

**UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO
CENTRO TECNOLÓGICO
PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA**

VÍTOR BARBIRATO AZEVEDO

**IDENTIFICAÇÃO E LOCALIZAÇÃO DOS LIMITES DA REGIÃO
TRAFEGÁVEL PARA NAVEGAÇÃO DE UM VEÍCULO AUTÔNOMO**

VITÓRIA

2012

VÍTOR BARBIRATO AZEVEDO

**IDENTIFICAÇÃO E LOCALIZAÇÃO DOS LIMITES DA REGIÃO
TRAFEGÁVEL PARA NAVEGAÇÃO DE UM VEÍCULO AUTÔNOMO**

Dissertação apresentada ao Programa de Pós-Graduação em Informática do Centro Tecnológico da Universidade Federal do Espírito Santo, como requisito parcial para obtenção do Grau de Mestre em Informática.

VITÓRIA

2012

VÍTOR BARBIRATO AZEVEDO

**IDENTIFICAÇÃO E LOCALIZAÇÃO DOS LIMITES DA REGIÃO
TRAFEGÁVEL PARA NAVEGAÇÃO DE UM VEÍCULO AUTÔNOMO**

COMISSÃO EXAMINADORA

Prof. Dr. Alberto Ferreira De Souza
Universidade Federal do Espírito Santo
Orientador

Profa. Dra. Claudine Santos Badue Gonçalves
Universidade Federal do Espírito Santo
Coorientadora

Prof. Dr. Edilson de Aguiar
Universidade Federal do Espírito Santo

Prof. Dr. Luiz Chaimowicz
Universidade Federal de Minas Gerais

Vitória, 10 de Dezembro de 2012.

Dedico este trabalho a meus familiares e amigos, pelo amor, apoio incondicional e carinho, não somente neste trabalho, mas em toda a vida.

AGRADECIMENTOS

Primeiramente a Deus, pela benção em minha vida. Ao meu orientador professor Alberto Ferreira de Souza, pela ajuda e ensinamentos que tornaram este trabalho possível. À minha coorientadora, professora Claudine Badue e à professora Mariella Berger pela rigorosa revisão técnica deste trabalho. À minha mãe Danuza que além de me oferecer todas as condições para que pudesse prosseguir com meus estudos e aprimorar meu conhecimento, ainda corrigiu os erros ortográficos do texto. Aos meus familiares e a minha namorada Renata pelo apoio, confiança e incentivo. Aos amigos do LCAD pelo companheirismo, incentivo e valioso apoio técnico na construção deste trabalho.

RESUMO

Neste trabalho é apresentado um sistema para mapeamento da região trafegável em tempo real que se baseia principalmente em uma câmera estéreo. Para cada novo par de imagens recebido da câmera estéreo, o sistema primeiro determina a posição da câmera com respeito ao plano da região trafegável usando algoritmos de visão estéreo e métodos probabilísticos, e então projeta a imagem de referência para o sistema de coordenadas do mundo através de mapeamento inverso de perspectiva. Depois disso, o sistema classifica os pixels da imagem projetada como trafegáveis ou não trafegáveis, gerando um mapa de ocupação instantâneo. Em seguida, o mapa de ocupação instantâneo é integrado em um mapa de ocupação probabilístico usando o filtro de Bayes, que estima a probabilidade de ocupação de cada localização do mapa. Finalmente, o mapa de ocupação probabilístico é utilizado para localizar os limites laterais da região trafegável.

O desempenho do sistema para mapeamento da região trafegável foi avaliado em comparação com imagens classificadas manualmente. Os resultados experimentais mostram que o sistema é capaz de identificar corretamente até 92,22% das localizações trafegáveis entre 20 e 35 m à frente do veículo com uma Taxa de Falsa Aceitação não superior a 3,57% no mapeamento do pavimento de paralelepípedos.

ABSTRACT

This work presents a computer vision system for real-time mapping of the traversable region ahead of an autonomous vehicle that relies mostly on a stereo camera. For each new image pair captured by the stereo camera, the system first determines the camera position with respect to the ground plane using stereo vision algorithms and probabilistic methods, and then projects the camera raw image to the world coordinate system using inverse perspective mapping. After that, the system classifies the pixels of the inverse perspective map as traversable or not traversable, generating an instantaneous occupancy map. Next, the instantaneous occupancy map is integrated to a probabilistic occupancy map, using the Bayes filter to estimate the occupancy probability of each map location. Finally, the probabilistic occupancy map is used to localize the lateral limits of the traversable region.

The performance of the system for mapping the traversable region was evaluated in comparison to manually classified images. The experimental results show that the system is able to correctly map up to 92.22 % of the traversable locations between 20 and 35 m ahead of an autonomous vehicle with a False Acceptance Rate not superior to 3.57 % when mapping parallel-piped pavements.

SUMÁRIO

1	INTRODUÇÃO.....	17
1.1	MOTIVAÇÃO	19
1.2	OBJETIVOS	19
1.3	CONTRIBUIÇÕES.....	20
1.4	ESTRUTURA DO TRABALHO	20
2	O PROBLEMA DE MAPEAMENTO	22
2.1	MAPEAMENTO EM <i>GRIDS</i> DE OCUPAÇÃO	24
2.2	MAPEAMENTO BASEADO EM VISÃO COMPUTACIONAL.....	26
2.2.1	<i>Introdução a visão estéreo.....</i>	<i>27</i>
2.2.2	<i>Técnicas de mapeamento baseado em visão computacional e visão estéreo</i>	<i>30</i>
3	ESTIMATIVA DO ESTADO DA CÂMERA E MAPEAMENTO INVERSO DE PERSPECTIVA. 36	
3.1	ESTIMATIVA DO ESTADO DA CÂMERA ESTÉREO.....	37
3.2	PROJEÇÃO DA IMAGEM DE REFERÊNCIA PARA O SISTEMA DE COORDENADAS DO MUNDO ATRAVÉS DE MAPEAMENTO INVERSO DE PERSPECTIVA.....	40
4	CLASSIFICAÇÃO, MAPEAMENTO E LOCALIZAÇÃO DOS LIMITES DA REGIÃO TRAFEGÁVEL	43
4.1	CLASSIFICAÇÃO DAS REGIÕES TRAFEGÁVEIS E NÃO TRAFEGÁVEIS	44
4.1.1	<i>Determinação das regiões seguras.....</i>	<i>46</i>
4.1.2	<i>Classificação das regiões trafegáveis.....</i>	<i>47</i>
4.2	INTEGRAÇÃO BAYESIANA DE MAPAS INSTANTÂNEOS DA REGIÃO TRAFEGÁVEL	50
4.3	LOCALIZAÇÃO DOS LIMITES LATERAIS DA REGIÃO TRAFEGÁVEL	52
5	METODOLOGIA.....	55
5.1	BASE DE DADOS.....	55
5.1.1	<i>Imagens estáticas.....</i>	<i>56</i>
5.2	HARDWARE.....	59
5.3	SOFTWARE.....	61
5.3.1	<i>Ambiente CARMEN.....</i>	<i>62</i>
5.3.2	<i>Módulos CARMEN.....</i>	<i>63</i>
5.3.3	<i>Sistema operacional e artefatos de software</i>	<i>64</i>
5.4	MÉTRICAS.....	66
6	EXPERIMENTOS E RESULTADOS	69
6.1	RESULTADOS DA ESTIMATIVA DE ESTADO DA CÂMERA E MAPEAMENTO INVERSO DE PERSPECTIVA	69
6.2	RESULTADOS DO MAPEAMENTO PARA IMAGENS ESTÁTICAS	71
6.2.1	<i>Limitações do classificador.....</i>	<i>79</i>
6.3	AVALIAÇÃO DE DESEMPENHO COMPUTACIONAL	80
7	DISCUSSÃO.....	83
7.1	TRABALHOS CORRELATOS	83
7.2	ANÁLISE CRÍTICA DESTE TRABALHO DE PESQUISA	86
8	CONCLUSÃO.....	88
8.1	SÍNTESE	88

8.2	CONCLUSÕES	88
8.3	TRABALHOS FUTUROS	90
9	REFERÊNCIAS BIBLIOGRÁFICAS.....	91

LISTA DE SIGLAS E ABREVIATURAS

BLAS	<i>Basic Linear Algebra Subprograms</i>
BSD	<i>Berkeley Software Distribution</i>
CADU	Carro Autônomo Desenvolvido na UFMG
CUDA	<i>Computer Unified Device Architecture</i>
DARPA	<i>Defense Advanced Research Projects Agency</i>
EKF	<i>Extended Kalman Filter</i>
FAR	<i>False Acceptance Rate</i>
FN	<i>False Negative</i>
FP	<i>False Positive</i>
FPS	<i>Frames per Second</i>
GPL	<i>General Public License</i>
GPS	<i>Global Positioning System</i>
GPU	<i>Graphics Processing Unit</i>
GSL	<i>GNU Scientific Library</i>
HFOV	<i>Horizontal Field of View</i>
HSL	<i>Hue/Saturation/Lightness</i>
HSV	<i>Hue/Saturation/Value</i>
IMU	<i>Inertial Measurement Unit</i>
IPC	<i>Interprocess Communication</i>
LCAD	Laboratório de Computação de Alto Desempenho
LIDAR	<i>Light Detection and Ranging</i>
OpenMP	<i>Open Multi-Processing</i>
RAID	<i>Redundant Array of Inexpensive Disks</i>
RDDF	<i>Route Description Data File</i>
RGB	<i>Red/Green/Blue</i>
SIFT	<i>Scale-invariant Features Transform</i>

SLAM	<i>Simultaneous Localization and Mapping</i>
SMA	<i>SubMiniature version A</i>
SSD	<i>Solid State Drive</i>
SURF	<i>Speed Up Robust Features</i>
TN	<i>True Negative</i>
TP	<i>True Positive</i>
TPR	<i>True Positive Rate</i>
UFES	Universidade Federal do Espírito Santo
USB	<i>Universal Serial Bus</i>
UTM	<i>Universal Transversal de Mercator</i>

LISTA DE FIGURAS

Figura 1 — Fluxo de processamento do sistema de mapeamento da região trafegável.....	18
Figura 2 — Mapeamento de ambiente. (a) Mapa obtido com dados não processados (posições fornecidas pela odometria). (b) Mapa de <i>grid</i> de ocupação (<i>Occupancy grid map</i>).	24
Figura 3 — Modelo gráfico do mapeamento com localizações conhecidas. As variáveis sombreadas (localizações xt e medições zt) são conhecidas. O objetivo do mapeamento é a construção do mapa, denotado pela variável não sombreada m	25
Figura 4 — Exemplos de mapas probabilísticos obtidos com mapeamento em <i>grids</i> de ocupação. A intensidade da cor nas células dos <i>grids</i> é inversamente proporcional à probabilidade de ocupação (quanto mais claro menos provável é a ocupação). Em (a) a presença de um anteparo próximo ao robô permite mapear com precisão uma região menor do que em (b), onde o anteparo está mais afastado da localização do robô.	26
Figura 5 — Configuração de um sensor de alcance. O sensor de alcance realiza leituras em um HFOV de 180°, através da emissão de vários feixes com um pequeno e constante espaçamento angular entre feixes vizinhos. O resultado do tempo de trajeto da emissão até a reflexão de cada feixe em um obstáculo é utilizado para cálculo das distâncias.	27
Figura 6 — Visão geral da configuração de uma câmera estéreo com <i>baseline</i> b e distância focal f em relação ao sistema de coordenadas adotado neste trabalho. Um ponto $P(X, Y, Z)$ no espaço é simultaneamente projetado sobre as imagens direita e esquerda da câmera estéreo, observando que a projeção na câmera esquerda ocorre mais à direita da imagem do que a projeção na câmera direita.	28
Figura 7 — Par estéreo de imagens retificadas e mapa de disparidades calculado a partir do par de imagens. (a) Imagem retificada capturada pela câmera esquerda. (b) Imagem retificada capturada pela câmera direita. (c) Mapa de disparidades calculado a partir do par de imagens (a) e (b), utilizando o algoritmo proposto por Geiger (Geiger, 2010).	30
Figura 8 — Localização e mapeamento simultâneos para construção de um mapa de <i>landmarks</i> . Um robô R, equipado com um sensor de detecção de <i>landmarks</i> , construindo um mapa de <i>landmarks</i> (cada estrela representa um <i>landmark</i> diferente). A última leitura do sensor identificou os <i>landmarks</i> 1 e 2 e, com base nessa informação, o robô R simultaneamente atualiza as informações do mapa e se localiza neste mapa (a incerteza da localização é representada graficamente como a elipse escura na qual está localizado o robô R).	31
Figura 9 — Par de imagens estéreo sintéticas utilizadas para testes. (a) Imagem sintética esquerda. (b) Imagem sintética direita. (c) Mapa de disparidades para o par de imagens sintéticas (a) e (b).	32
Figura 10 — Análise do mapa em v-disparidades para imagens sintéticas. (a) Mapa em v-disparidades extraído a partir do mapa de disparidades da Figura 9(c). (b) Resultado da localização dos segmentos de reta no mapa em v-disparidades. (c) Resultados da análise do mapa de disparidades da Figura 9(c) e do respectivo mapa em v-disparidades em (b), destacando o plano trafegável (cor azul), o plano do teto (cor azul) e os obstáculos (cor vermelha).	33
Figura 11 — Análise do mapa em v-disparidades. (a) Resultados da análise do mapa de disparidades da Figura 7(c) e do respectivo mapa em v-disparidades em (b), destacando o plano trafegável (cor verde) e a linha do horizonte (linha vermelha). (b) Mapa em v-	

disparidades extraído a partir do mapa de disparidades da Figura 7(c). (c) Segmento de reta que representa o perfil da região trafegável extraído a partir do mapa em v-disparidades (b) utilizando a Transformada de Hough (Duda, 1972).	33
Figura 12 — Posição da câmera no sistema de coordenadas do mundo. Os graus de liberdade são a altura h e o ângulo θ em relação ao plano.	38
Figura 13 — Mapeamento inverso de perspectiva. (a) Imagem de referência (capturada pela câmera direita). (b) Resultado do mapeamento inverso de perspectiva (<i>Bird's Eye View</i>) realizado sobre a imagem de referência (a).	42
Figura 14 — Resultados do sistema de mapeamento baseado em visão computacional proposto por Thrun et al. (THRUN, 2006). (a) Uma imagem da câmera monocular. (b) Imagem (a) com a janela de segurança identificada pelo quadrilátero com contorno azul. (c) Resultado da classificação da imagem (a) baseada em mistura de gaussianas.	46
Figura 15 — Localização das janelas de segurança. (a) Janela de segurança na imagem de referência, destacada em verde. (b) Janela de segurança validada projetada para a imagem obtida a partir do mapeamento inverso de perspectiva da imagem (a).	47
Figura 16 — Classificação das regiões trafegáveis. (a) Representação gráfica do resultado da classificação da Figura 13(b) baseado em mistura de gaussianas. (b) Mapa da região trafegável resultante da binarização da imagem (a).	49
Figura 17 — Mapa probabilístico obtido a partir da integração bayesiana de mapas instantâneos aplicando-se o algoritmo descrito na Tabela 3.	52
Figura 18 – Localização dos limites laterais da região trafegável. (a) Mapa de ocupação probabilístico da região trafegável após binarização. (b) Divisão do mapa probabilístico em blocos (cor azul). (c) Localização dos limites da região trafegável, representados como uma transição entre um bloco não trafegável (vermelho) e uma sequência de blocos trafegáveis (primeiro bloco da sequência representado na cor verde).	54
Figura 19 – Visão aérea do campus de Goiabeiras da UFES com o anel viário destacado em vermelho.	55
Figura 20 – Imagem estática 1. (a) Imagem de referência (capturada pela câmera direta do par estéreo). (b) Imagem editada manualmente, identificando em preto a região trafegável.	57
Figura 21 – Imagem estática 2. (a) Imagem de referência (capturada pela câmera direta do par estéreo). (b) Imagem editada manualmente, identificando em preto a região trafegável.	57
Figura 22 – Imagem estática 3. (a) Imagem de referência (capturada pela câmera direta do par estéreo). (b) Imagem editada manualmente, identificando em preto a região trafegável.	58
Figura 23 – Imagem estática 4. (a) Imagem de referência (capturada pela câmera direta do par estéreo). (b) Imagem editada manualmente, identificando em preto a região trafegável.	58
Figura 24 – Vista lateral do veículo experimental. Localizado acima do veículo, uma câmera estéreo, um sensor de medida inercial (IMU) e um <i>Laser Range Finder</i> (não utilizado pelo sistema de mapeamento).	59
Figura 25 – Câmera estéreo Bumblebee XB3.	60
Figura 26 – Unidade de medida inercial Xsens MTi-G.	60
Figura 27 – Sistema de Posicionamento Global Venus GPS 09133.	61
Figura 28 – Visão geral do sistema de navegação autônoma, com destaque para o subsistema de mapeamento. As dependências entre módulos e as direções dos fluxos de dados são representados por setas unidirecionais, enquanto sensores são representados por blocos amarelos, controladores de sensores (drivers) são representados por blocos vermelhos e	

módulos de processamento de dados são representados por blocos verdes. Blocos cinza não são relevantes para o subsistema de mapeamento.	63
Figura 29 — Tabela de Contingência com as quatro situações possíveis para um classificador binário.	67
Figura 30 – Resultados do mapeamento inverso de perspectiva. (a) Mapeamento obtido sem a utilização do EKF. (b) Mapeamento obtido com a utilização do EKF.	69
Figura 31 – Gráficos da variação dos graus de liberdade do estado da câmera. (a) Variação da altura h (em metros) em função do frame. (b) Variação do ângulo de orientação θ (em graus) em função do frame.	70
Figura 32 – Histogramas dos graus de liberdade do estado da câmera. (a) Histograma da altura h (em metros), com aproximação de 0.01 m. (b) Histograma do ângulo de orientação θ (em graus), com aproximação de 0.1°	70
Figura 33 – Gráficos dos resultados da comparação do sistema de mapeamento com o classificador humano. (a) Gráfico em barras dos resultados segundo a métrica TPR. (b) Gráfico em barras dos resultados segundo a métrica FAR.	73
Figura 34 – Resultados da comparação do sistema de mapeamento com o classificador humano para a imagem estática 1. (a) Imagem de referência classificada manualmente. (b) Resultado do classificador humano - mapeamento inverso de perspectiva da imagem de referência classificada manualmente. (c) Resultado do mapeamento bayesiano – mapa de <i>grid</i> de ocupação probabilístico. (d) a projeção para a imagem de referência da região trafegável obtida pelo sistema de mapeamento. (e) Comparação do resultado do classificador humano (b) com o resultado do mapeamento bayesiano (c) – as regiões em verde identificam os verdadeiros positivos e as regiões em vermelho identificam os falsos positivos. (f) Projeção para a imagem de referência da comparação dos resultados segundo as métricas de Verdadeiros Positivos e Falsos Positivos.	74
Figura 35 – Resultados da comparação do sistema de mapeamento com o classificador humano para a imagem estática 2. (a) Imagem de referência classificada manualmente. (b) Resultado do classificador humano - mapeamento inverso de perspectiva da imagem de referência classificada manualmente. (c) Resultado do mapeamento bayesiano – mapa de <i>grid</i> de ocupação probabilístico. (d) a projeção para a imagem de referência da região trafegável obtida pelo sistema de mapeamento. (e) Comparação do resultado do classificador humano (b) com o resultado do mapeamento bayesiano (c) – as regiões em verde identificam os verdadeiros positivos e as regiões em vermelho identificam os falsos positivos. (f) Projeção para a imagem de referência da comparação dos resultados segundo as métricas de Verdadeiros Positivos e Falsos Positivos.	75
Figura 36 – Resultados da comparação do sistema de mapeamento com o classificador humano para a imagem estática 3. (a) Imagem de referência classificada manualmente. (b) Resultado do classificador humano - mapeamento inverso de perspectiva da imagem de referência classificada manualmente. (c) Resultado do mapeamento bayesiano – mapa de <i>grid</i> de ocupação probabilístico. (d) a projeção para a imagem de referência da região trafegável obtida pelo sistema de mapeamento. (e) Comparação do resultado do classificador humano (b) com o resultado do mapeamento bayesiano (c) – as regiões em verde identificam os verdadeiros positivos e as regiões em vermelho identificam os falsos positivos. (f) Projeção para a imagem de referência da comparação dos resultados segundo as métricas de Verdadeiros Positivos e Falsos Positivos.	76
Figura 37 – Resultados da comparação do sistema de mapeamento com o classificador humano para a imagem estática 4. (a) Imagem de referência classificada manualmente. (b) Resultado do classificador humano - mapeamento inverso de perspectiva da imagem de	

referência classificada manualmente. (c) Resultado do mapeamento bayesiano – mapa de *grid* de ocupação probabilístico. (d) a projeção para a imagem de referência da região trafegável obtida pelo sistema de mapeamento. (e) Comparação do resultado do classificador humano (b) com o resultado do mapeamento bayesiano (c) – as regiões em verde identificam os verdadeiros positivos e as regiões em vermelho identificam os falsos positivos. (f) Projeção para a imagem de referência da comparação dos resultados segundo as métricas de Verdadeiros Positivos e Falsos Positivos.77

