, Particles Clustering, Vehicle Count, Vehicle Tracking, Video Processing, Signal Processing, Instrumentation.

LISTA DE ILUSTRAÇÕES

Figura 1: Exemplificação dos laços virtuais.....	13
Figura 2: Exemplo de câmera acima da pista.....	18
Figura 3: Exemplificação do método proposto por (KANHERE e BIRCHFIELD, 2008). (a) Linhas fornecidas pelo usuário; (b) Projeção tridimensional obtida a partir de (a); (c) Diferentes tipos de partículas; (d) Agrupamentos obtidos.....	19
Figura 4: (a) Agrupamentos e suas trajetórias em nível de partículas e (b) agrupamentos e trajetórias em nível de objetos, obtidas através do método proposto por (KIM, 2008).....	20
Figura 5: (a) Objetos conectados e (b) objetos identificados pelo método apresentado por (PURNAMA <i>et al.</i> , 2009).....	21
Figura 6: Exemplo de rastreamento utilizando-se câmera embarcada, por (MEI e LING, 2011).....	21
Figura 7: Exemplo das categorias de deslocamento dos veículos apresentado por (FERIS <i>et al.</i> , 2012).....	22
Figura 8: Utilização do espaçamento no histograma para dividir dois veículos identificados por (WU e JUANG, 2012).....	23
Figura 9: Exemplo de câmera lateral utilizada em (FERIS <i>et al.</i> , 2012).....	24
Figura 10: Exemplo de oclusão e região de separação entre dois veículos para (HUANG e LIAO, 2004).....	24
Figura 11: Rodas apresentadas como objetos conectados (em preto) dentro da região do objeto de interesse (sombra do veículo, em branco).....	25
Figura 12: Exemplo de oclusão de veículos em (MORRIS e TRIVEDI, 2008).....	26
Figura 13: Exemplo de rastreamento de faróis realizado por (CHEN <i>et al.</i> , 2011).....	26
Figura 14: (a) Reflexos causados pelos faróis; (b) Mapa de intensidade de reflexão; (c) Mapa de reflexão suprimida. De (ZHANG <i>et al.</i> , 2012).....	27
Figura 15: Exemplo de rastreamento através de câmeras no interior de veículos de (SIVARAMAN e TRIVEDI, 2010).....	28
Figura 16: (a) Diferença no fluxo óptico entre fundo e veículo; (b) Exemplo de mapa espaço-temporal; (c) Exemplificação da análise de um mapa espaço-temporal. De (JAZAYERI <i>et al.</i> , 2011).....	29
Figura 17: Cubo RGB.....	32
Figura 18: Exemplo de quantização de cores no cubo RGB.....	33
Figura 19: Exemplo de filtro de mediana temporal.....	36
Figura 20: Função densidade de probabilidade para o ruído impulsivo.....	37
Figura 21: Exemplificação da comparação de blocos.....	39
Figura 22: Amostras dos vídeos utilizados.....	43
Figura 23: Regiões de interesse em vermelho.....	43
Figura 24: Laços virtuais em vermelho com suas respectivas contagens em azul.....	43
Figura 25: Exemplos de fundos.....	44
Figura 26: Exemplos de mapas de partículas repetidas.....	44
Figura 27: Visão geral do método proposto.....	47

Figura 28: Similaridade de histogramas de 0,9679.	51
Figura 29: Similaridade de histogramas de 0,3514.	51
Figura 30: Fluxograma do agrupamento inicial.	52
Figura 31: Exemplo de divisão de agrupamento de partículas.	54
Figura 32: Exemplo de agrupamentos espaciais inadequados de partículas.	55
Figura 33: Exemplo de fusão de agrupamentos de partículas.	56
Figura 34: Fluxograma da detecção de veículos.	57
Figura 35: Forma do agrupamento.	58
Figura 36: Análise do fundo na região do veículo.	58
Figura 37: Resultado da identificação de veículos.	59
Figura 38: Fluxograma de rastreamento de veículos.	60
Figura 39: Similaridade de histograma de cor de 0,9840 entre duas partículas.	61
Figura 40: Similaridade de histograma de cor de 0,6494 entre duas partículas.	61
Figura 41: Processo de ativação do laço virtual e incremento do contador.	62
Figura 42: Passagem do veículo sobre mais de um laço virtual.	63
Figura 43: Quadro do vídeo Cam06 (a), ROI (b), laços virtuais (c) e fundo (d) utilizados.	65
Figura 44: Mapa de repetitividade das partículas iniciais da Cam06.	65
Figura 45: Rastreamento de veículos em situação de trânsito normal na Cam06.	66
Figura 46: Ativação dos laços virtuais e incremento dos respectivos contadores na Cam06.	66
Figura 47: Rastreamento em situação de proximidade e oclusão parcial devido à parada no semáforo na Cam06.	67
Figura 48: Rastreamento em situação de proximidade inicial dos veículos envolvidos na Cam06.	67
Figura 49: Quadro do vídeo Cam25 (a), ROI (b), laços virtuais (c) e fundo (d) utilizados.	70
Figura 50: Mapa de repetitividade das partículas iniciais da Cam25.	70
Figura 51: Rastreamento durante os primeiros quadros, com identificação inicial dos veículos na Cam25.	70
Figura 52: Rastreamento de uma motocicleta, identificada apenas na saída da ROI na Cam25.	71
Figura 53: Quadro do vídeo Cam41 (a), ROI (b), laços virtuais (c) e fundo (d) utilizados.	73
Figura 54: Borramento na obtenção das imagens da Cam41.	73
Figura 55: Mapa de repetitividade das partículas iniciais da Cam41.	74
Figura 56: Rastreamento de um conjunto caminhão – escavadeira na Cam41.	74
Figura 57: Rastreamento de um veículo grande na Cam41, inicialmente identificado como dois objetos.	74
Figura 58: Quadro do vídeo Cam64 (a), ROI (b), laços virtuais (c) e fundo (d) utilizados.	76
Figura 59: Mapa de repetitividade das partículas iniciais da Cam64.	77
Figura 60: Exemplo de veículos perdidos durante rastreamento na Cam64.	79
Figura 61: Separação dos veículos rastreados conforme início da movimentação na Cam64.	79
Figura 62: Quadro do vídeo Cam68 (a), ROI (b), laços virtuais (c) e fundo (d) utilizados.	80
Figura 63: Mapa de repetitividade das partículas iniciais da Cam68.	80
Figura 64: Exemplo de falha de detecção de motocicletas na Cam68.	82
Figura 65: Quadro do vídeo Cam73 (a), ROI (b), laços virtuais (c) e fundo (d) utilizados.	83
Figura 66: Mapa de repetitividade das partículas iniciais da Cam73.	83
Figura 67: Exemplo de falha de detecção e rastreamento na Cam73.	85
Figura 68: Exemplo de rastreamento para um veículo maior que a ROI na Cam73.	85

LISTA DE TABELAS

Tabela 1: Contagem de veículos – Cam06.	68
Tabela 2: Veículos não detectados – Cam06.	68
Tabela 3: Veículos multiplamente detectados – Cam06.	69
Tabela 4: Contagem de veículos – Cam25.	71
Tabela 5: Veículos não detectados – Cam25.	72
Tabela 6: Veículos multiplamente detectados – Cam25.	72
Tabela 7: Contagem de veículos – Cam41.	75
Tabela 8: Veículos não detectados – Cam41.	75
Tabela 9: Veículos multiplamente detectados – Cam41.	76
Tabela 10: Contagem de veículos – Cam64.	77
Tabela 11: Veículos não detectados – Cam64.	78
Tabela 12: Veículos multiplamente detectados – Cam64.	78
Tabela 13: Contagem de veículos – Cam68.	81
Tabela 14: Veículos não detectados – Cam68.	81
Tabela 15: Veículos multiplamente detectados – Cam68.	82
Tabela 16: Contagem de veículos – Cam73.	84
Tabela 17: Veículos não detectados – Cam73.	84
Tabela 18: Veículos multiplamente detectados – Cam73.	85
Tabela 19: Resultado geral para veículos perdidos.	86
Tabela 20: Resultado geral para veículos contados mais de uma vez.	86
Tabela 21: Resultado geral para contagem de veículos.	87

LISTA DE ABREVIATURAS

EPTC: Empresa Pública de Transporte e Circulação

HSI: Cor, Saturação e Brilho (*Hue, Saturation and Intensity*)

MAD: Média das Diferenças Absolutas

RGB: Vermelho, Verde e Azul (*Red, Green and Blue*)

ROI: Região de Interesse (*Region Of Interest*)

SUMÁRIO

1	INTRODUÇÃO	13
2	REVISÃO BIBLIOGRÁFICA	16
2.1	VISÃO GERAL DO MONITORAMENTO DE TRÁFEGO UTILIZANDO PROCESSAMENTO DE VÍDEO	16
2.2	RASTREAMENTO E CONTAGEM DE VEÍCULOS UTILIZANDO-SE CÂMERAS ESTÁTICAS...	18
2.2.1	Câmeras Acima das Pistas	18
2.2.2	Câmeras Laterais às Pistas	23
2.2.3	Rastreamento e Contagem Noturna	26
2.3	RASTREAMENTO DE VEÍCULOS UTILIZANDO-SE CÂMERAS MÓVEIS...	28
3	FUNDAMENTAÇÃO TEÓRICA	31
3.1	REPRESENTAÇÃO DE IMAGENS E VÍDEO	31
3.1.1	Métricas das Imagens Digitais	32
3.1.2	Comparação de Histogramas	33
3.2	OPERAÇÕES MORFOLÓGICAS	34
3.2.1	Dilatação	35
3.2.2	Erosão	35
3.2.3	Abertura e Fechamento	35
3.3	OBTENÇÃO DO FUNDO E OBJETOS DE INTERESSE	36
3.4	RUÍDO IMPULSIVO	37
3.5	DETECTOR DE CANTOS	38
3.6	VETORES DE MOVIMENTO	38
3.7	ALGORITMO K-MEANS	39
4	RASTREAMENTO VEICULAR EM VÍDEOS DE TRÁFEGO	42
4.1	VISÃO GERAL DO ALGORITMO PROPOSTO	43
4.2	PRÉ-PROCESSAMENTO	48
4.3	DETALHAMENTO DO ALGORITMO PROPOSTO	49
4.3.1	Agrupamento Inicial	52
4.3.2	Divisão de Agrupamentos	53
4.3.3	Fusão de Agrupamentos	54
4.3.4	Detecção dos Veículos	56
4.3.5	Rastreamento dos Veículos	59
4.3.6	Contagem dos Veículos	61
5	RESULTADOS EXPERIMENTAIS	64
5.1	RESULTADOS OBTIDOS	64
5.2	DISCUSSÃO	87
6	CONCLUSÃO	94
7	CONTRIBUIÇÕES DESTE TRABALHO	96
	REFERÊNCIAS	97
	APÊNDICE: DETALHAMENTO DOS DIAGRAMAS DE BLOCOS E FLUXOGRAMAS ATRAVÉS DE ALGORITMOS	101
	ALGORITMO DA VISÃO GERAL DO MÉTODO PROPOSTO	102

ALGORITMO D O AGRUPAMENTO INICIAL	103
ALGORITMO DA DETECÇÃO DE VEÍCULOS	103
ALGORITMO DO RASTREAMENTO DE VEÍCULOS	104

1 INTRODUÇÃO

Neste trabalho é apresentado um método baseado em agrupamento de partículas para realizar a contagem de veículos em vídeos de tráfego, que é o objetivo principal. A contagem é realizada em posições determinadas pelo usuário na pista, denominadas laços virtuais, conforme exemplificado na Figura 1. Como suporte ao algoritmo de contagem e objetivo secundário, desenvolveu-se uma etapa de rastreamento de veículos que utiliza filtro de partículas, baseado na comparação de histogramas de cores.

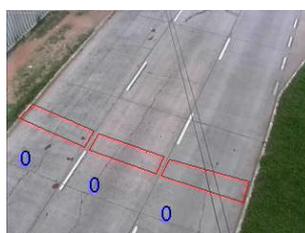


Figura 1: Exemplificação dos laços virtuais.

A contagem de veículos pode ser realizada sem a utilização de rastreamento (PURNAMA *et al.*, 2009), mas optou-se por implementá-lo porque o rastreamento pode fornecer informações adicionais ao longo de toda a pista, incluindo os dados fornecidos pelos sensores indutivos. Apesar da expansão da utilização de radares e vídeos, a tecnologia dominante para obtenção de dados referentes ao tráfego de veículos continua sendo os sensores indutivos, cujas informações fornecidas são a velocidade média, ocupação da via e número de veículos (WILLIAMS e GUIN, 2007).

Várias técnicas simples de processamento de imagens são aplicadas ao longo dos processos descritos. O método utiliza-se da coerência espaço-temporal de partículas pertencentes aos veículos. As partículas utilizadas são referentes às posições de cantos na imagem, e podem ser obtidas através do algoritmo de detecção de cantos de (HARRIS e STEPHENS, 1988) ou através dos mínimos autovalores de (SHI e TOMASI, 1994). Normalmente os métodos de rastreamento iniciam com a obtenção dos objetos de interesse, no caso os veículos. Um dos métodos utilizado é a eliminação do fundo estático da cena. Outro método é a subtração de quadros, onde apenas objetos que se movem ao longo de alguns quadros são identificados (TIAN *et al.*, 2011).

O método proposto utiliza-se do fundo para obter os objetos de interesse, mas apenas nas regiões em que há partículas válidas. As partículas são a primeira forma de identificação dos veículos, eliminando-se partículas persistentes ao longo de muitos quadros na mesma posição. A similaridade das regiões das partículas com o fundo também as eliminam. Com a aplicação desses dois processos espera-se que restem apenas partículas pertencentes a veículos.

Utilizando-se técnicas de clusterização (SEBER, 1984), as partículas são agrupadas de acordo com suas posições espaciais e vetores de movimentos. Dois modos de clusterização são utilizados, um inicial, em que não há qualquer informação sobre os agrupamentos, e um contínuo, em que os centroides dos agrupamentos anteriores são passados para os próximos quadros. Utilizando-se informação do fundo, os agrupamentos obtidos são divididos e/ou fundidos para aprimorar os resultados.

Os veículos são identificados através da morfologia das regiões convexas formadas pelas partículas dos agrupamentos e similaridade com o fundo, levando-se em consideração aspectos como área, taxa de ocupação de partículas e relação entre comprimento e largura dessas regiões. A similaridade com o fundo é analisada no interior das regiões, onde deve ser baixa, e na parte externa das regiões, em que deve ser alta. Os veículos identificados passam então a ser rastreados através de similaridade de histograma de cor das regiões das partículas.

Embora existam trabalhos que realizam o rastreamento de veículos a partir de câmeras instaladas no interior de outros veículos em movimento (SIVARAMAN e TRIVEDI, 2010; JAZAYERI *et al.*, 2011), ou em câmeras fixadas nas laterais das pistas detectando a passagem lateral dos veículos (HUANG e LIAO, 2004; IWASAKI e KUROGI, 2007), ou ainda dedicados ao rastreamento dos faróis durante períodos noturnos (CHEN *et al.*, 2011; O'MALLEY, GLAVIN e JONES, 2011; ZHANG *et al.*, 2012), o trabalho apresentado restringe-se à contagem e rastreamento dos veículos através de imagens de vídeo obtidos de câmeras fixadas acima das vias em questão, em que o movimento ocorre na direção vertical na imagem, não contemplando condições especiais de funcionamento para alterações climáticas, como chuva, neblina e neve, ou de iluminação, como incidência de sombras ou durante a noite.

O algoritmo resultante foi aplicado a vários vídeos, totalizando 81.000 quadros em seis cenas diferentes, incluindo situações desafiadoras como rápidas variações de luminosidade, aglomeração de veículos na pista devido à sinalização e veículos grandes em relação à área de interesse. Ainda assim obtiveram-se resultados satisfatórios, 95% de acerto

na contagem geral, coerente com os valores apresentados pelos trabalhos atuais (TSENG, LIN e SMITH, 2002; KANHERE e BIRCHFIELD, 2008; PURNAMA *et al.*, 2009). Os resultados obtidos também foram comparados com dois métodos da atualidade, (KIM, 2008) e (SÁNCHEZ *et al.*, 2011).

Nas comparações realizadas com os métodos atuais, o método proposto apresentou valores similares de veículos perdidos em relação à (KIM, 2008), 102 e 101 veículos respectivamente, que ficaram menores que os valores apresentados por (SÁNCHEZ *et al.*, 2011), 181 veículos. Em tratando-se de veículos contados mais de uma vez, o método proposto apresentou valores inferiores a ambos os métodos comparativos, 49 veículos contra 167 e 208 veículos para (KIM, 2008) e (SÁNCHEZ *et al.*, 2011) respectivamente. O método apresentado por (KIM, 2008) foi testado através de um aplicativo desenvolvido pelo autor, disponibilizado em sua página pessoal no seguinte endereço: <http://gateway.path.berkeley.edu/~zuwhan/>. Enquanto o método apresentado por (SÁNCHEZ *et al.*, 2011) foi implementado utilizando-se o artigo apresentado como referência.

O presente trabalho justifica-se devido à existência de um número elevado e crescente de veículos nas cidades e suas imediações, que demandam uma capacidade de gerenciamento de tráfego muito grande por parte das autoridades competentes para evitar congestionamentos e planejar modificações e ampliações nas vias. Neste contexto, o processamento de imagens apresenta-se como uma poderosa ferramenta, pois muitas informações acerca do tráfego podem ser automaticamente extraídas através das câmeras de monitoramento atualmente instaladas em diversas localidades e de novas câmeras que surgem constantemente devido à queda nos custos de instalação e manutenção (TIAN *et al.*, 2011).

Esses dados, como a contagem de veículos, podem ser utilizados para auxiliar nas ações de planejamento de ampliação de capacidade das vias, bem como servir de entrada em sistemas de sinalização automáticos (HOUNSELL *et al.*, 2009). Como é humanamente impossível analisar diversos monitores ao mesmo tempo, e economicamente inviável manter grandes equipes para tal atividade, informações pontuais como acidentes, veículos parados nas vias ou comportamentos suspeitos dos veículos, também podem ser obtidos automaticamente, através do rastreamento, e disparar alarmes em centrais de monitoramento para solicitar averiguações (WILLIAMS e GUIN, 2007; HOUNSELL *et al.*, 2009).

2 REVISÃO BIBLIOGRÁFICA

Esta seção apresenta, resumidamente, os trabalhos relacionados com o rastreamento e contagem de veículos utilizando imagens de vídeo que são referências para o estado da arte. Várias abordagens podem ser adotadas para a contagem de veículos, que é o objetivo principal do trabalho, portanto, apresentam-se trabalhos em que a contagem é realizada simultaneamente com o rastreamento, outros em que utiliza-se o resultado do rastreamento em regiões definidas para a realização da contagem, e trabalhos que realizam a contagem de veículos sem a necessidade de rastreamento.

Também são abordados alguns artigos que realizam o rastreamento ou contagem de veículos em situações diferentes das apresentadas neste trabalho, como o rastreamento a partir de câmeras no interior de outros veículos, rastreamento noturno utilizando os faróis dos veículos e contagem de veículos através de câmeras nas laterais das pistas. Apesar destes artigos não terem relação direta com o trabalho realizado, muitas técnicas utilizadas para suas implementações podem ser aproveitadas e sua apresentação ajuda a situar melhor o leitor.

2.1 VISÃO GERAL DO MONITORAMENTO DE TRÁFEGO UTILIZANDO PROCESSAMENTO DE VÍDEO

O gerenciamento de tráfego urbano deve possuir como entrada dados compreensíveis e precisos, conforme (HOUNSELL *et al.*, 2009). Considerando tais aspectos, o circuito fechado de televisão é um dos meios de aquisição de dados que fornecem informações mais ricas, podendo ser empregados em grandes intersecções e túneis.

A contagem de veículos é uma tarefa normalmente executada por sensores indutivos, que apesar de terem uma alta sensibilidade, são intrusivos à pista e conseqüentemente, possuem um alto custo de manutenção. Além disso, a utilização de imagens de vídeo possibilita a obtenção de outros parâmetros, como a classificação dos veículos (BUCH, VELASTIN e ORWELL, 2011).

Segundo (TIAN *et al.*, 2011), a detecção dos objetos de interesse é a primeira etapa nos métodos de monitoramento de tráfego que utilizam processamento de vídeo. Esta referência considera que os veículos estão em movimento e presta atenção especial à identificação e remoção de sombras. O trabalho apresentado por (KASTRINAKI,

ZERVAKIS e KALAITZAKIS, 2003) considera a detecção das pistas como ponto de partida para o monitoramento de tráfego e também aborda obstáculos e veículos parados na pista.

Já (BUCH, VELASTIN e ORWELL, 2011) considera duas abordagens para a análise dos vídeos. Uma de cima para baixo, em que a partir dos objetos de interesse os veículos são identificados, utilizando-se componentes conectados, por exemplo, classificados e rastreados. E outra de baixo para cima, em que partes do objeto são identificadas, classificadas, utilizando-se treinamento com modelos de classes, agrupadas e rastreadas.

O trabalho apresentado por (WANG, XIAO e GU, 2008) considera a seleção da Região de Interesse (*Region of Interest*, ROI) de várias maneiras, como métodos de subtração de quadros, métodos de atualização de fundo e métodos de laços virtuais. Destes, apenas o método de laços virtuais utiliza entrada manual do usuário, os outros são automáticos, mas analisam todo o quadro em questão.

Alguns métodos detectam os veículos através de suas sombras. O método proposto por (XU *et al.*, 2008) detecta as sombras sob os veículos e testa se as formas resultantes são retangulares, utilizando as posições de câmeras convencionais, ou seja, acima das vias em questão. Já o método proposto por (IWASAKI e KUROGI, 2007) também utiliza-se das sombras dos veículos para identificá-los, mas utiliza uma posição de câmera lateral à pista. Os dois métodos iniciam com a identificação dos objetos de interesse.

Para condições de luminosidade desfavoráveis, como a noite, alguns métodos utilizam-se da identificação e rastreamento dos faróis dos veículos (CHEN *et al.*, 2011; ZHANG *et al.*, 2012). Os faróis são a parte de maior luminosidade das imagens noturnas e normalmente aparecem em pares, tais aspectos são levados em conta por estes métodos.

A contagem de veículos pode ser realizada simultaneamente com o rastreamento. Nesse caso o valor da contagem é mantido junto com a caixa de identificação dos veículos (BARDET e CHATEAU, 2008; KANHERE e BIRCHFIELD, 2008). Utilizando-se do rastreamento para ativar contagens em locais específicos da pista (TSENG, LIN e SMITH, 2002). Ou sem a necessidade de rastreamento, em que a passagem dos veículos sobre laços virtuais incrementam seus contadores (PURNAMA *et al.*, 2009), inclusive à noite (MOSSI *et al.*, 2011).

Todos os métodos apresentados até o momento utilizam câmeras externas aos veículos, mas existem alguns métodos em que a câmera está localizada no interior do veículo. Para (KASTRINAKI, ZERVAKIS e KALAITZAKIS, 2003) esta abordagem visa identificar a posição do veículo em relação à pista e a posição dos outros veículos. Para (HOUNSELL *et*

al., 2009) câmeras instaladas no interior de veículos podem fornecer informações relativas às condições de estradas e tempos de viagens.

2.2 RASTREAMENTO E CONTAGEM DE VEÍCULOS UTILIZANDO-SE CÂMERAS ESTÁTICAS

Esta subseção apresenta algumas abordagens utilizadas por autores na implementação de métodos para rastreamento e contagem de veículos em vídeos obtidos de câmeras estáticas, externas aos veículos. Divide-se ainda em: (a) câmeras acima das pistas, (b) câmeras laterais às pistas, (c) rastreamento e contagem noturna.

2.2.1 Câmeras Acima das Pistas

São as situações mais presentes, inclusive as utilizadas pelo método proposto. As câmeras são localizadas acima das vias em questão, podendo ter leves inclinações. Normalmente os veículos deslocam-se na direção vertical e sentido de cima para baixo, conforme Figura 2, mas isto não exclui outras possibilidades.



Figura 2: Exemplo de câmera acima da pista.

Um dos vários métodos que utilizam partículas para realizar o rastreamento de veículos é o proposto por (KANHERE e BIRCHFIELD, 2008), exemplificado na Figura 3. Nesse método as partículas utilizadas são cantos, ou seja, o vértice ou o cruzamento presente entre duas bordas. Uma projeção tridimensional é obtida automaticamente, utilizando-se da entrada manual de três linhas retas pelo usuário, duas ao longo das faixas da estrada e uma perpendicular à direção do movimento dos veículos, conforme Figura 3 (a) e (b). Para eliminar regiões de sombra o método elimina as partículas muito próximas das bordas, na máscara dos objetos de interesse, conforme Figura 3 (c), em que os círculos pretos pertencem aos veículos, os círculos brancos pertencem ao fundo, e os quadrados brancos pertencem às bordas das sombras. As partículas são rastreadas utilizando-se o método de (SHI e TOMASI, 1994) e são separadas em estáveis e instáveis. As partículas estáveis são agrupadas no plano da estrada de acordo com suas posições na direção do movimento. As partículas instáveis são

agregadas a esses agrupamentos utilizando uma combinação de projeção de alinhamento e coerência de movimento. Após, os agrupamentos que não são identificados como veículos são eliminados. Os veículos são identificados quando os agrupamentos formam caixas, considerando as três dimensões espaciais, resultando na Figura 3 (d). Esse método apresenta resultados entre 91% e 97% de acerto na segmentação e rastreamento de veículos, para 11 cenas testadas.

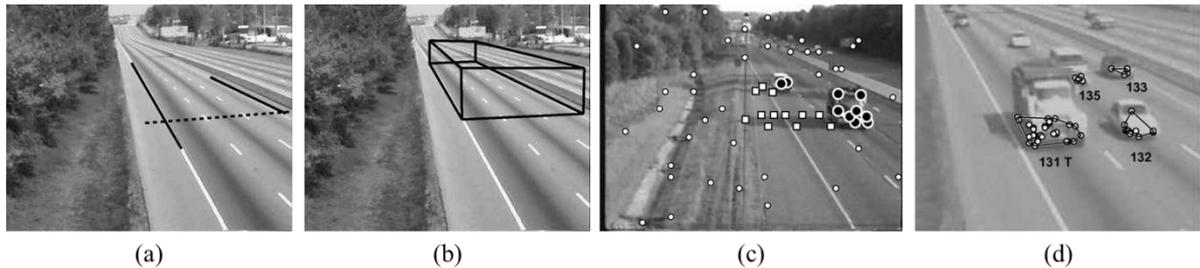


Figura 3: Exemplificação do método proposto por (KANHERE e BIRCHFIELD, 2008). (a) Linhas fornecidas pelo usuário; (b) Projeção tridimensional obtida a partir de (a); (c) Diferentes tipos de partículas; (d) Agrupamentos obtidos.

Outro método que utiliza partículas para rastrear os veículos é o proposto por (KIM, 2008), sendo um dos métodos utilizados nas comparações de resultados, através de um aplicativo disponibilizado no seguinte endereço: <http://gateway.path.berkeley.edu/~zuwhan/>. Inicialmente o método obtém as regiões de interesse através da aplicação de técnicas combinadas, como filtro de mediana temporal, filtro de Kalman e correção de iluminação. As partículas são obtidas utilizando-se autovalores das somas locais das derivadas (FORSTNER e GULCH, 1987), apenas nas regiões de interesse e rastreadas utilizando-se correlação entre janelas centradas nessas partículas e no quadro adjacente. Também é efetuada a correlação com a imagem de fundo e se alguma partícula tiver uma correlação alta é eliminada. Uma variação do algoritmo de maximização da expectativa (DEMPSTER, LAIRD e RUBIN, 1977) é aplicada para realizar pequenos agrupamentos de partículas, considerando posição, deslocamento, trajetória, agrupamento anterior e conjunto conectado de *pixels* pertencente, resultando na Figura 4 (a). Após, esses pequenos agrupamentos são agrupados em nível de objeto, tentando fixar uma forma elíptica, cujo tamanho mínimo e máximo são determinados por critérios tridimensionais para eliminar efeitos da perspectiva, resultando na Figura 4 (b). Esse método necessita de parâmetros de calibração da câmera, ou uma calibração inicial utilizando-se linhas retas definidas pelo usuário. Nos vídeos testados por esta dissertação este método apresentou uma diferença de 6% contados a mais que o valor real.



Figura 4: (a) Agrupamentos e suas trajetórias em nível de partículas e (b) agrupamentos e trajetórias em nível de objetos, obtidas através do método proposto por (KIM, 2008).

Sombras são utilizadas por (XU *et al.*, 2008) para detectar os veículos. Os autores consideram que não importa o nível de luminosidade da imagem, a parte inferior dos veículos será sempre mais escura que as demais, e utilizam tal informação para detectar os veículos. As regiões de sombra continuam tendo as mesmas cores e texturas que o fundo, apenas com menos luminosidade. Então, após obter os objetos de interesse, através da subtração do quadro processado do fundo, a região de sombra é obtida. Dependendo das geometrias das regiões conectadas e das regiões de sombra os veículos são definidos.