Figura 38 – Mapeamento inverso de perspectiva de imagens capturadas pouco antes da captura da imagem estática 4. (a), (b) e (c) Resultados com significativas distorções do mapeamento inverso de perspectiva realizado sobre imagens de referência capturadas menos de 2 segundos antes da captura da imagem estática 4. (d) Resultado do mapeamento inverso de perspectiva realizado sobre a imagem estática 4.79

Figura 39 – Limitações do classificador para identificação de faixas e sinalizações desenhadas sobre a região trafegável. (a) Imagem de referência (capturada pela câmera direita). (b) Resultado da classificação realizada sobre a imagem de referência após realizar o mapeamento inverso de perspectiva.80

Figura 40 – Gráfico de pizza com o percentual médio de tempo gasto por cada etapa do sistema de mapeamento, considerando a versão sequencial do classificador.81

Figura 41 – Gráfico de pizza com o percentual médio de tempo gasto por cada etapa do sistema de mapeamento, considerando a versão paralelizada do classificador.82

LISTA DE TABELAS

Tabela 1 — Algoritmo <i>Occupancy Grid Mapping</i>	26
Tabela 2 — Algoritmo para mapeamento inverso de perspectiva para um <i>grid</i> bidimensional de dimensões gx, gy e resolução rm	41
Tabela 3 — Algoritmo para integração bayesiana de mapas instantâneos em um mapa probabilístico.	51
Tabela 4 — Resultados da comparação do sistema de mapeamento com o classificador humano, divididos em diferentes alcances. Os números de 1, 2, 3 e 4 identificam as imagens estáticas utilizadas nos experimentos.	72
Tabela 5 — Resultados da comparação do sistema de mapeamento com o classificador humano obtidos por Thrun et al. (THRUN,2006) no mapeamento de regiões trafegáveis em desertos e montanhas.	72

1 INTRODUÇÃO

O problema de navegação autônoma de veículos de passeio tem obtido atenção crescente da comunidade científica, especialmente após os desafios organizados pela *Defense Advanced Research Projects Agency* – DARPA nos anos de 2004, 2005 e 2007 (BUEHLER,2005; BUEHLER,2008). Nesses desafios, o objetivo principal das equipes participantes era desenvolver veículos autônomos capazes de navegar através dos trajetos definidos pela DARPA como uma série de pontos de rota (*waypoints*). No entanto, esses pontos de rota não correspondiam precisamente às vias através das quais os veículos autônomos deveriam trafegar. Além disso, poderia haver obstáculos no caminho entre os pontos de rota. Em cenários complexos como nos desafios propostos pela DARPA, veículos autônomos necessitam construir mapas das regiões trafegáveis e não trafegáveis em suas proximidades. Além da capacidade de mapeamento, veículos autônomos precisam, em muitos casos, determinar os limites da região trafegável com o intuito de obedecer às leis de trânsito vigentes.

Várias técnicas podem ser utilizadas para a construção de mapas de regiões trafegáveis e não trafegáveis ao redor de um veículo. Entretanto, dependendo dos sensores empregados para realizar o mapeamento, o alcance máximo pode variar consideravelmente (THRUN,2006). Por exemplo, no trabalho de Thrun et al. (THRUN,2006) a aplicação de um conjunto de sensores de alcance *Light Detection And Ranging* – LIDAR limitava o mapeamento de regiões a até 22 m do veículo autônomo. Para estender o alcance limitado do mapeamento realizado por um conjunto de LIDARs, Thrun et al. (THRUN,2006) empregaram uma câmera monocular que permitiu capturar regiões mais afastadas.

Neste trabalho é apresentado um sistema para mapeamento da região trafegável em tempo real que se baseia principalmente em uma câmera estéreo. Para cada novo par de imagens recebido da câmera estéreo, o sistema primeiro determina a posição da câmera com respeito ao plano da região trafegável (plano do mundo real que contém a região trafegável) usando algoritmos de visão estéreo e métodos probabilísticos, e então projeta a imagem de referência para o sistema de coordenadas do mundo através de mapeamento inverso de perspectiva. Depois disso, o sistema classifica os pixels da imagem projetada como trafegáveis ou não trafegáveis, gerando um mapa de ocupação instantâneo. Em seguida, o mapa de ocupação instantâneo é integrado em um mapa de ocupação probabilístico usando o filtro de Bayes, que estima a pro-

babilidade de ocupação de cada localização do mapa. Finalmente, o mapa de ocupação probabilístico é utilizado para localizar os limites laterais da região trafegável. A Figura 1 apresenta um diagrama que ilustra o fluxo de processamento do sistema de mapeamento da região trafegável. Na Figura 1, as setas representam a sequencia, em ordem cronológica, das etapas do processamento de uma iteração do sistema de mapeamento.

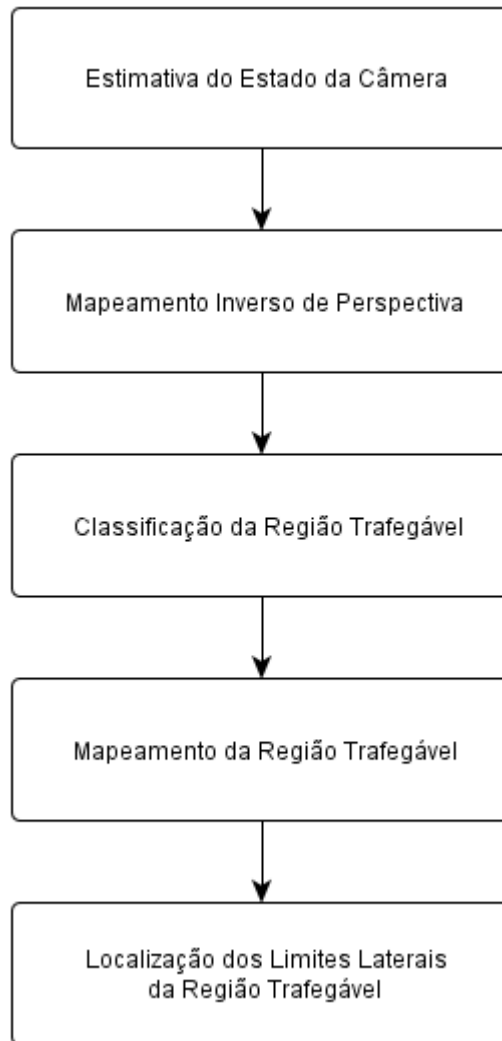


Figura 1 — Fluxo de processamento do sistema de mapeamento da região trafegável.

O desempenho do sistema para mapeamento da região trafegável foi avaliado em comparação com imagens classificadas manualmente. Os resultados experimentais mostram que o sistema é capaz de identificar corretamente até 92,22% das localizações trafegáveis entre 20 e 35 m à frente do veículo com uma Taxa de Falsa Aceitação (*False Acceptance Rate* – FAR) não superior a 3,57% no mapeamento do pavimento de paralelepípedos.

1.1 Motivação

A principal motivação para a realização deste trabalho é o interesse do Laboratório de Computação de Alto Desempenho (LCAD) da Universidade Federal do Espírito Santo (UFES) no desenvolvimento de um sistema de computação capaz de possibilitar que um carro de passeio adaptado trafegue de forma autônoma, seguindo trajetos pré-determinados, evitando colisões e obedecendo à legislação de trânsito.

Destaca-se também como grande motivação o interesse do grupo de pesquisadores do LCAD em aprofundar seus estudos sobre modelos matemático-computacionais da visão humana, a fim de compreender como nosso sistema visual viabiliza a cognição visual - compreensão do mundo e das ideias por meio da visão – em suas diversas facetas (percepção de profundidade, movimento e forma, reconhecimento de objetos, etc.) e, em particular, nossa capacidade de localização e mapeamento do ambiente. Isto é relevante no contexto de veículos autônomos porque somos capazes de dirigir um automóvel em alta velocidade e em condições desafiadoras utilizando nosso sistema visual para sensoriamento do ambiente e nossas habilidades cognitivas para processamento das informações visuais.

1.2 Objetivos

O objetivo principal deste trabalho é propor e avaliar um sistema de mapeamento em tempo real da região trafegável que se baseia principalmente nas informações visuais obtidas a partir de uma câmera estéreo frontal posicionada sobre um veículo autônomo. É também objetivo deste trabalho a localização dos limites laterais da região trafegável, com intuito de permitir que algoritmos de planejamento de caminhos utilizem a informação dos limites laterais para respeitar as leis de trânsito e minimizar as chances de se encontrar obstáculos à frente do veículo.

1.3 Contribuições

As principais contribuições deste trabalho são:

- Desenvolvimento de um método probabilístico para estimar com precisão o estado - posição e orientação - de uma câmera estéreo frontal localizada sobre um veículo autônomo em relação ao plano da região trafegável, baseando-se em informações visuais extraídas da própria câmera estéreo e processadas em tempo real por algoritmos de visão computacional;
- Extensão do método de classificação da região trafegável baseado em aprendizado de máquina proposto por Thrun et al. (THRUN,2006), com intuito de construir mapas de ocupação de alta qualidade da região trafegável empregando apenas uma câmera estéreo frontal como sensor de percepção do ambiente;
- Proposta e desenvolvimento de um algoritmo para localizar os limites laterais da região trafegável.

1.4 Estrutura do Trabalho

Esta dissertação está dividida da seguinte forma. No Capítulo 2 é apresentado o problema de mapeamento no campo da robótica, com especial destaque para o mapeamento baseado em visão computacional.

No Capítulo 3 são apresentadas as técnicas para estimativa do estado (posição e orientação) da câmera estéreo e projeção da imagem de referência para o sistema de coordenadas do mundo através de mapeamento inverso de perspectiva.

No Capítulo 4 é apresentada uma abordagem de aprendizado de máquina para classificação das regiões das imagens em trafegável e não trafegável, um método de integração bayesiana para geração de um mapa de ocupação probabilístico e um algoritmo para localização dos limites laterais da região trafegável.

No Capítulo 5 é apresentada a metodologia utilizada nos experimentos realizados para avaliar o desempenho do sistema de mapeamento, especificando as bases de dados, hardware, software e métricas utilizadas nesta avaliação.

No Capítulo 6 são apresentados e discutidos os resultados dos experimentos realizados para avaliar o sistema de mapeamento.

No Capítulo 7 são discutidos os principais trabalhos correlatos, bem como as principais limitações deste trabalho.

No Capítulo 8 é apresentado um breve sumário deste trabalho, suas principais conclusões e direções para trabalhos futuros.

2 O PROBLEMA DE MAPEAMENTO

Neste capítulo é apresentado o problema de mapeamento no campo da robótica, com especial destaque para o mapeamento baseado em visão computacional.

No campo da robótica móvel, o problema de Localização consiste em estimar a posição do robô dado que um mapa do ambiente é previamente conhecido. Essa situação é válida em um conjunto limitado de aplicações reais, nas quais um mapa do ambiente foi previamente construído e está disponível (THRUN,2005). Outros domínios de aplicação, no entanto, não apresentam o conhecimento prévio de um mapa do ambiente. Mesmo nos domínios nos quais é possível obter um mapa construído anteriormente, é provável que o mapa esteja desatualizado, em virtude de mudanças físicas no ambiente e elementos dinâmicos, como pessoas e outros robôs que se movimentam no ambiente.

O problema de Mapeamento pode ser definido como a capacidade de se construir um mapa dinamicamente. Tal capacidade pode reduzir os esforços envolvidos em se instalar um robô autônomo em um ambiente, e também permitir que o robô possa adaptar-se a mudanças sem a necessidade de supervisão humana. De fato, a capacidade de resolução do problema de Mapeamento no campo da robótica móvel é uma das competências mais fundamentais que deve estar presente em um robô verdadeiramente autônomo (THRUN,2005).

Segundo Thrun (THRUN,2005), a aquisição de mapas utilizando robôs móveis é um problema desafiador por uma série de razões:

- O espaço de hipóteses, isto é, o espaço que compreende todos os possíveis mapas, é enorme. Uma vez que mapas são definidos sobre um espaço contínuo, o espaço que compreende todos os possíveis mapas é infinito. Mesmo com aproximações discretas do espaço, como a aproximação em *grids* bidimensionais, mapas ainda podem ser descritos por 10^5 ou até mais variáveis;
- O aprendizado de mapas é um problema do tipo “o ovo ou a galinha”, razão pela qual ele é frequentemente denominado de Localização e Mapeamento Simultâneo (*Simultaneous Localization And Mapping* - SLAM). Primeiramente, um problema de Localização é introduzido: quando o robô movimenta-se através do ambiente, ele acumula erros de odometria, o que gradualmente torna a estimativa da sua posição no ambiente

mais incerta. Mesmo com toda a incerteza acumulada, existem métodos para determinar com precisão a localização, dado que um mapa do ambiente está disponível. A construção de um mapa nas situações em que a localização do robô é conhecida com precisão é também um problema relativamente fácil. Nas situações de ausência de um mapa inicial do ambiente e desconhecimento da posição exata, no entanto, o robô deve ser capaz de simultaneamente estimar o mapa do ambiente e se localizar neste mapa.

Nem todos os problemas de mapeamento são igualmente difíceis. Segundo Thrun (THRUN,2005), o grau de dificuldade de um problema de mapeamento resulta de uma correlação de fatores, dos quais os mais importantes são:

- Dimensão: quanto maior o tamanho do ambiente que deve ser mapeado, mais difícil torna-se a construção de um mapa;
- Ruídos na percepção e atuação: se os sensores e atuadores presentes nos robôs fossem totalmente precisos (ausência de erros), a construção de mapas seria simples. Quanto maior o ruído, mais difícil torna-se o problema de mapeamento;
- Ambiguidades na percepção: quando diferentes localizações são percebidas de maneira semelhante, é mais difícil estabelecer correspondências entre essas localizações. Ambiguidade é um fator complicado até mesmo para o ser humano, o que pode ser observado quando uma pessoa tenta localizar-se em um labirinto;
- Presença de ciclos: ciclos em um ambiente são particularmente difíceis de serem mapeados. Ciclos fazem com que os robôs retornem a uma mesma localização através de caminhos diferentes, e no fechamento do ciclo o erro de odometria acumulado pode ser bastante significativo.

A Figura 2 reforça a necessidade de se utilizar um bom algoritmo para resolução do problema de mapeamento, ilustrando mapas construídos a partir de um conjunto de dados coletados por um robô móvel em um amplo ambiente interno. O mapa de baixa qualidade da Figura 2(a) foi construído utilizando apenas as informações obtidas da odometria do robô (cada pixel preto da imagem representa um obstáculo detectado pelo sensor laser do robô). O mapa de melhor qualidade da Figura 2(b) representa um *grid* de ocupação construído ao aplicar um algoritmo probabilístico de mapeamento sobre o conjunto de dados coletados pelo robô móvel.

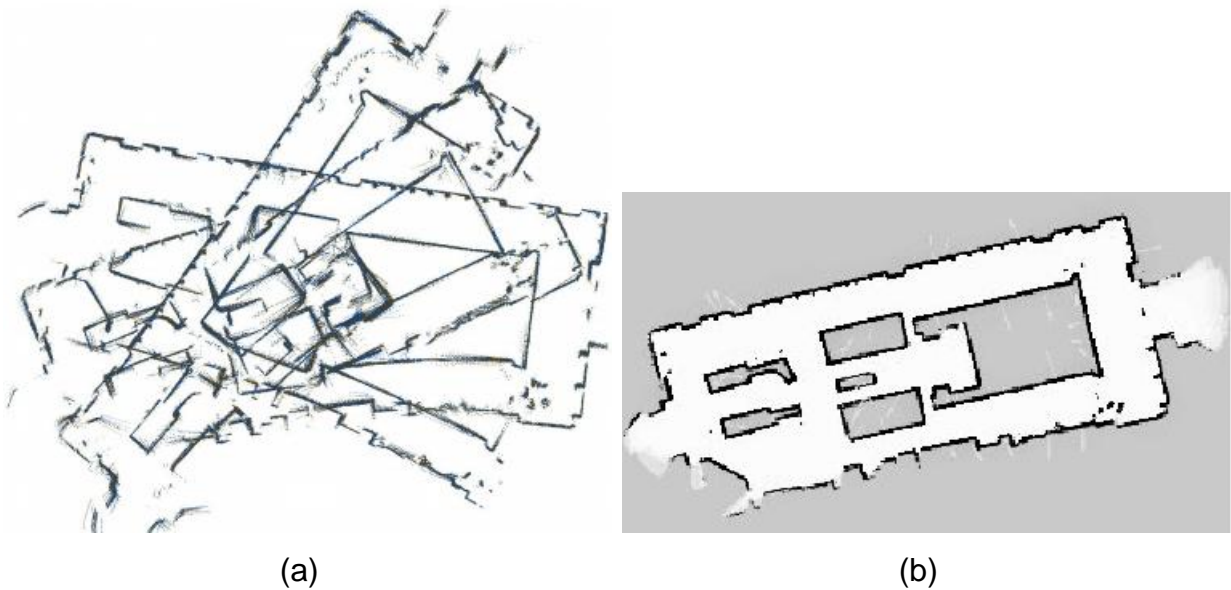


Figura 2 — Mapeamento de ambiente. (a) Mapa obtido com dados não processados (posições fornecidas pela odometria). (b) Mapa de *grid* de ocupação (*Occupancy grid map*).

Fonte: Adaptado de (THRUN,2005).

2.1 Mapeamento em *Grids* de Ocupação

A técnica para mapeamento em *grids* de ocupação foi introduzida por Elfes (ELFES,1987), ao propor um algoritmo para integrar medições de alcance oriundas de um sonar para mapeamento de um ambiente não estruturado e desconhecido. O trabalho de Elfes foi em seguida popularizado por Moravec (MORAVEC,1988), que formalizou alguns conceitos e generalizou a técnica para mapeamento com outros sensores. Segundo Thrun (THRUN,2005), mapeamento em *grids* de ocupação (*Occupancy Grid Mapping*) constitui uma família popular de algoritmos para o problema de mapeamento com localizações conhecidas (problema de mapeamento simplificado considerando que a localização exata do robô é sempre conhecida), que endereça o problema de geração de mapas consistentes a partir de informações incertas e com ruídos extraídas dos sensores. A Figura 3 apresenta o modelo gráfico de mapeamento com localizações conhecidas, onde a localização x_t do robô e a informação z_t obtida dos sensores do robô a cada instante de tempo t são conhecidas e utilizadas para a construção de um mapa denotado por m .

Um *grid* de ocupação representa o mapa m como um conjunto de células $\{m_i\}$, sendo que cada célula m_i possui uma probabilidade de ocupação $p(m_i) \in [0, 1]$. Para aprimorar a repre-

sentação das probabilidades de ocupação pode-se utilizar a representação *log odds*, que permite evitar instabilidades numéricas das probabilidades próximas dos valores limites zero e um (THRUN,2005). A Equação 1 define a representação l_i em *log odds* da probabilidade de ocupação $p(m_i)$ (THRUN,2005).

$$l_i = \log \frac{p(m_i)}{1 - p(m_i)} \quad \text{Equação 1}$$

A probabilidade de ocupação $p(m_i)$ da célula m_i pode ser recuperada a partir do valor l_i representado em *log odds* através da Equação 2 (THRUN,2005).

$$p(m_i) = 1 - \frac{1}{1 + e^{l_i}} \quad \text{Equação 2}$$

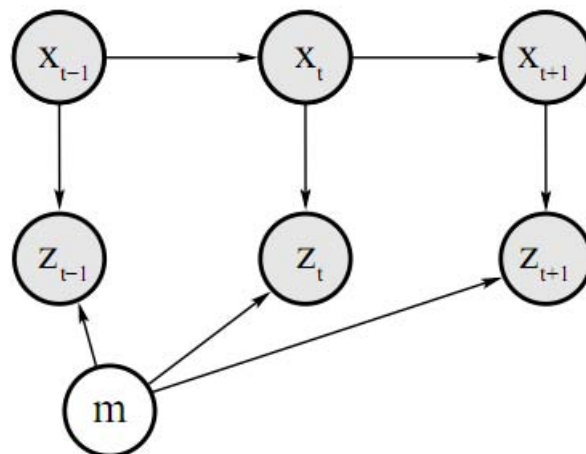


Figura 3 — Modelo gráfico do mapeamento com localizações conhecidas. As variáveis sombreadas (localizações x_t e medições z_t) são conhecidas. O objetivo do mapeamento é a construção do mapa, denotado pela variável não sombreada m .

Fonte: Adaptado de (THRUN,2005).

A Tabela 1 apresenta o algoritmo *Occupancy Grid Mapping* (THRUN,2005) para construção de um mapa de *grid* de ocupação. Este algoritmo pode ser facilmente adaptado para construção de um mapa utilizando sensores de alcance, tais como LIDARs ou sonares. A Figura 4 ilustra exemplos de mapas (*grids*) bidimensionais que são construídos utilizando técnicas de mapeamento em *grids* de ocupação. Na construção desses *grids* bidimensionais, tipicamente a informação (medição ou observação) obtida a partir dos sensores do robô é integrada ao longo do tempo. Dessa forma, a cada nova medição oriunda dos sensores, as localizações do mapa (células do *grid*) que estão no campo perceptivo do sensor são atualizadas. Se a medição informa que a localização é não ocupada, a probabilidade de ocupação da célula do *grid* é decrementada, e, caso contrário, a probabilidade de ocupação da célula do *grid* é incrementada.

A Figura 2(b) ilustra um mapa de *grid* de ocupação construído em um cenário real através da utilização do algoritmo *Occupancy Grid Mapping* apresentado na Tabela 1.

Tabela 1 — Algoritmo *Occupancy Grid Mapping*.

<p>l_i: representação em <i>log odds</i> da probabilidade de ocupação da célula m_i.</p> <p>x_t: localização do robô no instante t.</p> <p>z_t: informação (medição) obtida a partir dos sensores do robô no instante t.</p> <p>l_0: valor padrão para a probabilidade de ocupação. É definida por $l_0 = \log \frac{p(m_i=1)}{p(m_i=0)}$.</p> <p>occupancy_grid_mapping($\{l_i\}, x_t, z_t$)</p> <p>Para cada célula m_i do mapa $m = \{m_i\}$</p> <p>Se m_i está no campo perceptivo da medição z_t</p> $l_i = l_i + \log \frac{p(m_i z_t, x_t)}{1 - p(m_i z_t, x_t)} - l_0$
--

Fonte: Adaptado de (THRUN,2005).

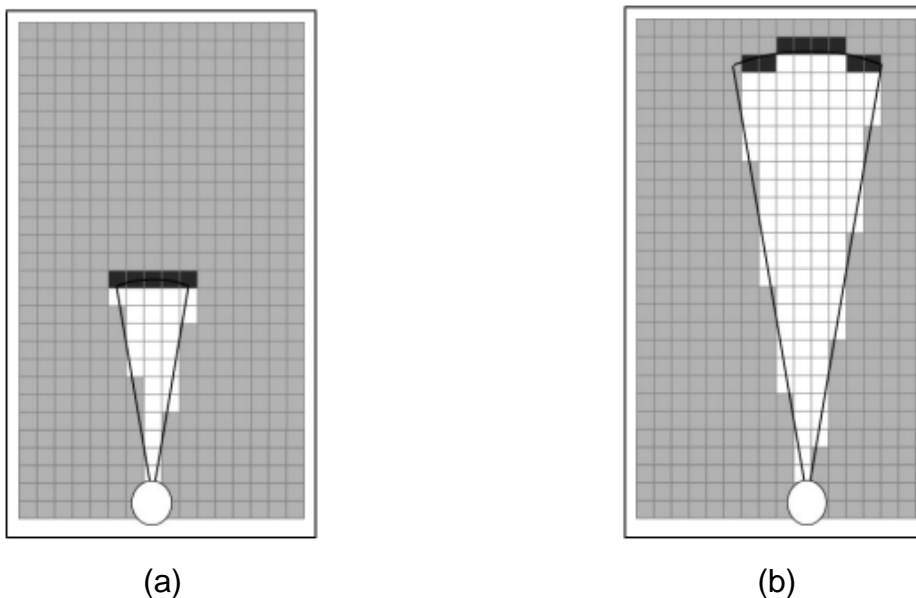


Figura 4 — Exemplos de mapas probabilísticos obtidos com mapeamento em *grids* de ocupação. A intensidade da cor nas células dos *grids* é inversamente proporcional à probabilidade de ocupação (quanto mais claro menos provável é a ocupação). Em (a) a presença de um anteparo próximo ao robô permite mapear com precisão uma região menor do que em (b), onde o anteparo está mais afastado da localização do robô.

Fonte: Adaptado de (THRUN,2005).

2.2 Mapeamento Baseado em Visão Computacional

A utilização de sensores de alcance, tais como Sonares e LIDARs, permite a construção de mapas de ocupação com qualidade, através da integração ao longo do tempo das informações

provenientes dos sensores. Uma grande vantagem dessa categoria de sensores é sem dúvida a simplicidade de interpretação dos dados, uma vez que a informação proveniente de sensores de alcance consiste basicamente em valores de distâncias, e possivelmente valores de intensidade de luz. Esses sensores tipicamente emitem um feixe de ondas, sendo as diferentes ondas emitidas de diferentes posições angulares, e assim que as ondas refletem em obstáculos, o sensor calcula o tempo de trajeto, desde a emissão até a detecção da onda refletida ou a diferença de fase das ondas, e com isso estima a distância do sensor até o obstáculo que causou a reflexão. A Figura 5 ilustra a configuração de um sensor de alcance com *Horizontal Field of View* (HFOV) de 180°. No caso dos sensores que capturam também os valores de intensidade de luz, como os LIDARs, as intensidades de luz dos obstáculos que causaram reflexão também podem ser calculadas.

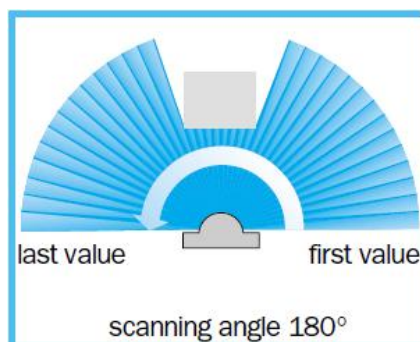


Figura 5 — Configuração de um sensor de alcance. O sensor de alcance realiza leituras em um HFOV de 180°, através da emissão de vários feixes com um pequeno e constante espaçamento angular entre feixes vizinhos. O resultado do tempo de trajeto da emissão até a reflexão de cada feixe em um obstáculo é utilizado para cálculo das distâncias.

Fonte: <http://www.sick-automation.ru/images/File/pdf/LMS%20Technical%20Description.pdf>.

Se a interpretação dos dados oriundos dos sensores de alcance é considerada simples, a interpretação dos dados provenientes de câmeras digitais, por outro lado, é considerada muito complexa, visto que a extração de distâncias a partir de imagens digitais não é trivial. Na tentativa de tratar essa complexidade, visão estéreo surge como um importante ramo de visão computacional capaz de extrair informações de alcance a partir de imagens digitais.

2.2.1 Introdução a visão estéreo

Na área de visão computacional, visão estéreo consiste em uma família de algoritmos que utilizam informações visuais obtidas a partir de um par de imagens digitais da mesma cena - o par estéreo - que foram capturadas no mesmo instante de tempo por um par de câmeras digi-

tais - a câmera estéreo - espaçadas por uma pequena distância horizontal, conhecida como *baseline*. A Figura 6 ilustra a configuração de uma câmera estéreo em relação ao sistema de coordenadas do mundo real adotado neste trabalho.

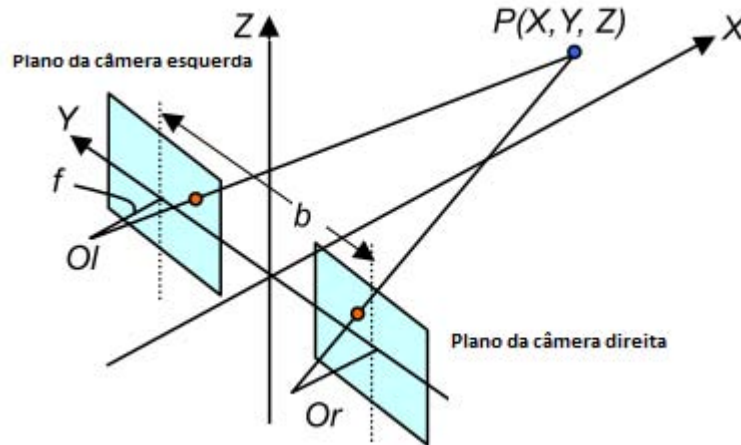


Figura 6 — Visão geral da configuração de uma câmera estéreo com *baseline* b e distância focal f em relação ao sistema de coordenadas adotado neste trabalho. Um ponto $P(X, Y, Z)$ no espaço é simultaneamente projetado sobre as imagens direita e esquerda da câmera estéreo, observando que a projeção na câmera esquerda ocorre mais à direita da imagem do que a projeção na câmera direita.

Fonte: Adaptado de (LAMOSA,2005).

Devido ao afastamento horizontal entre as câmeras monoculares que compõem uma câmera estéreo, a cena capturada por cada câmera é ligeiramente diferente, pois a câmera esquerda do par estéreo captura regiões mais a esquerda que não estarão presentes na imagem capturada pela câmera direita e, analogamente, a câmera direita captura regiões mais a direita que não estarão presentes na imagem capturada pela câmera esquerda. Essas regiões que não são simultaneamente capturadas por ambas as câmeras constituem regiões de oclusão que tipicamente não são interesse dos algoritmos de visão estéreo. Por outro lado, a região de interseção ou região comum presente em ambas as imagens, conhecida como região estéreo, permite a atuação dos algoritmos de visão estéreo.

O problema mais fundamental da visão estéreo é a identificação de correspondências entre os pixels que estão localizados na região estéreo (um pixel (x_r, y_r) da imagem direita que está localizado na região estéreo possui um correspondente (x_l, y_l) na imagem esquerda). Para resolvê-lo, primeiramente um par de imagens estéreo é retificado através do processo de retificação e, com isso, garante-se que as imagens estejam verticalmente alinhadas, ou seja, a relação $y_r = y_l$ é válida para pixels correspondentes. Devido à configuração das câmeras em virtude do afastamento horizontal, a relação $x_l > x_r$ também é válida. De posse dessas duas

restrições, o espaço para pesquisa por correspondências na região estéreo é limitado, simplificando o processamento dos algoritmos de visão estéreo, que identificam as correspondências na região estéreo e obtêm como resultado um mapa de disparidades. Uma disparidade é conceitualmente definida como a distância horizontal em pixels entre um ponto localizado na imagem direita e o seu equivalente localizado na imagem esquerda (BRADSKI,2008), e definida pela Equação 3 (LAMOSA,2005).

$$d = x_l - x_r \quad \text{Equação 3}$$

Um mapa de disparidades é uma imagem na qual a intensidade de cada pixel corresponde à disparidade do pixel (BRADSKI,2008). A Figura 7(c) apresenta um mapa de disparidades computado a partir de um par de imagens estéreo apresentado na Figura 7(a) e na Figura 7(b). A informação de disparidades permite a triangulação de um ponto (x_r, y_r) da imagem direita correspondente ao ponto (x_l, y_l) da imagem esquerda para uma localização tridimensional (X, Y, Z) no espaço tridimensional (sistema de coordenadas do mundo), através da Equação 4, Equação 5 e Equação 6 (LAMOSA, 2005; LIMA,2010), considerando uma câmera estéreo com afastamento horizontal b (em metros) e distância focal f (em pixels).

$$X = \frac{f * b}{x_l - x_r} \quad \text{Equação 4}$$

$$Y = \frac{x_r * X}{f} \quad \text{Equação 5}$$

$$Z = \frac{y_r * X}{f} \quad \text{Equação 6}$$

Ainda que a Equação 4, Equação 5 e Equação 6 possibilitem a extração de informações de distâncias a partir de informações visuais, a precisão obtida muitas vezes não é adequada, uma vez que visão estéreo enfrenta perda de precisão da profundidade à medida que os objetos presentes na cena capturada se afastam da câmera (GALLUP,2008). De fato, se a distância de um objeto para a câmera é multiplicada por um fator $n > 1$, a resolução da profundidade é diminuída na razão de $1/n^2$.

Embora a perda de precisão da visão estéreo com o aumento da profundidade limite consideravelmente a aplicabilidade de mapas de disparidades, visão estéreo e algoritmos derivados a partir de mapas de disparidades podem ser utilizados no problema de mapeamento.

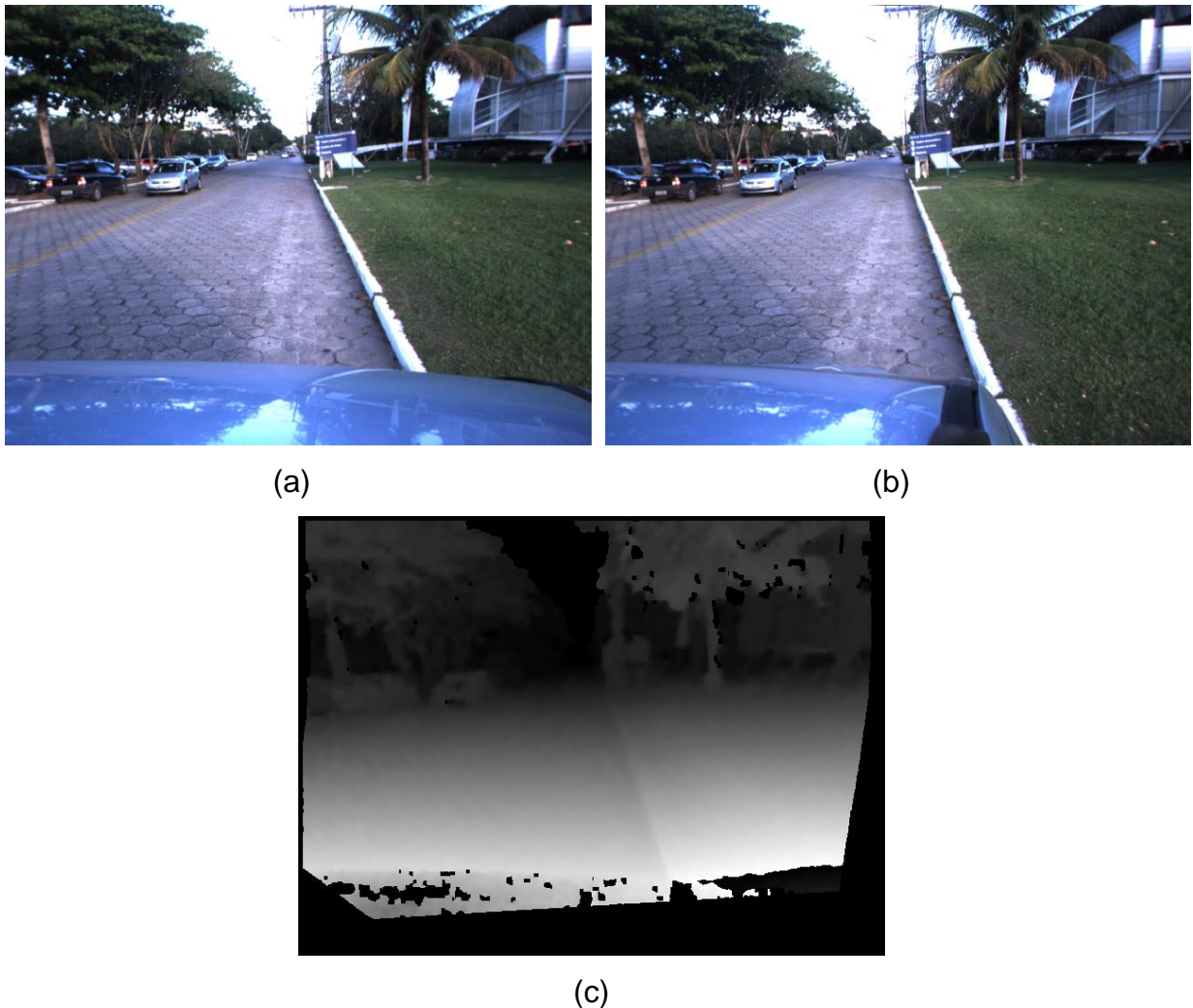


Figura 7 — Par estéreo de imagens retificadas e mapa de disparidades calculado a partir do par de imagens. (a) Imagem retificada capturada pela câmera esquerda. (b) Imagem retificada capturada pela câmera direita. (c) Mapa de disparidades calculado a partir do par de imagens (a) e (b), utilizando o algoritmo proposto por Geiger (Geiger, 2010).