O método apresentado por (PURNAMA *et al.*, 2009) não realiza rastreamento, apenas contagem. O sistema é muito simples e pode ser resumido em algumas etapas: inicialização, com a obtenção do fundo e definição da ROI, que é o local em que os veículos serão processados; obtenção da diferença entre o quadro processado e o fundo; aplicação de limiares até eliminar pequenas regiões; utilização de erosão e dilatação para separar as regiões (veículos); obtenção dos objetos conectados (veículos) e classificação, considerando-se suas áreas; atualização da contagem de veículos. Obviamente o método não trata oclusões e é extremamente dependente da qualidade do modelo de fundo utilizado. A Figura 5 apresenta os objetos conectados (a) e os objetos identificados (b) com este método. Os autores apresentam taxas de acerto na contagem de veículos que variam entre 80% e 95% quando as condições de luminosidade são boas (sem sombras), que caem para o intervalo entre 27% e 50% quando sombras estão presentes, conforme Figura 5 (b). Também são apresentados valores para contagem durante a noite, com taxas de acerto variando entre 60% e 90% para os vídeos testados pelos autores.



Figura 5: (a) Objetos conectados e (b) objetos identificados pelo método apresentado por (PURNAMA *et al.*, 2009).

O trabalho de (MEI e LING, 2011) trata o problema do rastreamento de veículos com uma abordagem baseada em esparsividade, obtida de modelos através da solução de um problema de mínimos quadrados regularizados em l_1 , sendo o candidato com o menor erro projetado assumido como alvo rastreado. O alvo pode ser inicializado através da introdução de suas coordenadas, em três pontos iniciais, ou através de um banco de modelos. O rastreamento é realizado através de inferência Bayesiana, utilizando a informação de deslocamento das partículas geradas no quadro anterior. O deslocamento é obtido através da solução de transformações afim entre dois quadros consecutivos. Ao longo do processo os modelos vão sendo atualizados e a aplicação de restrições, obtidas de modificações dos modelos, reduzem erros provocados por ruídos e variações de luminosidade. Este método pode ser aplicado a diversas posições de câmera, e também pode ser utilizado em câmeras embarcadas em outros veículos, conforme rastreamento apresentado na Figura 6.



Figura 6: Exemplo de rastreamento utilizando-se câmera embarcada, por (MEI e LING, 2011).

Em (SÁNCHEZ *et al.*, 2011), os objetos de interesse são obtidos por meio da subtração do fundo dos quadros em questão. O fundo pode ser um modelo fixo ou obtido adaptativamente, sendo cada *pixel*, que é o menor elemento da imagem, incrementado ou decrementado de acordo com a relação entre o *pixel* do quadro processado e o *pixel* do modelo atual de fundo. As regiões das pistas são delimitadas e utilizadas para resolver conflitos de identificação. São identificadas as regiões de entrada e saída de veículos, bem como as divisões entre pistas. Os veículos são rastreados através da sobreposição de regiões de *pixels* dos objetos de interesse conectadas. Este é um dos métodos utilizados nas comparações de resultados e foi implementado utilizando-se o artigo apresentado como

referência, apresentado uma diferença na contagem geral de 2,5% veículos contados a mais que o valor real. Para as sequências testadas pelos autores as taxas de acerto variam entre 85% e 100%.

Utilizando-se filtro de partículas tem-se o trabalho de (SCHARCANSKI *et al.*, 2011), que utiliza dois métodos de geração de partículas. Um quando não há oclusão, e partículas são geradas utilizando-se uma função densidade de probabilidade normal bivariada, e outro em que há oclusão, sendo utilizada uma função densidade de probabilidade normal bivariada e Rayleigh conjuntamente, levando-se em consideração o movimento principal do veículo rastreado. O veículo rastreado pode ser definido manualmente ou por algum método de detecção. Histogramas de cor locais são utilizados para rastrear os veículos, sendo atualizados quadro a quadro. Pesos maiores são atribuídos às partículas mais significativas e as partículas menos significativas vão sendo eliminadas ao longo do processo.

Em (FERIS *et al.*, 2012), um banco de dados com imagens de veículos é gerado de forma semiautomática. Os objetos de interesse em uma determinada ROI são obtidos através da subtração dos quadros do fundo. Objetos conectados que satisfizerem algumas características pré-definidas como comprimento, largura, relação comprimento / largura e direção de movimento são extraídos como sendo veículos. Os veículos adquiridos são classificados, de acordo com sua direção de deslocamento, em 12 categorias, conforme exemplificado na Figura 7. Oclusões sintéticas são geradas nesse banco de imagens para tratar tais problemas. Milhões de características como cor, textura e bordas são extraídas desse banco de imagens durante um período de treinamento. A identificação e rastreamento dos veículos são realizados baseados nos resultados desse treinamento, para uma posição de câmera específica. Qualquer posição de câmera pode ser usada com este método. A taxa média de acerto na classificação de veículos utilizando-se esse método é de 90,7% conforme os autores. Vídeos de exemplo podem ser acessados no seguinte endereço eletrônico: <http://rogerioferis.com/IEEEMultimedia/>.



Figura 7: Exemplo das categorias de deslocamento dos veículos apresentado por (FERIS *et al.*, 2012).

O método proposto por (WU e JUANG, 2012) inicia com uma normalização dos quadros através do processamento dos histogramas. Os objetos de interesse são obtidos através da aplicação de funções lógicas que comparam diferenças de *pixels* com limiares específicos para cada cor. Nesse sistema as cores cinza, do asfalto, e branco ou amarelo, das faixas, possuem valores menores. Os objetos de interesse são fundidos de acordo com a sobreposição de caixas externas a esses objetos. Atributos como comprimento, largura, relação comprimento / largura e densidade de objetos são testados para identificação de veículos. Esses atributos também determinam se alguma fusão de objetos acabou fundindo veículos diferentes, sendo então uma quebra de objetos efetuada, através da análise do histograma da imagem dos objetos em movimento. Na direção do movimento do veículo fundido incorretamente deve haver um espaçamento no histograma, conforme apresentado na Figura 8, esse espaçamento é a linha que dividirá o veículo em dois. Os objetos são rastreados através da correlação das informações comprimento, largura e deslocamento das caixas que envolvem os objetos identificados. Os autores apresentam resultados na contagem de veículos com taxas de acerto que variam entre 99,4% para dias ensolarados e 92% para dias chuvosos.

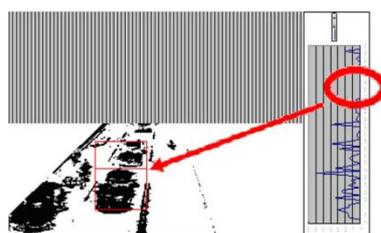


Figura 8: Utilização do espaçamento no histograma para dividir dois veículos identificados por (WU e JUANG, 2012).

2.2.2 Câmeras Laterais às Pistas

Alguns métodos consideram o rastreamento e contagem de veículos através de câmeras posicionadas lateralmente às pistas, mas dependendo da altura de instalação da câmera e do número de pistas monitoradas, esta abordagem pode tornar-se inviável tecnicamente. Em outras situações, essa abordagem pode facilitar a implementação de métodos de rastreamento, pois a perspectiva é reduzida em relação às câmeras posicionadas acima e a frente das pistas, em que as imagens dos veículos alteram-se drasticamente conforme aproximam-se ou afastam-se das câmeras. Alguns métodos apresentados na seção anterior também podem ser utilizados em câmeras laterais (MEI e LING, 2011; FERIS *et al.*, 2012), conforme Figura 9.

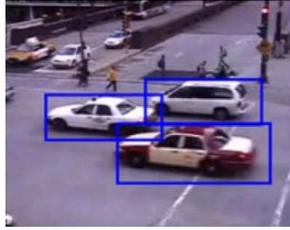


Figura 9: Exemplo de câmera lateral utilizada em (FERIS *et al.*, 2012).

Em (HUANG e LIAO, 2004), inicialmente são obtidos os objetos de interesse, através de subtração do fundo e aplicação de filtros para remover ruídos. As regiões dos objetos de interesse são então rotuladas e rastreadas de acordo com suas formas. Quando duas regiões são unidas, significa que uma oclusão está ocorrendo, conforme a Figura 10, e a região do quadro anterior não será mais localizada no quadro atual. Outra forma de detectar oclusões é através da variância dos vetores de movimento dos *pixels* das regiões. Quando a variância é elevada provavelmente está ocorrendo uma oclusão. Os vetores de movimento são obtidos através de comparação de blocos. Para separar os veículos em situação de oclusão são utilizados os vetores de movimento. Conjuntos de vetores de movimento similares são agrupados. Ao final do processo os veículos ainda são classificados de acordo com a silhueta de suas regiões de interesse em alguns tipos, como carros, vans e caminhões, obtendo-se taxas de acerto entre 66% e 98%, com média em 91,4%, para estas classificações.



Figura 10: Exemplo de oclusão e região de separação entre dois veículos para (HUANG e LIAO, 2004).

O método apresentado por (IWASAKI e KUROGI, 2007) utiliza as sombras sob os veículos para identificá-los e as distâncias entre as rodas traseiras e frontais para determinar o tamanho dos veículos. É capaz ainda de determinar em qual pista estão os veículos, desde que suas distâncias sejam conhecidas. Inicialmente os objetos de interesse, contendo os veículos e suas sombras, são obtidos através da subtração do fundo, obtido através de um método adaptativo. Na imagem dos objetos de interesse é realizada uma projeção vertical, alterando os *pixels* da base da região de identificação de preto para branco quando pelo menos um *pixel*

da respectiva coluna é branco. Quando pelo menos oito *pixels* de altura ao longo de quinze *pixels* de comprimento são brancos na base da região de identificação, um retângulo é gerado envolvendo tais *pixels*. Uma segunda projeção, mas na horizontal e considerando os *pixels* pretos, identifica o limite superior da sombra do veículo, quando existe um aumento abrupto no número de *pixels* pretos nessa projeção. As rodas apresentam-se como componentes conectados de *pixels* pretos no interior das regiões dos objetos de interesse, a sombra dos veículos, conforme a Figura 11. Esse método apresenta taxas de acerto para a detecção de veículos variando entre 91,1% e 98,8% em várias condições de imagens, como dia claro, chuvoso e períodos noturnos.

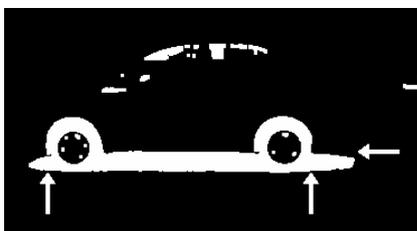


Figura 11: Rodas apresentadas como objetos conectados (em preto) dentro da região do objeto de interesse (sombra do veículo, em branco).

Em (MORRIS e TRIVEDI, 2008), os objetos de interesse são obtidos através da subtração do fundo, em que cada *pixel* é modelado como uma gaussiana com média μ (média da intensidade ao longo do tempo) e desvio padrão σ (desvio da intensidade ao longo do tempo). Operações morfológicas são efetuadas para melhorar os resultados na obtenção dos objetos de interesse. Cada componente conectado pertencente aos objetos de interesse (veículo) é rotulado e parâmetros como área, centroide e perímetro são obtidos. O rastreamento dos veículos é realizado utilizando-se filtro de Kalman e restrições baseadas em aparência, considerando-se que a aparência de um veículo não pode mudar drasticamente ao longo dos quadros. Um banco de imagens de treinamento é utilizado para classificar os veículos em vários tipos. Os veículos não são separados durante as oclusões, embora saiba-se que tal fato está ocorrendo, conforme Figura 12. Se a oclusão for de curta duração os rótulos dos veículos envolvidos podem ser recuperados, caso contrário são perdidos e novos rastreamentos são efetuados. Os resultados apresentados pelos autores correspondem ao fluxo de veículos em períodos de 30 segundos, obtendo-se erros menores que dois veículos por intervalo, cujo fluxo real de veículos varia de 5 a 13 veículos por período.



Figura 12: Exemplo de oclusão de veículos em (MORRIS e TRIVEDI, 2008).

2.2.3 Rastreamento e Contagem Noturna

Os métodos de rastreamento e contagem de veículos durante períodos noturnos concentram-se principalmente na identificação dos faróis dos veículos. Obviamente esses métodos falham caso algum veículo passe com os faróis desligados. Um dos principais desafios dessas abordagens é tratar as reflexões causadas pelos faróis nas pistas.

Em (CHEN *et al.*, 2011), inicialmente os objetos mais brilhantes são extraídos através da aplicação de múltiplos limiares baseados em histogramas. Após, uma análise espacial é efetuada sobre os componentes conectados resultantes, considerando que os faróis têm forma aproximadamente circular e área dentro de determinados valores, de acordo com a posição em relação às linhas de fuga. As relações de distâncias entre os faróis, tanto horizontalmente quanto verticalmente também são consideradas. Os objetos remanescentes são agrupados e rastreados de acordo com seu movimento. A Figura 13 exemplifica o rastreamento dos faróis. O artigo apresenta taxas de acerto da ordem de 97% para o rastreamento de faróis nos vídeos testados pelos autores.



Figura 13: Exemplo de rastreamento de faróis realizado por (CHEN *et al.*, 2011).

Outro método que considera a utilização dos faróis é o proposto por (MOSSI *et al.*, 2011). Diferentemente dos outros métodos, este trabalho apenas realiza a contagem de veículos, quando faróis identificados passam sobre laços virtuais. Os faróis são identificados através da aplicação da operação morfológica *TopHat*, que é a abertura da imagem seguido da subtração dessa imagem aberta da imagem original, utilizando elementos estruturantes elípticos em quatro direções. Como o reflexo gerado pelos faróis no asfalto geralmente

possuem formas alongadas na direção de movimento do veículo, a aplicação dessa operação elimina tais formas alongadas, mantendo apenas faróis, que tendem a ser arredondados. Para ativar a contagem, a detecção positiva de faróis deve se suceder durante três quadros. Quando após algum farol tiver sido detectado, dois quadros consecutivos passarem sem identificação positiva, a contagem é finalizada e um período de inibição é ativado, inversamente proporcional à velocidade do veículo, que pode ser obtida através da quantidade de quadros em que a detecção dos faróis ficou ativada. Segundo os autores, em sequências de vídeo com a passagem de 653 veículos houve 61 contagens incorretas.

Em (ZHANG *et al.*, 2012) dois mapas são utilizados para corretamente identificar os faróis dos veículos, um mapa de intensidade de reflexão e um mapa de reflexão suprimida. Considerando-se a Figura 14 (a), que identifica reflexões causadas pelos faróis, no mapa de intensidade de reflexão as partes da imagem que contém estas reflexões são identificadas através da aplicação de filtros baseados na lei de atenuação exponencial de Bourguer, enquanto a luminosidade proveniente dos faróis é atenuada, conforme a Figura 14 (b). Já no mapa de reflexão suprimida, um filtro Laplaciano de Gaussiana é utilizado para suprimir as reflexões e realçar os faróis, conforme Figura 14 (c). Os faróis são identificados utilizando-se campos aleatórios de Markov. Para o pareamento e rastreamento dos faróis, são obtidos os pontos de fuga da imagem, utilizando-se um rastreamento inicial dos faróis através da área e distância dos centroides durante alguns quadros. Após essa inicialização os faróis passam a ser rastreados utilizando-se um algoritmo que considera as distâncias entre centroides, mudanças de área e desvios dos pontos de fuga. Os resultados para identificação e rastreamento de veículos alcançam taxas de 95,2% e 88,2% respectivamente, para um total de 372 veículos, segundo os autores.

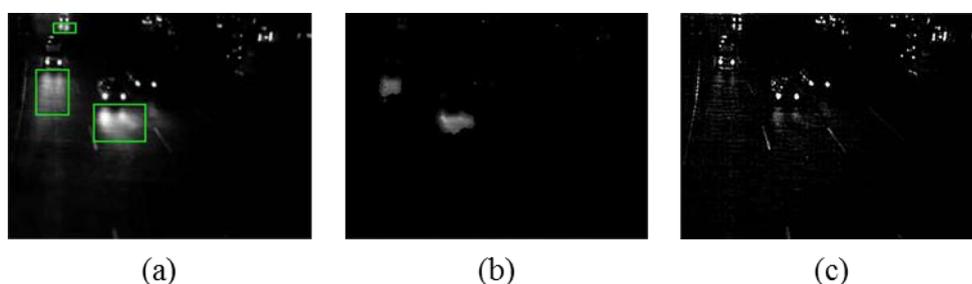


Figura 14: (a) Reflexos causados pelos faróis; (b) Mapa de intensidade de reflexão; (c) Mapa de reflexão suprimida. De (ZHANG *et al.*, 2012).

2.3 RASTREAMENTO DE VEÍCULOS UTILIZANDO-SE CÂMERAS MÓVEIS

Esta seção apresenta alguns métodos de rastreamento baseados em câmeras móveis, ou seja, localizadas no interior de veículos em movimento, conforme a Figura 15. Algumas abordagens sofrem alterações, como a utilização de fundo. Nessas situações o fundo está em mudança constante. Outras técnicas continuam sendo aplicadas, como filtros de partículas e métodos de reconhecimento baseados em modelos. Uma das principais justificativas para efetuar-se rastreamento de veículos a partir de outros veículos é a implantação de sistemas de assistência aos motoristas (SUN, BEBIS e MILLER, 2006).



Figura 15: Exemplo de rastreamento através de câmeras no interior de veículos de (SIVARAMAN e TRIVEDI, 2010).

Em (SIVARAMAN e TRIVEDI, 2010), um método de aprendizado ativo é utilizado para detectar os veículos. Inicialmente um treinamento é realizado utilizando imagens referentes a veículos e não-veículos. Após, utilizando outro conjunto de imagens o algoritmo apresenta seus resultados, que são confirmados por um usuário e utilizados para novos treinamentos. Um classificador *AdaBoost* (FREUND e SCHAPIRE, 1999) é utilizado em características retangulares obtidas das imagens. O rastreamento é realizado utilizando-se filtro de partículas, através da propagação de novas partículas baseadas nas condições de partículas anteriores. O rastreamento também é confirmado através da detecção dos veículos pelo classificador. Os resultados apresentados apontam o correto reconhecimento de veículos variando entre 91,7% e 99,8%.

O método proposto por (JAZAYERI *et al.*, 2011) é baseado em movimento e não em formas. Ele utiliza a diferença do fluxo ótico (movimento da imagem) existente entre o fundo, que parte do centro em direção às bordas da imagem, e os veículos, que tendem a ter direções diferentes das do fundo, conforme Figura 16 (a). As características utilizadas para obter o fluxo ótico são cantos, bordas horizontais e pontos luminosos. Um mapa espaço-temporal é gerado através da projeção unidimensional das informações de intensidade / cor das partículas de cada quadro, com o objetivo de condensar as informações e melhorar o desempenho do

processo de análise da trajetória dos veículos, conforme Figura 16 (b). Essa projeção ignora as características verticais obtidas, atendo-se apenas nas horizontais, ou movimento horizontal ao longo dos quadros. O rastreamento das trajetórias é realizado através de diferenciação horizontal desse mapa espaço-temporal. As projeções cujas trajetórias partem da lateral deste mapa em direção à parte inferior, de forma hiperbólica, são consideradas partes do fundo. As projeções que mantêm a trajetória aproximadamente no centro do mapa são os veículos rastreados, conforme exemplificado na Figura 16 (c). Essa distinção de trajetórias é realizada utilizando-se cadeias ocultas de Markov. O método apresenta uma assertiva na identificação e rastreamento de veículos de 86,6%.

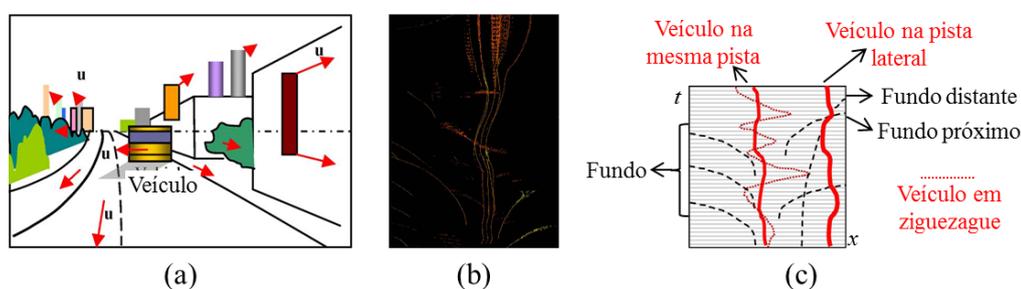


Figura 16: (a) Diferença no fluxo óptico entre fundo e veículo; (b) Exemplo de mapa espaço-temporal; (c) Exemplificação da análise de um mapa espaço-temporal. De (JAZAYERI *et al.*, 2011).

Em (O'MALLEY, GLAVIN e JONES, 2011) a identificação e o rastreamento são realizados durante a noite, através da identificação das lanternas traseiras dos veículos, utilizando-se câmeras com baixa exposição, o que reduz a introdução de artefatos nas imagens gerados por reflexão. As lanternas são obtidas através de crescimento de regiões, que iniciam em sementes escolhidas a partir da aplicação de um valor de limiar alto na imagem, resultando apenas nos pontos mais brilhantes. Limiares menores vão sendo aplicados sequencialmente durante o crescimento das regiões, até que os *pixels* atinjam bordas obtidas na imagem original através de um detector de bordas de Sobel. O fechamento morfológico das regiões é utilizado para eliminar ruídos e reconectar regiões próximas separadas. As sementes localizadas na terça parte superior da imagem são eliminadas para evitar a inclusão das lâmpadas dos postes no processo. Para a detecção dos veículos, uma abordagem baseada em pareamento de lanternas é utilizada, sendo que a relação entre áreas mínimas e máximas de candidatos, e o ângulo máximo do centroide entre esses dois candidatos são avaliados. Uma correção da perspectiva é aplicada, utilizando-se uma matriz de transformação projetiva

que transforma as regiões quadriláteras que enquadram os pares de lanternas em regiões retangulares. Finaliza-se o pareamento através de correlação cruzada utilizando-se a imagem de uma das lanternas, espelhada horizontalmente, como modelo. O rastreamento dos veículos é realizado através de filtro de Kalman, utilizando-se dados dos retângulos dos pares de lanternas identificadas, como posição e tamanho. Segundo os autores, em testes realizados utilizando-se aproximadamente 30.000 frames, a taxa de acerto na identificação de veículos é de 92,9%.

O método proposto por (CHAN *et al.*, 2012) utiliza filtro de partículas e clusterização para rastrear veículos com câmeras embarcadas. O rastreamento é realizado utilizando-se características como posição, tamanho e deslocamento das amostras de partículas. Os veículos são identificados através de: análise de suas bordas verticais, obtidas através do detector de bordas de Sobel; das sombras sob os veículos, obtidas através de restrições horizontais das bordas e da aplicação de limiares que mantêm apenas as bordas mais escuras; dos faróis, obtidos por um método descrito por (WANG, HUANG e FU, 2005) em que apenas faróis e lâmpadas com luz branca são identificados; e da simetria horizontal existente na traseira dos veículos. Essas características são fundidas, dando origem a um único descritor para as partículas, capaz de identificar veículos em várias situações de luminosidade e condições climáticas. Inicialmente as partículas geradas são agrupadas utilizando o método descrito em (THEODORIDIS e KOUTROUMBAS, 2003), para após serem rastreadas utilizando-se filtro de Kalman. As taxas de acerto na detecção de veículos variam entre 72,5% para condições de baixa luminosidade, como túneis, até 98,6% para ambientes ensolarados, com uma taxa média de 84,9% de veículos corretamente detectados.

3 FUNDAMENTAÇÃO TEÓRICA

Este capítulo apresenta as definições teóricas das técnicas utilizadas no desenvolvimento do método proposto, como obtenção do fundo, obtenção das partículas, vetores de movimento, clusterização, histogramas de cor, entre outras. São apresentadas as equações envolvidas e fontes de referência.

3.1 REPRESENTAÇÃO DE IMAGENS E VÍDEO

Imagens digitais podem ser definidas como funções de duas dimensões, $f(x, y)$, onde x e y são coordenadas espaciais da intensidade, ou escala de cinza, discretas de uma imagem. Também podem ser representadas como matrizes de números reais de M linhas e N colunas, com a origem $(x, y) = (0,0)$, em que cada elemento da matriz é um *pixel* da imagem, como exemplificado a seguir na Eq. (1) (GONZALEZ e WOODS, 2002).

$$f(x, y) = \begin{bmatrix} f(0,0) & f(0,1) & \dots & f(0, N - 1) \\ f(1,0) & f(1,1) & \dots & f(1, N - 1) \\ \vdots & \vdots & \ddots & \vdots \\ f(M - 1,0) & f(M - 1,1) & \dots & f(M - 1, N - 1) \end{bmatrix} \quad (1)$$

A quantidade de tons de cinza que um *pixel* de uma imagem pode assumir depende do número de *bits* utilizados para sua representação, através da função $L = 2^k$, onde L é o número de níveis de cinza e k é o número de *bits* utilizados na representação do *pixel* (GONZALEZ e WOODS, 2002). Utilizando-se 8 *bits*, que é o caso do presente trabalho, um *pixel* pode assumir um entre 256 níveis de cinza. A quantidade total de *bits*, b , para armazenar uma imagem em tons de cinza pode ser obtida através da função $b = M \times N \times k$. As imagens utilizadas no presente trabalho apresentam $M = 240$ e $N = 320$, logo $b = 614,4$ kB.

As imagens coloridas podem ser representadas utilizando-se vários espaços de cores, sendo adotado no presente trabalho a representação baseada no sistema de cores Vermelho, Verde e Azul (RGB, *Red, Green and Blue*). Esse sistema é baseado na composição de três imagens, uma para cada cor primária. O número de *bits* utilizado para representar cada *pixel* no cubo RGB é chamado de profundidade de *pixel* (GONZALEZ e WOODS, 2002). Neste trabalho utiliza-se imagens coloridas compostas por 24 *bits* por *pixel*, ou 8 *bits* por canal de cor RGB, o que corresponde a 16,7 milhões de cores. Frequentemente utiliza-se um cubo de

cores normalizado para representar o espaço de cores RGB, conforme a Figura 17. A linha pontilhada que vai do preto ao branco possui quantidades iguais de cada cor primária, resultando nos tons de cinza.

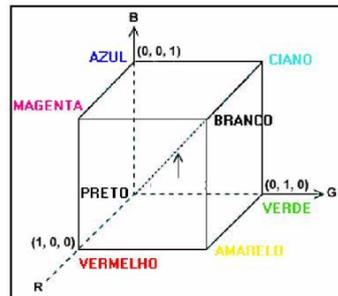


Figura 17: Cubo RGB.

A conversão de imagens coloridas para tons de cinza pode ser realizada através da obtenção da média dos três canais de cores para cada pixel, $I = (R + G + B)/3$, que é a conversão utilizada do espaço de cores RGB para o espaço de cores que representa a Cor, a Saturação e o Brilho (HSI, *Hue, Saturation and Intensity*), selecionando-se apenas o brilho. Não é possível obter as imagens coloridas originais a partir de imagens obtidas ou convertidas em tons de cinza sem informações adicionais, apenas acrescentar pseudo-cores (GONZALEZ e WOODS, 2002).

Vídeos nada mais são do que sequências de imagens ao longo do tempo, coloridas ou em tons de cinza. Cada imagem é chamada de quadro e a quantidade de quadros apresentados por segundo é conhecida como frequência de quadros. Neste trabalho a frequência de quadros não é um parâmetro importante, podendo variar livremente, mas os vídeos utilizados apresentam 25 ou 30 quadros por segundo.

3.1.1 Métricas das Imagens Digitais

Distâncias entre pontos da imagem são obtidas através da aplicação de algumas métricas definidas. Considerando-se os *pixels* p , q e r , qualquer função D que satisfaça as condições seguintes é considerada uma métrica e pode ser utilizada para determinar distâncias entre estes *pixels* (SONKA, HLAVAC e BOYLE, 2007).

$$D(p, q) \geq 0, \text{ para } (D(p, q) = 0, \text{ se e somente se } p = q) \quad \text{Identidade}$$

$$D(p, q) = D(q, p) \quad \text{Simetria}$$

$$D(p, r) \leq D(p, q) + D(q, r)$$

Desigualdade do triângulo

A distância Euclidiana pode ser usada para obter a distância entre a posição de dois *pixels*, conforme a Eq. (2). Outra distância que pode ser utilizada é a *city block*, conforme a Eq. (3).

$$D_e(p, q) = \sqrt{(x - s)^2 + (y - t)^2} \quad (2)$$

$$D_{cb}(p, q) = |x - s| + |y - t| \quad (3)$$

sendo $D_e(p, q)$ a distância Euclidiana e $D_{cb}(p, q)$ a distância *city block* entre os *pixels* p e q nas respectivas posições espaciais (x, y) e (s, t) .

3.1.2 Comparação de Histogramas

O histograma de uma imagem em tons de cinza no intervalo $[0, L - 1]$ é uma função discreta $h(\gamma_k) = n_k$, onde γ_k é o k -ésimo nível de cinza e n_k é o número de *pixels* da imagem que possui o nível de cinza γ_k . Para normalizar o histograma, basta dividir n_k pelo número total de *pixels* da imagem, denotado por n , obtendo-se $p(\gamma_k) = n_k/n$. O histograma normalizado, $p(\gamma_k)$, fornece uma estimativa da probabilidade de ocorrência do tom de cinza γ_k (GONZALEZ e WOODS, 2002).

Para imagens coloridas, pode-se requantizar cada canal de cor da imagem para reduzir a quantidade de cores. Aplicando tal processo, o γ_k passa a ser o k -ésimo hipercubo do cubo de cores RGB. A Figura 18 apresenta um exemplo em que cada canal de cor é requantizado de 256 níveis, totalizando 16,7 milhões de cores, para 4 níveis, resultando em 64 cores, ou hipercubos.

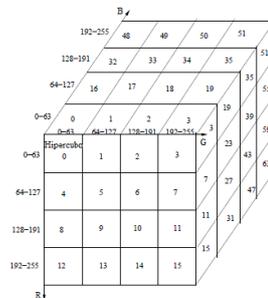


Figura 18: Exemplo de quantização de cores no cubo RGB.

Utilizando-se os coeficientes de Bhattacharyya (COMANICIU, RAMESH e MEER, 2000), conforme a Eq. (4) é possível verificar a semelhança entre dois histogramas. O presente trabalho utiliza essa comparação para eliminar partículas pertencentes ao fundo e para rastrear as partículas pertencentes aos veículos identificados. Para tanto, uma janela de 9×9 *pixels* é centrada em cada partícula e comparada com janelas centradas nas partículas do fundo ou dos veículos dos quadros adjacentes.

$$s(\hat{h}1(y), \hat{h}2) = \sum_{i=1}^L \sqrt{\hat{h}1_i(y) \hat{h}2_i} \quad (4)$$

onde s é a similaridade entre os histogramas normalizados $\hat{h}1(y)$ e $\hat{h}2$, e L é a quantidade de tons de cinza, para imagens desse tipo, ou o número de hipercubos, no caso de imagens coloridas. A similaridade pode assumir valores entre zero (histogramas diferentes) e um (histogramas iguais).

A interpretação geométrica da Eq. (4) é o cosseno do ângulo entre os vetores unitários $\left(\sqrt{\hat{h}1_1}, \dots, \sqrt{\hat{h}1_L}\right)^T$ e $\left(\sqrt{\hat{h}2_1}, \dots, \sqrt{\hat{h}2_L}\right)^T$, considerando as estimativas de densidades discretas $\hat{h}1(y) = \{\hat{h}1_i\}_{i=1 \dots L}$, com $\sum_{i=1}^L \hat{h}1_i = 1$, em uma dada posição y do histograma do candidato (janela da partícula processada), e $\hat{h}2 = \{\hat{h}2_i\}_{i=1 \dots L}$, com $\sum_{i=1}^L \hat{h}2_i = 1$, o histograma do modelo (janela da partícula de referência – fundo ou partícula de veículo rastreado).

3.2 OPERAÇÕES MORFOLÓGICAS

As operações morfológicas são uma parte separada da análise de imagens, baseadas na álgebra de operadores não lineares. São utilizadas para segmentação utilizando formas dos objetos, e para a quantificação de objetos. Utilizam-se de operadores aplicados aos conjuntos de pixels, considerando suas conectividades e formas (SONKA, HLAVAC e BOYLE, 2007).

União e intersecção são operações de conjuntos que podem ser aplicados às imagens binárias. A união entre as imagens A e B , $C = A \cup B$, resulta na imagem C contendo elementos unitários pertencentes a A , B ou ambos, sendo a aplicação da operação lógica *OR* (OU). A intersecção entre as imagens A e B , $C = A \cap B$, resulta na imagem C contendo elementos unitários pertencentes a A e B , sendo equivalente à aplicação da operação lógica *AND* (E) (GONZALEZ e WOODS, 2002).

A operação lógica *NOT* (NÃO) alterna os valores dos *pixels*, zeros tornam-se uns e vice-versa. A operação *XOR* (OU exclusivo) entre A e B retorna elementos unitários em C nas posições de A e B que são diferentes, ou seja, em A é zero e em B é um ou vice-versa (GONZALEZ e WOODS, 2002).

3.2.1 Dilatação

Segundo (GONZALEZ e WOODS, 2002), a dilatação de A pelo elemento estruturante B , sendo conjuntos em Z^2 , denotada por $A \oplus B$ é definida conforme a Eq. (5):

$$A \oplus B = \{z | (\hat{B})_z \cap A \neq \emptyset\} \quad (5)$$

resultando na dilatação por z nas partes de A que houver sobreposição com \hat{B} .

3.2.2 Erosão

Já a erosão, segundo (GONZALEZ e WOODS, 2002), de A pelo elemento estruturante B , sendo conjuntos em Z^2 , denotada por $A \ominus B$ é definida conforme a Eq. (6):

$$A \ominus B = \{z | (B)_z \subseteq A\} \quad (6)$$

resultando na erosão de A por z quando B , transladado por z , está contido em A .

3.2.3 Abertura e Fechamento

A abertura, denotada por $A \circ B$, conforme a Eq. (7) é a aplicação da erosão de A pelo elemento estruturante B , seguido da dilatação desse resultado pelo mesmo elemento estruturante B . Enquanto o fechamento, denotado por $A \bullet B$, conforme a Eq. (8) é a aplicação da dilatação de A pelo elemento estruturante B , seguido pela erosão do resultado pelo mesmo elemento estruturante B (GONZALEZ e WOODS, 2002).

$$A \circ B = (A \ominus B) \oplus B \quad (7)$$

$$A \bullet B = (A \oplus B) \ominus B \quad (8)$$

A abertura suaviza os contornos dos objetos, quebrando ligações finas entre objetos e eliminando pontas e protuberâncias finas. O fechamento também suaviza as bordas dos

objetos, mas ao contrário da abertura, une objetos separados por pequenas distâncias, fecha pequenos buracos no interior dos objetos e preenche entradas no contorno dos objetos.

3.3 OBTENÇÃO DO FUNDO E OBJETOS DE INTERESSE

O fundo é uma imagem que apresenta apenas elementos fixos da cena, ou que alteram-se lentamente, ao longo da passagem dos quadros. Através da subtração do fundo dos quadros processados e da aplicação de uma limiarização a essa diferença é possível obter os objetos que estão movendo-se, ou objetos de interesse (BUCH, VELASTIN e ORWELL, 2011). Uma máscara binária é gerada contendo os objetos de interesse representado por *pixels* brancos (1) e o fundo representado por *pixels* pretos (0).

O fundo modelado no presente trabalho utiliza a aplicação do filtro de mediana temporal, que é um filtro estatístico não linear. Sua aplicação é implementada ordenando-se os *pixels* de cada coordenada (x, y) ao longo do tempo e tomando-se os valores dos *pixels* das posições centrais desse ordenamento para compor o fundo, conforme a Figura 19.

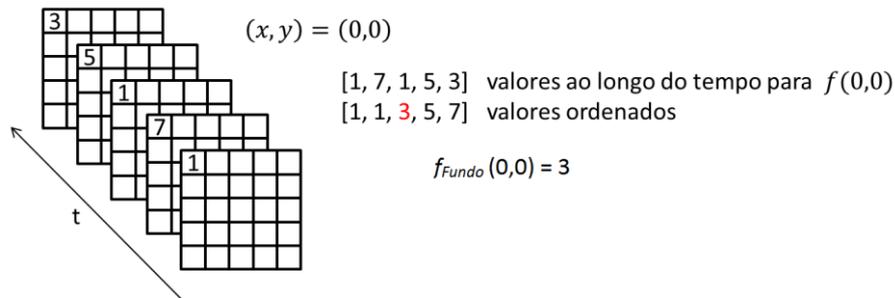


Figura 19: Exemplo de filtro de mediana temporal.

A limiarização da distância Euclidiana entre intensidades de *pixels* do quadro processado e do fundo é utilizada para obter os objetos de interesse a partir de imagens coloridas, conforme as Eqs. (9) e (10):

$$D_{\text{Cor}}(p, q) = \sqrt{(R_p - R_q)^2 + (G_p - G_q)^2 + (B_p - B_q)^2} \quad (9)$$

$$g(x, y) = \begin{cases} 0 & \text{se } f(x, y) < T \\ 1 & \text{se } f(x, y) \geq T \end{cases} \quad (10)$$

onde $D_{\text{Cor}}(p, q)$ é a distância Euclidiana entre as cores dos *pixels* p e q . R , G e B são os canais de cor dos respectivos *pixels*. A imagem binária $g(x, y)$ é resultante da aplicação da limiarização de $f(x, y)$, que é obtida através das distâncias Euclidianas entre as cores dos

pixels do quadro processado e do fundo, utilizando-se o valor de limiar T . O processo de limiarização é o mesmo para imagens em tons de cinza.

3.4 RUÍDO IMPULSIVO

As imagens são degradadas por ruídos principalmente durante a aquisição e transmissão (GONZALEZ e WOODS, 2002). O presente trabalho não trata tais questões, mas utiliza-se do ruído impulsivo para gerar partículas aleatórias em determinadas circunstâncias. Uma imagem ruidosa, $g(x, y)$, pode ser obtida através do modelamento do ruído, $\eta(x, y)$, aplicado à imagem original, $f(x, y)$, conforme a Eq. (11) (GONZALEZ e WOODS, 2002).

$$g(x, y) = f(x, y) + \eta(x, y) \quad (11)$$

O ruído impulsivo pode ser modelado como uma variável aleatória com densidade de probabilidade impulsiva, conforme a Figura 20 e Eq. (12).

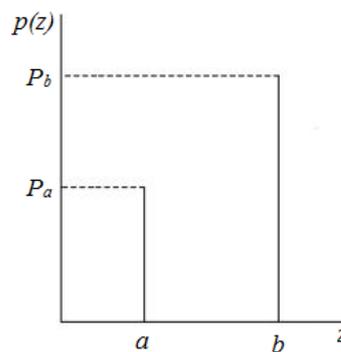


Figura 20: Função densidade de probabilidade para o ruído impulsivo.

$$p(z) = \begin{cases} P_a & \text{para } z = a \\ P_b & \text{para } z = b \\ 0 & \text{outros casos} \end{cases} \quad (12)$$

Se $b > a$, os tons de cinza b aparecerão como pontos brilhantes na imagem, enquanto os tons de cinza a aparecerão como pontos escuros. Se P_a ou P_b é igual a zero, o ruído impulsivo é chamado unipolar. Se P_a e P_b não são zero e possuem aproximadamente o mesmo valor, o ruído é dito sal e pimenta, em alusão aos pontos brancos e pretos.

3.5 DETECTOR DE CANTOS

O detector de cantos utilizado é o apresentado em (SHI e TOMASI, 1994), que considera a identificação de mínimos autovalores, $\min(\lambda_1, \lambda_2)$, de matrizes de segundo momento A , conforme formulação seguinte. Considerando uma imagem em tons de cinza, I , e uma região dessa imagem, w , sobre a área (u, v) deslocada por (x, y) , pode-se obter a soma das diferenças quadradas entre essas imagens, S , conforme a Eq. (13).

$$S(x, y) = \sum_u \sum_v w(u, v) (I(u + x, v + y) - I(u, v))^2 \quad (13)$$

Considerando-se I_x e I_y derivadas parciais de I , $I(u + x, v + y)$ pode ser aproximada por uma expansão de Taylor, conforme a Eq. (14) e S pode ser reescrita utilizando-se essa aproximação, conforme a Eq. (15):

$$I(u + x, v + y) \approx I(u, v) + I_x(u, v)x + I_y(u, v)y \quad (14)$$

$$S(x, y) = \sum_u \sum_v w(u, v) (I_x(u, v)x + I_y(u, v)y)^2 \quad (15)$$

Utilizando-se a matriz de segundo momento A , definida na Eq. (16), pode-se reescrever S na forma matricial, conforme a Eq. (17):

$$A = \sum_u \sum_v w(u, v) \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} = \begin{bmatrix} \langle I_x^2 \rangle & \langle I_x I_y \rangle \\ \langle I_x I_y \rangle & \langle I_y^2 \rangle \end{bmatrix} \quad (16)$$

$$S(x, y) \approx \begin{pmatrix} x & y \end{pmatrix} A \begin{pmatrix} x \\ y \end{pmatrix} \quad (17)$$

Quando $\min(\lambda_1, \lambda_2)$ da matriz A para $S(x, y)$ possui valor acima de limiares definidos, um canto é identificado na imagem $I(x, y)$.

3.6 VETORES DE MOVIMENTO

Cada *pixel* de um quadro em uma sequência de vídeo pode ter associado a si um vetor de movimento, ou seja, valores em (x, y) que o *pixel* moveu-se em tais direções entre dois quadros adjacentes. Neste trabalho, são obtidos os vetores de movimento das partículas utilizadas (obtidas com o detector de cantos) através de comparação de blocos e busca exaustiva (DUFAUX e MOSCHENI, 1995).

Janelas centradas nas partículas (bloco atual) são comparadas com janelas centradas em cada *pixel* de uma região de busca (bloco de busca) no quadro seguinte (em torno das posições espaciais destas partículas). A Figura 21 exemplifica o bloco e sua área de busca, e a Eq. (18) apresenta a função de custo Média das Diferenças Absolutas (*MDA*), que quando minimizada retorna os vetores de movimento.

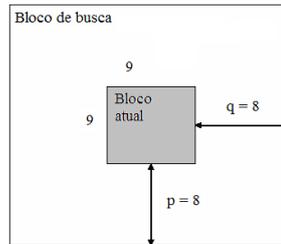


Figura 21: Exemplificação da comparação de blocos.

$$MDA(p, q) = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N |C_{ij} - R_{i+p, j+q}| \quad (18)$$

onde $p, q = -8, -7, \dots, 7, 8$ são os deslocamentos do bloco atual C , cujo tamanho em *pixels* é N , no bloco de busca R e (i, j) são as posições de *pixel* do bloco atual. Os valores de (p, q) que retornam o menor valor de *MDA* são os componentes do vetor de movimento estimado.

3.7 ALGORITMO K-MEANS

O *k-means* (k-médias) é um algoritmo de classificação de dados que divide n observações em k partições, de acordo com a distância das características desses dados com os centroides das partições, que são os valores médios das observações pertencentes a tais partições (SEBER, 1984). Diversas métricas podem ser empregadas para a determinação de tais distâncias.

Definindo-se matematicamente, dado um conjunto de observações $x = x_1, x_2, \dots, x_n$, onde cada observação é um vetor real com d dimensões, o *k-means* tenta dividir as n observações em k partições, com $k \leq n$, de forma a minimizar a soma dos quadrados das distâncias das observações dos agrupamentos, $S = \{S_1, S_2, \dots, S_k\}$, aos respectivos centroides, μ , conforme a Eq. (19).

$$\arg \min \sum_{i=1}^k \sum_{x_j \in S_i} \|x_j - \mu_i\|^2 \quad (19)$$

onde μ_i é a média dos pontos (centroides) de S_i (agrupamentos).

Dado um conjunto de observações x e k médias iniciais $m_1^{(1)}, \dots, m_k^{(1)}$, o algoritmo *k-means* pode ser executado nos seguintes passos para determinar os agrupamentos, S_i , em cada iteração, t :

- 1 Alocar cada observação x_p para o agrupamento S_i se a distância de x_p para a média m_i (centroide do agrupamento i) for menor que a distância de x_p para a média m_j (centroide do agrupamento j), conforme a Eq. (20);

$$S_i^{(t)} = \{x_p : \|x_p - m_i^{(t)}\| \leq \|x_p - m_j^{(t)}\| \quad \forall 1 \leq j \leq k\} \quad (20)$$

- 2 Atualizar os centroides, m , através da média das observações dos agrupamentos, S , conforme a Eq. (21);

$$m_i^{(t+1)} = \frac{1}{|S_i^{(t)}|} \sum_{x_j \in S_i^{(t)}} x_j \quad (21)$$

- 3 Repetir passo 1 e 2 até que não haja mais mudanças de observações x_p entre os agrupamentos S_i .

Dois métricas são utilizadas neste trabalho para gerar os agrupamentos. A *city block*, que é a soma das diferenças absolutas, conforme a Eq. (22), e os centroides são as medianas dos componentes dos respectivos clusters. E a distância do cosseno, definida como um menos o cosseno do ângulo formado pelos pontos, tratados como vetores, conforme a Eq. (23), em que os centroides são as médias dos pontos de cada agrupamento, após normalização desses pontos em relação a uma unidade Euclidiana. Além dessas duas métricas utilizadas para a formação de agrupamentos, uma terceira métrica é utilizada para selecionar agrupamentos quando partículas pertencem às regiões de interseção entre dois veículos rastreados, a distância de Mahalanobis (MAHALANOBIS, 1936), conforme a Eq. (24).

Dada uma matriz de dados X , de m linhas por n colunas, em que cada linha representa um vetor, x_1, x_2, \dots, x_m , as distâncias utilizadas neste trabalho para a execução do *k-means* podem ser computadas como:

$$d_{st} = \sum_{j=1}^n |x_{sj} - x_{tj}| \quad (22)$$

$$d_{st} = 1 - \frac{x_s x_t'}{\sqrt{(x_s x_s')(x_t x_t')}} \quad (23)$$

$$d_{st} = (x_s - x_t)C^{-1}(x_s - x_t)' \quad (24)$$

onde d_{st} é a distância (*city block*, cosseno ou Mahalanobis) entre dois vetores da matriz X , x_s e x_t , sendo x_{sj} e x_{tj} elementos destes vetores respectivamente, podendo representar uma observação e um centroide, por exemplo. C é a matriz de covariância de X .

4 RASTREAMENTO VEICULAR EM VÍDEOS DE TRÁFEGO

Nesta seção é apresentada de forma detalhada a implementação do algoritmo de rastreamento e contagem de veículos juntamente com exemplos de situações que ocorrem ao longo do processo. As técnicas de processamento de imagens empregadas são justificadas por sua simplicidade e eficiência para a obtenção dos resultados desejados, conforme exposto ao longo do texto.

Para avaliação do algoritmo foram utilizados seis vídeos diferentes cedidos pela Empresa Pública de Transporte e Circulação (EPTC), de Porto Alegre - RS. Os vídeos foram divididos em blocos de 3.000 quadros por questões de gerenciamento de memória, resultando-se em quatro ou cinco partes por vídeo, 81.000 quadros no total. Cinco dos seis vídeos possuem três pistas, e um apresenta duas pistas. Além disso, quatro vídeos possuem sinalização de trânsito que forcem a parada dos veículos, tornando a tarefa de rastreamento desafiadora devido à sobreposição de veículos.

Outra questão que dificulta a solução do rastreamento é a variedade de tamanhos dos veículos, ocorrendo desde a passagem de motocicletas até grandes caminhões, alguns deles ocupando praticamente toda a imagem. A posição das câmeras também varia de acordo com o vídeo, mas todas estão posicionadas acima da via, com leves inclinações para a esquerda, de forma que o fluxo de veículos ocorre na direção vertical, de cima para baixo na imagem. Um dos vídeos apresenta uma parte do fluxo de veículos no sentido contrário, mas apenas a parte semelhante aos outros vídeos é processada. Outro vídeo possui a câmera instalada de forma que os veículos passem quase que perpendicularmente à câmera, de cima para baixo. Tais vídeos são apresentados na Figura 22.

O algoritmo foi desenvolvido em ambiente Matlab 7.13 de 64 *bits*, utilizando um computador equipado com processador Intel i5-2400 de 3,10 GHz e 32,0 GB de memória, localizado no Laboratório de Sinais e Instrumentação. Com tal configuração não seria necessário particionar os vídeos utilizados, mas o início da implementação ocorreu utilizando-se um computador portátil equipado com apenas 4,0 GB de memória. Outra atitude tomada inicialmente e mantida posteriormente foi o redimensionamento dos vídeos para 240x320 *pixels*, originalmente eles variavam entre 480x640 e 960x1280 *pixels*.



Figura 22: Amostras dos vídeos utilizados.

4.1 VISÃO GERAL DO ALGORITMO PROPOSTO

Conforme apresentado no resumo e introdução do trabalho, o objetivo do algoritmo é realizar a contagem de veículos em laços virtuais definidos pelo usuário, utilizando-se das informações do rastreamento efetuado. Para tanto, inicialmente o algoritmo abre um vídeo e solicita a inclusão manual de uma ROI, que é a região da imagem em que o algoritmo atua, as demais regiões não são processadas. Outra informação solicitada é a inclusão dos laços virtuais, que são as regiões em que os veículos são efetivamente contados. A Figura 23 e a Figura 24 apresentam as ROI e os laços virtuais utilizados neste trabalho.

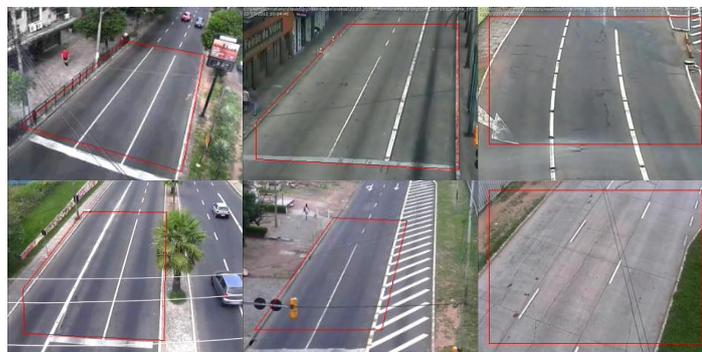


Figura 23: Regiões de interesse em vermelho.

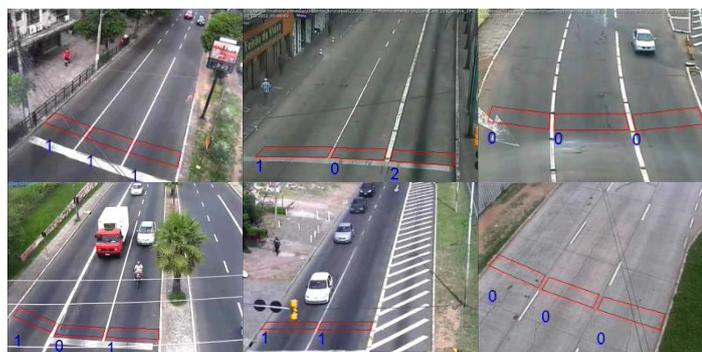


Figura 24: Laços virtuais em vermelho com suas respectivas contagens em azul.

Apesar das imagens das ROI não estarem presentes nas imagens dos laços virtuais, deve-se levá-las em consideração para sua colocação. Os laços devem ser posicionados no interior das ROI para o correto funcionamento da contagem. Como o processo de rastreamento só é efetuado dentro das ROI e a contagem é realizada a partir do rastreamento, caso os laços sejam posicionados fora dessas regiões, os veículos rastreados nunca passarão por eles, inutilizando suas colocações.

Após essa etapa de inicialização que necessita do usuário, o restante do processo é automático e só termina quando todo o vídeo é processado, retornando vídeos com o resultado do rastreamento e das contagens nos laços virtuais. Antes de entrar no laço principal do algoritmo, que processa quadro a quadro todo o vídeo, informações iniciais são geradas, como a imagem de fundo e uma imagem representativa da repetição das partículas em suas respectivas posições espaciais, através da intensidade luminosa nessas posições. Apenas as partículas obtidas através da extração dos cantos são consideradas nessa etapa. Como será visto posteriormente, novas partículas são geradas ao longo do processo, não sendo necessariamente cantos. A Figura 25 e a Figura 26 apresentam respectivamente exemplos de fundos e de mapas de partículas repetidas.

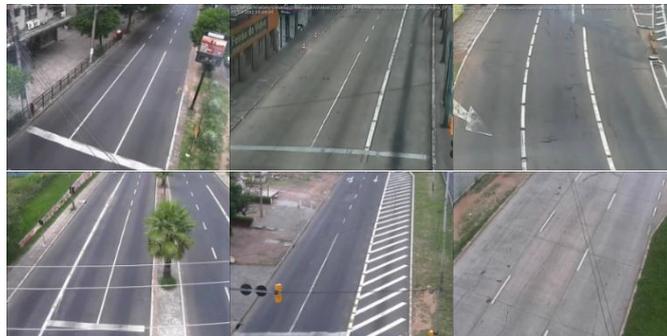


Figura 25: Exemplos de fundos.

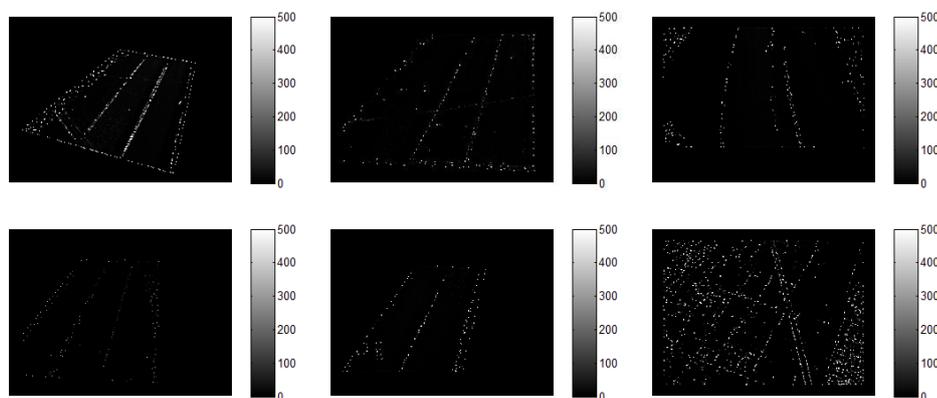


Figura 26: Exemplos de mapas de partículas repetidas.

Ao entrar no laço principal do algoritmo, que vai do segundo até o penúltimo quadro, obtém-se dois quadros do vídeo em questão, o $quadro_t$, que é o quadro na posição temporal atual, e o $quadro_{t+1}$, que é o quadro adjacente na próxima posição temporal. Os dois quadros são necessários para a obtenção dos vetores de movimento das partículas. No presente trabalho a obtenção dos vetores de movimento foi implementada desta maneira, mas nada impede a utilização do quadro anterior, $quadro_{t-1}$, para essa tarefa, mas o laço principal do algoritmo começaria no segundo quadro e iria até o último quadro.

Aos dois quadros obtidos são aplicadas compensações de cores, utilizadas para reduzir variações drásticas de iluminação que ocorrem nos vídeos utilizados. Isto é feito utilizando-se as cores médias do fundo e dos quadros em questão. Em vídeos cuja iluminação seja mais estável ou se a obtenção do fundo for executada adaptativamente às variações de luminosidade, esta etapa pode ser eliminada. Imagens em tons de cinza também são obtidas a partir destes quadros para utilização futura.

Em seguida, é verificado se há alguma região pertencente a veículos rastreados ou agrupamentos de partículas, provenientes do processamento dos quadros anteriores. Cada uma dessas regiões é tratada de forma diferente ao longo do processo, bem como as regiões em que não há veículos rastreados e agrupamentos de partículas. Nas regiões rastreadas ou sob os agrupamentos, são geradas partículas aleatórias, para minimizar a eliminação de veículos rastreados e agrupamentos formados.

Após a verificação das regiões do quadro anterior e criação de partículas aleatórias em seu interior, as partículas que representam cantos da imagem são obtidas. Utilizando-se os mapas de repetição de partículas e a imagem do fundo, as partículas são filtradas de acordo com a região à qual pertencem, levando-se em consideração suas repetitividades e similaridades com o fundo. Para as partículas resultantes dessa filtragem, são obtidos os vetores de movimento, através de comparação de blocos. Cada partícula terá então quatro variáveis que serão passadas para a próxima etapa, as posições espaciais em x e y e os vetores de movimento formados por v_x e v_y .

O próximo passo então verifica se existem centroides referentes a agrupamentos ou veículos rastreados, resultantes do processamento do quadro anterior, ou se o número de partículas é inferior a um número determinado. Se não houver centroides e o número de partículas for inferior ao mínimo especificado, a iteração encerra-se sem alterações. Isso ocorre quando não há nenhum veículo passando. Se não houver centroides, mas o número de partículas for superior ao mínimo especificado, um agrupamento inicial das partículas é

executado. Esse processo ocorre nos primeiros quadros, quando os agrupamentos ainda não foram completamente formados, ou quando ocorre a entrada de veículos na ROI.

Se existem centroides remanescentes do processamento do quadro anterior, significa que pode haver veículos rastreados ou agrupamentos formados, que futuramente poderão tornar-se veículos rastreados. O algoritmo então identifica as regiões pertencentes a essas situações e à região livre, e trata-as de forma diferente. Às partículas das regiões livres, ou seja, não pertencem a veículos rastreados ou agrupamentos formados, é aplicado o agrupamento inicial. Nas regiões dos agrupamentos formados, aplica-se o algoritmo de agrupamento, mas passando-se os centroides anteriores como inicialização. Desses dois processos de agrupamento resultam clusters que serão divididos e fundidos de acordo com critérios explanados nas Subseções 4.3.2 e 4.3.3. Os agrupamentos resultantes são verificados em busca de veículos.

A identificação dos veículos é realizada através da análise morfológica das regiões convexas formadas pelas partículas dos agrupamentos formadas, juntamente com a análise da similaridade dessas regiões com o fundo. Fatores como área, taxa de ocupação de partículas, proporcionalidade entre comprimento e largura, similaridade da região com o fundo, tanto internamente como externamente, são avaliados e pontuados. Se estes fatores obtiverem uma pontuação que seja suficiente, de acordo com um limiar estabelecido experimentalmente, o agrupamento em questão é considerado um veículo identificado, e será rastreado através de similaridade de histogramas de cor nos próximos quadros.

As regiões dos veículos identificados são deslocadas de acordo com os vetores de movimento de suas partículas nos frames seguintes, e cada partícula possuindo uma janela centrada em si, é comparada com as partículas do frame anterior, através de similaridade de histogramas de cor. As partículas cuja similaridade é elevada são mantidas, enquanto as partículas cuja similaridade é baixa são eliminadas. Cada novo veículo rastreado recebe um número identificador único, assim como, as partículas que a ele pertencem. Esse número acompanha o respectivo veículo até sua saída da ROI.