2.2.2 Técnicas de mapeamento baseado em visão computacional e visão estéreo

Visão computacional e particularmente visão estéreo têm sido aplicadas com sucesso ao problema de mapeamento para identificação e localização de pontos de referência (construção de mapas de pontos de referência), identificação e localização de obstáculos (construção de mapas de obstáculos) ou identificação de planos trafegáveis.

Thrun (THRUN,2005) apresenta algoritmos genéricos para SLAM baseados na capacidade de um sensor em identificar e localizar pontos de referência, conhecidos como *landmarks*. Nesses algoritmos, o mapa é definido como um conjunto de pontos de referência, e o robô deve se

localizar no mapa de acordo com a sua posição em relação a esses pontos de referência. A Figura 8 ilustra uma situação de localização e mapeamento simultâneos baseado em mapas de *landmarks*. Embora esses algoritmos não estejam restritos ao uso de um sensor específico, a utilização de câmeras digitais é possível, uma vez que técnicas como SURF (*Speed Up Robust Features*) (BRADSKI,2008) ou SIFT (*Scale-invariant Features Transform*) (BRADSKI,2008) podem ser empregadas na identificação dos pontos de referência, enquanto visão estéreo (mapas de disparidades e a Equação 4, Equação 5 e Equação 6) pode ser empregada para a localização dos pontos de referência no espaço tridimensional.

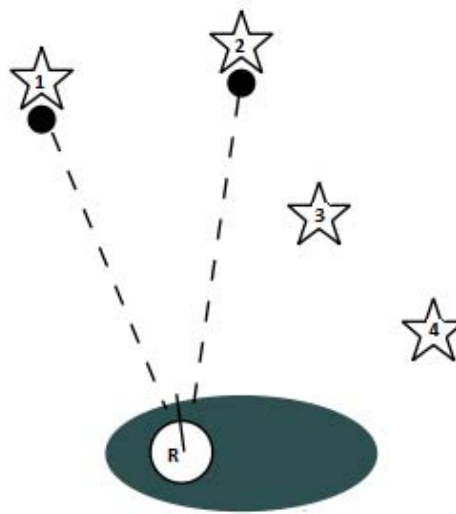


Figura 8 — Localização e mapeamento simultâneos para construção de um mapa de *landmarks*. Um robô R, equipado com um sensor de detecção de *landmarks*, construindo um mapa de *landmarks* (cada estrela representa um *landmark* diferente). A última leitura do sensor identificou os *landmarks* 1 e 2 e, com base nessa informação, o robô R simultaneamente atualiza as informações do mapa e se localiza neste mapa (a incerteza da localização é representada graficamente como a elipse escura na qual está localizado o robô R).

Fonte: Adaptado de (THRUN,2005).

Para identificação do plano trafegável e identificação e localização de obstáculos, Labayrade (LABAYRADE, 2002) introduz o conceito de mapa em v-disparidades, ou mapa de disparidades “V”, que é derivado diretamente a partir do mapa de disparidades. Por definição, um mapa em v-disparidades é uma matriz bidimensional na qual cada linha y corresponde a uma linha do mapa de disparidades, cada coluna x corresponde a um valor de disparidade, e cada elemento (x, y) é dado pela quantidade total de pixels na linha y que possuem disparidade igual a x (LABAYRADE,2002; LAMOSA,2005).

Um mapa em v-disparidades pode ser analisado para identificar características importantes da geometria da cena no mundo real, tais como o plano trafegável, a localização de obstáculos e

a linha do horizonte (LABAYRADE,2002; LAMOSA,2005). O plano trafegável pode ser definido como o plano do mundo real que contém as localizações com altura $Z = -h$, sendo h a altura da câmera estéreo em relação à região trafegável. A região trafegável corresponde ao subconjunto de localizações no plano trafegável que podem ser trafegados com segurança por um veículo.

No domínio de um mapa em v -disparidades, segmentos de reta contínuos representam o perfil da maioria dos tipos de superfícies que se apresentam no sistema de coordenadas do mundo real (LAMOSA,2005), e dentre essas superfícies destacam-se o plano trafegável e os obstáculos. A Figura 9 apresenta um par de imagens estéreo sintéticas e o respectivo mapa de disparidades utilizados por Lamosa (LAMOSA,2005) em seus experimentos. A Figura 10(c) apresenta os resultados obtidos por Lamosa ao analisar o mapa de disparidades da Figura 9(c) e o respectivo mapa em v -disparidades da Figura 10(a). A Figura 10(b) apresenta os resultados da detecção de segmentos de reta na Figura 10(b), sendo que os segmentos de reta inclinados (2) e (11) representam, respectivamente, o chão e o teto, enquanto os demais segmentos de reta com inclinação aproximada de $\pi/2$ representam obstáculos. A Figura 11(a) apresenta o resultado da análise do mapa de disparidades da Figura 7(c) e do respectivo mapa em v -disparidades da Figura 11(b) para identificação da região trafegável e da linha do horizonte. A Figura 11(c) apresenta o resultado da detecção do segmento de reta contínuo e inclinado que denota o perfil da região trafegável a partir do mapa em v -disparidades da Figura 11(b). A detecção do segmento foi realizada através da Transformada de Hough (DUDA,1972).

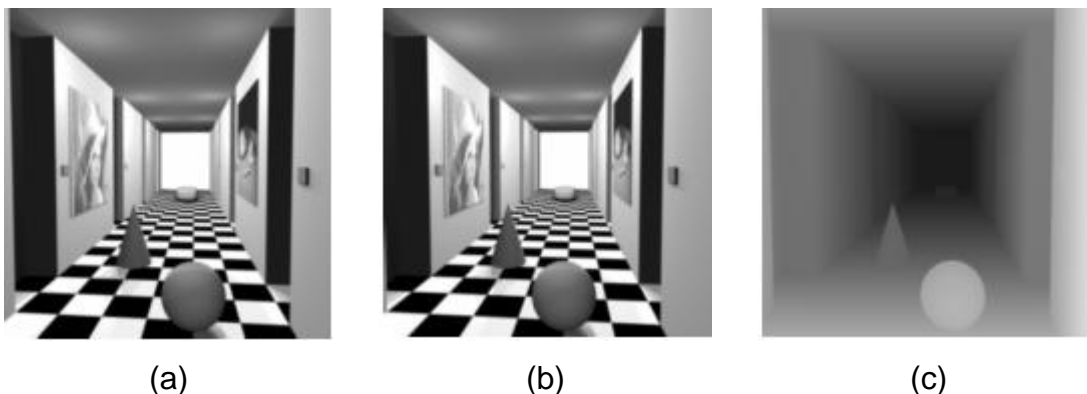


Figura 9 — Par de imagens estéreo sintéticas utilizadas para testes. (a) Imagem sintética esquerda. (b) Imagem sintética direita. (c) Mapa de disparidades para o par de imagens sintéticas (a) e (b).

Fonte: Adaptado de (LAMOSA,2005).

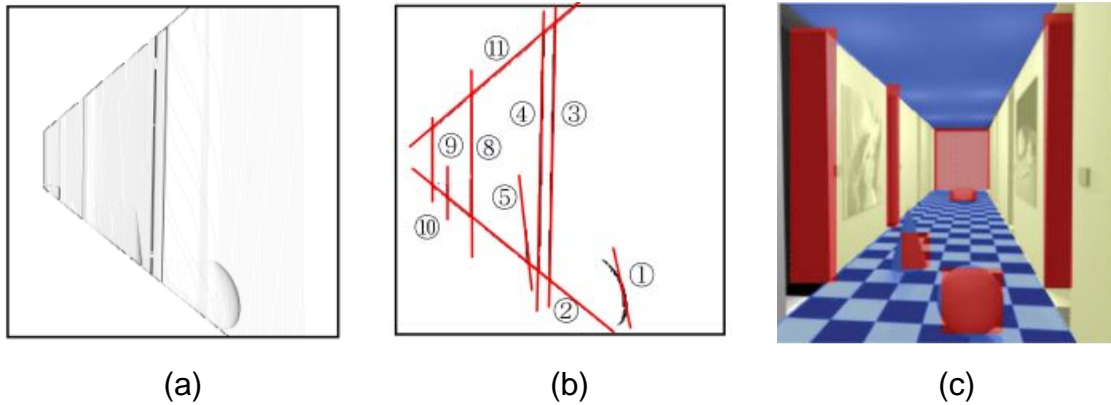


Figura 10 — Análise do mapa em v-disparidades para imagens sintéticas. (a) Mapa em v-disparidades extraído a partir do mapa de disparidades da Figura 9(c). (b) Resultado da localização dos segmentos de reta no mapa em v-disparidades. (c) Resultados da análise do mapa de disparidades da Figura 9(c) e do respectivo mapa em v-disparidades em (b), destacando o plano trafegável (cor azul), o plano do teto (cor azul) e os obstáculos (cor vermelha).

Fonte: Adaptado de (LAMOSA,2005).

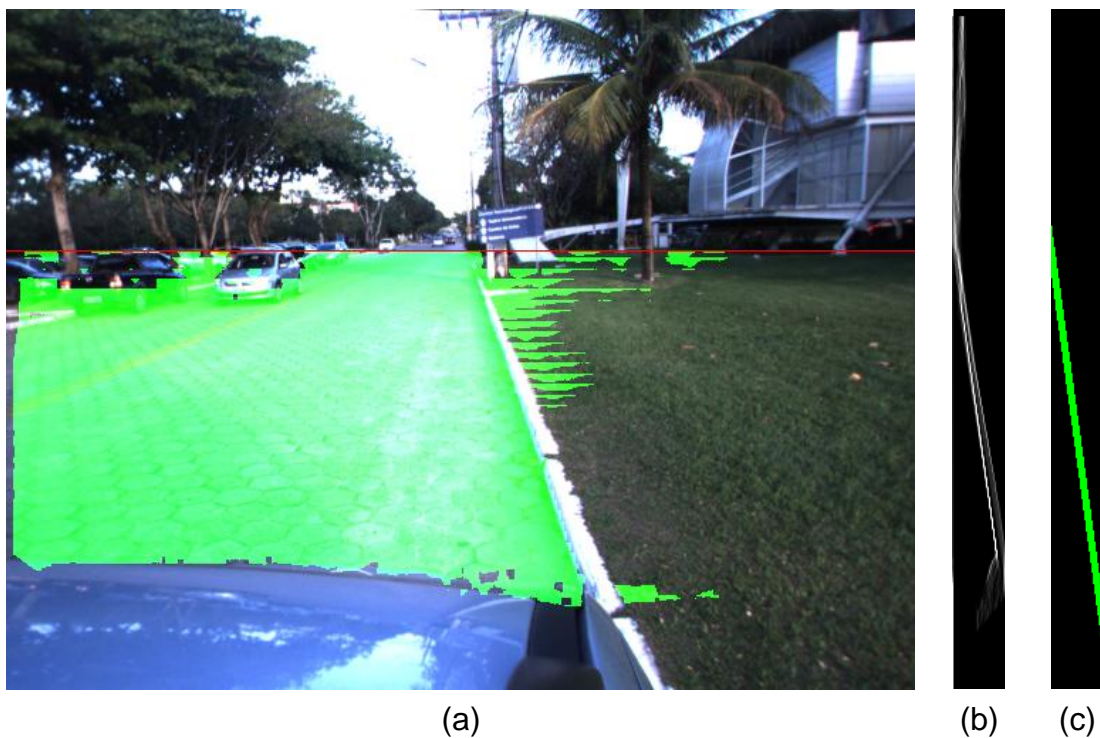


Figura 11 — Análise do mapa em v-disparidades. (a) Resultados da análise do mapa de disparidades da Figura 7(c) e do respectivo mapa em v-disparidades em (b), destacando o plano trafegável (cor verde) e a linha do horizonte (linha vermelha). (b) Mapa em v-disparidades extraído a partir do mapa de disparidades da Figura 7(c). (c) Segmento de reta que representa o perfil da região trafegável extraído a partir do mapa em v-disparidades (b) utilizando a Transformada de Hough (Duda, 1972).

No domínio de um mapa em v-disparidades, o plano trafegável é projetado como um segmento de reta contínuo e inclinado com um ângulo superior a $\pi/2$ (LAMOSA,2005). Este segmen-

to de reta, que denota o perfil da região trafegável, apresenta-se dessa forma no mapa em v-disparidades porque tipicamente as regiões trafegáveis apresentam variações graduais na disparidade desde as proximidades da câmera até o horizonte, com o mesmo valor em uma linha do mapa de disparidades (LIMA,2010). Além disso, a região trafegável ocupa a maior parte da cena capturada pela câmera estéreo frontal e, conseqüentemente, também preenche a maior parte do mapa de disparidades. Assim, seu perfil acaba por dominar também o mapa em v-disparidades.

No domínio do sistema de coordenadas do mundo real, obstáculos localizados a frente do veículo se apresentam como superfícies verticais. Quando esses obstáculos são representados no domínio de um mapa em v-disparidades, eles se apresentam como segmentos de reta contínuos verticais com inclinação igual a $\pi/2$ (LAMOSA,2005). Essa situação ocorre porque tipicamente uma superfície vertical do mundo real apresenta altas intensidades das mesmas disparidades em linhas adjacentes do mapa de disparidades. Com isso, a análise de mapas em v-disparidades possibilita ao menos duas formas de identificação de obstáculos: através da localização de segmentos de reta verticais no mapa em v-disparidades ou através da análise do perfil do plano trafegável, considerando que as regiões que não pertencem ao plano trafegável constituem potenciais obstáculos. Enquanto Lima (LIMA,2010) empregou a primeira abordagem em seu trabalho, Labayrade (LABAYRADE,2002) empregou a segunda.

A análise de mapas em v-disparidades também pode ser empregada para uma localização aproximada da linha do horizonte. Por definição, os pontos da imagem localizados na linha do horizonte estão a uma distância infinita (LABAYRADE,2002). Conseqüentemente, pela Equação 4, a disparidade desses pontos é nula e, além disso, esses pontos pertencem à região trafegável (LABAYRADE,2002). No domínio do mapa em v-disparidades, a localização aproximada da linha do horizonte pode ser obtida a partir da análise do segmento de reta que denota o perfil da região trafegável, e corresponde à interseção do segmento de reta com o eixo das ordenadas no mapa em v-disparidades. Esta interseção ocorre na linha do mapa em v-disparidades a partir da qual o valor de disparidade zero passa a ter intensidade dominante.

A utilização de mapas em v-disparidades para navegação de veículos autônomos terrestres já foi empregada em aplicações reais. Lima (LIMA,2010) emprega mapas em v-disparidades para detecção de obstáculos e utiliza essa informação para permitir planejamento de caminho e desvio de obstáculos no Carro Autônomo Desenvolvido na UFMG (CADU). Hummel

(HUMMEL,2006) apresenta um método de planejamento de caminhos em ambientes não estruturados baseado em mapas obtidos a partir de visão computacional utilizando mapas em v-disparidades, que foi empregado no veículo autônomo Desert Buckeyes durante o desafio DARPA *Grand Challenge 2005* (BUEHLER,2005). Com esse planejamento de caminho, o Desert Buckeyes foi capaz de evitar todos os obstáculos da etapa de qualificação e de dirigir autonomamente em terrenos completamente desconhecidos a distância de aproximadamente 50 km na corrida final da DARPA *Grand Challenge 2005* (HUMMEL,2006).

3 ESTIMATIVA DO ESTADO DA CÂMERA E MAPEAMENTO INVERSO DE PERSPECTIVA

Neste capítulo são apresentadas as técnicas para estimativa do estado (posição e orientação) da câmera estéreo e projeção da imagem de referência para o sistema de coordenadas do mundo através de mapeamento inverso de perspectiva.

O sistema proposto neste trabalho tem por objetivo a construção de um mapa de *grid* bidimensional que represente o plano da região trafegável. Bertozzi (BERTOZZI,1998) utiliza a técnica de mapeamento inverso de perspectiva (*Inverse Perspective Mapping*) como o primeiro passo para obtenção de um mapa do plano trafegável a partir das imagens originais de uma câmera. Segundo Bertozzi (BERTOZZI,1998), para a realização deste tipo de mapeamento é necessário prévio conhecimento das condições específicas de aquisição da imagem (posição e orientação da câmera e parâmetros intrínsecos da câmera). Além disso, por questões de simplificação, é comum assumir também que a cena capturada representa uma região plana (BERTOZZI,1998).

Os parâmetros intrínsecos da câmera podem ser obtidos através de métodos de calibração como, por exemplo, o método proposto por Zhang (ZHANG,2000) e cuja implementação está disponível na ferramenta *Camera Calibration Toolbox for Matlab* (http://www.vision.caltech.edu/bouguetj/calib_doc/) e na biblioteca de código aberto *OpenCV Computer Vision Library* (BRADSKI,2008). Também, a restrição de que a cena capturada em um cenário de navegação de um veículo autônomo representa uma região plana constitui uma boa hipótese, uma vez que as vias (ruas, avenidas, estradas e rodovias) são geralmente planas, pelo menos nas regiões próximas ao veículo tipicamente capturadas na imagem. Com isso, o principal desafio para realizar o mapeamento inverso de perspectiva é a obtenção do estado da câmera.

A estimativa do estado da câmera é obtida utilizando visão computacional e filtros de *Kalman* (KALMAN,1960). Para isso, mapas de disparidades são computados a partir das imagens originais da câmera estéreo e, a partir dos mapas de disparidades, são construídos os mapas em *v*-disparidades. O passo seguinte é detectar, no mapa em *v*-disparidades, o segmento de reta que denota o perfil da região trafegável, isto é, o segmento de reta contínuo e inclinado

com ângulo superior a $\pi/2$. Por fim, uma variante dos filtros probabilísticos de *Kalman*, o *Extended Kalman Filter* (EKF), é aplicado para estimar o estado da câmera, utilizando informações obtidas a partir da fusão de sensores no passo de predição do filtro e informações derivadas do mapa em v-disparidades no passo de correção do filtro.

De posse do estado estimado da câmera estéreo, a imagem original de uma das câmeras do par estéreo (neste trabalho a imagem de referência é a capturada pela câmera direita) é projetada para o sistema de coordenadas do mundo real através de mapeamento inverso do sistema de coordenadas do mundo para o sistema de coordenadas da câmera. Para isso, primeiramente é assumido que todas as posições no sistema de coordenadas do mundo estão localizadas no plano trafegável e, utilizando a estimativa do estado da câmera, essas posições do mundo real são mapeadas para os pontos correspondentes na imagem de referência. Por fim, um *grid* bidimensional é construído, sendo cada célula do *grid* correspondente a uma posição no mundo real associada com um ponto na imagem de referência, e a intensidade do pixel na imagem de referência é utilizada para representar a intensidade do pixel no *grid* bidimensional. Esta representação das regiões próximas e à frente do veículo autônomo em um *grid* bidimensional, conhecida formalmente na literatura como mapeamento inverso de perspectiva, é muitas vezes denominada *Bird's Eye View*, em referência à perspectiva que um pássaro sobrevoando o ambiente teria das cenas capturadas.

3.1 Estimativa do Estado da Câmera Estéreo

Na maioria das situações de navegação de um veículo, a altura do veículo e a orientação do veículo com respeito ao plano da região trafegável (ângulo *pitch*) não se mantêm constantes durante todo o trajeto. Isto ocorre porque frequentemente o plano da região trafegável não é uma superfície plana, o número de passageiros e o peso total do veículo variam de viagem para viagem, e o sistema de amortecimento e suspensão oscila em diferentes graus de acordo com o tipo de terreno, a velocidade e o peso do veículo. O sistema visual humano é suficientemente robusto para tolerar essas e outras variações, permitindo aos seres humanos a criação de suas próprias representações do mundo real, ou, no caso de interesse, excelentes mapas da região trafegável. Esses mapas tornam a tarefa de dirigir possível e até mesmo simples para a maioria das pessoas.

Para permitir que um veículo autônomo construa uma representação apropriada do mundo real, um sistema de inteligência artificial deve ser capaz de tolerar as variações que ocorrem no estado da câmera. Embora o estado da câmera possua muitos graus de liberdade, neste trabalho o estado da câmera é definido como sendo as duas variáveis a seguir:

- h : altura da câmera com relação ao plano da região trafegável;
- θ : ângulo de orientação *pitch*, formado entre o eixo ótico da câmera e o plano da região trafegável.

A Figura 12 ilustra os dois graus de liberdade que definem o estado da câmera. Os outros graus de liberdade não são essenciais para construir um mapa da região trafegável com a precisão desejada neste trabalho.

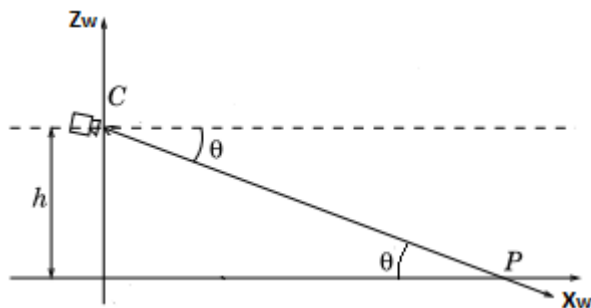


Figura 12 — Posição da câmera no sistema de coordenadas do mundo. Os graus de liberdade são a altura h e o ângulo θ em relação ao plano.

Como o estado da câmera não se mantém constante ao longo do tempo durante um trajeto, não é conveniente calcular esse estado previamente e usá-lo para construir um mapa. Para resolver este problema, o sistema proposto neste trabalho utiliza o filtro de Kalman Estendido (*Extended Kalman Filter* - EKF) a fim de obter uma aproximação robusta do estado da câmera.

Um filtro de Kalman tipicamente opera em dois passos: predição e correção. No passo de predição, um mecanismo, tipicamente um sensor, é empregado para prever o valor atual das variáveis de interesse. No passo de correção, um mecanismo diferente, tipicamente outro sensor, é utilizado para melhorar a estimativa dos valores das variáveis calculados no passo de predição. Neste trabalho foram utilizadas para a estimativa de estado da câmera informações de localização fornecidas pela fusão de diferentes sensores para realizar o passo de predição, e informações derivadas de visão estéreo para realizar o passo de correção. A combinação dos dois passos do filtro é essencial porque por um lado as localizações fornecidas pela fusão de

sensores não são diretamente relacionadas com o estado da câmera, mas sim com a posição geográfica e orientação do veículo, enquanto que por outro lado as informações derivadas de visão estéreo são precisamente relacionadas ao estado da câmera, mas apresentam incertezas consideráveis.

No passo de predição, informações de localização fornecidas pela fusão de sensores permitem calcular variações no ângulo de orientação e na altitude do veículo. O EKF adiciona essas variações ao estado da câmera, produzindo um estado predito.

No passo de correção, apenas informações derivadas a partir de algoritmos de visão estéreo são utilizadas. Para isso, primeiramente um mapa de disparidades é computado a partir do par de imagens capturado pela câmera estéreo. É importante observar que o procedimento independe do algoritmo utilizado para cálculo do mapa de disparidades, embora critérios como eficiência (para permitir o processamento em tempo real) e qualidade (para obter estimativas de estado superiores) sejam fundamentais. Em seguida, a partir de um mapa de disparidades, é computado um mapa em v-disparidades e, então, o segmento de reta que representa o perfil da região trafegável é detectado neste mapa através da Transformada de Hough (DUDA,1972). Deste segmento são calculados o seu coeficiente angular C_r e seu ponto de interseção $(0, v_{0r})$ com o eixo vertical. A partir de C_r e v_{0r} é possível estimar através da Equação 7 e Equação 8 (LABAYRADE,2002) as variáveis de interesse ângulo de orientação θ e altura h da câmera. Por fim, o EKF emprega os valores estimados das variáveis θ e h no passo de correção.

$$\theta = \text{atan} \frac{(v_0 - v_{0r})}{\alpha} \quad \text{Equação 7}$$

$$h = b * \frac{\cos \theta}{C_r} \quad \text{Equação 8}$$

Na Equação 7 e Equação 8, (u_0, v_0) representam as coordenadas, em pixels, da projeção do eixo ótico da câmera de referência (câmera direita) na imagem capturada; α é um parâmetro derivado definido por $\alpha = \frac{f}{t}$, onde f é a distância focal, em metros, da câmera de referência e t é o tamanho de um pixel, em metros; e b é o afastamento horizontal (*baseline*), em metros, entre as câmeras do par estéreo.

3.2 Projeção da Imagem de Referência para o Sistema de Coordenadas do Mundo Através de Mapeamento Inverso de Perspectiva

A construção de um mapa de *grid* de ocupação utilizando sensores de alcance é relativamente simples se a localização do sensor com relação ao sistema de coordenadas do mundo real é conhecida (THRUN,2005). Em contraste, a construção de mapas de *grid* de ocupação utilizando câmeras é considerada muito mais complexa, uma vez que a extração de informação de profundidades a partir das imagens capturadas não é trivial.

Quando se trabalha com câmeras, observa-se inicialmente um problema dimensional: o sistema de coordenadas do mundo real é tridimensional, enquanto o sistema de coordenadas das imagens de câmeras é bidimensional, ou seja, o processo de captura de imagens utilizando câmeras “perde” as informações de profundidade da cena do mundo real. Visão estéreo busca solucionar essa questão ao adicionar a dimensão da disparidade ao sistema de coordenadas da câmera estéreo. Entretanto, visão estéreo perde resolução de profundidade à medida que os objetos presentes na cena capturada se afastam da câmera estéreo (GALLUP,2008). Em virtude da baixa resolução de profundidade da câmera estéreo utilizada nos experimentos deste trabalho (Point Grey Bumblebee XB3 stereo camera), no mapeamento de regiões a mais de 10 metros os mapas de disparidades são utilizados apenas para o cálculo dos mapas em v-disparidades que são empregados na estimativa do estado da câmera.

Para projetar a imagem da câmera de referência em um *grid* bidimensional são utilizadas a Equação 9, Equação 10 e Equação 11 (SE,2002) para converter a partir do sistema de coordenadas da câmera, onde um ponto bidimensional é denotado por (x_c, y_c) , para o sistema de coordenadas do mundo real, onde um ponto tridimensional é denotado por (X_w, Y_w, Z_w) .

$$Z_w = -h \quad \text{Equação 9}$$

$$X_w = \frac{-Z_w * (\alpha * \cos \theta + (v_0 - y_c) * \sen \theta)}{\alpha * \sen \theta + (y_c - v_0) * \cos \theta} \quad \text{Equação 10}$$

$$Y_w = \frac{(x_c - u_0) * (X_w * \cos \theta - Z_w * \sen \theta)}{\alpha} \quad \text{Equação 11}$$

A partir da Equação 9, Equação 10 e Equação 11 é possível derivar a Equação 12 e Equação

13 que convertem do sistema de coordenadas do mundo real para o sistema de coordenadas da câmera.

$$x_c = u_0 + \frac{\alpha * Y_w}{X_w * \cos \theta - Z_w * \sin \theta} \quad \text{Equação 12}$$

$$y_c = v_0 + \frac{-\alpha * (X_w * \sin \theta + Z_w * \cos \theta)}{X_w * \cos \theta - Z_w * \sin \theta} \quad \text{Equação 13}$$

De posse da Equação 12 e Equação 13, o mapeamento inverso de perspectiva é simples e pode ser descrito como a seguir. Primeiro, um *grid* bidimensional é definido nas coordenadas do mundo real (no caso da região trafegável, um plano que está à mesma altura $Z_w = -h$ da origem do eixo ótico da câmera) em frente à câmera, onde todas as células de um *grid* bidimensional de dimensões (g_x, g_y) representam localizações do mundo real com $r_m \times r_m$ m² de área, ou seja, o plano da região trafegável é discretizado em unidades quadradas de dimensão r_m metros. Então, para cada célula do *grid*, denotada em coordenadas do mundo por $(X_w, Y_w, -h)$, a Equação 12 e Equação 13 são utilizadas para obter o pixel na imagem de referência, denotado por (x_c, y_c) , que corresponde a célula do *grid* e o valor da célula do *grid* recebe o valor da intensidade do pixel (x_c, y_c) . O algoritmo de mapeamento inverso de perspectiva é apresentado na Tabela 2. A Figura 13(b) apresenta o resultado da aplicação do algoritmo de mapeamento inverso de perspectiva sobre a imagem de referência da Figura 13(a).

Tabela 2 — Algoritmo para mapeamento inverso de perspectiva para um *grid* bidimensional de dimensões (g_x, g_y) e resolução r_m .

$I_M(x, y)$: intensidade do pixel (x, y) na imagem projetada (*grid*).

$I_R(x, y)$: intensidade do pixel (x, y) na imagem de referência.

(g_x, g_y) : dimensões do *grid* bidimensional nas dimensões x e y

h: altura da câmera com relação ao plano da região trafegável

θ : ângulo formado entre o eixo ótico da câmera e o plano da região trafegável.

$(u_0, v_0), \alpha$: parâmetros intrínsecos da câmera de referência

$$Z_w = -h$$

$$\text{Para } X_w = 0, X_w < g_x * r_m, X_w = X_w + r_m$$

$$y_c = v_0 + \frac{-\alpha * (X_w * \sin \theta + Z_w * \cos \theta)}{X_w * \cos \theta - Z_w * \sin \theta}$$

$$\text{Para } Y_w = \frac{-g_y * r_m}{2}, Y_w \leq \frac{g_y * r_m}{2}, Y_w = Y_w + r_m$$

$$x_c = u_0 + \frac{\alpha * Y_w}{X_w * \cos \theta - Z_w * \sin \theta}$$

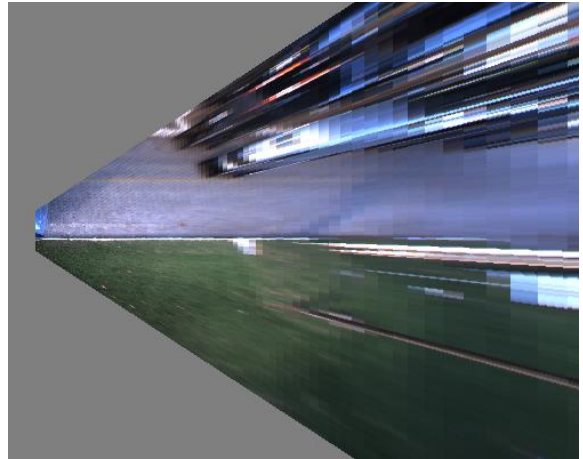
$$x_m = \frac{X_w}{r_m}$$

$$y_m = \frac{Y_w}{r_m} + \frac{g_y}{2}$$

$$I_M(x_m, y_m) = I_R(x_c, y_c)$$



(a)



(b)

Figura 13 — Mapeamento inverso de perspectiva. (a) Imagem de referência (capturada pela câmera direita). (b) Resultado do mapeamento inverso de perspectiva (*Bird's Eye View*) realizado sobre a imagem de referência (a).

4 CLASSIFICAÇÃO, MAPEAMENTO E LOCALIZAÇÃO DOS LIMITES DA REGIÃO TRAFEGÁVEL

Neste capítulo é apresentada uma abordagem de aprendizado de máquina para classificação das regiões das imagens em trafegável e não trafegável, um método de integração bayesiana para geração de um mapa de ocupação probabilístico e um algoritmo para localização dos limites laterais da região trafegável.

Para permitir navegação autônoma de um veículo, a presença de sensores que permitam a percepção do ambiente é essencial. No entanto, a escolha de quais sensores serão utilizados é uma decisão técnica importante, que deve considerar as condições específicas nas quais o veículo autônomo será utilizado. No desenvolvimento do veículo Stanley, vencedor do desafio *DARPA Grand Challenge 2005* (BUEHLER,2005) , Thrun et al. (THRUN,2006) optaram por utilizar um conjunto de LIDARs (*Light Detection And Ranging*) para detecção de obstáculos nas proximidades dos veículos. No entanto, devido ao alcance limitado dos LIDARs, esse conjunto permitia que o veículo autônomo detectasse obstáculos a no máximo 22 m, o que possibilitava que Stanley trafegasse com segurança a velocidades de até 25 milhas/h, ou aproximadamente 40 km/h (THRUN,2006). Para contornar essa limitação de alcance dos sensores LIDARs, Thrun et al. optaram por adicionar ao sistema uma câmera digital monocular que captura imagens em cores, e utilizaram essa câmera para construir um mapa baseado em visão computacional capaz de identificar corretamente regiões trafegáveis a mais de 70 m de distância.

Neste capítulo, será detalhado um método para classificação de regiões trafegáveis baseado em aprendizado de máquina, que é uma extensão do método proposto originalmente por Crisman e Thorpe (CRISMAN,1993) e adaptado por Thrun et al. (THRUN,2006) para o desenvolvimento do veículo vencedor do desafio *DARPA Grand Challenge 2005*. Na sequência, é apresentado um algoritmo para construção de um mapa de ocupação probabilístico da região trafegável, baseado em métodos probabilísticos que integram observações instantâneas ao longo do tempo. Por fim, é apresentado um método capaz de detectar os limites da região trafegável no mapa de ocupação probabilístico.

4.1 Classificação das Regiões Trafegáveis e Não Trafegáveis

Com intuito de mapear corretamente a região trafegável, um sistema robusto deve ser capaz de endereçar ao menos as seguintes situações e necessidades:

- Ausência de linhas de sinalização nas vias;
- Contornos das vias degradados;
- Vias acidentadas, com buracos e desníveis;
- Forte presença de sombras;
- Adaptação em tempo real para diferentes situações de iluminação e tipos de superfície (asfalto, estrada de paralelepípedo, estrada de terra, etc.).

Na tentativa de enfrentar os vários desafios que existem para a detecção de regiões trafegáveis em imagens, Crisman e Thorpe (CRISMAN,1993) desenvolveram um algoritmo de classificação baseado em aprendizado de máquina que filtra as imagens de entrada e produz imagens de saída nas quais é possível identificar as regiões trafegáveis e as regiões não trafegáveis. O algoritmo de aprendizado de máquina proposto por Crisman e Thorpe opera em duas fases: treinamento (realizado previamente, antes da navegação autônoma) e classificação. Durante a fase de treinamento, o algoritmo modela as cores da região trafegável como uma coleção de mistura de gaussianas e modela também as cores das regiões não trafegáveis como outra coleção de mistura de gaussianas. Na fase de classificação, o algoritmo calcula a probabilidade de cada pixel pertencer ou não à região trafegável, baseado em quão bem a coloração do pixel corresponde aos modelos de cores.

O algoritmo proposto por Crisman e Thorpe (CRISMAN,1993) foi adaptado por Thrun et al. (Thrun,2006) para utilização no veículo autônomo Stanley, vencedor do desafio *DARPA Grand Challenge 2005*. Thrun et al. utilizaram um conjunto de LIDARs, um Sistema de Posicionamento Global (*Global Positioning System – GPS*) e uma unidade de medida inercial (*Inertial Measurement Unit - IMU*) localizadas acima do veículo Stanley para determinar o estado - localização e orientação - do veículo, mapear as regiões próximas e imediatamente a frente do veículo e determinar quais segmentos destas regiões pertencem à região trafegável – uma janela de segurança.

Para estender o alcance limitado do mapeamento realizado pelo conjunto de LIDARs, Thrun et al. empregaram uma câmera monocular que permite capturar regiões mais afastadas em frente ao carro. Baseando-se no estado estimado a partir de outros sensores (GPS, IMU e odômetro) e utilizando projeções geométricas simples, a janela de segurança é localizada na imagem da câmera. A Figura 14(b) ilustra a localização da janela de segurança na Figura 14(a) obtida a partir da câmera monocular.