Uma divisão dos agrupamentos referentes aos veículos rastreados também é executada, pois na entrada da ROI, devido à perspectiva, veículos sobrepostos podem ser considerados como sendo apenas um. Conforme deslocam-se ao longo da ROI, podem ficar mais distantes, revelando segmentos de fundo entre eles. Esses segmentos de fundo são comparados com a região do veículo rastreado e utilizados para dividi-lo, caso necessário.

A última etapa do processo consiste na contagem dos veículos nos laços virtuais. Esses laços podem ser colocados em qualquer posição dentro da ROI, mas a sua colocação próxima à saída dos veículos da ROI garante uma contagem mais precisa, pois assim os veículos terão passado por um número maior de quadros e a possibilidade de uma correta identificação será maior. O incremento dos contadores dos respectivos laços ocorre quando há intersecção das áreas do veículo rastreado com a região do laço virtual. Cada veículo poderá acionar mais de um laço ao longo do processo, mas apenas o primeiro laço acionado terá seu contador incrementado. Além disso, apenas um laço pode ser acionado por veículo em um determinado quadro, aquele em que a área de intersecção for maior.

A Figura 27 apresenta um diagrama de blocos com uma visão geral do método proposto. Há um Apêndice junto à dissertação com o algoritmo referente à visão geral do método proposto, sendo uma ferramenta que detalha melhor o processo. As subseções a seguir apresentam detalhadamente cada etapa descrita na apresentação da visão geral do algoritmo. Detalhes dos métodos utilizados podem ser observados na Seção 3.

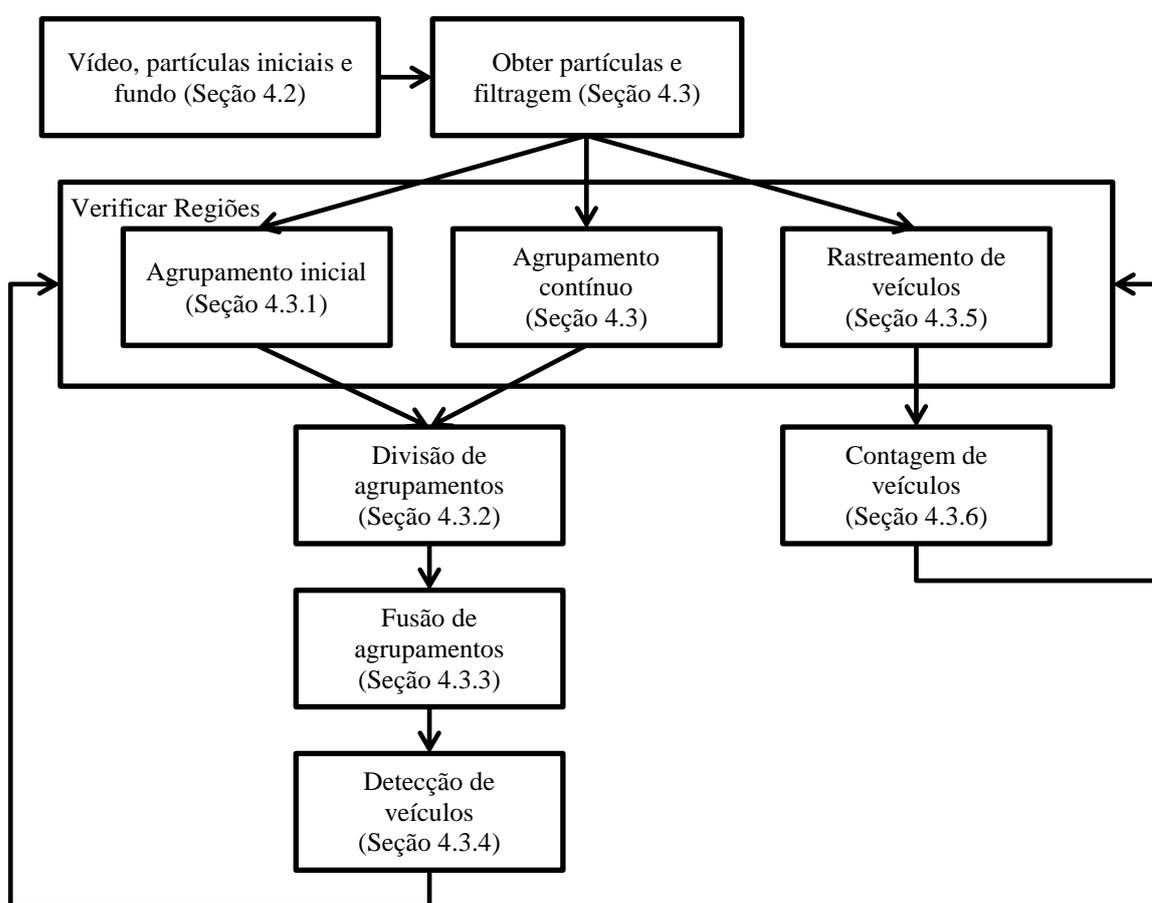


Figura 27: Visão geral do método proposto.

4.2 PRÉ-PROCESSAMENTO

O pré-processamento corresponde às atividades executadas antes do algoritmo entrar no laço principal de execução. Compreendem a inicialização de variáveis e obtenção de parâmetros fornecidos pelo usuário. Entre as variáveis inicializadas está o número mínimo de pontos para a criação de clusters, as listas de regiões e veículos rastreados, os centroides, e as contagens dos laços virtuais.

Os parâmetros fornecidos pelo usuário são o vídeo utilizado, a ROI e os laços virtuais. A ROI e os laços virtuais são obtidos através da chamada de funções que operam sobre o vídeo selecionado, abrindo um quadro e solicitando a entrada de linhas pelo usuário. Inicialmente é solicitada a ROI, que consiste de uma máscara binária em que o interior da ROI é composto por *pixels* unitários enquanto a região externa é composta por *pixels* zerados. O processamento do vídeo apenas na região da ROI deve-se à aplicação da operação *AND* dos quadros com a ROI.

A inclusão dos laços virtuais ocorre de forma semelhante à ROI, mas a cada laço incluído é solicitado se novos laços serão incluídos ou não. Enquanto a opção *sim* for selecionada novos laços podem ser adicionados. Quando a opção *não* é selecionada a operação é finalizada e uma máscara em que cada laço contém uma numeração é gerada, variando de um até o número de laços gerados.

O fundo é obtido durante essa etapa, através da aplicação do filtro de mediana temporal, conforme explicado na Subseção 3.3, para cada canal de cor RGB. Para reduzir a influência por variações rápidas de iluminação, a diferença da cor média do fundo para cada quadro é compensada, somando-se essa diferença em cada quadro. E nessa etapa que a cor média do fundo é obtida, através da Eq. (25).

$$CMF_C = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N p(i, j, C) \quad (25)$$

onde CMF_C é a cor média do fundo para cada canal de cor RGB C , M e N são as dimensões do quadro em *pixels*, e $p(i, j, C)$ é um *pixel* do fundo na posição (i, j) e no canal de cor C .

O conjunto de partículas iniciais, denotado por *partículas_iniciais*, é obtido utilizando-se o método dos mínimos autovalores apresentado por (SHI e TOMASI, 1994), conforme apresentado na Subseção 3.5, varrendo-se todo o vídeo. Tal conjunto de partículas é utilizado na primeira etapa de filtragem de partículas, adaptada de (MOSSI *et al.*, 2011), executada no

laço principal do algoritmo, que é apresentado na próxima seção. Considerando-se $S_t(x, y)$ uma imagem binária, contendo as partículas obtidas do $quadro_t$, a imagem $particulas_iniciais(x, y)$ representa o número de partículas detectadas em cada posição (x, y) do quadro para todo o vídeo, conforme as Eqs. (26) e (27).

$$S_t(x, y) = \begin{cases} 1 & \text{se } S_t \text{ contém uma partículas em } (x, y) \\ 0 & \text{caso contrário} \end{cases} \quad (26)$$

$$particulas_iniciais(x, y) = \sum_{t=1}^{nrQuadros} S_t(x, y) \quad (27)$$

onde $nrQuadros$ é o número total de quadros do vídeo processado. A Figura 26, na Subseção 4.1, apresenta as $particulas_iniciais$ para os vídeos utilizados.

4.3 DETALHAMENTO DO ALGORITMO PROPOSTO

O laço principal compreende a parte do algoritmo que iterativamente processa todos os quadros do vídeo. Cada iteração utiliza tanto informações da iteração anterior, como regiões de partículas, centroides e veículos rastreados, como informações do quadro seguinte, para obtenção dos vetores de movimento das partículas.

Inicialmente são obtidos os quadros atual e seguinte, denotados por $quadro_t$ e $quadro_{t+1}$ e suas cores são compensadas utilizando-se a cor média do fundo, CMF . Essa compensação é realizada através da soma da diferença da cor média do fundo para a cor média do $quadro_t$ e $quadro_{t+1}$, CMQ , nos respectivos quadros, conforme as Eqs. (28) e (29).

$$CMQ_C = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N p(i, j, C) \quad (28)$$

onde CMQ_C é a cor média do $quadro_t$ ou $quadro_{t+1}$ para cada canal de cor RGB, C , M e N são as dimensões do quadro em *pixels*, e $p(i, j, C)$ é um *pixel* do quadro na posição (i, j) e no canal de cor C .

$$DifCor_C = CMF_C - CMQ_C \quad (29)$$

Essas diferenças de cores médias, $DifCor_C$, uma para o $quadro_t$ e uma para o $quadro_{t+1}$ são somadas em cada *pixel* dos respectivos quadros, nos seus respectivos canais

de cor, C . Até este ponto, os quadros foram adequados para o início do processamento dos agrupamentos. Várias outras técnicas podem ser aplicadas, como redimensionamento e filtragens. A partir daqui o processamento deve ocorrer sobre as imagens sem novas alterações.

As regiões são verificadas de acordo com as partículas do quadro anterior, denotado por $quadro_{t-1}$. Na primeira passagem pelo laço principal do algoritmo, não existem agrupamentos anteriores nem veículos rastreados, logo, não há centroides anteriores e todas as partículas obtidas nessa passagem serão consideradas partículas livres. Caso houvesse agrupamentos de partículas ou veículos rastreados no $quadro_{t-1}$, essas partículas seriam deslocadas de seus vetores de movimento para o $quadro_t$, e partículas aleatórias seriam geradas no interior das regiões convexas dilatadas formadas pelas partículas desses agrupamentos e veículos. A geração dessas partículas aleatórias utiliza-se de ruído impulsivo, conforme descrito na Subseção 3.4.

Do $quadro_t$ são obtidas as partículas através da detecção de cantos, conforme Subseção 3.5, que são acrescentadas às eventuais partículas geradas com o ruído impulsivo. Considerando-se as partículas como *pixels* unitários em uma imagem binária, denominada $Particulas(x, y)$, nas posições (x, y) , são então filtradas utilizando-se o conjunto de partículas iniciais, conforme a Eq. (30).

$$Particulas(x, y) = \begin{cases} 1 & \text{se } particulas_iniciais(x, y) < T \\ 0 & \text{se } particulas_iniciais(x, y) \geq T \end{cases} \quad (30)$$

onde T é um limiar definido de acordo com a repetitividade das partículas na posição (x, y) , tal limiar depende do tempo que os veículos ficam parados no mesmo lugar. Neste trabalho utilizou-se $T = 50$, obtido experimentalmente. Valores muito pequenos de T podem eliminar partículas pertencentes a veículos lentos ou parados, enquanto valores muito elevados de T podem manter partículas do fundo.

O limiar T pode eliminar partículas de veículos parados, mas sempre haverá partículas disponíveis nas regiões de agrupamentos formados e veículos rastreados, devido à geração de partículas aleatórias, exposto anteriormente.

A segunda etapa de filtragem das partículas considera a similaridade de histogramas de cor, conforme a Subseção 3.1.2. Janelas de 9×9 *pixels*, centradas nas partículas do $quadro_t$, são comparadas com janelas nas mesmas posições do fundo. Quando as similaridades entre os histogramas de cor dessas janelas são inferiores a 0,5 as partículas são

mantidas, caso contrário são eliminadas. Este limiar de similaridade foi obtido experimentalmente. A Figura 28 apresenta um exemplo em que uma partícula detectada no $quadro_t$ apresenta uma similaridade de 0,9679 com o fundo, sendo eliminada, enquanto a Figura 29 apresenta uma partícula do $quadro_t$ com uma similaridade de 0,3514 com o fundo, sendo mantida.

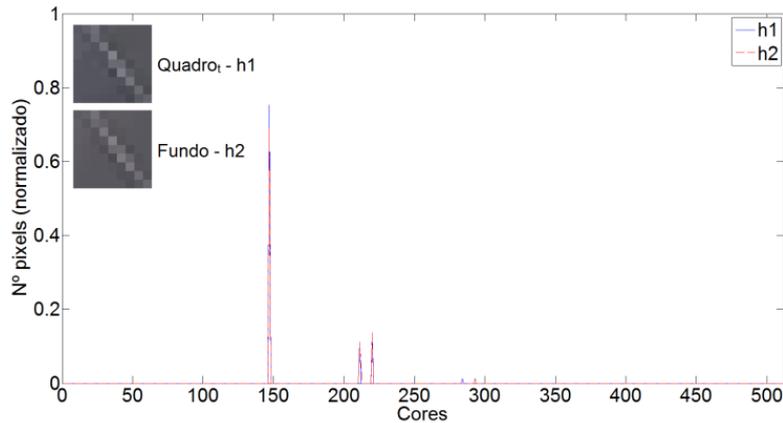


Figura 28: Similaridade de histogramas de 0,9679.

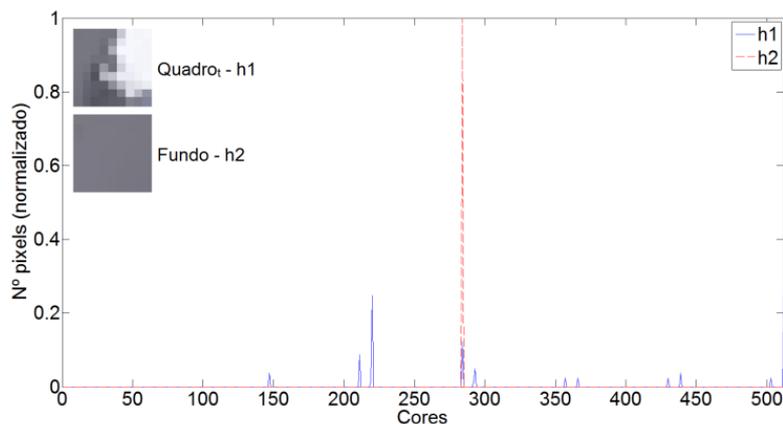


Figura 29: Similaridade de histogramas de 0,3514.

Com as partículas remanescentes dos processos de filtragem o processo continua, sendo obtidos os vetores de movimento dessas partículas. Conforme a Subseção 3.6, os vetores de movimento são obtidos através de comparação de blocos com busca exaustiva. Janelas de 9×9 *pixels* centradas nas partículas do $quadro_t$ são comparadas com regiões de busca de 17×17 *pixels* centradas nas mesmas posições, mas no $quadro_{t+1}$, ou seja, cada componente do vetor de movimento pode assumir valores inteiros entre -8 e 8.

Após a obtenção das partículas e dos vetores de movimento, testa-se a existência de centroides na iteração anterior. Se não houver centroides e houver um número suficiente de partículas para a geração de agrupamentos, um agrupamento inicial é gerado. Tal situação ocorre no processamento dos primeiros quadros ou quando veículos entram em cena, após um período sem veículos passando.

Quando existem centroides passados pela iteração anterior, significa que há agrupamentos formados e/ou veículos identificados, mas também pode haver partículas que não pertencem a nenhum desses grupos. Essas partículas são denotadas por *partículas_livres*, e também passam pelo processo de agrupamento inicial.

4.3.1 Agrupamento Inicial

O agrupamento inicial utiliza-se do *k-means*, apresentado na Subseção 3.7, aplicado às partículas, considerando suas posições espaciais e vetores de movimento. Nenhuma informação inicial a respeito de agrupamentos é utilizada nesta etapa, utilizam-se apenas as informações das partículas e o número mínimo de partículas que um agrupamento deve possuir. A Figura 30 apresenta o fluxograma do agrupamento inicial, o algoritmo referente à tal processo encontra-se no Apêndice desta dissertação.

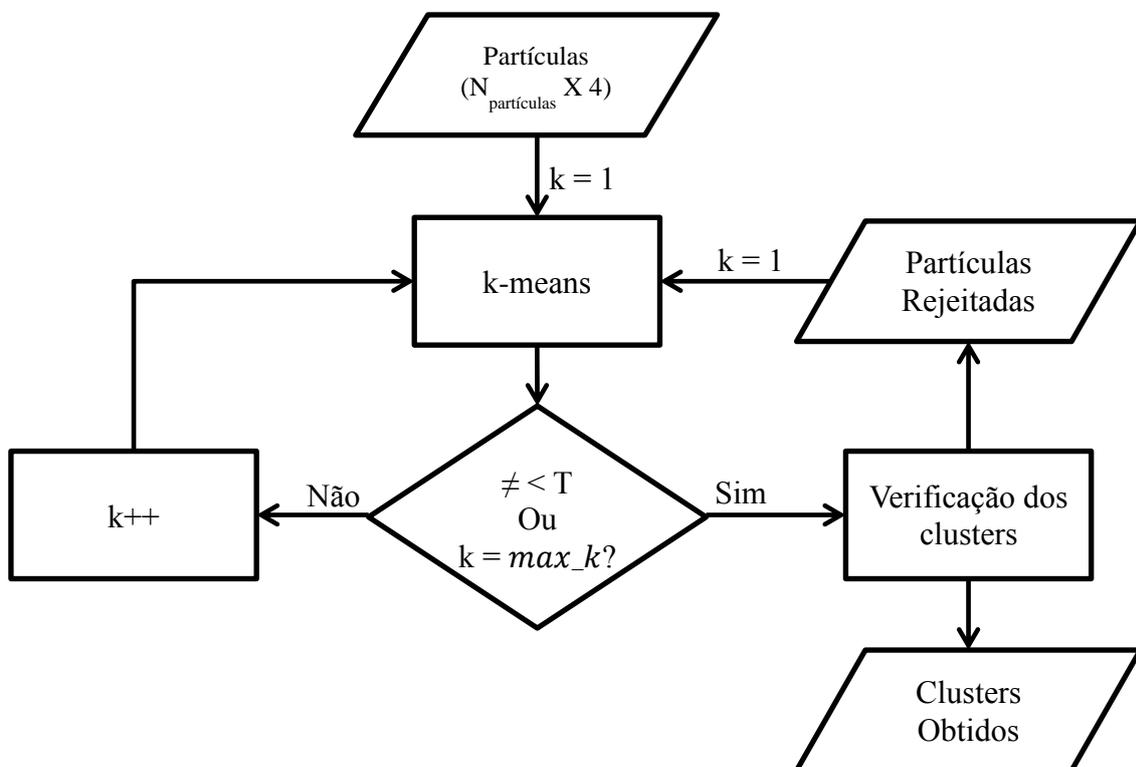


Figura 30: Fluxograma do agrupamento inicial.

Inicialmente, obtém-se o número máximo de agrupamentos possíveis, Max_k , através do arredondamento para baixo da razão $nrParticulas/Min_pontos$. O k -means é aplicado às partículas utilizando-se apenas um agrupamento. Após, iterativamente, de dois agrupamentos até no máximo Max_k , o k -means é aplicado às partículas. Se a diferença entre as distâncias das partículas aos respectivos centroides é menor que o limiar T , entre duas iterações adjacentes, o processo é interrompido e o resultado com $k - 1$ agrupamentos é retornado. De acordo com observações experimentais, a métrica para obter as distâncias das partículas aos centroides que formam os melhores agrupamentos, é a medida do cosseno, apresentada na Subseção 3.7.

Os agrupamentos são testados em função de gráficos de silhueta, que determinam o quão parecido as partículas são de seus agrupamentos e dos outros agrupamentos, através da Eq. (31):

$$S(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}} \quad (31)$$

onde $S(i)$ é a medida de similaridade da i -ésima partícula, $a(i)$ é a distância média da i -ésima partícula em relação às partículas de seu próprio agrupamento e $b(i)$ é a distância média da i -ésima partícula para as partículas dos outros agrupamentos. Partículas com similaridades negativas são eliminadas dos seus respectivos agrupamentos. A métrica utilizada é a *city block*, apresentada na Subseção 3.7.

As partículas eliminadas dos agrupamentos são processadas recursivamente em busca de novos agrupamentos possíveis. Os agrupamentos resultantes são concatenados com os agrupamentos formados anteriormente. Nos agrupamentos resultantes do agrupamento inicial é aplicada uma função para dividi-los caso necessário, conforme explicado a seguir.

4.3.2 Divisão de Agrupamentos

Os agrupamentos resultantes do *agrupamento_inicial* e/ou a união entre os agrupamentos resultantes do *agrupamento_inicial* e os agrupamentos do *agrupamento_cont* passam por uma divisão caso necessário, pois veículos próximos e movendo-se com a mesma velocidade e direção podem ser agrupados como um único objeto. A informação do fundo é utilizada para verificar se uma divisão do agrupamento deve ser realizada.

Considerando a imagem convexa formada pelas partículas de um determinado agrupamento, a distância Euclidiana entre cores de cada *pixel* do $quadro_t$ e do fundo é

comparada com um limiar superior T , resultando em uma máscara correspondente a similaridade do $quadro_t$ com o fundo para essa região. A essa máscara é aplicado a operação de abertura, conforme a Subseção 3.2.1, utilizando-se um elemento estruturante SE apresentado na Eq. (32), para remover pequenos conjuntos de $pixels$.

$$SE = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad (32)$$

Se a área de similaridade com o fundo for grande em relação à área da região convexa ($> 15\%$) a divisão do agrupamento é efetuada, através da aplicação do k -means com 2 clusters nas partículas pertencentes ao agrupamento processado. A Figura 31 exemplifica o processo, onde dois veículos inicialmente agrupados como um único objeto são corretamente divididos, apresentando as seguintes imagens: (a) detalhe dos veículos próximos e deslocando-se com mesma velocidade e direção; (b) único agrupamento gerado com as partículas pertencentes aos veículos; (c) fundo da região em detalhe; (d) região convexa formada pelas partículas do agrupamento; (e) imagem da distância Euclidiana entre as cores dos $pixels$ do $quadro_t$ e fundo para a região convexa; (f) limiarização da imagem (e), utilizando-se $T = 50$; (g) resultado da aplicação da operação de abertura, utilizando-se SE ; (h) resultado da divisão do agrupamento.

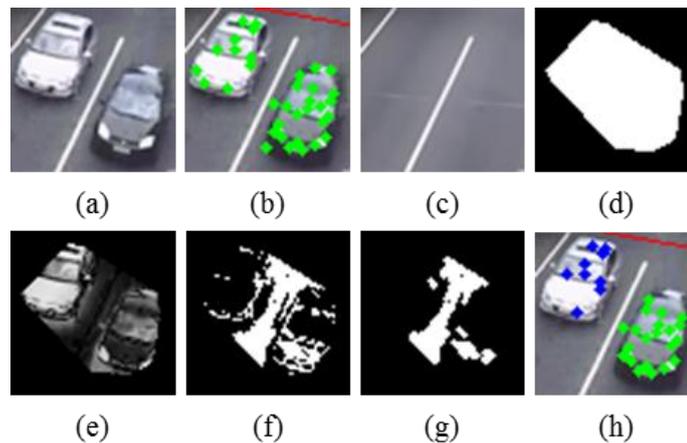


Figura 31: Exemplo de divisão de agrupamento de partículas.

4.3.3 Fusão de Agrupamentos

Assim como veículos próximos e andando com mesma velocidade e direção podem ser agrupados como um mesmo objeto, veículos grandes e com regiões de pouca textura, o

que gera poucas partículas, podem resultar em mais de um agrupamento. Vetores de movimento diferentes também são gerados em veículos grandes, devido à perspectiva, o que contribui para o incorreto agrupamento. A Figura 32 apresenta um exemplo em que as partículas apresentam dois agrupamentos distantes, além de vetores de movimento diferentes ao longo do veículo.

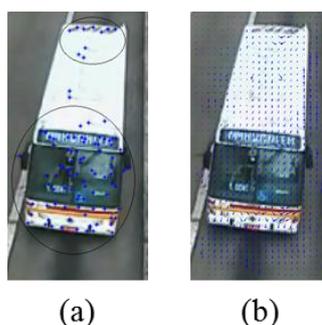


Figura 32: Exemplo de agrupamentos espaciais inadequados de partículas.

Na Figura 32, a imagem (a) apresenta as posições das partículas obtidas para o veículo em detalhe, enquanto a imagem (b) apresenta os vetores de movimento, diferentes ao longo do veículo. Nestes casos a fusão de agrupamentos pode agrupar corretamente as partículas em um único agrupamento.

A fusão de agrupamentos ocorre de forma semelhante à divisão. Quando a distância Euclidiana entre dois centroides é menor que a soma das maiores distâncias entre as partículas e os centroides dos dois agrupamentos envolvidos, o processo de fusão é inicializado. Inicialmente são tomadas as distâncias Euclidianas entre cores dos *pixels* do $quadro_t$ e do fundo, para a região convexa formada pelas partículas dos dois agrupamentos. Essas distâncias são limiarizadas com um valor inferior $T = 25$, e tratadas com o operador abertura com o mesmo elemento estruturante SE aplicado na divisão de clusters, Eq. (32), para eliminar pequenos agrupamentos de *pixels*.

Quando a região de similaridade com o fundo for pequena em relação à área da região convexa formada pelas partículas dos dois agrupamentos ($< 15\%$), a fusão desses dois agrupamentos é efetuada. A Figura 33 apresenta um exemplo de fusão de agrupamentos, onde um veículo apresenta dois agrupamentos de partículas inicialmente, e após a aplicação do processo os agrupamentos são fundidos.

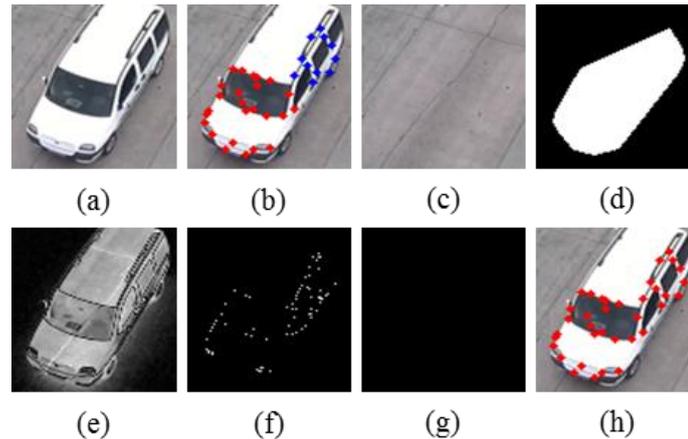


Figura 33: Exemplo de fusão de agrupamentos de partículas.

A Figura 33 apresenta as seguintes imagens: (a) detalhe do veículo processado; (b) dois agrupamentos gerados com as partículas pertencentes a um único veículo; (c) fundo da região em detalhe; (d) região convexa formada pelas partículas dos agrupamentos; (e) imagem da distância Euclidiana entre as cores dos *pixels* do $quadro_t$ e fundo para a região em detalhe; (f) limiarização da imagem (e), utilizando-se $T = 25$; (g) resultado da aplicação da operação de abertura, utilizando-se SE ; (h) resultado da fusão dos agrupamentos.

4.3.4 Detecção dos Veículos

A geração de agrupamentos de partículas não garante a existência de veículos, portanto, uma etapa de detecção dos veículos é implementada. Tal detecção é baseada em características morfológicas da região convexa formada pelas partículas pertencentes ao agrupamento analisado e na similaridade do interior e exterior dessa região com o fundo.

A morfologia das regiões convexas considera os aspectos: área, que deve ser grande o suficiente para compor um veículo; relação entre eixo maior e eixo menor da região, garantindo formas retangulares mais proporcionais, com o eixo maior no sentido do movimento; e taxa de ocupação de partículas, tendo o veículo partículas espalhadas ao longo de sua região. A similaridade com o fundo é analisada em duas regiões: na região interna ao veículo, devendo ser baixa; na região externa ao veículo, devendo ser alta, pois trata-se do fundo que circunda o veículo.

A Figura 34 fornece uma visão do processo de detecção de veículos, para um maior detalhamento o algoritmo referente à detecção de veículos encontra-se no Apêndice desta dissertação. Cada agrupamento é testado individualmente, sendo inicialmente obtidas as partículas pertencentes ao agrupamento analisado. Uma região convexa é gerada

considerando-se tais partículas, a qual será utilizada para a obtenção da área e da relação entre os eixos maior e menor da região. As partículas dilatadas, obtidas conforme explicado na Subseção 3.2.1, utilizando-se um elemento estruturante SE conforme a Eq. (33), e a região convexa formada a partir das partículas dilatadas são utilizadas para obter-se a taxa de ocupação de partículas.