Na fase de treinamento, a cada nova imagem capturada pela câmera monocular, os pixels pertencentes à janela de segurança são utilizados como novas amostras para construir uma coleção de gaussianas. Cada gaussiana multidimensional é representada no sistema de cores vermelho/verde/azul (*red/green/blue* - RGB), e identificada por uma média μ_i (vetor de tamanho três), por uma covariância Σ_i (matriz quadrada de dimensão três) e por um valor inteiro m_i que denota o total de pixels utilizados como amostra na construção da gaussiana. Dessa maneira, a fase de treinamento do algoritmo ocorre a todo o momento durante a navegação autônoma, a cada nova imagem capturada pela câmera monocular. As novas gaussianas formadas a partir da janela de segurança são incorporadas à coleção, substituindo gaussianas já existentes (a gaussiana descartada é sempre a menos representativa, ou seja, a gaussiana com menor número de amostras m_i) ou mesclando as amostras das novas gaussianas com gaussianas semelhantes já existentes na coleção. A métrica utilizada para determinar semelhança de gaussianas é a Distância de *Mahalanobis*, definida pela Equação 14 como a distância escalar $d_{i,j}$ entre uma gaussiana de média μ_i e matriz de covariância Σ_i e outra gaussiana de média μ_j e matriz de covariância Σ_j (THRUN,2006).

$$d_{i,j} = (\mu_i - \mu_j)^T (\Sigma_i + \Sigma_j)^{-1} (\mu_i - \mu_j) \quad \text{Equação 14}$$

Na fase de classificação, a imagem é classificada usando as gaussianas aprendidas da seguinte forma: os pixels da imagem cujo valor é próximo a uma ou mais das gaussianas aprendidas são classificados como trafegáveis, e todos os outros pixels são classificados como não trafegáveis (THRUN,2006). A Figura 14(c) apresenta o resultado obtido por Thrun et al. (THRUN,2006) ao classificar a Figura 14(a) utilizando a coleção de mistura de gaussianas construída na fase de treinamento.

A utilização dessa abordagem de treinamento em tempo real e classificação baseada em uma coleção de mistura de gaussianas possibilitou que Stanley expandisse o mapa construído a

partir do conjunto de LIDARs, com alcance de aproximadamente 22 metros, para um mapa baseado em visão computacional com alcance muitas vezes superior a 70 metros, permitindo que o veículo autônomo trafegasse com segurança a velocidades mais altas (THRUN,2006).

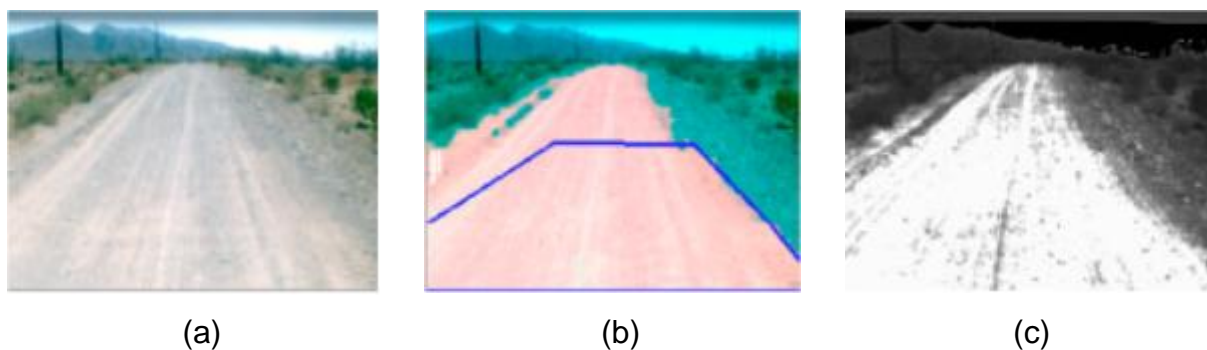


Figura 14 — Resultados do sistema de mapeamento baseado em visão computacional proposto por Thrun et al.(THRUN,2006). (a) Uma imagem da câmera monocular. (b) Imagem (a) com a janela de segurança identificada pelo quadrilátero com contorno azul. (c) Resultado da classificação da imagem (a) baseada em mistura de gaussianas.

Fonte: Adaptado de (THRUN,2006).

O método proposto neste trabalho para reconhecimento da região trafegável é similar ao método empregado por Thrun et al. (THRUN,2006), mas possui uma importante adaptação: o sistema de mapeamento se baseia principalmente em uma câmera estéreo frontal, em contraste com a abordagem de Thrun et al. que se baseava principalmente em um conjunto de LIDARs para a determinação de janelas de segurança.

4.1.1 Determinação das regiões seguras

O problema de classificar e reconhecer a região trafegável em imagens é simplificado se for conhecida previamente uma janela de segurança - uma pequena região que certamente é trafegável. Isso ocorre porque as propriedades do modelo de cores presentes na janela de segurança podem ser analisadas e utilizadas para classificar regiões remotas. Enquanto Thrun et al. (THRUN,2006) emprega vários sensores para permitir detectar uma janela de segurança, Katramados (KATRAMADOS,2009) define estaticamente a janela de segurança como uma pequena região retangular a frente do veículo autônomo.

O método proposto neste trabalho para detecção da janela de segurança utiliza inicialmente a abordagem da pequena região fixa à frente do veículo (um exemplo da localização da janela

de segurança na imagem de referência é ilustrado na Figura 15(a)). Mas uma vez que obstáculos poderiam estar localizados nesta região fixa, a análise de mapas em v -disparidades é utilizada para validação dos pixels. Como a análise do mapa em v -disparidades fornece informação sobre quais pixels pertencem ao plano da região trafegável, um pixel é aceito para formar a janela de segurança se ele simultaneamente: (i) está localizado na pequena região fixa em frente ao veículo, e (ii) pertence ao plano da região trafegável. Neste processo, pixels que não pertencem ao plano da região trafegável são imediatamente descartados, e os demais pixels localizados na pequena região fixa formam a janela de segurança que será utilizada para evoluir um classificador baseado em coleção de mistura de gaussianas.

A janela de segurança na imagem de referência é projetada para a imagem obtida a partir do mapeamento inverso de perspectiva da imagem de referência. A Figura 15(b) ilustra a projeção da janela de segurança para uma imagem obtida a partir do mapeamento inverso de perspectiva da imagem da Figura 15(a). Os pixels localizados na janela de segurança da imagem projetada são utilizados como amostras para a criação de uma nova gaussiana, que será incorporada à coleção de mistura de gaussianas.

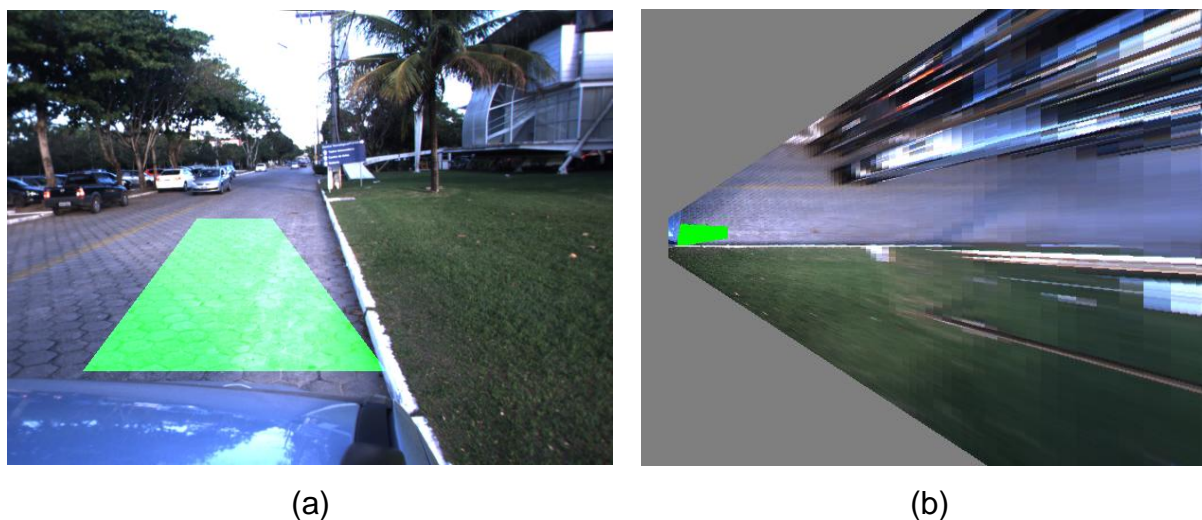


Figura 15 — Localização das janelas de segurança. (a) Janela de segurança na imagem de referência, destacada em verde. (b) Janela de segurança validada projetada para a imagem obtida a partir do mapeamento inverso de perspectiva da imagem (a).

4.1.2 Classificação das regiões trafegáveis

Para cada novo par de imagens recebido da câmera estéreo, são realizadas as fases de treinamento e de classificação. Na fase de treinamento, a coleção de mistura de gaussianas que re-

presenta as propriedades da região trafegável é evoluída através da incorporação de novas amostras extraídas da janela de segurança, de maneira semelhante ao trabalho de Thrun et al. (THRUN,2006). A coleção de mistura de gaussianas é denotada pelo conjunto das gaussianas multivariadas $G_i = (\mu_i, \Sigma_i)$ com média μ_i (vetor de tamanho três) e covariância Σ_i (matriz quadrada de dimensão três). Na fase de classificação, os pixels da imagem são classificados com base na coleção de mistura de gaussianas indicando o grau em que cada pixel, individualmente, se aproxima da coleção.

A classificação é realizada pixel a pixel na imagem obtida a partir do mapeamento inverso de perspectiva da imagem de referência. Para cada pixel $x = (P_R, P_G, P_B)^T$, com componentes no sistema de cores RGB denotados por P_R , P_G e P_B , representando respectivamente os componentes Vermelho, Verde e Azul do pixel, é encontrada a gaussiana que melhor se relaciona com os componentes do pixel, isto é, a gaussiana da coleção que maximiza o valor da função de densidade de probabilidade aplicada à variável $x = (P_R, P_G, P_B)^T$. A função de densidade de probabilidade para uma gaussiana multivariada em três dimensões com média μ e matriz de covariância Σ é dada pela Equação 15.

$$f(x, \mu, \Sigma) = \frac{1}{(2 * \pi)^{\frac{3}{2}} * |\Sigma|^{\frac{1}{2}}} * e^{-\frac{(x-\mu)^T \Sigma^{-1} (x-\mu)}{2}} \quad \text{Equação 15}$$

Na Equação 15, Σ^{-1} e $|\Sigma|$ denotam, respectivamente, a matriz inversa e o valor absoluto do determinante da matriz de covariância.

Denotando por $G_* = (\mu_*, \Sigma_*)$ a gaussiana que maximiza o resultado da função de densidade de probabilidade definida pela Equação 15 aplicada ao pixel $x = (P_R, P_G, P_B)^T$, o resultado da classificação do pixel x é obtido ao normalizar o valor da função com base no valor de pico da gaussiana G_* (valor da função de densidade de probabilidade aplicada à média da gaussiana). O fator de normalização é definido pela Equação 16 e o resultado da classificação de x pela gaussiana $G_* = (\mu_*, \Sigma_*)$ é definido pela Equação 17.

$$\eta_* = \frac{1}{f(\mu_*, \mu_*, \Sigma_*)} \quad \text{Equação 16}$$

$$v = \eta_* * f(x, \mu_*, \Sigma_*) \quad \text{Equação 17}$$

O fator de normalização $\eta_* \geq 1$ acentua diferenças entre resultados de classificação de pixels semelhantes mas mantém os valores dos resultados no intervalo $[0,1]$. Após a normalização definida pela Equação 16 e Equação 17, pixels x com intensidade igual à média μ_* terão re-

sultado da classificação igual a um, enquanto pixels $x \neq \mu_*$ terão resultado da classificação menor que um. O resultado da classificação de cada pixel é interpretado como a probabilidade do pixel pertencer a uma região trafegável. Desta forma, o resultado da classificação é uma matriz bidimensional M , de mesma dimensão da imagem, onde os elementos $M_{i,j}$ pertencem ao intervalo $[0,1]$ e representam a probabilidade de que o pixel (i, j) da imagem seja trafegável.

Para evitar trabalhar diretamente com valores de probabilidades no intervalo $[0,1]$, o resultado da classificação é transformado em uma representação gráfica - uma imagem em tons de cinza - na qual a intensidade $I_{i,j}$ do pixel (i, j) é diretamente proporcional ao valor da probabilidade $M_{i,j}$ do pixel, e é dada pela Equação 18.

$$I_{i,j} = M_{i,j} * 255 \quad \text{Equação 18}$$

A Figura 16(a) ilustra uma representação gráfica em tons de cinza do resultado da classificação baseada em mistura de gaussianas realizada na Figura 13(b).

Para permitir a construção de um mapa da região trafegável a partir do resultado da classificação, é desejável obter uma resposta do tipo Sim ou Não para a pergunta “O pixel (i, j) da imagem representa uma localização trafegável?”. Isso é feito através da binarização do resultado da classificação, utilizando um limiar de binarização. Caso o resultado da classificação seja maior ou igual ao limiar, o pixel é considerado trafegável; caso contrário, o pixel é considerado não trafegável. A Figura 16(b) ilustra o resultado da classificação após a binarização da Figura 16(a) utilizando o valor de intensidade de pixel 8 como limiar de binarização.

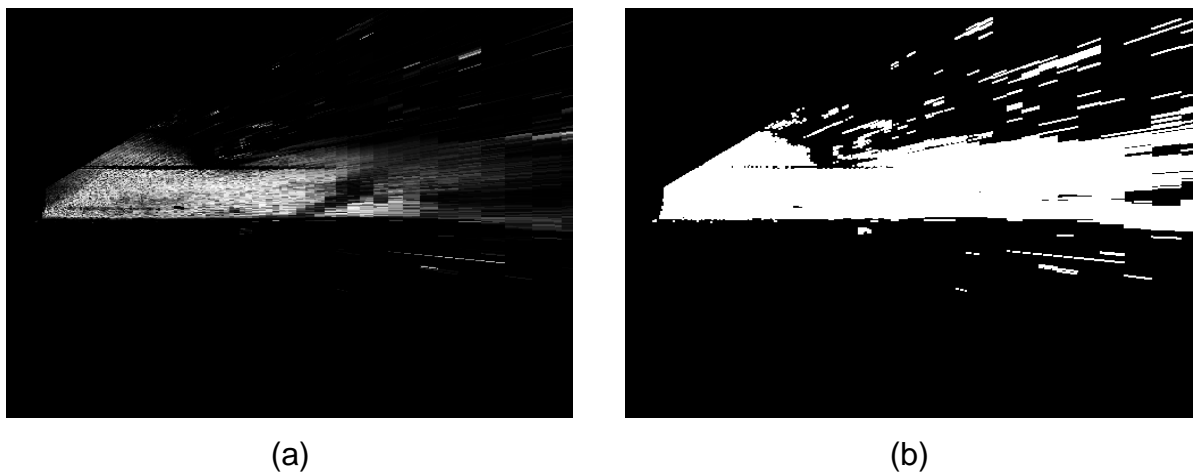


Figura 16 — Classificação das regiões trafegáveis. (a) Representação gráfica do resultado da classificação da Figura 13(b) baseado em mistura de gaussianas. (b) Mapa da região trafegável resultante da binarização da imagem (a).

4.2 Integração Bayesiana de Mapas Instantâneos da Região Trafegável

Mapas instantâneos, como o mapa da região trafegável ilustrado na Figura 16(b), embora possam ser empregados para navegação de um veículo autônomo, podem conter erros se a classificação da região trafegável não obtiver resultados de qualidade. No caso do método de classificação baseado em mistura de gaussianas, mapas instantâneos de baixa qualidade surgem tipicamente nos casos em que o modelo de cores da região trafegável não é bem representado pela coleção de mistura de gaussianas no instante em que é realizada a classificação da imagem.

Uma solução para permitir a construção de mapas de maior qualidade e menos sensíveis a erros de classificação instantânea é uma integração baseada no Teorema de Bayes (THRUN,2005) e no filtro de Bayes para estimar a ocupação de cada localização do mapa. Com esta integração bayesiana os mapas instantâneos ou observações instantâneas são integrados em um mapa de *grid* de ocupação probabilístico. O modelo gráfico da integração bayesiana para construção de um mapa de ocupação é apresentado na Figura 3. Para construção do mapa de ocupação probabilístico, é assumido que para cada localização ou célula de um mapa de *grid* existe uma probabilidade de que a localização seja trafegável. A informação de um mapa instantâneo é integrada no mapa de ocupação probabilístico, atualizando a probabilidade de ocupação das regiões que estão presentes no campo perceptivo da câmera de referência. A probabilidade de ocupação da célula é decrementada ou incrementada se a classificação da célula no mapa instantâneo indica que a célula representa uma região trafegável ou não trafegável, respectivamente. Esse processo permite a convergência da probabilidade de ocupação das localizações frequentemente classificadas de maneira idêntica (frequentemente trafegáveis ou frequentemente não trafegáveis) e, além disso, permite ao sistema de mapeamento se recuperar de erros de observação instantânea, ou seja, uma localização classificada equivocadamente em um mapa instantâneo pode convergir para a classificação correta se for classificada adequadamente em outras oportunidades.

Neste trabalho, o mapa de *grid* de ocupação probabilístico é construído utilizando o algoritmo *Occupancy Grid Mapping* (THRUN,2005) para resolver o problema de mapeamento com localizações conhecidas. As localizações fornecidas para o algoritmo *Occupancy Grid Map-*

ping são estimadas através da fusão de informações de diferentes fontes (GPS, IMU e odometria visual).

A Tabela 3 apresenta o algoritmo utilizado neste trabalho para integração bayesiana de mapas instantâneos z_t em um mapa probabilístico $m = \{m_i\}$. Este algoritmo é uma adaptação do algoritmo *Occupancy Grid Mapping* apresentado na Tabela 1. Um exemplo do mapa construído através da aplicação do algoritmo é apresentado na Figura 17, na qual a última observação incorporada foi o mapa instantâneo da Figura 16(b).

Tabela 3 — Algoritmo para integração bayesiana de mapas instantâneos em um mapa probabilístico.

<p>l_{occ}: fator de importância de uma observação do tipo célula ocupada. l_{free}: fator de importância de uma observação do tipo célula não ocupada. l_{min}: limiar de probabilidade mínima de ocupação na representação <i>log odds</i> l_{max}: limiar de probabilidade máxima de ocupação na representação <i>log odds</i></p> <p>Para todas as células m_i do mapa $m = \{m_i\}$ que estão no campo perceptivo do mapa instantâneo $z_t = \{z_{t,i}\}$</p> <p>Se $z_{t,i}$ é uma localização ocupada</p> $l_i = l_i + l_{occ}$ <p>Senão</p> $l_i = l_i + l_{free}$ <p>Se $l_i > l_{max}$</p> $l_i = l_{max}$ <p>Se $l_i < l_{min}$</p> $l_i = l_{min}$ $p(m_i) = 1 - \frac{1}{1+e^{l_i}}$
--

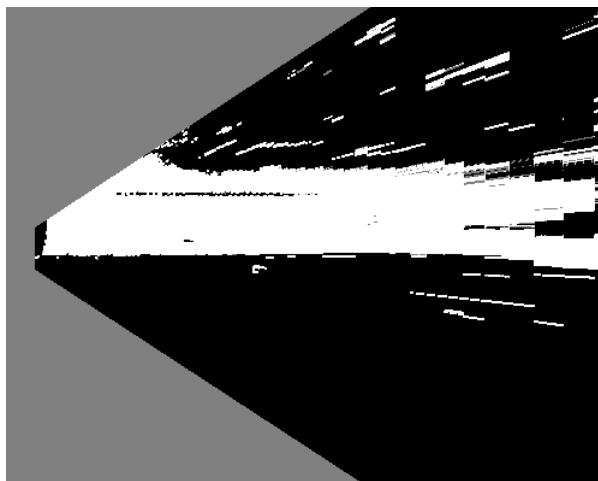


Figura 17 — Mapa probabilístico obtido a partir da integração bayesiana de mapas instantâneos aplicando-se o algoritmo descrito na Tabela 3.