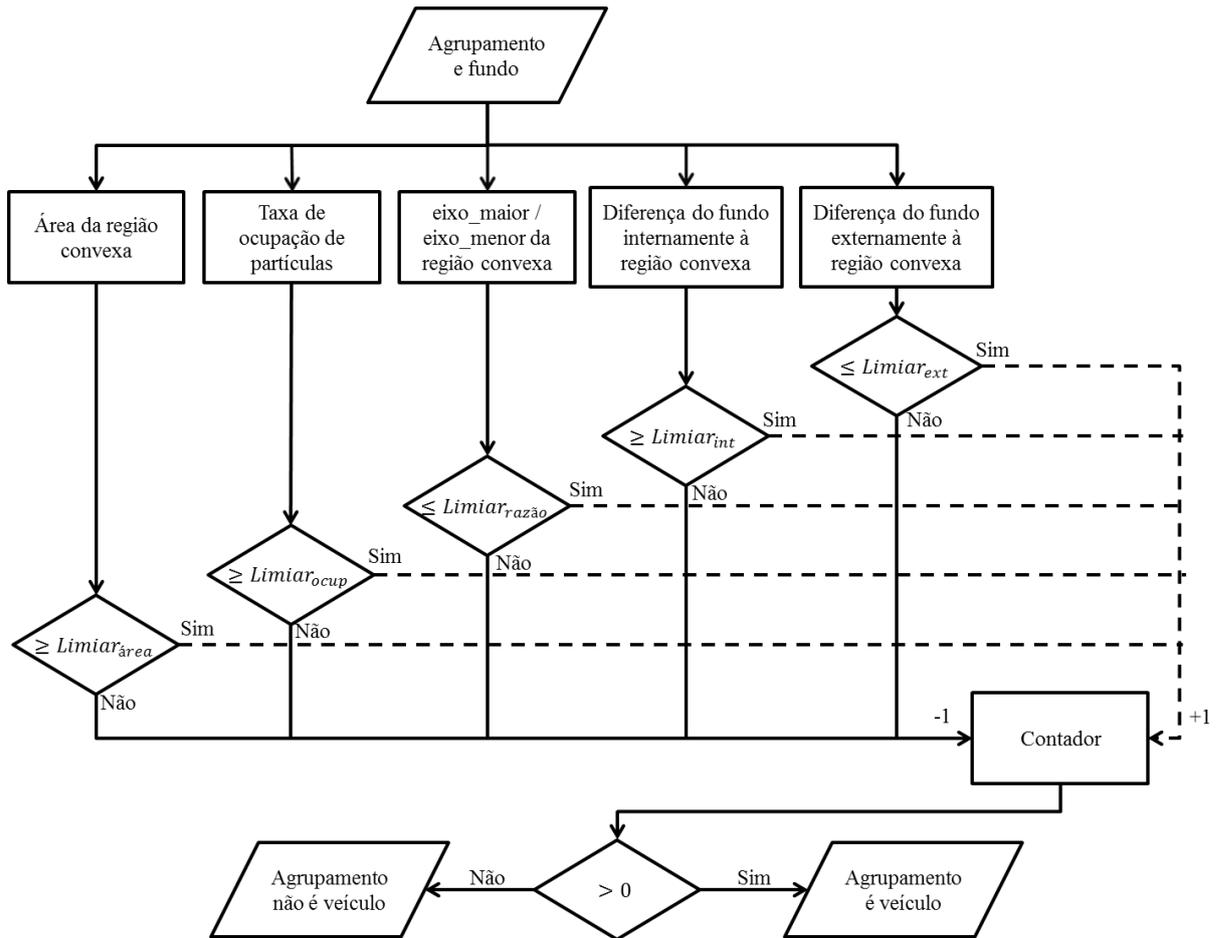


Figura 34: Fluxograma da detecção de veículos.

$$SE = \begin{bmatrix} 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \end{bmatrix} \quad (33)$$

A Figura 35 exemplifica o processo de obtenção dos parâmetros morfológicos para a identificação dos veículos, enquanto a Figura 36 exemplifica a comparação com o fundo nas regiões interna e externa à imagem convexa formada pelas partículas do agrupamento analisado.

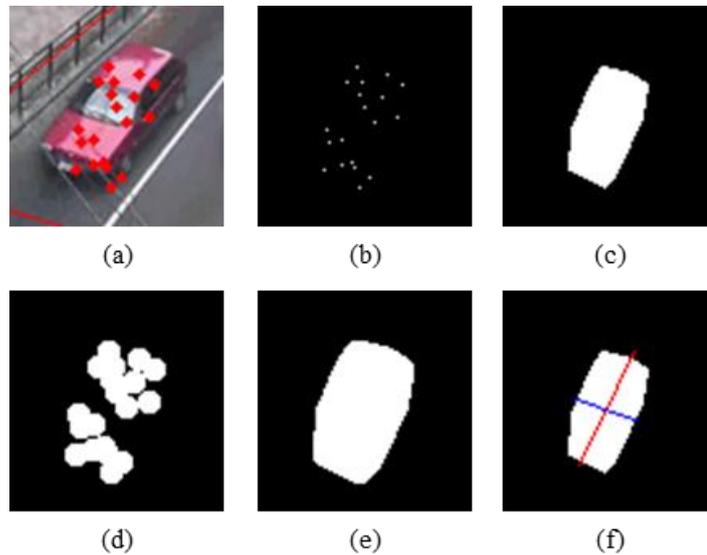


Figura 35: Forma do agrupamento.

Na Figura 35 as imagens representam: (a) agrupamento das partículas; (b) máscara binária com as partículas do agrupamento; (c) região convexa formada pelas partículas; (d) partículas dilatadas; (e) região convexa formada pelas partículas dilatadas; (f) eixos maior e menor da região convexa formada pelas partículas.

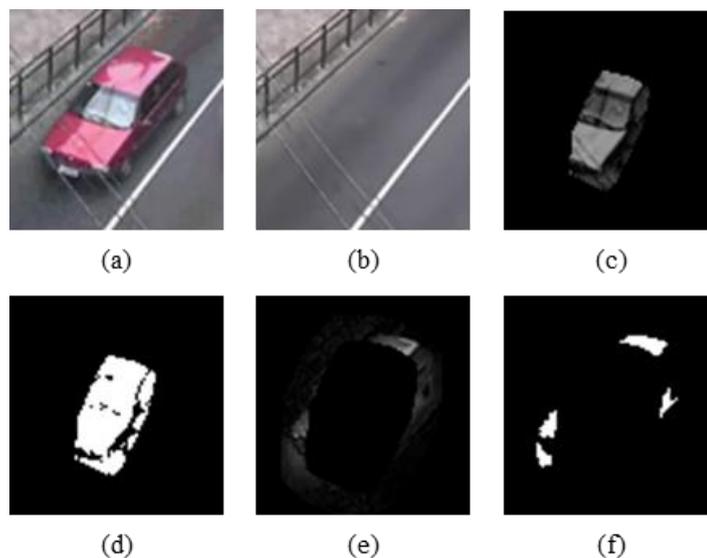


Figura 36: Análise do fundo na região do veículo.

Na Figura 36 as imagens representam: (a) veículo analisado; (b) fundo da região; (c) diferença do fundo na região convexa formada pelas partículas do agrupamento; (d) limiarização da diferença da região convexa; (e) diferença do fundo na região externa à região convexa; (f) limiarização da diferença da região externa à região convexa.

Os aspectos morfológicos da região convexa formada pelas partículas e as relações entre áreas de similaridade com o fundo com a área da região convexa são pontuadas positiva ou negativamente, conforme apresentado na Figura 34. Se a pontuação final ficar acima de zero um alvo é identificado, e uma caixa com um número de identificação é gerada, conforme a Figura 37.



Figura 37: Resultado da identificação de veículos.

4.3.5 Rastreamento dos Veículos

Os veículos detectados são rastreados através de similaridade de histogramas de cores das regiões em torno das partículas, conforme apresentado na Subseção 3.1.2. São geradas janelas de 9×9 *pixels* centradas nas partículas pertencentes aos agrupamentos identificados como veículos no $quadro_{t-1}$ e no $quadro_t$. As cores dessas janelas são requantizadas de 24 *bits* para 9 *bits*, reduzindo-se de 16,7 milhões para 512 cores, e então seus histogramas são tomados e comparados.

Cada partícula dos veículos detectados no $quadro_t$ é comparada com partículas dos veículos detectados no $quadro_{t-1}$. Partículas com similaridade superior a 0,8 são mantidas, enquanto partículas com similaridades inferiores a esse valor são eliminadas. Os identificadores de cada veículo vão sendo passados conforme as partículas pertencentes à mesma região convexa, deslocada de seus vetores de movimento, do $quadro_{t-1}$ para o $quadro_t$.

As partículas que pertencerem a intersecções entre duas regiões, são testadas em função da distância de Mahalanobis, conforme descrito na Subseção 3.7, para os centroides

dos agrupamentos dessas regiões. A menor distância seleciona o agrupamento ao qual a partícula testada pertence. A Figura 38 apresenta o processo de rastreamento de veículos baseado na comparação de histogramas de cor. Para maior detalhamento verificar o algoritmo referente ao rastreamento de veículos presente no Apêndice desta dissertação. A Figura 39 e Figura 40 apresentam respectivamente os histogramas de cor de dois pares de partículas selecionadas com similaridades de 0,9840, sendo consideradas a mesma partícula, e 0,6494, sendo consideradas partículas diferentes.

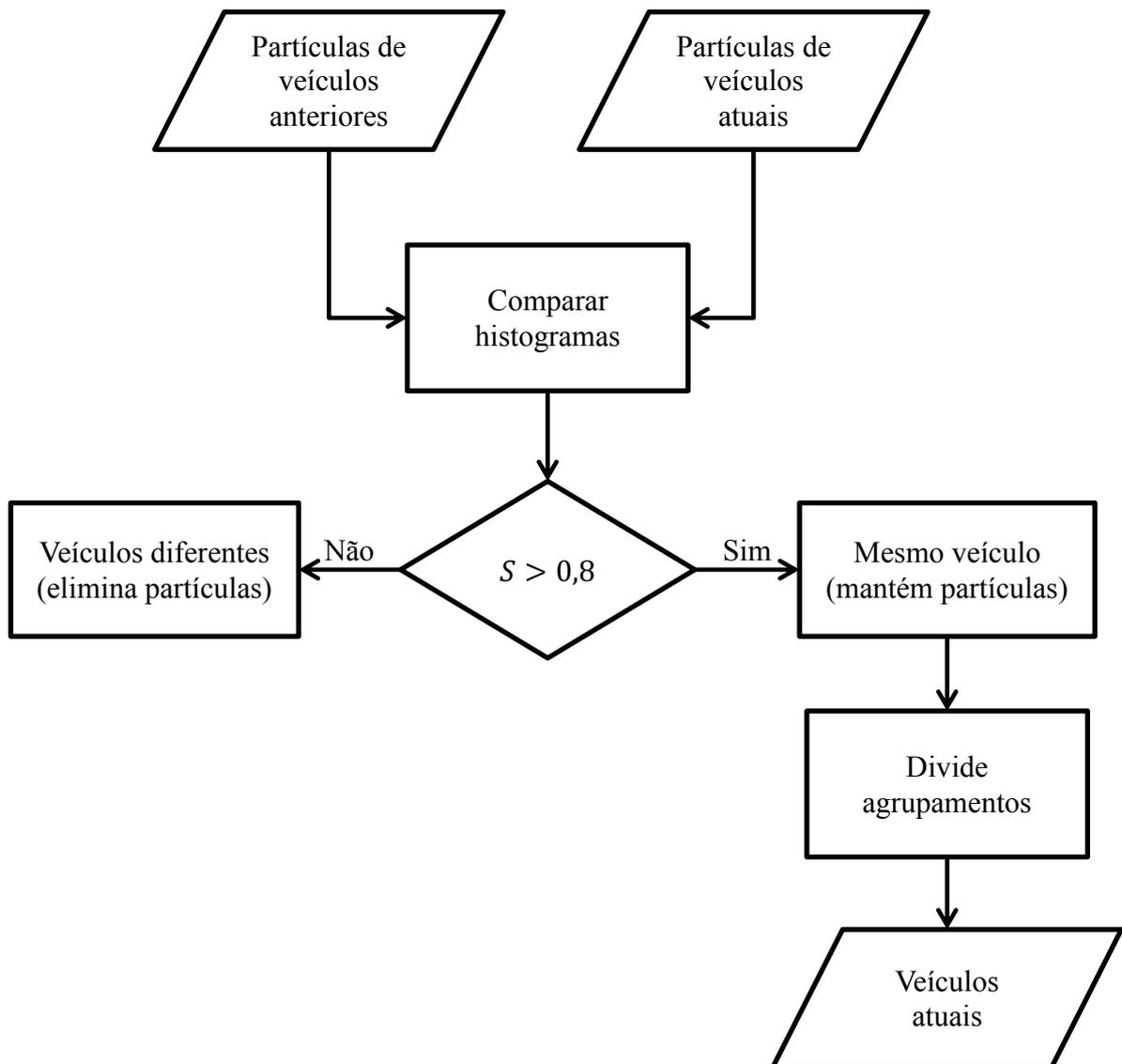


Figura 38: Fluxograma de rastreamento de veículos.

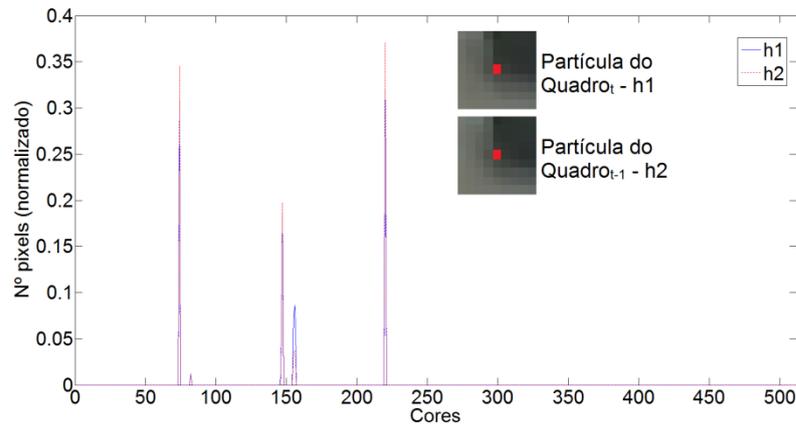


Figura 39: Similaridade de histograma de cor de 0,9840 entre duas partículas.

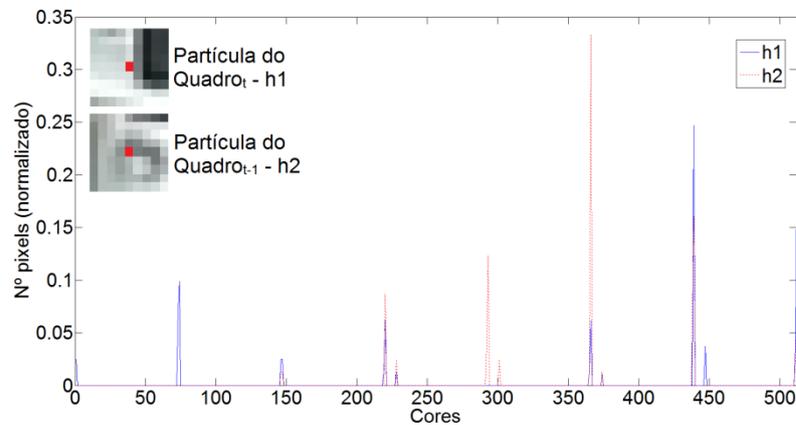


Figura 40: Similaridade de histograma de cor de 0,6494 entre duas partículas.

4.3.6 Contagem dos Veículos

A contagem dos veículos é baseada no resultado do rastreamento, ou seja, apenas veículos que foram detectados são contados. Os agrupamentos de partículas, mesmo que persistentes, não são contados se não forem identificados como veículos. A contagem é efetuada em laços virtuais definidos pelo usuário. Os laços podem ser desenhados em qualquer posição do quadro, mas levando-se em consideração que apenas veículos detectados são contados, a melhor posição para colocá-los é na região onde os veículos estão saindo da cena. Assim, a chance dos veículos serem detectados é maior, pois passaram por um número maior de quadros processados e foram validados antes de passar pelo laço virtual.

A contagem é baseada na intersecção das regiões convexas formadas pelas partículas rastreadas com as regiões dos laços virtuais. O laço virtual que possuir a maior área de

intersecção, para um determinado grupo de partículas rastreadas, tem seu contador incrementado, e essas partículas rastreadas não ativam mais os outros laços. Ou seja, cada agrupamento de partículas rastreadas, identificadas como veículos, é contado uma única vez, mesmo passando por mais de um laço. A Figura 41 exemplifica o processo de incremento do contador, e quando há intersecção das regiões o respectivo laço altera sua cor de vermelho para verde, facilitando a identificação de qual laço foi ativado. Na Figura 42 ilustra-se a ativação do laço virtual quando o veículo passa sobre mais de um laço.

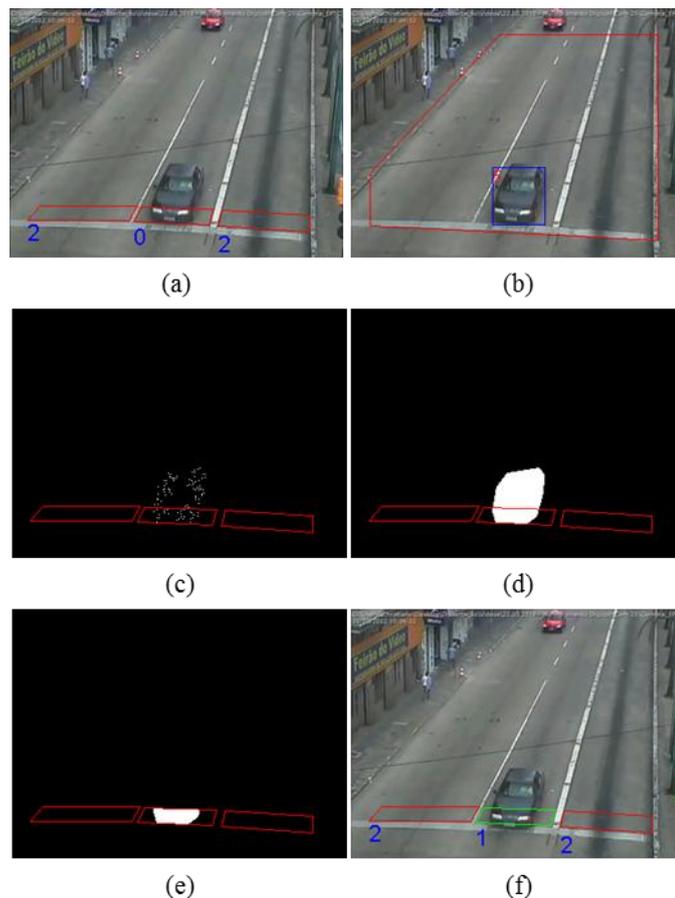


Figura 41: Processo de ativação do laço virtual e incremento do contador.

Na Figura 41 apresentam-se as seguintes imagens: (a) veículo sobre o laço virtual; (b) caixa de identificação do veículo rastreado; (c) partículas do veículo rastreado; (d) região convexa formada pelas partículas; (e) área de intersecção da região convexa com o laço virtual; (f) ativação do laço virtual e incremento do seu respectivo contador.

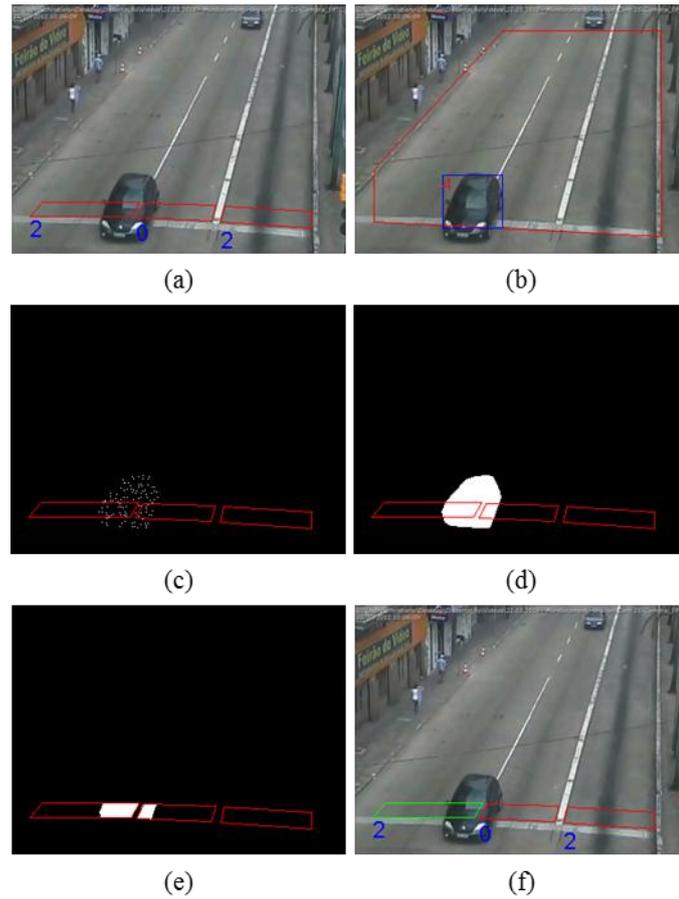


Figura 42: Passagem do veículo sobre mais de um laço virtual.

Na Figura 42 apresentam-se as seguintes imagens: (a) veículo sobre os laços virtuais; (b) caixa de identificação do veículo rastreado; (c) partículas do veículo rastreado; (d) região convexa formada pelas partículas; (e) áreas de intersecção da região convexa com os laços virtuais; (f) ativação do laço virtual com maior área de intersecção, não há incremento do contador porque o veículo já havia sido contado nos quadros anteriores.

5 RESULTADOS EXPERIMENTAIS

Conforme apresentado ao longo da dissertação, o presente trabalho consiste na contagem de veículos através da passagem de veículos rastreados por laços virtuais. Compararam-se os resultados obtidos na contagem dos veículos com um trabalho de abordagem semelhante, ou seja, utiliza partículas e informações do fundo (KIM, 2008) e um trabalho que utiliza uma abordagem baseada no rastreamento de componentes conectados formados através da obtenção dos objetos de interesse (SÁNCHEZ *et al.*, 2011).

A seguir, a Subseção 5.1 apresenta os resultados de rastreamento e contagem para cada um dos seis vídeos utilizados. São apresentadas imagens dos vídeos utilizados, seus respectivos fundos, e exemplos dos rastreamentos efetuados. Também são apresentadas tabelas com os resultados das contagens efetuadas para cada um dos métodos, contendo as diferenças numéricas e relativas à contagem efetuada por inspeção visual. Além da contagem dos veículos, são apresentadas mais duas tabelas, uma contendo o número de veículos que não foram contados e outra contendo o número de veículos que foram multiplamente contados. Os dados para a formação destas tabelas são obtidos apenas nas posições dos laços virtuais, para todos os métodos.

A Subseção 5.2 apresenta discussões em relação aos resultados apresentados, expondo justificativas para resultados adversos e propondo soluções para melhorar tais resultados.

5.1 RESULTADOS OBTIDOS

A Figura 43 apresenta um quadro do vídeo Cam06, com sua ROI, posição dos laços virtuais e seus respectivos contadores, e o fundo utilizado. A ROI foi definida de modo a abranger todas as pistas, evitando-se estender muito para a parte em que os veículos entram em cena, para evitar oclusões devido à perspectiva. Os laços virtuais foram posicionados dentro da ROI, um para cada pista, mas de forma que toda a extensão da via seja coberta, o mais próximo possível da saída dos veículos da ROI. O fundo, como era esperado, apresenta a imagem estática do vídeo em questão, com um aspecto mais limpo, propiciado pelo método de obtenção do mesmo, através do filtro de mediana do domínio temporal.

A Figura 44 apresenta o mapa das partículas iniciais, utilizadas na primeira etapa de filtragem das partículas da Cam06. Percebe-se que apenas as partículas na ROI são aparentes, e nas posições das faixas e dos limites da via a intensidade das partículas é maior,

significando que muitas partículas são obtidas na mesma posição espacial ao longo dos quadros e que, portanto, fazem parte do fundo, e como tal são eliminadas na primeira etapa de filtragem.

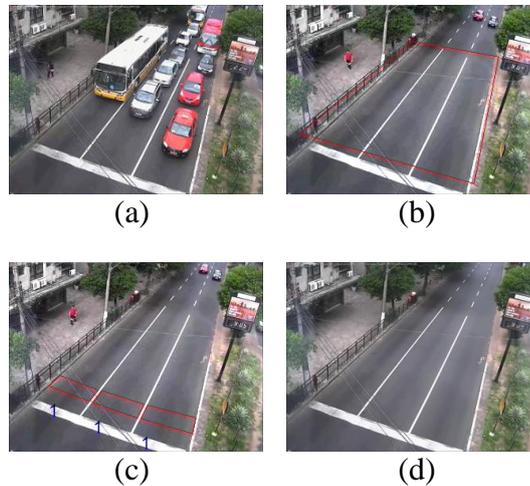


Figura 43: Quadro do vídeo Cam06 (a), ROI (b), laços virtuais (c) e fundo (d) utilizados.

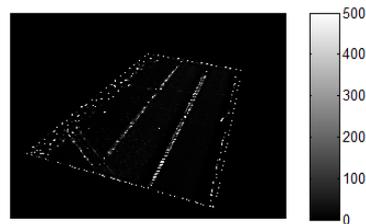


Figura 44: Mapa de repetitividade das partículas iniciais da Cam06.

A Figura 45 apresenta uma situação normal de rastreamento da Cam06, em que vários veículos entram e saem da ROI, em vermelho. Todos são veículos de porte médio (carros de passeio) e apresentam distâncias entre si que facilitam a detecção individual. São apresentados quadros desde a entrada de alguns veículos na ROI até a sua saída completa, com suas respectivas caixas de identificação em azul com números identificadores em vermelho. A Figura 46 apresenta a ativação (mudança para verde) dos laços virtuais (vermelhos) e suas respectivas contagens (em azul) para alguns veículos rastreados da Figura 45, desde suas entradas nos laços até suas saídas completas.

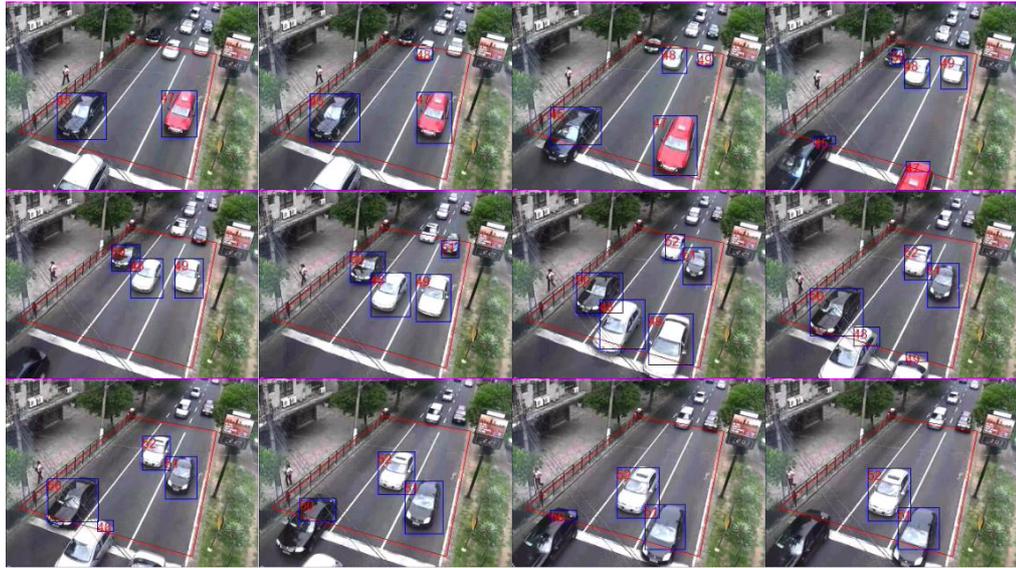


Figura 45: Rastreamento de veículos em situação de trânsito normal na Cam06.



Figura 46: Ativação dos laços virtuais e incremento dos respectivos contadores na Cam06.

A Figura 47 apresenta uma situação, durante o rastreamento na Cam06, em que os veículos acabam gerando oclusões devido à parada no semáforo, obtendo-se apenas uma caixa de identificação para todos os veículos na ROI. A situação normaliza-se quando os veículos retomam o movimento, gerando novas caixas de identificação. A Figura 48 apresenta uma situação semelhante à anterior na Cam06, em que vários veículos são considerados um só, mas neste caso a ocorrência é devido à proximidade dos veículos.

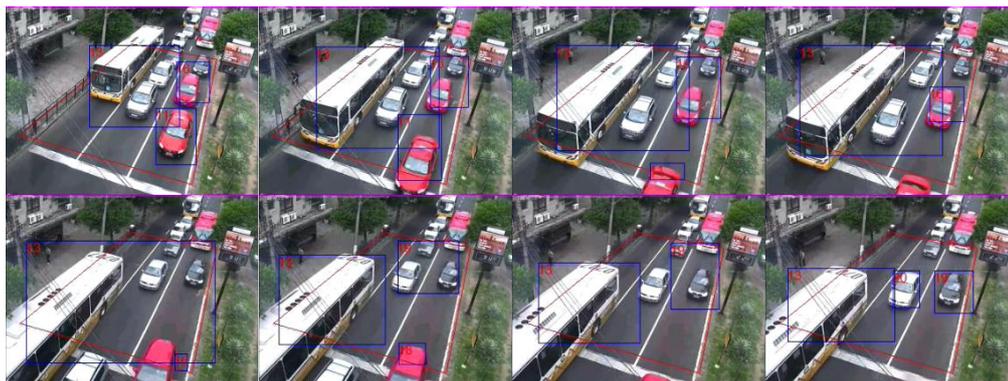


Figura 47: Rastreamento em situação de proximidade e oclusão parcial devido à parada no semáforo na Cam06.