4.3 Localização dos Limites Laterais da Região Trafegável

A localização dos limites laterais da região trafegável é uma informação importante que pode ser utilizada por algoritmos de planejamento de caminhos para respeitar as leis de trânsito, trafegando na faixa da direita em vias de mão dupla e estacionando corretamente em uma parada regulamentar e também para minimizar as chances de se encontrar obstáculos à frente do veículo. Thrun et al. (THRUN,2006) alegam que em cenários como a navegação no deserto a maioria dos obstáculos pode ser evitada utilizando a informação de localização dos limites laterais para planejar caminhos que conduzam o veículo pelo centro da região trafegável. Em cenários como navegação em vias de mão dupla, trafegar próximo aos limites laterais à direita da região trafegável minimiza as chances de colisão com outros veículos que trafegam no sentido oposto.

Neste trabalho, o mapa de *grid* de ocupação probabilístico obtido através da integração bayesiana dos mapas instantâneos é utilizado para localizar os limites laterais à direita da região trafegável. Primeiramente, a representação gráfica do mapa de ocupação probabilístico, uma imagem em tons de cinza, é binarizada utilizando o limiar de probabilidade $\delta = 0.5$. A binarização maximiza o valor das células do mapa com probabilidade de ocupação baixa - regiões trafegáveis - e minimiza o valor das células do mapa com probabilidade de ocupação alta - regiões não trafegáveis. A Figura 18(a) apresenta o resultado da binarização do mapa probabilístico da Figura 17.

Após a binarização, o mapa é subdividido em pequenos blocos de 4 m². Essa divisão é ilustrada na Figura 18(b). Horizontalmente, o mapa é dividido em faixas de quatro metros de profundidade, e cada faixa horizontal é novamente dividida, agora verticalmente, em blocos de um metro de largura. Na sequência, a intensidade I_k de cada bloco k é calculada pela Equação 19 como sendo o somatório da intensidade $I_{i,j}$ dos pixels (i, j) que pertencem ao bloco k .

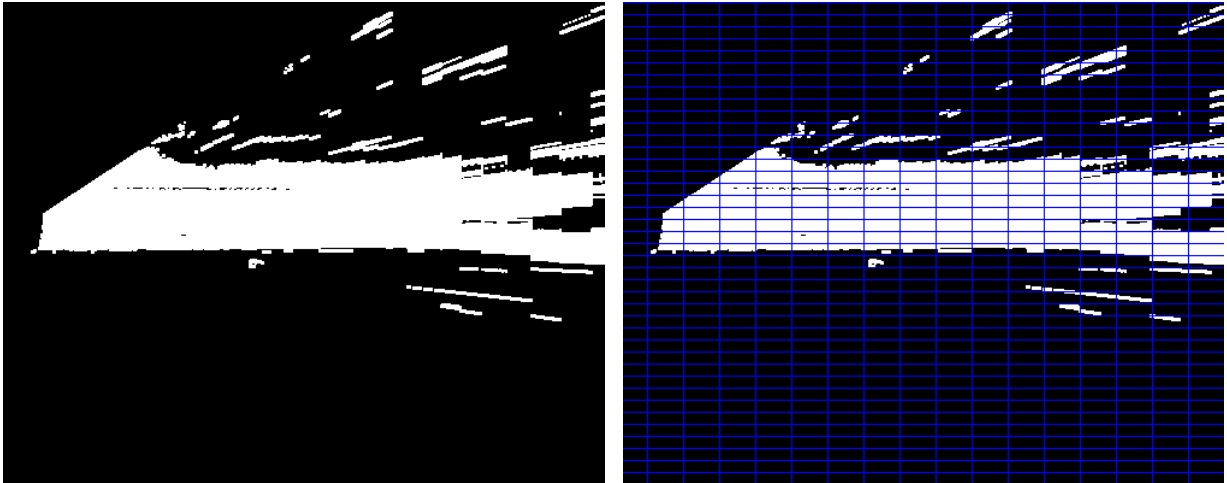
$$I_k = \sum_{i,j \in k} I_{i,j} \quad \text{Equação 19}$$

Definida pela Equação 19, a intensidade de um bloco é inversamente proporcional à probabilidade de ocupação do bloco.

Finalmente, a localização dos limites laterais à direita da região trafegável é realizada analisando as variações de intensidade entre blocos adjacentes que estão localizados na mesma faixa horizontal. Em uma situação típica, é esperado que o limite da região trafegável esteja localizado entre um bloco de baixa intensidade (região não trafegável) que é adjacente a vários blocos de alta intensidade (regiões trafegáveis). A classificação entre blocos de baixa intensidade e blocos de alta intensidade pode ser realizada através da definição de um limiar de intensidade Δ . Para cálculo do limiar de intensidade, é feita uma análise estatística das intensidades de blocos localizados em uma mesma faixa horizontal, determinando-se a média μ e o desvio padrão δ das intensidades dos blocos. O limiar de intensidade de um bloco é definido pela Equação 20.

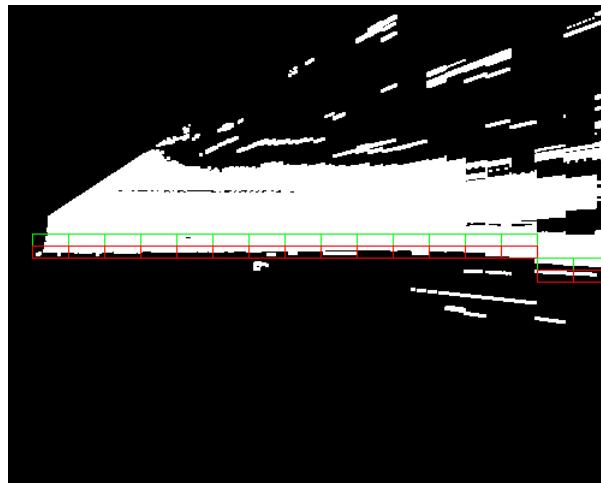
$$\Delta = \mu - \alpha * \delta \quad \text{Equação 20}$$

Para maior robustez na detecção dos limites laterais, o parâmetro α da Equação 20 é variado discretamente em 0.1 no intervalo de 1.0 a 2.0. De posse do limiar de intensidade Δ são considerados trafegáveis os blocos com intensidade $I_k \geq \Delta$, enquanto os blocos com intensidade $I_k < \Delta$ são considerados não trafegáveis. A Figura 18(c) apresenta o resultado da localização dos limites laterais da região trafegável no mapa probabilístico da Figura 17.



(a)

(b)



(c)

Figura 18 – Localização dos limites laterais da região trafegável. (a) Mapa de ocupação probabilístico da região trafegável após binarização. (b) Divisão do mapa probabilístico em blocos (cor azul). (c) Localização dos limites da região trafegável, representados como uma transição entre um bloco não trafegável (vermelho) e uma sequência de blocos trafegáveis (primeiro bloco da sequência representado na cor verde).

5 METODOLOGIA

Neste capítulo é apresentada a metodologia utilizada nos experimentos realizados para avaliar o desempenho do sistema de mapeamento, especificando as bases de dados, hardware, software e métricas utilizadas nesta avaliação.

5.1 Base de Dados

Para avaliação do sistema de mapeamento da região trafegável baseado em visão computacional, foi construída uma base de dados a partir de um cenário real ao dirigir um carro de passeio equipado com sensores pelo anel viário do campus de Goiabeiras da Universidade Federal do Espírito Santo (UFES). O campus possui um anel viário com 3.570 metros (LYRIO,2012). A Figura 19 mostra a visão aérea do campus de Goiabeiras da UFES com o anel viário destacado em vermelho. As informações obtidas dos sensores foram capturadas e gravadas, ordenadas cronologicamente, em um arquivo de *log*. O arquivo de *log* está disponível no endereço http://www.lcad.inf.ufes.br/log/log_voltadaufes-20120711-1.txt, e contém dados obtidos de uma câmera estéreo frontal, de um sistema de posicionamento global, de uma unidade de medida inercial e de outros sensores não relevantes para este trabalho. As especificações do *hardware* dos sensores e do *software* utilizado para gravação e posterior execução do arquivo de *log* são apresentadas nas seções 5.2 e 5.3, respectivamente.



Figura 19 – Visão aérea do campus de Goiabeiras da UFES com o anel viário destacado em vermelho.

Fonte: Adaptado de (LYRIO,2012).

5.1.1 Imagens estáticas

O conjunto de teste é constituído de um pequeno conjunto de imagens estáticas extraídas a partir do arquivo de *log* http://www.lcad.inf.ufes.br/log/log_voltadaufes-20120711-1.txt, que representam a imagem de referência capturada pela câmera direita do par estéreo. Estas imagens foram editadas manualmente para identificar a região trafegável. Todas as imagens possuem o mesmo tamanho (640x480 *pixels*) e estão codificadas em três canais no sistema de cores RGB. As imagens estáticas selecionadas e as respectivas imagens editadas manualmente são descritas a seguir.

- **Imagem estática 1:** representa uma situação de mapeamento em pavimento de paralelepípedos, na qual o veículo ainda não se encontra em movimento e, portanto, os erros acumulados de localização são mínimos. A principal dificuldade desta situação é a presença de um veículo próximo trafegando em outra faixa da via, no sentido oposto. A Figura 20(a) apresenta a **imagem estática 1** e a Figura 20(b) apresenta a respectiva imagem editada manualmente para identificação da região trafegável.
- **Imagem estática 2:** representa uma situação de mapeamento em pavimento de paralelepípedos, na qual o veículo encontra-se em movimento e, portanto, temos a presença de erros de localização acumulados. Como dificuldades adicionais desta situação, destaca-se ainda a presença de um veículo distante trafegando em outra faixa da via no sentido oposto e de um ciclista próximo trafegando na mesma faixa da via e no mesmo sentido do veículo. A Figura 21(a) apresenta a **imagem estática 2** e a Figura 21(b) apresenta a respectiva imagem editada manualmente para identificação da região trafegável.
- **Imagem estática 3:** representa uma situação de mapeamento em pavimento de paralelepípedos, na qual o veículo encontra-se em movimento e, portanto, temos a presença de erros de localização acumulados. Como dificuldades adicionais desta situação, destaca-se ainda a presença de um veículo distante trafegando em outra faixa da via no sentido oposto e a presença de fortes sombras no campo de visão da imagem. A Figura 22(a) apresenta a **imagem estática 3** e a Figura 22(b) apresenta a respectiva imagem editada manualmente para identificação da região trafegável.
- **Imagem estática 4:** representa uma situação de mapeamento em pavimento asfaltado, na qual o veículo encontra-se em movimento e, portanto, temos a presença de erros de localização acumulados. Como dificuldades adicionais desta situação, destaca-se ainda

a recente transição de pavimentos (transição do pavimento de paralelepípedo para o pavimento asfaltado) e a presença de outros veículos próximos trafegando na mesma faixa e no mesmo sentido. A Figura 23(a) apresenta a **imagem estática 4** e a Figura 23(b) apresenta a respectiva imagem editada manualmente para identificação da região trafegável.

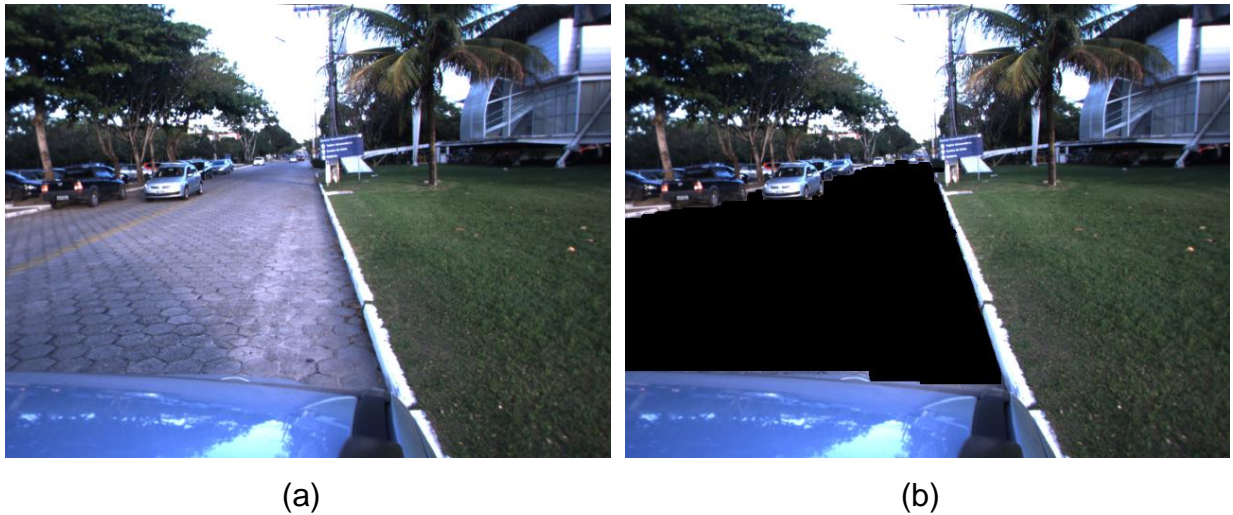


Figura 20 – Imagem estática 1. (a) Imagem de referência (capturada pela câmera direta do par estéreo). (b) Imagem editada manualmente, identificando em preto a região trafegável.



Figura 21 – Imagem estática 2. (a) Imagem de referência (capturada pela câmera direta do par estéreo). (b) Imagem editada manualmente, identificando em preto a região trafegável.



(a)

(b)

Figura 22 – Imagem estática 3. (a) Imagem de referência (capturada pela câmera direta do par estéreo). (b) Imagem editada manualmente, identificando em preto a região trafegável.



(a)

(b)

Figura 23 – Imagem estática 4. (a) Imagem de referência (capturada pela câmera direta do par estéreo). (b) Imagem editada manualmente, identificando em preto a região trafegável.

O conjunto de treino do sistema de aprendizado de máquina baseado em uma coleção de mistura de gaussianas é também extraído a partir do arquivo de *log*. Para representar uma situação real, cada uma das imagens estáticas do conjunto de teste possui o seu próprio conjunto de treino, que é definido como o conjunto de informações do *log* até o instante em que foi capturada a imagem estática. Dessa forma, na fase de treinamento a metodologia utilizada para avaliação emprega o mesmo comportamento que o sistema de mapeamento executado em uma situação real.

5.2 Hardware

No desenvolvimento deste trabalho, foram usados dois ambientes de trabalho, um para a construção de um arquivo de *log* e outro para a execução do arquivo de *log* e avaliação do sistema de mapeamento.

Para construção do arquivo de *log*, foi utilizado um veículo de passeio comum dirigido por uma pessoa, equipado com uma câmera estéreo frontal, um sistema de posicionamento global, uma unidade de medida inercial e outros sensores de alcance não relevantes, cujas informações não foram utilizadas. A Figura 24 apresenta o veículo experimental utilizado e a configuração dos sensores. Foi utilizado também um computador de alto desempenho equipado com um disco SSD e tecnologia RAID 0, que permite fluxo de dados de entrada e saída de até 500 MB/s. Durante as sessões de construção dos arquivos de *log*, o sistema (computador e sensores) foi alimentado por um único *nobreak*.



Figura 24 – Vista lateral do veículo experimental. Localizado acima do veículo, uma câmera estéreo, um sensor de medida inercial (IMU) e um *Laser Range Finder* (não utilizado pelo sistema de mapeamento).

A câmera estéreo utilizada nos experimentos deste trabalho foi a câmera estéreo *multi-baseline Point Grey Research Inc. Bumblebee XB3* (ilustrada na Figura 25) modelo BBX3-13S2C-38, configurada com a *baseline* máximo de 0,24 m para capturar imagens de 640x480

pixels a 16 *Frames per Second* - FPS, com 66° de HFOV. A câmera é conectada ao computador através de uma interface *firewire* IEEE-1394b, que permite fluxo de dados de 800 Mb/s. A documentação técnica da câmera *Bumblebee XB3* é disponibilizada *online* pela *Point Grey Research Inc.*, no endereço http://www.ptgrey.com/products/bbxb3/bumblebeeXB3_stereo_camera.asp.



Figura 25 – Câmera estéreo Bumblebee XB3.

Fonte: http://www.ptgrey.com/products/bbxb3/bumblebeeXB3_stereo_camera.asp.

A unidade de medida inercial utilizada nos experimentos deste trabalho foi o *Xsens North America Inc. Inertial Measurement Unit MTi-G* (ilustrado na Figura 26), que provê informações a uma frequência de até 120 Hz. O sensor foi conectado ao computador através de uma interface USB. A documentação técnica do *Xsens MTi-G* é disponibilizada *online* pela *Xsens North America Inc.*, no endereço <http://www.xsens.com/en/general/mti-g>.



Figura 26 – Unidade de medida inercial Xsens MTi-G.

Fonte: <http://www.xsens.com/en/general/mti>.

O sistema de posicionamento global utilizado nos experimentos deste trabalho foi o *Venus GPS 09133* (ilustrado na Figura 27), que provê informações a uma frequência de até 10 Hz. O

sensor foi conectado ao computador através de uma interface USB, utilizando um adaptador para o conector SMA. A documentação técnica é disponibilizada *online* pela *Sparkfun Electronics* no endereço <https://www.sparkfun.com/products/9133>.



Figura 27 – Sistema de Posicionamento Global Venus GPS 09133.

Fonte: <https://www.sparkfun.com/products/9133>.

Para execução em tempo real do sistema de mapeamento foi utilizado um computador desktop Dell *Alienware* com processador Intel Core i7 930 2.80 GHz e 12 *Gigabytes* de memória RAM, equipado com placa gráfica GTX-480 *CUDA-Enabled* Nvidia com 480 CUDA cores. Embora recomendado para execução em tempo real, não existe a necessidade de um computador habilitado na tecnologia *Computer Unified Device Architecture* - CUDA para execução do sistema de mapeamento. A construção do sistema foi realizada utilizando um notebook Dell XPS, com processador Intel Core i5 M480 2.66 GHz e 4 *Gigabytes* de memória RAM. Durante a construção dos algoritmos, o arquivo de *log* era reproduzido em velocidade reduzida, diminuindo a taxa de recebimento de informações da câmera estéreo de 16 FPS (*frame rate* máximo da *Bumblebee XB3*) para aproximadamente 2 FPS.