Figura 48: Rastreamento em situação de proximidade inicial dos veículos envolvidos na Cam06.

A Tabela 1 apresenta o resultado da contagem (em unidades) efetuada nos laços virtuais pelo método proposto para a Cam06, juntamente com as contagens realizadas pelos métodos comparativos e a contagem real, obtida por inspeção visual. As diferenças numéricas (em unidades) entre as contagens obtidas pelos respectivos métodos e as contagens reais, bem como as diferenças relativas (em porcentagem), que são as razões entre as diferenças numéricas e as contagens reais, também são apresentadas. As diferenças totais correspondem às diferenças das somas dos resultados das contagens nos laços virtuais, para cada método, em relação ao número total de veículos que passaram pela ROI durante o processamento.

Tabela 1: Contagem de veículos – Cam06.

		Pista 1	Pista 2	Pista 3	Totais
Inspeção Visual	Contagem (un.)	44	92	89	225
Método Proposto	Contagem (un.)	38	90	89	217
	Diferenças (un.)	-6	-2	0	-8
	Dif. Relativas (%)	-13,6	-2,2	0	-3,6
(KIM, 2008)	Contagem (un.)	83	103	80	266
	Diferenças (un.)	39	11	-9	41
	Dif. Relativas (%)	88,6	11,9	-10,1	18,2
(SÁNCHEZ, et al., 2011)	Contagem (un.)	52	100	96	248
	Diferenças (un.)	8	8	7	23
	Dif. Relativas (%)	18,2	8,7	7,9	10,2

A Tabela 2 apresenta o número de veículos perdidos por cada um dos métodos em cada pista da Cam06, enquanto a Tabela 3 apresenta o número de veículos que foram contados além do valor real. Apesar de os métodos não fazerem tal distinção, os veículos destas tabelas foram classificados em motos (veículos de duas ou três rodas), carros (veículos de passeio e caminhonetes) e caminhões (ônibus, caminhões e furgões), para que a análise dos resultados seja mais aprofundada.

Tabela 2: Veículos não detectados – Cam06.

		Motos	Carros	Caminhões	Totais
Método Proposto	Pista 1	4	3	1	8
	Pista 2	1	4	1	6
	Pista 3	2	1	0	3
	Totais	7	8	2	17
(KIM, 2008)	Pista 1	1	0	0	1
	Pista 2	0	4	0	4
	Pista 3	2	9	0	11
	Totais	3	13	0	16
(SÁNCHEZ, et al., 2011)	Pista 1	5	3	2	10
	Pista 2	0	5	1	6
	Pista 3	0	9	1	10
	Totais	5	17	4	26

Tabela 3: Veículos multiplamente detectados – Cam06.

		Motos	Carros	Caminhões	Totais
Método Proposto	Pista 1	0	1	1	2
	Pista 2	0	2	2	4
	Pista 3	0	3	0	3
	Totais	0	6	3	9
(KIM, 2008)	Pista 1	0	8	32	40
	Pista 2	0	8	7	15
	Pista 3	0	1	1	2
	Totais	0	17	40	57
(SÁNCHEZ, et al., 2011)	Pista 1	0	11	7	18
	Pista 2	1	5	8	14
	Pista 3	1	14	2	17
	Totais	2	30	17	49

A Figura 49 apresenta um quadro do vídeo Cam25, com sua ROI, posição dos laços virtuais e seus respectivos contadores, e o fundo utilizado. A ROI foi definida de modo a abranger todas as pistas e sua extensão cobre praticamente toda a cena, excetuando-se os passeios laterais. Os laços virtuais foram posicionados dentro da ROI, um para cada pista, mas de forma que toda a extensão da via seja coberta, o mais próximo possível da saída dos veículos da ROI. O fundo, como era esperado, apresenta a imagem estática do vídeo em questão, com um aspecto mais limpo, propiciado pelo método de obtenção do mesmo, através do filtro de mediana do domínio temporal.

A Figura 50 apresenta o mapa das partículas iniciais, utilizadas na primeira etapa de filtragem das partículas. Percebe-se que apenas as partículas na ROI são aparentes, e nas posições das faixas e dos limites da via, bem como ao longo de um cabo aéreo que cruza a via, a intensidade das partículas é maior, significando que muitas partículas são obtidas na mesma posição espacial ao longo dos quadros e que, portanto, fazem parte do fundo, e como tal são eliminadas na primeira etapa de filtragem.

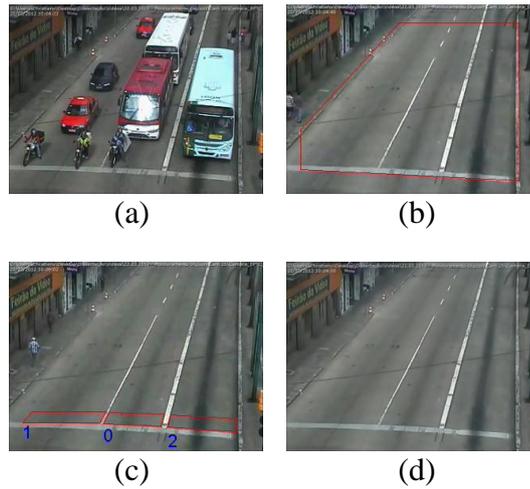


Figura 49: Quadro do vídeo Cam25 (a), ROI (b), laços virtuais (c) e fundo (d) utilizados.

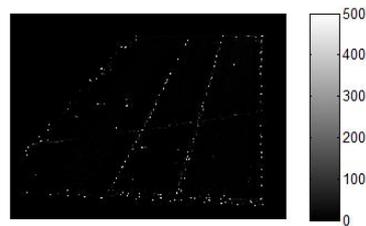


Figura 50: Mapa de repetitividade das partículas iniciais da Cam25.

A Figura 51 apresenta o rastreamento na Cam25 quando a situação inicial é os veículos estarem parados devido à sinalização de trânsito. Já a Figura 52 apresenta a detecção de uma motocicleta na Cam25, que somente ocorre quando ela está deixando a ROI.

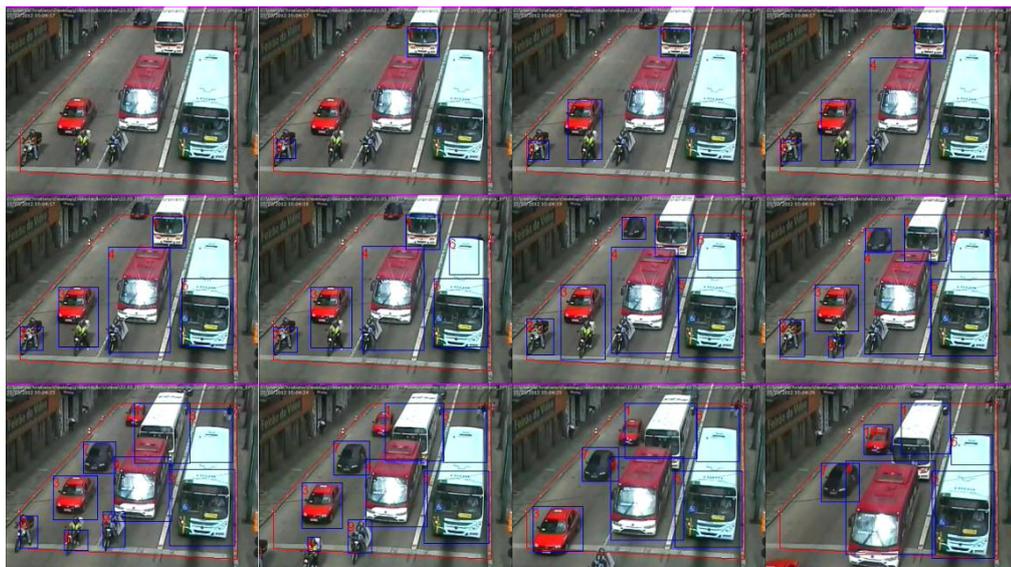


Figura 51: Rastreamento durante os primeiros quadros, com identificação inicial dos veículos na Cam25.

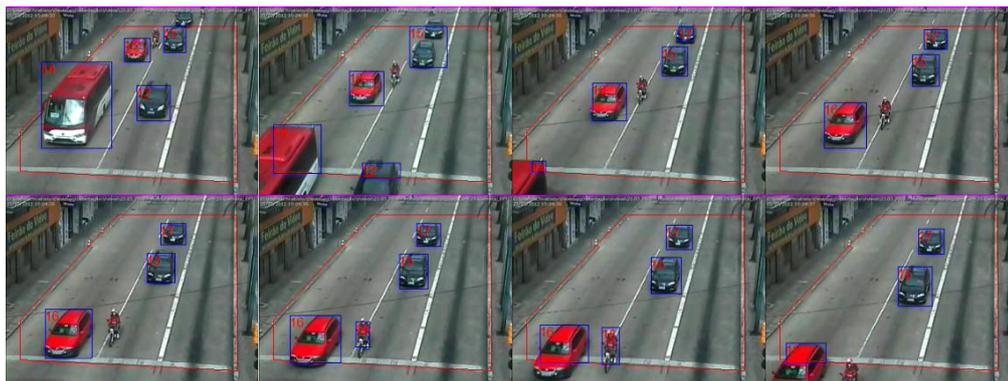


Figura 52: Rastreamento de uma motocicleta, identificada apenas na saída da ROI na Cam25.

A Tabela 4 apresenta o resultado da contagem (em unidades) efetuada nos laços virtuais pelo método proposto para a Cam25, juntamente com as contagens realizadas pelos métodos comparativos e a contagem real, obtida por inspeção visual. As diferenças numéricas (em unidades) entre as contagens obtidas pelos respectivos métodos e as contagens reais, bem como as diferenças relativas (em porcentagem), que são as razões entre as diferenças numéricas e as contagens reais, também são apresentadas. As diferenças totais correspondem às diferenças das somas dos resultados das contagens nos laços virtuais, para cada método, em relação ao número total de veículos que passaram pela ROI durante o processamento.

Tabela 4: Contagem de veículos – Cam25.

		Pista 1	Pista 2	Pista 3	Totais
Inspeção Visual	Contagem (un.)	97	68	24	189
Método Proposto	Contagem (un.)	95	67	24	186
	Diferenças (un.)	-2	-1	0	-3
	Dif. Relativas (%)	-2,1	-1,5	0	-1,6
(KIM, 2008)	Contagem (un.)	85	65	39	189
	Diferenças (un.)	-12	-3	15	0
	Dif. Relativas (%)	-12,4	-4,4	62,5	0
(SÁNCHEZ, et al., 2011)	Contagem (un.)	118	74	22	214
	Diferenças (un.)	21	6	-2	25
	Dif. Relativas (%)	21,6	8,8	-8,3	13,2

A Tabela 5 apresenta o número de veículos perdidos por cada um dos métodos em cada pista, enquanto a Tabela 6 apresenta o número de veículos que foram contados além do valor real. Apesar de os métodos não fazerem tal distinção, os veículos destas tabelas foram classificados em motos (veículos de duas ou três rodas), carros (veículos de passeio e

caminhonetes) e caminhões (ônibus, caminhões e furgões), para que a análise dos resultados seja mais aprofundada.

Tabela 5: Veículos não detectados – Cam25.

		Motos	Carros	Caminhões	Totais
Método Proposto	Pista 1	2	3	0	5
	Pista 2	0	2	1	3
	Pista 3	0	0	1	1
	Totais	2	5	2	9
(KIM, 2008)	Pista 1	9	4	2	15
	Pista 2	1	3	1	5
	Pista 3	0	0	0	0
	Totais	10	7	3	20
(SÁNCHEZ, et al., 2011)	Pista 1	4	16	1	21
	Pista 2	0	9	2	11
	Pista 3	0	0	5	5
	Totais	4	25	8	37

Tabela 6: Veículos múltiplemente detectados – Cam25.

		Motos	Carros	Caminhões	Totais
Método Proposto	Pista 1	3	0	0	3
	Pista 2	0	2	0	2
	Pista 3	0	0	1	1
	Totais	3	2	1	6
(KIM, 2008)	Pista 1	1	0	2	3
	Pista 2	0	1	1	2
	Pista 3	0	0	15	15
	Totais	1	1	18	20
(SÁNCHEZ, et al., 2011)	Pista 1	3	34	5	42
	Pista 2	4	10	3	17
	Pista 3	0	1	2	3
	Totais	7	45	10	62

A Figura 53 apresenta um quadro do vídeo Cam41, com sua ROI, posição dos laços virtuais e seus respectivos contadores, e o fundo utilizado. A ROI foi definida de modo a abranger todas as pistas e sua extensão cobre praticamente toda a cena. Não utilizou-se toda a cena como ROI devido a um defeito na obtenção das imagens que causa um borrimento na parte inferior da cena, conforme apresentado na Figura 54.

Os laços virtuais foram posicionados dentro da ROI, um para cada pista, mas de forma que toda a extensão da via seja coberta, o mais próximo possível da saída dos veículos da ROI, mas sem sobrepor-se a parte defeituosa da imagem. O fundo, como era esperado, apresenta a imagem de fundo do vídeo em questão, com um aspecto mais limpo, propiciado pelo método de obtenção do mesmo, através do filtro de mediana do domínio temporal.

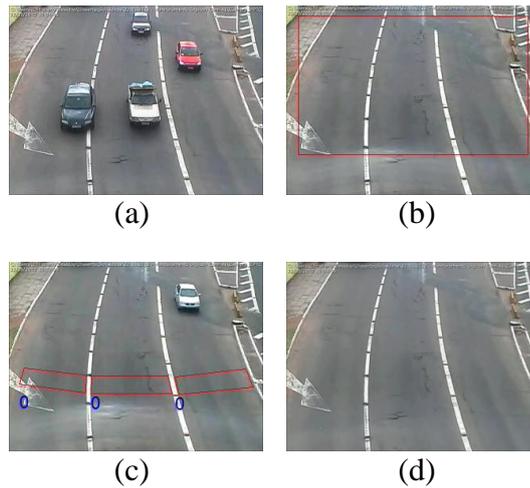


Figura 53: Quadro do vídeo Cam41 (a), ROI (b), laços virtuais (c) e fundo (d) utilizados.



Figura 54: Borrramento na obtenção das imagens da Cam41.

A Figura 55 apresenta o mapa das partículas iniciais da Cam41, utilizadas na primeira etapa de filtragem das partículas. Percebe-se que apenas as partículas na ROI são aparentes, e nas posições das faixas e dos limites da via, bem como no passeio lateral e na divisória das pistas, a intensidade das partículas é maior, significando que muitas partículas são obtidas na mesma posição espacial ao longo dos quadros e que, portanto, fazem parte do fundo, e como tal são eliminadas na primeira etapa de filtragem.

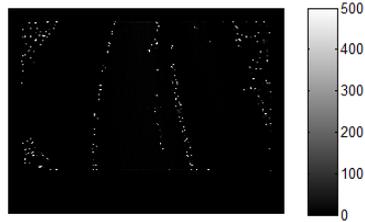


Figura 55: Mapa de repetitividade das partículas iniciais da Cam41.

A Figura 56 apresenta uma situação da Cam41 em que um caminhão carregando uma escavadeira entra em cena, sendo rastreado como um único objeto, apesar da irregularidade da forma do conjunto e do tamanho do objeto cobrir praticamente toda a ROI. A Figura 57 apresenta outra situação em que um veículo grande entra em cena, mas é inicialmente dividido em dois objetos, sendo que ao longo do processo um dos objetos é eliminado.

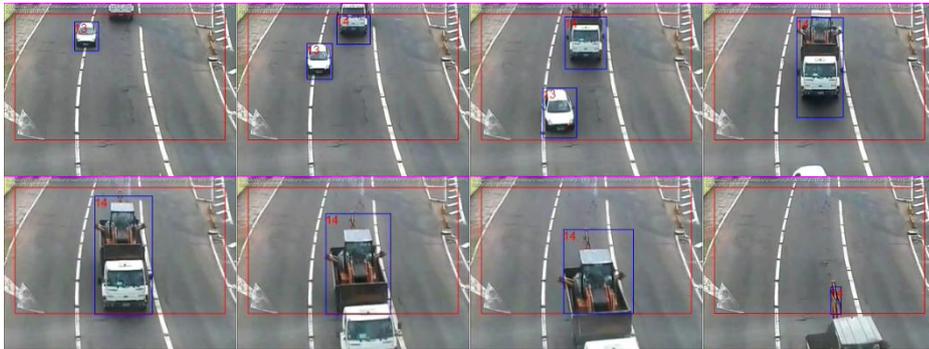


Figura 56: Rastreamento de um conjunto caminhão – escavadeira na Cam41.



Figura 57: Rastreamento de um veículo grande na Cam41, inicialmente identificado como dois objetos.

A Tabela 7 apresenta o resultado da contagem (em unidades) efetuada nos laços virtuais pelo método proposto para a Cam41, juntamente com as contagens realizadas pelos métodos comparativos e a contagem real, obtida por inspeção visual. As diferenças numéricas (em unidades) entre as contagens obtidas pelos respectivos métodos e as contagens reais, bem

como as diferenças relativas (em porcentagem), que são as razões entre as diferenças numéricas e as contagens reais, também são apresentadas. As diferenças totais correspondem às diferenças das somas dos resultados das contagens nos laços virtuais, para cada método, em relação ao número total de veículos que passaram pela ROI durante o processamento.

Tabela 7: Contagem de veículos – Cam41.

		Pista 1	Pista 2	Pista 3	Totais
Inspeção Visual	Contagem (un.)	10	157	84	251
Método Proposto	Contagem (un.)	10	164	84	258
	Diferenças (un.)	0	7	0	7
	Dif. Relativas (%)	0	4,5	0	2,8
(KIM, 2008)	Contagem (un.)	10	185	81	276
	Diferenças (un.)	0	28	-3	25
	Dif. Relativas (%)	0	17,8	-3,6	9,9
(SÁNCHEZ, et al., 2011)	Contagem (un.)	11	132	83	226
	Diferenças (un.)	1	-25	-1	-25
	Dif. Relativas (%)	10,0	-15,9	-1,2	-9,9

A Tabela 8 apresenta o número de veículos perdidos por cada um dos métodos em cada pista, enquanto a Tabela 9 apresenta o número de veículos que foram contados além do valor real. Apesar de os métodos não fazerem tal distinção, os veículos destas tabelas foram classificados em motos (veículos de duas ou três rodas), carros (veículos de passeio e caminhonetes) e caminhões (ônibus, caminhões e furgões), para que a análise dos resultados seja mais aprofundada.

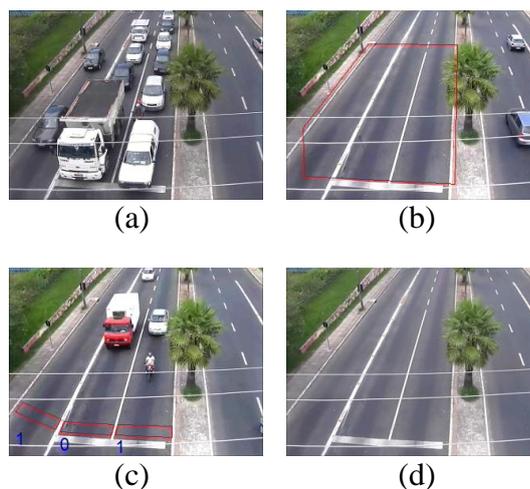
Tabela 8: Veículos não detectados – Cam41.

		Motos	Carros	Caminhões	Totais
Método Proposto	Pista 1	0	0	0	0
	Pista 2	1	0	0	1
	Pista 3	0	1	0	1
	Totais	1	1	0	2
(KIM, 2008)	Pista 1	0	0	0	0
	Pista 2	2	5	0	7
	Pista 3	0	4	1	5
	Totais	2	9	1	12
(SÁNCHEZ, et al., 2011)	Pista 1	0	0	0	0
	Pista 2	0	18	7	25
	Pista 3	0	1	0	1
	Totais	0	19	7	26

Tabela 9: Veículos multiplamente detectados – Cam41.

		Motos	Carros	Caminhões	Totais
Método Proposto	Pista 1	0	0	0	0
	Pista 2	0	2	6	8
	Pista 3	0	0	1	1
	Totais	0	2	7	9
(KIM, 2008)	Pista 1	0	0	0	0
	Pista 2	0	1	34	35
	Pista 3	0	0	2	2
	Totais	0	1	36	37
(SÁNCHEZ, et al., 2011)	Pista 1	0	0	1	1
	Pista 2	0	0	0	0
	Pista 3	0	0	0	0
	Totais	0	0	1	1

A Figura 58 apresenta um quadro do vídeo Cam64, com sua ROI, posição dos laços virtuais e seus respectivos contadores, e o fundo utilizado. A ROI foi definida de modo a abranger todas as pistas do lado esquerdo, ou seja, pegando o fluxo de veículos no sentido de cima para baixo na cena. Os laços virtuais foram posicionados dentro da ROI, um para cada pista, mas de forma que toda a extensão da via seja coberta, o mais próximo possível da saída dos veículos da ROI. O fundo, como era esperado, apresenta a imagem de fundo do vídeo em questão, com um aspecto mais limpo, propiciado pelo método de obtenção do mesmo, através do filtro de mediana do domínio temporal.

**Figura 58: Quadro do vídeo Cam64 (a), ROI (b), laços virtuais (c) e fundo (d) utilizados.**

A Figura 59 apresenta o mapa das partículas iniciais da Cam64, utilizadas na primeira etapa de filtragem das partículas. Percebe-se que apenas as partículas na ROI são aparentes, e

nas posições das faixas e dos limites da via, bem como nas folhas da árvore que invadem a ROI, a intensidade das partículas é maior, significando que muitas partículas são obtidas na mesma posição espacial ao longo dos frames e que, portanto, fazem parte do fundo, e como tal são eliminadas na primeira etapa de filtragem.



Figura 59: Mapa de repetitividade das partículas iniciais da Cam64.

A Tabela 10 apresenta o resultado da contagem (em unidades) efetuada nos laços virtuais pelo método proposto para a Cam64, juntamente com as contagens realizadas pelos métodos comparativos e a contagem real, obtida por inspeção visual. As diferenças numéricas (em unidades) entre as contagens obtidas pelos respectivos métodos e as contagens reais, bem como as diferenças relativas (em porcentagem), que são as razões entre as diferenças numéricas e as contagens reais, também são apresentadas. As diferenças totais correspondem às diferenças das somas dos resultados das contagens nos laços virtuais, para cada método, em relação ao número total de veículos que passaram pela ROI durante o processamento.

Tabela 10: Contagem de veículos – Cam64.

		Pista 1	Pista 2	Pista 3	Totais
Ispeção Visual	Contagem (un.)	74	36	55	165
Método Proposto	Contagem (un.)	47	34	52	133
	Diferenças (un.)	-27	-2	-3	-32
	Dif. Relativas (%)	-36,5	-5,6	-5,4	-19,4
(KIM, 2008)	Contagem (un.)	69	48	44	161
	Diferenças (un.)	-5	12	-11	-4
	Dif. Relativas (%)	-6,8	33,3	-20,0	-2,4
(SÁNCHEZ, et al., 2011)	Contagem (un.)	69	33	62	164
	Diferenças (un.)	-5	-3	7	-1
	Dif. Relativas (%)	-6,8	-8,3	12,7	-0,6

A Tabela 11 apresenta o número de veículos perdidos por cada um dos métodos em cada pista para a Cam64, enquanto a

Tabela 12 apresenta o número de veículos que foram contados além do valor real. Apesar de os métodos não fazerem tal distinção, os veículos destas tabelas foram classificados em motos (veículos de duas ou três rodas), carros (veículos de passeio e caminhonetes) e caminhões (ônibus, caminhões e furgões), para que a análise dos resultados seja mais aprofundada.

Tabela 11: Veículos não detectados – Cam64.

		Motos	Carros	Caminhões	Totais
Método Proposto	Pista 1	7	21	4	32
	Pista 2	2	1	2	5
	Pista 3	6	3	0	9
	Totais	15	25	6	46
(KIM, 2008)	Pista 1	2	8	0	10
	Pista 2	0	2	0	2
	Pista 3	3	8	1	12
	Totais	5	18	1	24
(SÁNCHEZ, et al., 2011)	Pista 1	8	13	0	21
	Pista 2	1	6	3	10
	Pista 3	7	7	0	14
	Totais	16	26	3	45

Tabela 12: Veículos multiplamente detectados – Cam64.

		Motos	Carros	Caminhões	Totais
Método Proposto	Pista 1	0	1	4	5
	Pista 2	0	2	1	3
	Pista 3	0	6	0	6
	Totais	0	9	5	14
(KIM, 2008)	Pista 1	0	1	4	5
	Pista 2	0	5	9	14
	Pista 3	0	1	0	1
	Totais	0	7	13	20
(SÁNCHEZ, et al., 2011)	Pista 1	1	8	7	16
	Pista 2	0	5	2	7
	Pista 3	2	17	2	21
	Totais	3	30	11	44

A Figura 60 apresenta um exemplo de rastreamento na Cam64 em que vários veículos são perdidos na pista da esquerda, devido à oclusão por um ônibus parado na pista central, ao

mesmo tempo o ônibus, juntamente com os veículos parados próximos, é tratado como um único veículo.

A Figura 61 apresenta a separação dos veículos do ônibus rastreado na Figura 60, conforme ocorre a movimentação. Notar que todos os veículos acabam sendo rastreados separadamente ao final do processo, com exceção da motocicleta que é um veículo menor e acaba ficando muito próxima do ônibus.

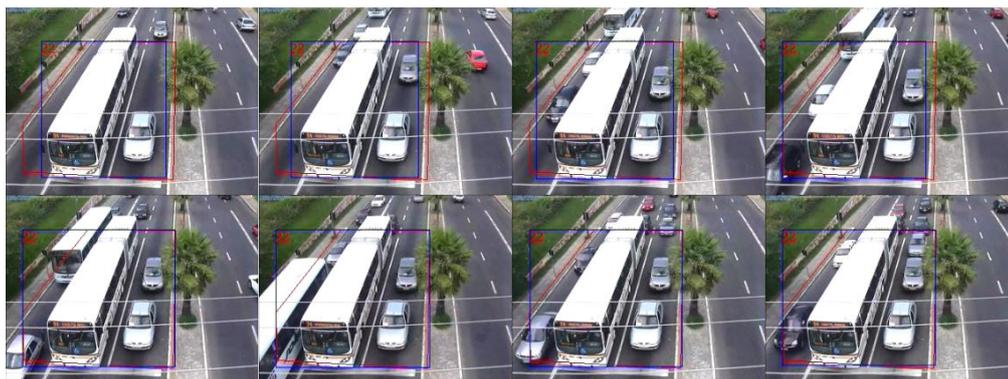


Figura 60: Exemplo de veículos perdidos durante rastreamento na Cam64.

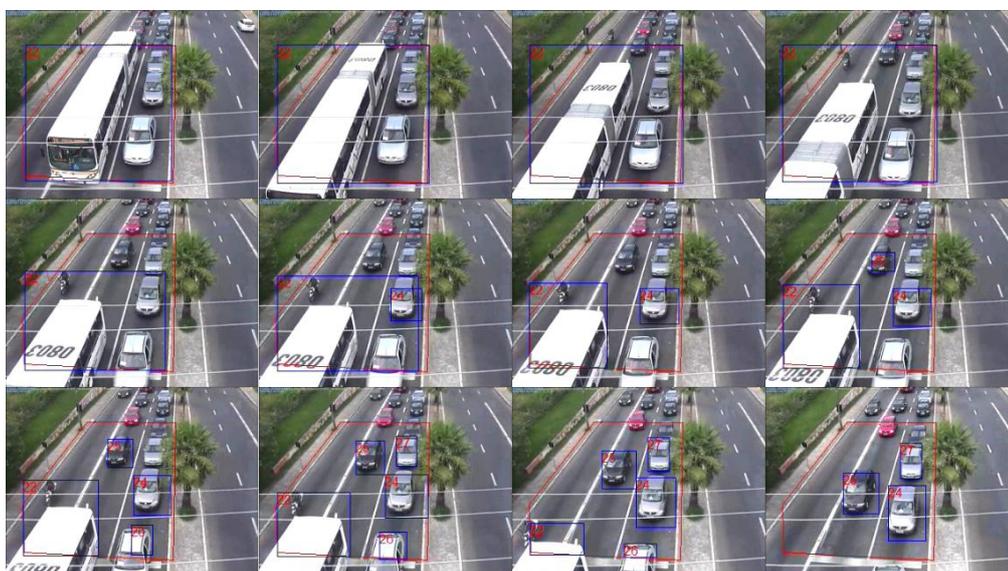


Figura 61: Separação dos veículos rastreados conforme início da movimentação na Cam64.

A Figura 62 apresenta um quadro do vídeo Cam68, com sua ROI, posição dos laços virtuais e seus respectivos contadores, e o fundo utilizado. A ROI foi definida de modo a abranger as duas pistas. Os laços virtuais foram posicionados dentro da ROI, um para cada pista, mas de forma que toda a extensão da via seja coberta, o mais próximo possível da saída dos veículos da ROI. O fundo, como era esperado, apresenta a imagem de fundo do vídeo em

questão, com um aspecto mais limpo, propiciado pelo método de obtenção do mesmo, através do filtro de mediana do domínio temporal.

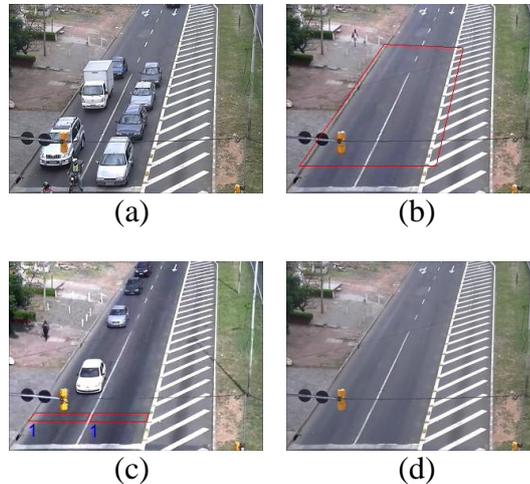


Figura 62: Quadro do vídeo Cam68 (a), ROI (b), laços virtuais (c) e fundo (d) utilizados.

A Figura 63 apresenta o mapa das partículas iniciais da Cam68, utilizadas na primeira etapa de filtragem das partículas. Percebe-se que apenas as partículas na ROI são aparentes, e nas posições das faixas e dos limites da via, a intensidade das partículas é maior, significando que muitas partículas são obtidas na mesma posição espacial ao longo dos quadros e que, portanto, fazem parte do fundo, e como tal são eliminadas na primeira etapa de filtragem.

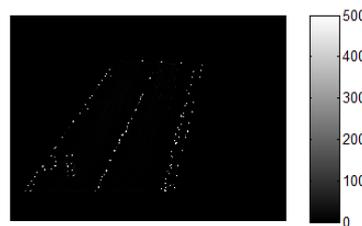


Figura 63: Mapa de repetitividade das partículas iniciais da Cam68.

A Tabela 13 apresenta o resultado da contagem (em unidades) efetuada nos laços virtuais pelo método proposto para a Cam68, juntamente com as contagens realizadas pelos métodos comparativos e a contagem real, obtida por inspeção visual. As diferenças numéricas (em unidades) entre as contagens obtidas pelos respectivos métodos e as contagens reais, bem como as diferenças relativas (em porcentagem), que são as razões entre as diferenças numéricas e as contagens reais, também são apresentadas. As diferenças totais correspondem às diferenças das somas dos resultados das contagens nos laços virtuais, para cada método, em relação ao número total de veículos que passaram pela ROI durante o processamento.

Tabela 13: Contagem de veículos – Cam68.

		Pista 1	Pista 2	Totais
Inspeção Visual	Contagem (un.)	66	88	154
Método Proposto	Contagem (un.)	63	72	135
	Diferenças (un.)	-3	-16	-19
	Dif. Relativas (%)	-4,6	-18,2	-12,3
(KIM, 2008)	Contagem (un.)	83	91	174
	Diferenças (un.)	17	3	20
	Dif. Relativas (%)	25,8	3,4	13,0
(SÁNCHEZ, et al., 2011)	Contagem (un.)	66	96	162
	Diferenças (un.)	0	8	8
	Dif. Relativas (%)	0	9,1	5,2

A Tabela 14 apresenta o número de veículos perdidos por cada um dos métodos em cada pista para a Cam68, enquanto a Tabela 15 apresenta o número de veículos que foram contados além do valor real. Apesar de os métodos não fazerem tal distinção, os veículos destas tabelas foram classificados em motos (veículos de duas ou três rodas), carros (veículos de passeio e caminhonetes) e caminhões (ônibus, caminhões e furgões), para que a análise dos resultados seja mais aprofundada.

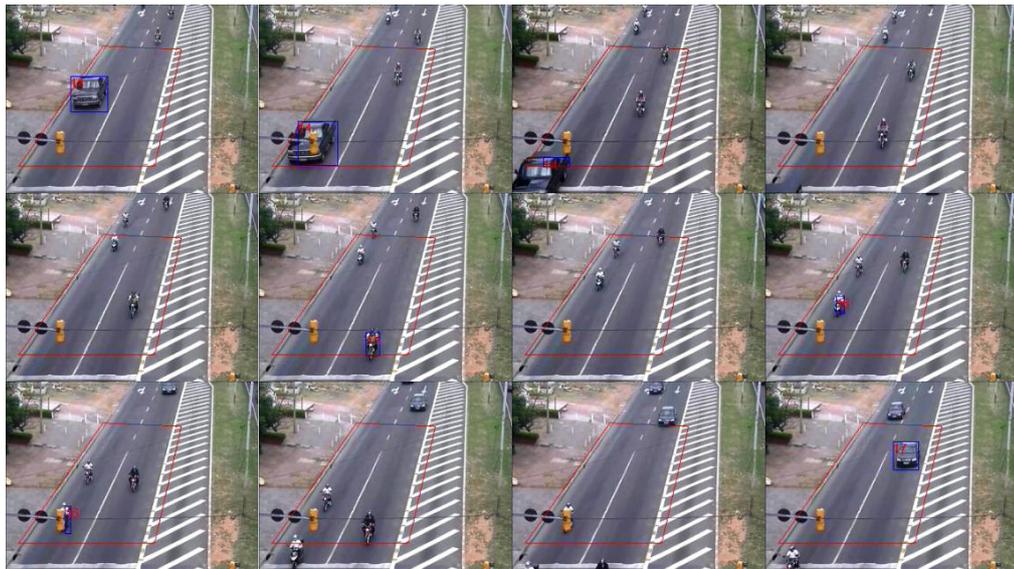
Tabela 14: Veículos não detectados – Cam68.

		Motos	Carros	Caminhões	Totais
Método Proposto	Pista 1	5	4	0	9
	Pista 2	14	1	1	16
	Totais	19	5	1	25
(KIM, 2008)	Pista 1	1	0	1	2
	Pista 2	2	3	0	5
	Totais	3	3	1	7
(SÁNCHEZ, et al., 2011)	Pista 1	4	14	2	20
	Pista 2	11	8	2	21
	Totais	15	22	4	41

Tabela 15: Veículos multiplamente detectados – Cam68.

		Motos	Carros	Caminhões	Totais
Método Proposto	Pista 1	0	3	3	6
	Pista 2	0	0	0	0
	Totais	0	3	3	6
(KIM, 2008)	Pista 1	0	5	14	19
	Pista 2	0	0	8	8
	Totais	0	5	22	27
(SÁNCHEZ, et al., 2011)	Pista 1	3	11	6	20
	Pista 2	0	17	12	29
	Totais	3	28	18	49

A Figura 64 apresenta uma situação da Cam68 em que várias motocicletas passam pela ROI, mas apenas duas são detectadas na saída da mesma.

**Figura 64: Exemplo de falha de detecção de motocicletas na Cam68.**

A Figura 65 apresenta um quadro do vídeo Cam73, com sua ROI, posição dos laços virtuais e seus respectivos contadores, e o fundo utilizado. A ROI foi definida de modo a abranger todas as pistas e sua extensão cobre praticamente toda a cena, mantendo-se apenas uma pequena distância das bordas da imagem por motivos de processamento. Os laços virtuais foram posicionados dentro da ROI, um para cada pista, mas de forma que toda a extensão da via seja coberta.

Diferentemente dos outros vídeos, os laços virtuais foram posicionados quase no centro da imagem, para que fossem mantidos em linha e os laços das pistas externas não

saíssem da ROI. O fundo, como era esperado, apresenta a imagem estática do vídeo em questão, com um aspecto mais limpo, propiciado pelo método de obtenção do mesmo, através do filtro de mediana do domínio temporal.

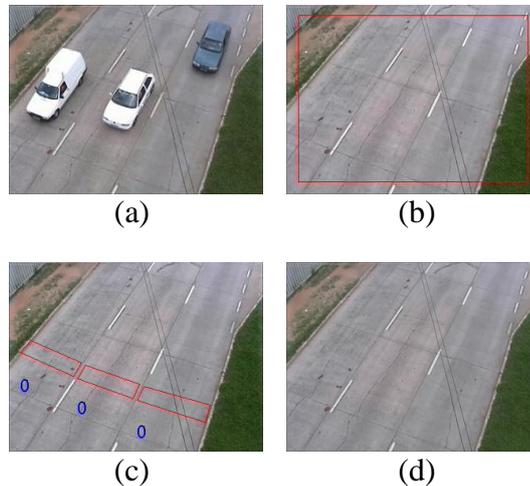


Figura 65: Quadro do vídeo Cam73 (a), ROI (b), laços virtuais (c) e fundo (d) utilizados.

A Figura 66 apresenta o mapa das partículas iniciais da Cam73, utilizadas na primeira etapa de filtragem das partículas. Percebe-se que apenas as partículas na ROI são aparentes, e nas posições das faixas, ranhuras e limites da via, bem como no gramado e passeio de terra, a intensidade das partículas é maior, significando que muitas partículas são obtidas na mesma posição espacial ao longo dos quadros e que, portanto, fazem parte do fundo, e como tal são eliminadas na primeira etapa de filtragem.

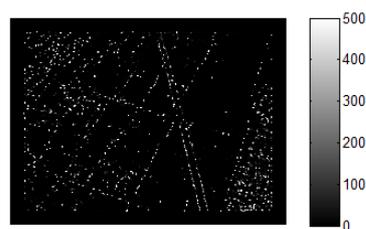


Figura 66: Mapa de repetitividade das partículas iniciais da Cam73.

A Tabela 16 apresenta o resultado da contagem (em unidades) efetuada nos laços virtuais pelo método proposto para a Cam73, juntamente com as contagens realizadas pelos métodos comparativos e a contagem real, obtida por inspeção visual. As diferenças numéricas (em unidades) entre as contagens obtidas pelos respectivos métodos e as contagens reais, bem como as diferenças relativas (em porcentagem), que são as razões entre as diferenças numéricas e as contagens reais, também são apresentadas. As diferenças totais correspondem

às diferenças das somas dos resultados das contagens nos laços virtuais, para cada método, em relação ao número total de veículos que passaram pela ROI durante o processamento.

Tabela 16: Contagem de veículos – Cam73.

		Pista 1	Pista 2	Pista 3	Totais
Inspeção Visual	Contagem (un.)	25	48	28	101
Método Proposto	Contagem (un.)	23	51	29	103
	Diferenças (un.)	-2	3	1	2
	Dif. Relativas (%)	-8,0	6,2	3,6	2,0
(KIM, 2008)	Contagem (un.)	24	43	18	85
	Diferenças (un.)	-1	-5	-10	-16
	Dif. Relativas (%)	-4,0	-10,4	-35,7	-15,8
(SÁNCHEZ, et al., 2011)	Contagem (un.)	21	49	28	98
	Diferenças (un.)	-4	1	0	-3
	Dif. Relativas (%)	-16,0	2,1	0	-3,0

A Tabela 17 apresenta o número de veículos perdidos por cada um dos métodos em cada pista da Cam73, enquanto a Tabela 18 apresenta o número de veículos que foram contados além do valor real. Apesar de os métodos não fazerem tal distinção, os veículos destas tabelas foram classificados em motos (veículos de duas ou três rodas), carros (veículos de passeio e caminhonetes) e caminhões (ônibus, caminhões e furgões), para que a análise dos resultados seja mais aprofundada.

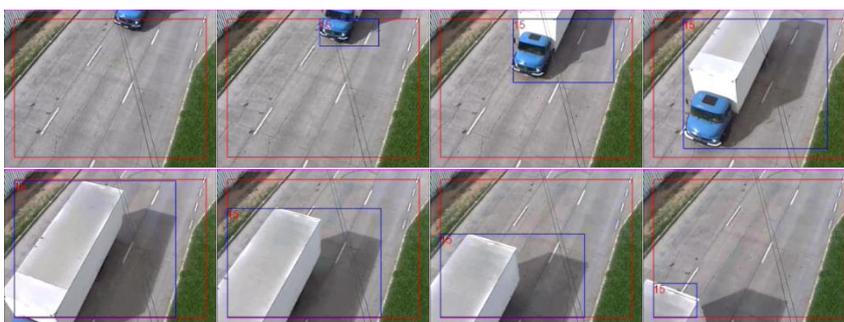
Tabela 17: Veículos não detectados – Cam73.

		Motos	Carros	Caminhões	Totais
Método Proposto	Pista 1	0	1	1	2
	Pista 2	0	0	0	0
	Pista 3	0	1	0	1
	Totais	0	2	1	3
(KIM, 2008)	Pista 1	0	5	0	5
	Pista 2	1	6	0	7
	Pista 3	0	10	0	10
	Totais	1	21	0	22
(SÁNCHEZ, et al., 2011)	Pista 1	0	1	3	4
	Pista 2	1	0	0	1
	Pista 3	0	1	0	1
	Totais	1	2	3	6

Tabela 18: Veículos multiplamente detectados – Cam73.

		Motos	Carros	Caminhões	Totais
Método Proposto	Pista 1	0	0	1	1
	Pista 2	0	1	1	2
	Pista 3	0	2	0	2
	Totais	0	3	2	5
(KIM, 2008)	Pista 1	0	0	4	4
	Pista 2	0	0	2	2
	Pista 3	0	0	0	0
	Totais	0	0	6	6
(SÁNCHEZ, et al., 2011)	Pista 1	0	0	0	0
	Pista 2	0	0	2	2
	Pista 3	0	0	1	1
	Totais	0	0	3	3

A Figura 67 apresenta uma situação de rastreamento na Cam73 em que um carro na Pista 1 entra em cena muito próxima de um caminhão na Pista 2, e os veículos acabam sendo rastreados como apenas um. Já a Figura 68 apresenta uma situação em que um veículo grande, maior que a ROI, entra em cena e é corretamente identificado, apesar de sua sombra ser rastreada junto com o veículo em boa parte do processo.

**Figura 67: Exemplo de falha de detecção e rastreamento na Cam73.****Figura 68: Exemplo de rastreamento para um veículo maior que a ROI na Cam73.**

A Tabela 19 condensa os resultados para veículos não identificados e a Tabela 20 os veículos identificados além do valor real, para todos os vídeos e todos os métodos. A Tabela 21 condensa os resultados de contagem obtidos para todos os vídeos e todos os métodos, e inclui um resultado geral.

Tabela 19: Resultado geral para veículos perdidos.

		Cam06	Cam25	Cam41	Cam64	Cam68	Cam73	Geral
Método Proposto	Pista 1	8	5	0	32	9	2	56
	Pista 2	6	3	1	5	16	0	31
	Pista 3	3	1	1	9	*	1	15
	Totais	17	9	2	46	25	3	102
(KIM, 2008)	Pista 1	1	15	0	10	2	5	33
	Pista 2	4	5	7	2	5	7	30
	Pista 3	11	10	5	12	*	10	38
	Totais	16	20	12	24	7	22	101
(SÁNCHEZ, et al., 2011)	Pista 1	10	21	0	21	20	4	76
	Pista 2	6	11	25	10	21	1	74
	Pista 3	10	5	1	14	*	1	31
	Totais	26	37	26	45	41	6	181

Tabela 20: Resultado geral para veículos contados mais de uma vez.

		Cam06	Cam25	Cam41	Cam64	Cam68	Cam73	Geral
Método Proposto	Pista 1	2	3	0	5	6	1	17
	Pista 2	4	2	8	3	0	2	19
	Pista 3	3	1	1	6	*	2	13
	Totais	9	6	9	14	6	5	49
(KIM, 2008)	Pista 1	40	3	0	5	19	4	71
	Pista 2	15	2	35	14	8	2	76
	Pista 3	2	15	2	1	*	0	20
	Totais	57	20	37	20	27	6	167
(SÁNCHEZ, et al., 2011)	Pista 1	18	42	1	16	20	0	97
	Pista 2	14	17	0	7	29	2	69
	Pista 3	17	3	0	21	*	1	42
	Totais	49	62	1	44	49	3	208

Tabela 21: Resultado geral para contagem de veículos.

Vídeos e Pistas	Inspeção Visual	Método Proposto			(KIM, 2008)			(SÁNCHEZ, et al., 2011)			
		Cont.	Dif.	Dif. (%)	Cont.	Dif.	Dif. (%)	Cont.	Dif.	Dif. (%)	
Cam06	Pista 1	44	38	-6	-13,6	83	39	88,6	52	8	18,2
	Pista 2	92	90	-2	-2,2	103	11	12,0	100	8	8,7
	Pista 3	89	89	0	0	80	-9	-10,1	96	7	7,9
	Totais	225	217	-8	-3,6	266	41	18,2	248	23	10,2
Cam25	Pista 1	97	95	-2	-2,1	85	-12	-12,4	118	21	21,6
	Pista 2	68	67	-1	-1,5	65	-3	-4,4	74	6	8,8
	Pista 3	24	24	0	0	39	15	62,5	22	-2	-8,3
	Totais	189	186	-3	-1,6	189	0	0	214	25	13,2
Cam41	Pista 1	10	10	0	0	10	0	0	11	1	10,0
	Pista 2	157	164	7	4,5	185	28	17,8	132	-25	-15,9
	Pista 3	84	84	0	0	81	-3	-3,6	83	-1	-1,2
	Totais	251	258	7	2,8	276	25	10,0	226	-25	-10,0
Cam64	Pista 1	74	47	-27	-36,5	69	-5	-6,8	69	-5	-6,8
	Pista 2	36	34	-2	-5,6	48	12	33,3	33	-3	-8,3
	Pista 3	55	52	-3	-5,5	44	-11	-20,0	62	7	12,7
	Totais	165	133	-32	-19,4	161	-4	-2,4	164	-1	-0,6
Cam68	Pista 1	66	63	-3	-4,5	83	17	25,8	66	0	0
	Pista 2	88	72	-16	-18,2	91	3	3,4	96	8	9,1
	Totais	154	135	-19	-12,3	174	20	13,0	162	8	5,2
Cam73	Pista 1	25	23	-2	-8,0	24	-1	-4,0	21	-4	-16,0
	Pista 2	48	51	3	6,3	43	-5	-10,4	49	1	2,1
	Pista 3	28	29	1	3,6	18	-10	-35,7	28	0	0
	Totais	101	103	2	2,0	85	-16	-15,8	98	-3	-3,0
GERAL	1085	1032	-53	-4,9	1151	66	6,1	1112	27	2,5	

5.2 DISCUSSÃO

Esta subseção analisa os resultados experimentais obtidos para cada cena e o resultado geral. Justifica situações que degradam ou melhoram os resultados e propõe soluções.

A Figura 47 apresenta uma situação, durante o rastreamento na Cam06, em que os veículos acabam gerando oclusões devido à parada no semáforo. Enquanto os veículos estão parados, apenas três objetos são rastreados, embora haja sete veículos na ROI. Conforme os veículos se movem, novos objetos vão sendo criados, devido à divisão dos veículos

detectados, por estes apresentarem grandes regiões similares ao fundo. Ao final da sequência de quadros, apenas um veículo, localizado ao lado do ônibus, não é detectado.

A Figura 48 apresenta uma situação semelhante à anterior na Cam06, em que vários veículos são considerados um só, mas neste caso a ocorrência é devido à proximidade dos veículos. Da mesma forma que no caso anterior, os veículos são separados ao longo do processo. Tais situações são mais comuns quando envolvem veículos de grandes dimensões, porque a região similar com o fundo, pertencente ao objeto rastreado, deve ser grande o suficiente para que a divisão do mesmo ocorra.

Analisando-se a Tabela 1, percebe-se que o método proposto possui uma diferença de contagem menor que os métodos comparativos para a Cam06, e esta diferença é negativa, ou seja, a contagem foi inferior ao valor da contagem real. Nota-se ainda que a Pista 1, ou seja, a mais a esquerda, é a que possui a maior diferença, e isto ocorre em todos os métodos. Considerando-se a Tabela 2, o método proposto e o primeiro método comparativo (KIM, 2008), perdem praticamente a mesma quantidade de veículos para a Cam06, sendo que o método proposto perde muitas motos, enquanto o método comparativo perde mais carros. O segundo método comparativo (SÁNCHEZ *et al.*, 2011) perde mais veículos, sendo a maioria carros.

Já segundo a Tabela 3, nota-se que a quantidade de veículos contados a mais que o número real é muito inferior para o método proposto em relação aos métodos comparativos para a Cam06, e a maioria desses veículos é composta por carros de passeio. No método comparativo (KIM, 2008) a maioria dos veículos multiplamente detectados são caminhões, porque este método tende a dividir veículos grandes em vários objetos menores. No método comparativo (SÁNCHEZ *et al.*, 2011) os tipos de veículos multiplamente detectados são mais equilibrados, incluindo motos.

A múltipla detecção para o segundo método comparativo (SÁNCHEZ *et al.*, 2011) na Cam06, ocorre devido à forma de obtenção do fundo, adaptativamente, que considera os veículos parados muito tempo sobre os laços virtuais como fundo. Quando esses veículos voltam a locomover-se, são novamente contados. Além disso, muitos veículos são divididos em dois, pois os para-brisas são muito semelhantes ao asfalto, e este método só considera os componentes conectados, formados por *pixels* diferentes do fundo para efetuar o rastreamento.

Concluindo-se a análise dos resultados do método proposto para a Cam06, a contagem de veículos obteve uma diferença absoluta inferior aos métodos comparativos, e esta

diferença foi negativa porque o número de veículos não detectados é superior ao número de veículos detectados além do valor real. A perda de veículos ocorre principalmente na pista esquerda, em que a possibilidade de oclusão é maior devido à posição da câmera. A perda é maior entre motos, que são veículos menores, e as regiões formadas por suas partículas muitas vezes não satisfazem os requisitos para a identificação de um veículo.

A Figura 51 apresenta o rastreamento na Cam25, em que a situação inicial é os veículos estarem parados devido à sinalização de trânsito. Os veículos vão sendo identificados ao longo do processo e os alvos detectados vão sendo divididos conforme inicia a movimentação. Já na Figura 52, a detecção de uma motocicleta na Cam25 somente ocorre quando ela está deixando a ROI, pois devido à perspectiva, é na saída da ROI que mais partículas são detectadas e a região convexa formada por essas partículas torna-se suficientemente grande para satisfazer os critérios de detecção de veículos.

Conforme observado na Tabela 4, o método proposto e o primeiro método comparativo (KIM, 2008), possuem diferenças de contagem gerais muito baixas para a Cam25. Entretanto, o método comparativo (KIM, 2008) apresenta uma diferença de contagem muito elevada na terceira pista, mais a direita, onde passam apenas ônibus. Pela Tabela 6 verifica-se que a maioria dos veículos múltiplemente detectados na Cam25 pelo método comparativo (KIM, 2008) é composta por caminhões (ônibus) na terceira pista, fato devido ao método em questão dividir veículos grandes em objetos menores. O segundo método comparativo apresenta diferenças de contagem bem superiores, que poderiam ser maiores ainda caso o número de veículos não detectados e veículos múltiplemente detectados fossem mais desequilibrados.

Tanto as diferenças de contagem quanto os números de veículos não detectados e veículos múltiplemente detectados ficaram baixos para o método proposto na Cam25, apesar de haver um semáforo e os veículos sobreporem-se e ocluam-se quando este está fechado. As pistas largas e a pouca inclinação da câmera facilitam a separação dos veículos quando entram em movimento.

A Cam41, juntamente com a Cam73, é uma das cenas que apresenta os melhores resultados. Isto ocorre devido ao fato das pistas serem amplas, haver pouca perspectiva lateral e não haver interrupção no fluxo de veículos. Não só para o método proposto, mas os resultados apresentados por todos os métodos para o vídeo Cam41 são melhores que os resultados dos outros vídeos. Conforme observado na Tabela 7, as maiores diferenças de contagem concentram-se na pista central, que possui o maior fluxo de veículos.

O número de veículos não detectados pelo método proposto é extremamente baixo para a Cam41, conforme a Tabela 8, sendo apenas um carro e uma motocicleta. O primeiro método comparativo (KIM, 2008) possui muitos veículos detectados mais de uma vez, sendo a maioria formada por caminhões, enquanto o segundo método comparativo (SÁNCHEZ *et al.*, 2011) possui muitos veículos não detectados, sendo principalmente veículos trafegando atrás dos caminhões.

A Cam64, juntamente com a Cam68, é uma das cenas que apresenta os piores resultados, pois nestes vídeos existe parada de veículos devido à sinalização de trânsito, gerando aglomeração e oclusão de veículos. No caso específico da Cam64, o tempo de parada é muito grande, representando 70% do tempo total. Além disso, devido à perspectiva, os veículos que possuem passagem livre na pista da esquerda acabam sendo ocluídos pelos veículos grandes (ônibus e caminhões) que ficam parados na pista central.

A Figura 60 apresenta um exemplo de rastreamento na Cam64 em que vários veículos são perdidos na pista da esquerda. Isto ocorre devido à oclusão por um ônibus parado na pista central. As partículas dos veículos que entram em cena não sendo agregadas ao ônibus devido à proximidade, impedindo a criação de novos agrupamentos de partículas e conseqüentemente a identificação dos veículos que passam na lateral do ônibus. Ao mesmo tempo o ônibus, juntamente com os veículos parados próximos, é tratado como um único veículo porque há pouca região de fundo entre o ônibus e os veículos.

A Figura 61 apresenta a separação dos veículos do ônibus rastreado na Figura 60. Conforme ocorre a movimentação, a região de fundo entre os veículos aumenta e a divisão do alvo é efetuada. Todos os veículos acabam sendo rastreados separadamente ao final do processo, com exceção da motocicleta que é um veículo menor e acaba ficando muito próxima do ônibus.

Como observado na Tabela 10, o método proposto foi o que apresentou os piores resultados de contagem geral entre os três métodos, não significando, porém, que tenha tido o pior desempenho. Os resultados para a Pista 2 e Pista 3 foram superiores aos métodos comparativos, e o segundo método comparativo (SÁNCHEZ *et al.*, 2011) somente obteve o melhor resultado de contagem geral porque o número de veículos perdidos e de veículos multiplamente detectados ficaram muito equilibrados, embora sejam os mais elevados entre os três métodos, conforme apresentado pelas Tabela 11 e Tabela 12.

O primeiro método comparativo (KIM, 2008) trata melhor as situações de oclusão e obteve os melhores resultados para os veículos perdidos. O método proposto perde muitos

veículos na Pista 1, da esquerda, devido à oclusão, por outro lado, é o método que menos contou veículos além do valor real.

A Cam68, juntamente com a Cam64, é uma das cenas que apresenta os piores resultados, pois nestes vídeos existe parada de veículos devido à sinalização de trânsito, gerando aglomeração e oclusão de veículos. No caso específico da Cam68, além das paradas devido à sinalização de trânsito, a captura das imagens é mais distante que os outros vídeos, resultando que várias motocicletas passam pelas vias sem serem identificadas, pois as regiões convexas formadas por suas partículas não satisfazem as condições de identificação de veículos. O número de partículas poderia ser aumentado, mas acarretaria maior custo computacional no processamento. Outra solução seria diminuir o número mínimo de partículas para a criação de agrupamentos, mas poderia resultar na criação de vários agrupamentos para cada veículo, forçando a execução de muitos processos de fusão de agrupamentos, nem sempre corretamente.

Assim como para a Cam64, o melhor resultado na contagem geral para a Cam68 obtido pelo segundo método comparativo (SÁNCHEZ *et al.*, 2011), conforme a Tabela 13, não significa que esse método tenha sido melhor nessa cena. Considerando-se as Tabela 14 e Tabela 15, observa-se que o segundo método comparativo (SÁNCHEZ *et al.*, 2011) apresenta os maiores números de veículos perdidos e veículos contados acima do valor real, mas estes números são mais equilibrados que para os outros métodos.

O primeiro método comparativo (KIM, 2008) obteve os menores números para veículos perdidos, conforme a Tabela 14, pois este método trata melhor os veículos pequenos, como as motocicletas. Em compensação a quantidade de veículos multiplamente detectados é elevada, principalmente para os veículos grandes, como os caminhões, conforme a Tabela 15. O método proposto apresenta os menores valores para veículos multiplamente detectados, e o número elevado de veículos não detectados deve-se principalmente a veículos pequenos, como as motocicletas na Figura 64.

Tais motocicletas, na Figura 64, não são detectadas porque não geram partículas suficientes para a criação de agrupamentos, ou a região convexa formada pelas partículas não satisfaz os critérios de detecção de veículos. Poderia-se reduzir o número de partículas para a geração de agrupamentos e flexibilizar as regras para detecção de veículos, mas tais procedimentos poderiam acarretar na múltipla detecção dos veículos maiores.

A Cam73, juntamente com a Cam41, é uma das cenas que apresenta os melhores resultados, pois nestes vídeos não existe parada de veículos e a perspectiva é muito pequena em relação aos outros vídeos.

Conforme apresentado pela Tabela 16, a diferença de contagem geral, bem como a diferença de contagem para cada pista da Cam73, do método proposto é inferior às diferenças dos métodos comparativos. A diferença de contagem para cada pista também mostrou-se mais estável que as diferenças apresentadas pelos métodos comparativos.

A Tabela 17 mostra muitos veículos perdidos pelo primeiro método comparativo (KIM, 2008) na terceira pista, e diferentemente dos outros vídeos, em que o número de veículos multiplamente detectados, conforme a Tabela 18, é maior ou próxima ao número de veículos não detectados, para este vídeo o número de veículos não detectados é muito maior que o número de veículos detectados além do valor real.

As Tabela 16, Tabela 17 e Tabela 18 demonstram ainda o bom desempenho do segundo método comparativo, tanto em nível de veículos perdidos como de veículos multiplamente detectados, obtendo resultados próximos aos do método proposto.

No vídeo Cam73, para o método proposto, a área de busca para a obtenção dos vetores de movimento precisou ser aumentada, devido à velocidade com que os veículos passam nessa cena, fazendo com que as partículas sejam perdidas se a área de busca for muito pequena. Tal alteração, utilizando-se o método proposto, poderia ser aplicada para todos os outros vídeos sem alteração de resultados, apenas aumentando o custo computacional envolvido.

Conforme observado na Tabela 21 o método proposto é o único que possui uma diferença de contagem negativa, ou seja, o valor da contagem efetuada pelo método é inferior ao valor real. Os métodos comparativos obtiveram contagens superiores ao valor real. Em valores absolutos, o método proposto é superior ao método de (KIM, 2008) e inferior ao método de (SÁNCHEZ *et al.*, 2011), considerando-se a contagem geral. Porém, observando-se os resultados de cada cena, ou mesmo cada pista, percebe-se que os resultados individuais apresentados pelo método proposto possuem diferenças absolutas de contagens menores em relação aos métodos comparativos.

Em um total de 17 pistas, em 11 a diferença absoluta de contagem do método proposto é menor que para os dois métodos comparativos, e em 4 a diferença é menor que pelo menos um dos dois métodos comparativos, perdendo para ambos os métodos comparativos em apenas 2 pistas. Observa-se também que a Pista 1 da Cam64 e a Pista 2 da Cam68 são

responsáveis por 47% do valor de veículos contados inferiormente ao número real de veículos.

Analisando-se a Tabela 19 e a Tabela 20 percebe-se que a diferença de contagem inferior apresentado por (SÁNCHEZ *et al.*, 2011) na Tabela 21 não significa que o método é superior aos demais, mas apenas que a diferença absoluta entre veículos contados acima do valor real e veículos perdidos é menor, sendo de 27 veículos, enquanto para o método proposto tal diferença fica em 53 veículos e para o primeiro método comparativo (KIM, 2008) fica em 66 veículos.

O segundo método comparativo (SÁNCHEZ *et al.*, 2011) apresenta as maiores diferenças, tanto em relação a veículos perdidos, 181, quanto a veículos contados além do valor real, 208. O número de veículos perdidos ficou muito próximo para o método proposto e (KIM, 2008), 102 e 101 respectivamente, sendo frisado novamente que a Pista 1 da Cam64 e a Pista 2 da Cam68 são responsáveis por um grande número de veículos perdidos pelo método proposto, 47%.

Em tratando-se de múltipla detecção, o método proposto apresenta os melhores resultados, 49 veículos contra 167 veículos de (KIM, 2008) e 208 veículos de (SÁNCHEZ *et al.*, 2011). Os números de veículos multiplamente detectados ficaram mais constantes em função das cenas e pistas, não apresentando valores elevados como para o caso de veículos não detectados. Os melhores resultados, tanto em relação a veículos perdidos como para veículos multiplamente detectados, para o método proposto, são apresentados pelas Cam41 e Cam73, que são as cenas em que o fluxo de veículos nunca é interrompido.

6 CONCLUSÃO

O trabalho apresenta uma alternativa viável para o rastreamento e contagem de veículos baseada em imagens de vídeo, utilizando-se da integração de técnicas bem conhecidas de processamento de imagens e clusterização. Futuramente, espera-se que com a otimização do código para uma plataforma mais rápida que a utilizada nos testes (Matlab), seja possível rodar o algoritmo em tempo real, pois todo o processamento pode ser realizado sequencialmente, e não em blocos como está sendo executado no presente trabalho.

Por tratar-se de um método que utiliza informação do movimento das partículas, situações em que o fluxo de veículos é interrompido podem degradar os resultados, como nas Cams06, 25, 64 e 68, normalizando-se com o retorno do movimento. Oclusão é outro fator de degradação, explicitado nos resultados para a Cam64 na Pista 1, devendo esse problema ser tratado com a utilização de outro método de rastreamento de partículas, que seja adaptativo à oclusões. A possibilidade de instalações personalizadas das câmeras também podem evitar situações de oclusões, pois seria possível instalar as câmeras de forma centralizada às pistas.

Veículos muito pequenos podem ser perdidos, pois o número de partículas muitas vezes é insuficiente para gerar um cluster, como ocorre com as motocicletas da Cam68, que possui o plano de imagem das pistas mais afastado da câmera que para as outras cenas. Tal situação poderia ser resolvida de duas maneiras: reduzindo-se a qualidade das partículas que podem ser obtidas, o que aumentaria o número de partículas e tempo de processamento; ou diminuindo o número mínimo de partículas para geração de agrupamentos, o que poderia gerar muitos agrupamentos de poucas partículas, que onerariam o processamento na fase de fusão de agrupamentos.

Em compensação, veículos grandes, muitas vezes maiores que a região de interesse, são corretamente identificados, diferentemente dos métodos comparativos. A fusão de agrupamentos, que leva em conta a informação de fundo presente entre agrupamentos, garante que agrupamentos próximos pertencentes ao mesmo veículo sejam fundidos. Tais situações são mais evidentes nas Cams41 e 73, em que mesmo os veículos que ocupam praticamente toda a ROI são corretamente identificados.

Poucas alterações foram necessárias para rodar cada vídeo, demonstrando a versatilidade do método. Apenas a área de busca para os vetores de movimento da Cam73 foi aumentada, pois os veículos passam muito rápido nesse vídeo, mas essa alteração poderia ser

aplicada aos outros vídeos sem alteração de resultados, apenas teria um custo computacional maior. Outras diferenças para cada vídeo são as seleções das regiões de interesse e a localização dos laços virtuais.

O objetivo principal foi atingido e o secundário foi atingido em parte. Apesar de não haver uma avaliação objetiva do rastreamento, sabe-se que este falha em situações de oclusão e aglomeração dos veículos. Em compensação, da forma como a contagem foi implementada, corrigem-se vários problemas do rastreamento, pois apenas os veículos que estão sendo rastreados na saída das ROI são contados, dando tempo para o algoritmo de rastreamento obter os melhores resultados. A taxa de acerto na contagem dos veículos ficou em 96,1%, compatível com os melhores métodos disponíveis atualmente.

Na comparação com os métodos atuais testados, a taxa de acerto na contagem de veículos ficou melhor em 3 dos 6 vídeos testados. Quando considera-se as pistas individualmente, o método proposto é melhor em 11 das 17 pistas, obtendo os piores resultados em apenas 2 pistas. O número de veículos perdidos pelo método proposto, 102, ficou similar ao valor obtido por (KIM, 2008), 101, e melhor que o valor obtido por (SÁNCHEZ *et al.*, 2011), 181. Já o número de veículos contados mais de uma vez ficou melhor para o método proposto, 49, em relação aos métodos comparativos, 167 e 208 para (KIM, 2008) e (SÁNCHEZ *et al.*, 2011) respectivamente.

A utilização de sistemas automáticos de monitoramento e controle de tráfego pode ser capaz de diminuir os problemas de congestionamento, além de fornecer dados importantes para a gestão da manutenção e ampliação das vias existentes. Com os resultados obtidos e melhorias futuras espera-se contribuir para o desenvolvimento de tais sistemas.

7 CONTRIBUIÇÕES DESTE TRABALHO

O presente trabalho apresentou contribuições ao projeto de Rastreamento Veicular, coordenado pelo Professor Jacob Scharcanski e mantido em conjunto pela UFRGS e Grupo Digicon S.A., gerando transferência de tecnologia da Universidade para a Indústria.

O trabalho desenvolvido também gerou duas submissões. Uma para a Conferência Internacional de Instrumentação e Tecnologias de Medições (I2MTC, *International Instrumentation and Measurement Technology Conference*) do Instituto dos Engenheiros Eletricistas e Eletrônicos (IEEE, *Institute of Electrical and Electronic Engineers*) sob o título *Tracking and Counting Vehicles in Traffic Video Sequences Using Particle Filtering* (BOUVIÉ *et al.*, 2013). E outra para a revista Controle & Automação, da Sociedade Brasileira de Automática (SBA), sob o título *Counting and Tracking of Vehicles in Traffic Videos* (BOUVIÉ, ESCOUTO e SCHARCANSKI).

REFERÊNCIAS

- BARDET, F.; CHATEAU, T. MCMC Particle Filter for Real-Time Visual Tracking of Vehicles. In: CONFERENCE ON INTELLIGENT TRANSPORTATION SYSTEMS INTERNATIONAL (ITSC), 11., 2008, Beijing. **Proceedings...** Piscataway: IEEE, 2008. p. 539-544.
- BOUVIÉ, C. et al. Tracking and Counting Vehicles in Traffic Video Sequences Using Particle Filtering. In: INSTRUMENTATION AND MEASUREMENT TECHNOLOGY CONFERENCE (I2MTC), 2013, Minneapolis. **Proceedings...** Piscataway: IEEE, 2013. Submetido.
- BOUVIÉ, C.; ESCOUTO, F. L.; SCHARCANSKI, J. Counting and Tracking of Vehicles in Traffic Videos. **Controle & Automação**, Campinas. Submetido.
- BUCH, N.; VELASTIN, S. A.; ORWELL, J. A Review of Computer Vision Techniques for the Analysis of Urban Traffic. **Intelligent Transportation Systems, IEEE Transactions on**, New York, v. 12, n. 3, p. 920-939, Sept. 2011.
- CHAN, Y.-M. et al. Vehicle Detection and Tracking Under Various Lighting Conditions Using a Particle Filter. **Intelligent Transport Systems, IET**, Stevenage, v. 6, n. 1, p. 1-8, Mar. 2012.
- CHEN, Y.-L. et al. A Real-Time Vision System for Nighttime Vehicle Detection and Traffic Surveillance. **Industrial Electronics, IEEE Transactions on**, New York, v. 58, n. 5, p. 2030-2044, May 2011.
- COMANICIU, D.; RAMESH, V.; MEER, P. Real-Time Tracking of Non-Rigid Objects Using Mean Shift. In: CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION, 2000, Hilton Head Island. **Proceedings...** Los Alamitos: IEEE Computer Society, 2000. p. 142-149.
- DEMPSTER, A. P.; LAIRD, N. M.; RUBIN, D. B. Maximum Likelihood from Incomplete Data via the EM Algorithm. **Journal of the Royal Statistical Society, Series B**, London, v. 39, n. 1, p. 1-38, 1977.
- DUFAUX, F.; MOSCHENI, F. Motion Estimation Techniques for Digital TV: A Review and a New Contribution. **Proceedings of the IEEE**, Piscataway, v. 83, n. 6, p. 858-876, Jun. 1995.
- FERIS, R. S. et al. Large-Scale Vehicle Detection, Indexing, and Search in Urban Surveillance Videos. **Multimedia, IEEE Transactions on**, Piscataway, v. 14, n. 1, p. 28-42, Feb. 2012.

FORSTNER, W.; GULCH, E. A fast operator for detection and precise location of distinct points, corners, and centers of circular features. In: INTERCOMMISSION CONFERENCE ON FAST PROCESSING OF PHOTOGRAMMETRIC DATA (ISPRS), 1987, Interlaken. **Proceedings...** Berlin, Heidelberg: Springer Berlin Heidelberg, 1987. p. 281-305.

FREUND, Y.; SCHAPIRE, R. E. A Short Introduction to Boosting. **Japanese Society for Artificial Intelligence**, Tokyo, v. 14, n. 5, p. 771-780, 1999.

GONZALEZ, R. C.; WOODS, R. E. **Digital Image Processing**. 2nd. ed. New Jersey: Prentice-Hall, 2002.

HARRIS, C.; STEPHENS, M. J. A Combined Corner and Edge Detection. In: ALVEY VISION CONFERENCE, 4., 1988, Manchester. **Proceedings...** Sheffield: University of Sheffield Printing Unit, 1988. p. 147-151.

HOUNSELL, N. B. et al. Review of urban traffic management and the impacts of new vehicle technologies. **Intelligent Transport Systems, IET**, Stevenage, v. 3, n. 4, p. 419-428, Dec. 2009.

HUANG, C.-L.; LIAO, W.-C. A Vision-Based Vehicle Identification System. In: INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION (ICPR), 17., 2004, Cambridge. **Proceedings...** Los Alamitos: IEEE Computer Society, 2004. p. 364- 367.

IWASAKI, Y.; KUROGI, Y. Real-Time Robust Vehicle Detection Through the Same Algorithm Both Day and Night. In: INTERNATIONAL CONFERENCE ON WAVELET ANALYSIS AND PATTERN RECOGNITION (ICWAPR '07), 2007, Beijing. **Proceedings...** Piscataway: IEEE, 2007. p. 1008-1014.

JAZAYERI, A. et al. Vehicle Detection and Tracking in Car Video Based on Motion Model. **Intelligent Transportation Systems, IEEE Transactions on**, New York, v. 12, n. 2, p. 583-595, June 2011.

KANHERE, N. K.; BIRCHFIELD, S. T. Real-Time Incremental Segmentation and Tracking of Vehicles at Low Camera Angles Using Stable Features. **Intelligent Transportation Systems, IEEE Transactions on**, New York v. 9, n. 1, p. 148-160, Mar. 2008.

KASTRINAKI, V.; ZERVAKIS, M.; KALAITZAKIS, K. A Survey of Video Processing Techniques for Traffic Applications. **Image and Vision Computing**, Amsterdam, v. 21, n. 4, p. 359-381, Apr. 2003.

KIM, Z. W. Real Time Object Tracking Based on Dynamic Feature Grouping with Background Subtraction. In: CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION (CVPR), 2008, Providence. **Proceedings...** New York: IEEE, 2008. p. 1-8.

MAHALANOBIS, P. C. On the Generalised Distance in Statistics. **Proceedings of the Indian National Science Academy**, New Delhi, v. 2, n. 1, p. 49-55. 1936.

MEI, X.; LING, H. Robust Visual Tracking and Vehicle Classification via Sparse Representation. **Pattern Analysis and Machine Intelligence, IEEE Transactions on**, New York, v. 33, n. 11, p. 2259-2272, Nov. 2011.

MORRIS, B. T.; TRIVEDI, M. M. Learning, Modeling, and Classification of Vehicle Track Patterns from Live Video. **Intelligent Transportation Systems, IEEE Transactions on**, New York, v. 9, n. 3, p. 425-437, Sept. 2008.

MOSSI, J. M. et al. Real-Time Traffic Analysis at Night-Time. In: INTERNATIONAL CONFERENCE ON IMAGE PROCESSING (ICIP), 18., 2011, Brussels. **Proceedings...** Piscataway: IEEE, 2011. p. 2941-2944.

O'MALLEY, R.; GLAVIN, M.; JONES, E. Vision-Based Detection and Tracking of Vehicles to the Rear With Perspective Correction in Low-Light Conditions. **Intelligent Transport Systems, IET**, Stevenage, v. 5, n. 1, p. 1-10, Mar. 2011.

PURNAMA, I. K. E. et al. Real Time Vehicle Counter System for Intelligent Transportation System. In: INTERNATIONAL CONFERENCE ON INSTRUMENTATION, COMMUNICATIONS, INFORMATION TECHNOLOGY, AND BIOMEDICAL ENGINEERING (ICICI-BME), 2009, Bandung. **Proceedings...** Piscataway: IEEE, 2009. p. 1-4.

SÁNCHEZ, A. et al. Video-Based Distance Traffic Analysis: Application to Vehicle Tracking and Counting. **Computing in Science & Engineering**, Los Alamitos, v. 13, n. 3, p. 38-45, May 2011.

SCHARCANSKI, J. et al. A Particle-Filtering Approach for Vehicular Tracking Adaptive to Occlusions. **Vehicular Technology, IEEE Transactions on**, New York, v. 60, n. 2, p. 381-389, Feb. 2011.

SEBER, G. A. F. **Multivariate Observations**. Hoboken: John Wiley & Sons, 1984.

SHI, J.; TOMASI, C. Good features to track. In: CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION (CVPR '94), 1994, Seattle. **Proceedings...** Los Alamitos: IEEE Computer Society, 1994. p. 593-600.

SIVARAMAN, S.; TRIVEDI, M. M. A General Active-Learning Framework for On-Road Vehicle Recognition and Tracking. **Intelligent Transportation Systems, IEEE Transactions on**, New York, v. 11, n. 2, p. 267-276, June 2010.

SONKA, M.; HLAVAC, V.; BOYLE, R. **Image Processing, Analysis, and Machine Vision**. 3rd. ed. Toronto: Thomson Engineering, 2007.

SUN, Z.; BEBIS, G.; MILLER, R. On-Road Vehicle Detection: A Review. **Pattern Analysis and Machine Intelligence, IEEE Transactions on**, New York, v. 28, n. 5, p. 694-711, May 2006.

THEODORIDIS, S.; KOUTROUMBAS, K. **Pattern Recognition**. San Diego: Academic Press, 2003.

TIAN, B. et al. Video Processing Techniques for Traffic Flow Monitoring: A Survey. In: INTERNATIONAL CONFERENCE ON INTELLIGENT TRANSPORTATION SYSTEMS (ITSC), 14., 2011, Washington. **Proceedings...** Piscataway: IEEE, 2011. p. 1103-1108.

TSENG, B. L.; LIN, C.-Y.; SMITH, J. R. Real-Time Video Surveillance for Traffic Monitoring Using Virtual Line Analysis. In: CONFERENCE ON MULTIMEDIA AND EXPO INTERNATIONAL (ICME '02), 2002, Lausanne. **Proceedings...** Piscataway: IEEE, 2002. p. 541- 544.

WANG, C.-C.; HUANG, S.-S.; FU, L.-C. Driver Assistance System for Lane Detection and Vehicle Recognition With Night Vision. In: INTERNATIONAL CONFERENCE ON INTELLIGENT ROBOTS AND SYSTEMS (IROS 2005), 2005, Vilamoura. **Proceedings...** New York: IEEE, 2005. p. 3530-3535.

WANG, G.; XIAO, D.; GU, J. Review on Vehicle Detection Based on Video for Traffic Surveillance. In: INTERNATIONAL CONFERENCE ON AUTOMATION AND LOGISTICS (ICAL 2008), 2008, Qingdao. **Proceedings...** New York: IEEE, 2008. p. 2961 - 2966.

WILLIAMS, B. M.; GUIN, A. Traffic Management Center Use of Incident Detection Algorithms: Findings of a Nationwide Survey. **Intelligent Transportation Systems, IEEE Transactions on**, New York, v. 8, n. 2, p. 351-358, June 2007.

WU, B.-F.; JUANG, J.-H. Adaptive Vehicle Detector Approach for Complex Environments. **Intelligent Transportation Systems, IEEE Transactions on**, New York, v. 13, n. 2, p. 817-827, June 2012.

XU, S. et al. Vehicle Detection Algorithm Based on Shadow Feature. In: INTERNATIONAL COLLOQUIUM ON COMPUTING, COMMUNICATION, CONTROL, AND MANAGEMENT, (CCCM '08), 2008, Guangzhou. **Proceedings...** Los Alamitos: IEEE Computer Society, 2008. p. 105-109.

ZHANG, W. et al. Tracking and Pairing Vehicle Headlight in Night Scenes. **Intelligent Transportation Systems, IEEE Transactions on**, New York, v. 13, n. 1, p. 140-153, Mar. 2012.

APÊNDICE: Detalhamento dos Diagramas de
Blocos e Fluxogramas Através de Algoritmos

APÊNDICE: DETALHAMENTO DOS DIAGRAMAS DE BLOCOS E FLUXOGRAMAS ATRAVÉS DE ALGORITMOS

Este anexo apresenta algoritmos que detalham os diagramas de blocos e fluxogramas presentes na dissertação.

ALGORITMO DA VISÃO GERAL DO MÉTODO PROPOSTO

Início

Variáveis:

Regioes = \emptyset ; Centroides = \emptyset ; Agrupamentos = \emptyset ; Veiculos = \emptyset ; Contagem = \emptyset ;

Constantes:

Video_in; Min_pontos = 8; nrQuadros = 3000;

ROI = criar_roi(Video_in);

Fundo = obter_fundo(Video_in);

Cor_media = obter_cor_media(Fundo);

Particulas_iniciais = obter_particulas_iniciais(Video_in, ROI);

Laços_virtuais = criar_Laços_virtuais(Video_in, ROI);

Para t = 2 até nrQuadros - 1, faça:

 Quadro_t = obter_quadro(Video_in, t);

 Quadro_t = compensa_cor(Quadro_t, Cor_media);

 Quadro_{t+1} = obter_quadro(Video_in, t+1);

 Quadro_{t+1} = compensa_cor(Quadro_{t+1}, Cor_media);

 Particulas_t = obter_particulas(Quadro_t, ROI);

 Particulas_t = filtrar_particulas1(Particulas_t, Partículas_iniciais);

 Particulas_t = filtrar_particulas2(Particulas_t, Quadro_t, Fundo, Regioes_{t-1});

 Vetores_movimento_t = obter_vetores_movimento(Particulas_t, Quadro_t, Quadro_{t+1});

 Se Centroides_{t-1} == \emptyset ou num_particulas < Min_pontos, então:

 Agrupamentos = agrupamento_inicial(Particulas_t, Vetores_movimento_t, Min_pontos);

 Agrupamentos = dividir_agrupamentos(Agrupamentos, Fundo, Min_pontos);

 Agrupamentos = fundir_agrupamentos(Agrupamentos, Fundo);

 Centroides_t = obter_centroides(Agrupamentos);

 Alvos = identificar_veiculos(Agrupamentos, Quadro_t, Fundo);

 Senão:

 [Part_livres, Part_agrup, Part_veic] = verifica_regioes(Particulas_t, Regioes_{t-1}, Vetores_movimento_{t-1});

 Agrupamentos1 = agrupamento_inicial(Part_livres, Vetores_movimento_t, Min_pontos);

 Agrupamentos2 = agrupamento_cont(Part_agrup, Vetores_movimento_t, Centroides_{t-1}, Min_pontos);

 Agrupamentos = união(Agrupamentos1, Agrupamentos2);

 Agrupamentos = dividir_agrupamentos(Agrupamentos, Fundo, Min_pontos);

 Agrupamentos = fundir_agrupamentos(Agrupamentos, Fundo);

 Centroides1 = obter_centroides(Agrupamentos);

 Alvos = identificar_veiculos(Agrupamentos, Quadro_t, Fundo);

 Alvos = dividir_alvos(Alvos, Fundo, Min_pontos);

 Centroides2 = obter_centroides(Alvos);

 Veiculos_t = rastrear_veiculos(Veiculos_{t-1}, Quadro_{t-1}, Particulas_t, Quadro_t);

 Centroides3 = obter_centroides(Veiculos_t);

 Centroides = união(Centroides1, Centroides2, Centroides3);

 Fim Se

 Contagem_t = contar_veiculos(Veiculos, Laços_virtuais, t);

Fim Para t

Fim

ALGORITMO DO AGRUPAMENTO INICIAL

```

Início
Variáveis:
Agrupamentos = Ø; Agrupamentos_temp = Ø; k = 1; Difer = 0;
Constantes:
Particulas; Min_pontos = 8; T = 0.1;
Max_k = fix(nrParticulas / Min_pontos);
Agrupamentos(k) = kmeans(Particulas , k);
Para k = 2 até Max_k, faça:
    Agrupamentos(k) = kmeans(Particulas , k);
    Difer = abs(Dist(Agrupamentos(k)) - Dist(Agrupamentos(k - 1)));
    Se Difer < T, então:
        Agrupamentos = Agrupamentos(k - 1);
        break();
    Senão Se k == Max_k, então:
        Agrupamentos = Agrupamentos(k);
    Fim Se
Fim Para k
[Agrupamentos , Particulas] = verifica_agrupamentos(Agrupamentos); // Utilizando Eq. ( 31 )
// Agrupamento inicial recorrente com as partículas descartadas:
Agrupamentos_temp = initial_clustering(Particulas);
Agrupamentos = união(Agrupamentos , Agrupamentos_temp);
Fim

```

ALGORITMO DA DETECÇÃO DE VEÍCULOS

```

Início
Variáveis:
Alvos = Ø; Pontos = 0;
Constantes:
Agrupamentos; Fundo; Quadro; Limiar_area = 500; Limiar_ocupacao = 0,8;
Limiar_proporcao = 3; Limiar_razao_int = 0,9; Limiar_razao_ext = 0,1;
Para i = 1 até nrAgrupamentos, faça:
    Particulas = extraí_particulas(Agrupamentos(i));
    Regiao_convexa = obter_regiao_convexa (Particulas);
    Part_dilatadas = dilatar_imagem(Particulas);
    Area_particulas = obter_area(Particulas_dilatadas);
    Regiao_convexa_dilatada = obter_regiao_convexa (Particulas_dilatadas);
    Eixo_maior = obter_eixo_maior(Regiao_convexa);
    Eixo_menor = obter_eixo_menor(Regiao_convexa);
    Dif_fundo_int = obter_dif(Quadro , Fundo, Regiao_convexa);
    Dif_fundo_int = binarizar(Dif_fundo_int);
    Regiao_ext = dilatar_imagem(Regiao_convexa) – Regiao_convexa;
    Dif_fundo_ext = obter_dif(Quadro, Fundo, Regiao_ext);
    Dif_fundo_ext = binarizar(Dif_fundo_ext);
    // Obter parâmetros:
    Area_regiao = obter_area(Regiao_convexa);
    Area_regiao_dilatada = obter_area(Regiao_convexa_dilatada);
    Ocupacao = Area_particulas / Area_regiao_dilatada;
    Proporcao = Eixo_maior / Eixo_menor;
    Razao_int = Dif_fundo_int / Regiao_convexa;

```

```

Razao_ext = Dif_fundo_ext / Regiao_ext;
// Pontuar parâmetros:
Pontos = 0;
Se Area_regiao > Limiar_area, então: Pontos = Pontos + 1;
Senão: Pontos = Pontos - 1;
Fim Se
Se Ocupacao > Limiar_ocupacao, então: Pontos = Pontos + 1;
Senão: Pontos = Pontos - 1;
Fim Se
Se Proporcao < Limiar_proporcao, então: Pontos = Pontos + 1;
Senão: Pontos = Pontos - 1;
Fim Se
Se Razao_int > Limiar_razao_int, então: Pontos = Pontos + 1;
Senão: Pontos = Pontos - 1;
Fim Se
Se Razao_ext < Limiar_razao_ext, então: Pontos = Pontos + 1;
Senão: Pontos = Pontos - 1;
Fim Se
Se Pontos > 0, então: Alvos(i) = Agrupamentos(i);
Fim Se
Fim Para
Fim

```

ALGORITMO DO RASTREAMENTO DE VEÍCULOS

```

Início
Variáveis:
Veiculos_atuais = Ø; Regiao_atual = Ø; Particulas_atuais = Particulas(t);
Constantes:
Veiculos_ant = Veiculos(t-1); Quadro(t-1); Quadro(t);
Para i = 1 até nrVeiculos_ant, faça:
    Particulas_ant = extrai_particulas(Veiculos_ant(i));
    Vetores_mov_ant = extrai_vetores(Particulas_ant);
    Regiao_atual = obter_regiao_convexa(Particulas_ant, Vetores_mov_ant);
    Regiao_atual = dilatar_imagem(Regiao_atual);
    Particulas_regiao_atual = filtrar_part(Particulas_atuais, Regiao_atual);
    Para j = 1 até nrParticulas_regiao_atual, faça:
        Janela_atual = gerar_janela(Particulas_regiao_atual(j), Quadro(t));
        Janela_atual = requant_cores(Janela_atual);
        Sim = 0;
        Para k = 1 até nrParticulas_ant, faça:
            Janela_ant = gerar_janela(Particulas_ant(k), Quadro(t-1));
            Janela_ant = requant_cores(Janela_ant);
            H1 = obter_hist(Janela_atual); H2 = obter_hist(Janela_ant);
            Sim_temp = compara_hists(H1,H2); // Utilizando a Eq. ( 4 )
            Se Sim_temp > Sim, então:
                Sim = Sim_temp;
            Fim Se
        Fim Para k
    Se Sim < 0.8, então:
        Particulas_regiao_atual(j) = Ø;
    Fim Se
Fim Para j
Veiculos_atuais(i) = Particulas_regiao_atual;
Fim Para i
Particulas_atuais = extrai_particulas(Veiculos_atuais);

```

```
Para i = 1 até nrVeiculos_atuais - 1, faça:
  Particulas1 = extrai_particulas(Veiculos_atuais(i));
  Para j = 2 até nrVeiculos_atuais, faça:
    Particulas2 = extrai_particulas(Veiculos_atuais(j));
    Para m = 1 : nrParticulas1, faça:
      Para n = 1 até nrParticulas2, faça:
        Se Particulas1(m) == Particulas2(n), então:
          // Utilizando a Eq. ( 24 )
          M1 = Mahal(Particulas1,m);
          M2 = Mahal(Particulas2,n);
          Se M1 <= M2, então:
            Particulas2(n) = Ø;
          Senão:
            Particulas1(m) = Ø;
        Fim Se
      Fim Para n
    Fim Para m
  Veiculos(j) = Particulas2;
Fim Para j
Veiculos(i) = Particulas1;
Fim Para i
Fim
```