5.3 Software

Neste trabalho foram utilizados somente softwares de livre distribuição (*freeware*). Para desenvolvimento do sistema de mapeamento, foi utilizado um ambiente de código aberto para construção de algoritmos de robótica, o *Carnegie Mellon Robot Navigation Toolkit* - *CARMEN* (MONTEMERLO, 2003), que possui documentação disponível *online* no endereço

<http://carmen.sourceforge.net>. O ambiente CARMEN foi utilizado em conjunto com a distribuição Linux Fedora Core e com outros artefatos de software de livre distribuição.

5.3.1 Ambiente CARMEN

A estratégia adotada neste trabalho foi reutilizar as facilidades que já estavam presentes no ambiente de CARMEN, adicionando novos módulos necessários para controle de sensores específicos e processamento de informações visuais, além de evoluir alguns módulos já existentes. Dentre as facilidades presentes em CARMEN, já estavam disponíveis um módulo *logger* para gravação de um arquivo de *log* com dados oriundos de sensores e um módulo *player* para reprodução de um arquivo de *log*.

Ao utilizar a arquitetura de software modular de CARMEN, o processamento das informações é distribuído entre módulos executando assincronamente, que compartilham apenas o Inter-process Communication (IPC) e bibliotecas de funções utilitárias. A Figura 28 apresenta um diagrama de blocos que ilustra a visão geral do sistema de navegação autônoma em desenvolvimento no LCAD, com destaque para o subsistema de mapeamento desenvolvido neste trabalho, mostrando as direções dos fluxos de dados e as dependências entre sensores, controladores de sensores (*drivers*) e módulos de processamento de dados. Neste diagrama, o bloco “Road Mapper” representa o módulo CARMEN desenvolvido neste trabalho.

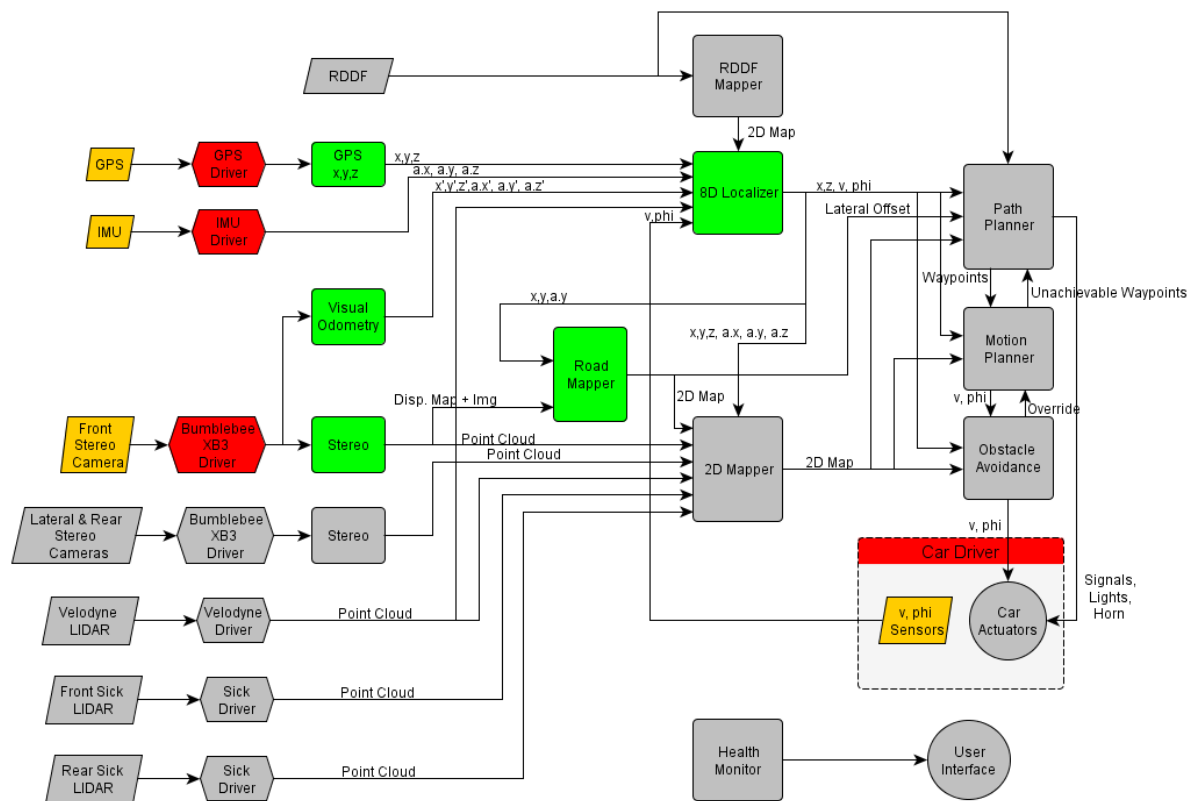


Figura 28 – Visão geral do sistema de navegação autônoma, com destaque para o subsistema de mapeamento. As dependências entre módulos e as direções dos fluxos de dados são representados por setas unidirecionais, enquanto sensores são representados por blocos amarelos, controladores de sensores (drivers) são representados por blocos vermelhos e módulos de processamento de dados são representados por blocos verdes. Blocos cinza não são relevantes para o subsistema de mapeamento.

5.3.2 Módulos CARMEN

Nesta seção é descrita de maneira sucinta a funcionalidade dos módulos CARMEN que foram desenvolvidos pelo LCAD e foram utilizados como parte do sistema de mapeamento baseado em visão computacional proposto neste trabalho.

5.3.2.1 GPS x,y,z

Módulo responsável pela conversão dos dados de coordenadas geográficas do sistema *Universal Transversal de Mercator* – UTM obtidas do GPS para o sistema de coordenadas métrico adotado como padrão neste trabalho.

5.3.2.2 Visual Odometry

Módulo responsável por disponibilizar informações de odometria com seis graus de liberdade (translação nas dimensões X, Y e Z e rotação nas dimensões *roll*, *pitch* e *yaw*). As informa-

ções de odometria são calculadas a partir das imagens obtidas de uma câmera estéreo frontal. Neste módulo é utilizado o método proposto por Geiger (GEIGER,2011) que está disponível na biblioteca *Library for VISual Odometry 2* - LIBVISO2 (<http://www.cvlibs.net/software/libviso2.html>). A adaptação da biblioteca LIBVISO2 para construção de um módulo na arquitetura do ambiente CARMEN foi uma das contribuições do trabalho de Lyrio (LYRIO,2012).

5.3.2.3 Stereo

Módulo responsável por receber os pares de imagens estéreo da câmera estéreo frontal e por disponibilizar os respectivos mapas de disparidades. O módulo “Stereo” suporta, através de parâmetros de inicialização, a seleção de diferentes algoritmos para cálculo do mapa de disparidades. O algoritmo utilizado no desenvolvimento deste trabalho foi proposto por Geiger (GEIGER,2010) e está disponível na biblioteca *Library for Efficient Large-Scale Stereo Matching* – LIBELAS (<http://www.cvlibs.net/software/libelas.html>).

5.3.2.4 8D Localizer

Módulo responsável por realizar a integração bayesiana dos dados oriundos de diferentes sensores, tais como um sistema de posicionamento global, uma unidade de medida inercial e odometria visual sobre as imagens da câmera estéreo frontal. A integração é obtida utilizando um filtro de partículas para estimativa do estado do veículo composto por oito graus de liberdade, sendo três de posição (latitude, longitude e altitude), três de orientação (ângulos Eulerianos *roll*, *pitch* e *yaw*) e dois dos atuadores do veículo (velocidade linear e ângulo do volante).

5.3.3 Sistema operacional e artefatos de software

Nesta seção, são descritos de maneira sucinta o sistema operacional e os principais artefatos de *software* utilizados como ferramentas para viabilizar a construção do sistema de mapeamento proposto neste trabalho.

5.3.3.1 Sistema Operacional Fedora Core 15

No desenvolvimento deste trabalho foi utilizada a distribuição Linux Fedora Core, na versão 15. Uma distribuição mais antiga do Fedora Core, a versão 11, já era utilizada nos computa-

dores do LCAD antes do início do desenvolvimento deste trabalho, e como a instalação do ambiente CARMEN na versão 15 não apresentou problemas, a decisão de evoluir da versão 11 para a versão 15 do Fedora Core mostrou-se adequada. É importante ressaltar que enquanto a instalação padrão da distribuição foi suficiente para realizar a construção dos algoritmos (codificação), foi necessária a utilização de um *kernel* Linux compilado com um *patch* de “*real time*” para a captura de informações dos sensores durante as sessões de *log*, porque a frequência com que alguns sensores disponibilizavam os seus dados era superior à velocidade de processamento e gravação do módulo responsável pelo *log*. A solução técnica para a instalação de um *kernel* com a opção “*real time*” foi documentada no *wiki* do LCAD, no endereço http://www.lcad.inf.ufes.br/wiki/index.php/Criar_Kernel_RT.

O sistema operacional Fedora Core 15 está disponível em <http://fedoraproject.org/>.

5.3.3.2 OpenCV – Open Source Computer Vision Library

A OpenCV é uma biblioteca de processamento de imagens de distribuição gratuita (*freeware*) e código fonte aberto, distribuída sob a licença BSD. A ferramenta foi escolhida pela facilidade de uso e abrangência de algoritmos de visão computacional que estão disponíveis na biblioteca, além da alta qualidade e diversidade de documentação técnica e exemplos disponíveis na internet. Neste trabalho, foram utilizadas a implementação da Transformada de Hough para detecção de linhas no mapa em v-disparidades e do Filtro de Kalman Estendido para estimativa do estado da câmera estéreo.

A biblioteca OpenCV está disponível em <http://opencv.org/>.

5.3.3.3 GSL – GNU Scientific Library

A GSL é uma biblioteca numérica de distribuição gratuita (*freeware*) e código fonte aberto, distribuída sob a licença GNU GPL, para desenvolvimento nas linguagens C e C++. A biblioteca disponibiliza uma vasta gama de rotinas matemáticas, incluindo suporte à interface de programação *Basic Linear Algebra Subprograms* - BLAS. Neste trabalho, foram utilizadas a implementação de algumas operações matriciais e vetoriais, tais como multiplicação e inversão de matrizes, produto interno e cálculo de determinantes, para manter e evoluir a coleção de mistura de gaussianas e para classificar uma imagem baseada em uma coleção de mistura de gaussianas.

A biblioteca GSL está disponível em <http://www.gnu.org/software/gsl/>.

5.3.3.4 LIBELAS – Library for Efficient Large-Scale Stereo Matching

A LIBELAS (GEIGER, 2010) é uma biblioteca C++ independente de plataforma, de distribuição gratuita (*freeware*) e código fonte aberto, distribuída sob a licença GNU GPL, que permite computar mapas de disparidades a partir de pares de imagens estéreo retificadas. A biblioteca LIBELAS foi escolhida porque se mostrou adequada nos requisitos de qualidade e desempenho, que eram fundamentais para o algoritmo de visão estéreo utilizado na computação do mapa de disparidades para estimativa do estado da câmera estéreo. Pelo fato de ter o código fonte aberto, foi possível evoluir a implementação da biblioteca para uma versão paralelizada utilizando *OpenMP* que permite processamento em até 20 FPS, frequência superior aos 16 FPS fornecidos pela câmera estéreo *Bumblebee XB3* que foi utilizada nos experimentos deste trabalho.

A biblioteca LIBELAS está disponível em <http://www.cvlibs.net/software/libelas.html>.

5.4 Métricas

O principal objetivo de um sistema de mapeamento é a capacidade de identificar objetos e regiões no mapa. No caso específico de um mapa de ocupação da região trafegável, a capacidade de distinguir entre regiões trafegáveis e regiões não trafegáveis representa sua característica mais importante. Um sistema de mapeamento deve retornar uma resposta binária, identificando se uma localização é ou não é trafegável. Nesse cenário de classificação binária podem ocorrer quatro situações (FAWCETT, 2006):

a — Verdadeiro Positivo (*True Positive* — TP)

A localização é trafegável e o classificador (sistema de mapeamento) a classifica como trafegável.

b — Verdadeiro Negativo (*True Negative* — TN)

A localização é não trafegável e o classificador (sistema de mapeamento) a classifica como não trafegável.

c — Falso Positivo (*False Positive* — FP)

A localização é não trafegável, mas o classificador (sistema de mapeamento) a classifica como trafegável. Neste caso uma localização não trafegável é mapeada como sendo trafegável, representando um risco para o veículo autônomo que for utilizar o mapa para planejamento de suas ações.

d — Falso Negativo (*False Negative* — FN)

A localização é trafegável, mas o classificador (sistema de mapeamento) a classifica como não trafegável. Neste caso uma localização trafegável é mapeada como sendo não trafegável, representando uma limitação inexistente no mundo real para o veículo autônomo que for utilizar o mapa para planejamento de suas ações.

A Figura 29 apresenta a Tabela de Contingência que ilustra graficamente as situações possíveis para um classificador binário.

		Classe Real	
		Trafegável	Não Trafegável
Classe Sugerida pelo Classificador	Trafegável	Verdadeiro Positivo	Falso Positivo
	Não Trafegável	Falso Negativo	Verdadeiro Negativo

Figura 29 — Tabela de Contingência com as quatro situações possíveis para um classificador binário.

Fonte: Adaptado de (FAWCETT, 2006).

A partir da Tabela de Contingência da Figura 29 é possível computar duas métricas padrão usadas para medir o desempenho de classificadores binários:

- Taxa de Verdadeiros Positivos (*True Positive Rate* - TPR): É a taxa de localizações trafegáveis consideradas trafegáveis. É dada pela divisão do total dos verdadeiros positivos (TP) pelo número total de positivos (FN + TP), sendo definida pela Equação 21.

$$TPR = \frac{TP}{FN + TP}$$

Equação 21

• Taxa de Falsa Aceitação (*False Acceptance Rate* - FAR): É a taxa de localizações não trafegáveis consideradas trafegáveis. É dada pela divisão do total de falsos positivos (FP) pelo número total de negativos (FP + TN), sendo definida pela Equação 22.

$$FAR = \frac{FP}{FP + TN}$$

Equação 22

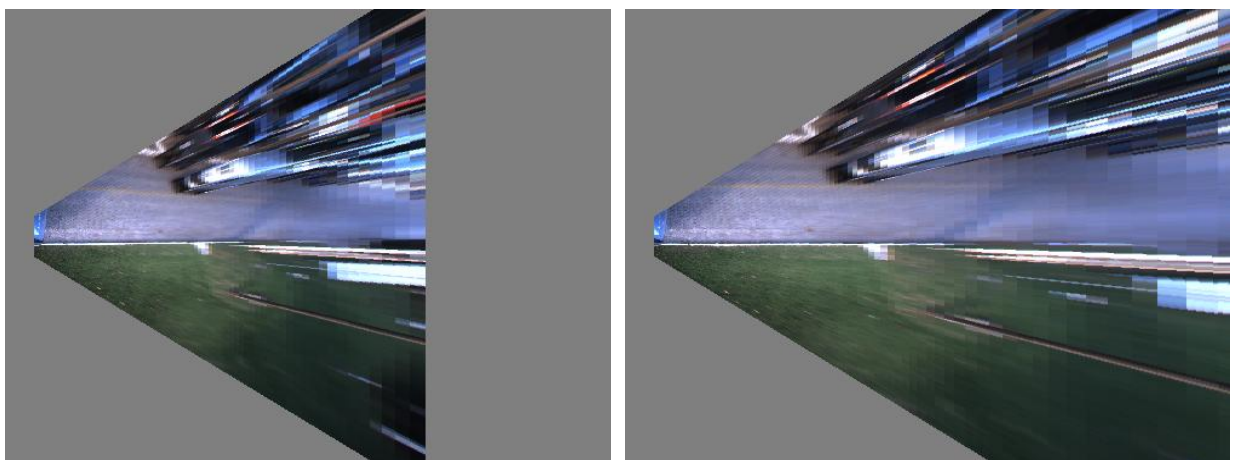
Um sistema de mapeamento ideal ou perfeito seria capaz de atribuir probabilidade de ocupação mínima a uma localização trafegável e probabilidade de ocupação máxima a uma localidade não trafegável. Em outras palavras, o classificador ideal é capaz de maximizar a TPR (Equação 21) e de minimizar a FAR (Equação 22). Essas métricas são fundamentais e serão empregadas para a avaliação do desempenho do sistema de mapeamento proposto neste trabalho.

6 EXPERIMENTOS E RESULTADOS

Neste capítulo são apresentados e discutidos os resultados dos experimentos realizados para avaliar o sistema de mapeamento. O vídeo disponível no endereço <http://www.youtube.com/watch?v=jfZ5IDAfOTI&feature=plcp> ilustra o funcionamento do sistema de mapeamento.

6.1 Resultados da Estimativa de Estado da Câmera e Mapeamento Inverso de Perspectiva

A projeção da imagem de referência para o sistema de coordenadas do mundo real obtida através do mapeamento inverso de perspectiva é avaliada através da comparação dos resultados obtidos ao utilizar o EKF para estimar o estado da câmera com os resultados obtidos sem a utilização do EKF. Os resultados obtidos sem o uso do EKF utilizam um estado da câmera estático calculado previamente. A Figura 30(a) apresenta um resultado ruim do mapeamento inverso de perspectiva obtido sem a utilização do EKF para estimar o estado da câmera e a Figura 30(b) apresenta um bom resultado para o mapeamento inverso de perspectiva obtido ao utilizar o EKF. Os resultados apresentados na Figura 30 ilustram a importância de se obter um estado da câmera confiável para realizar o mapeamento inverso de perspectiva.



(a)

(b)

Figura 30 – Resultados do mapeamento inverso de perspectiva. (a) Mapeamento obtido sem a utilização do EKF. (b) Mapeamento obtido com a utilização do EKF.

Com intuito de observar a natureza dinâmica do estado da câmera, que varia constantemente durante uma situação de navegação do veículo, foram catalogados os valores dos graus de liberdade do estado da câmera durante uma sequência de 900 frames obtidos a partir do arquivo de *log*. A Figura 31(a) e a Figura 31(b) apresentam os gráficos da variação dos graus de liberdade altura e ângulo de orientação do estado da câmera em função do número do frame. A Figura 32(a) e a Figura 32(b) apresentam os histogramas dos graus de liberdade altura h e ângulo de orientação θ do estado câmera. Para a construção do histograma dos valores da altura apresentado na Figura 32(a) foi considerada uma aproximação de 0.01 m e para a construção do histograma dos valores do ângulo de orientação apresentado na Figura 32(b) foi considerada uma aproximação de 0.1° .

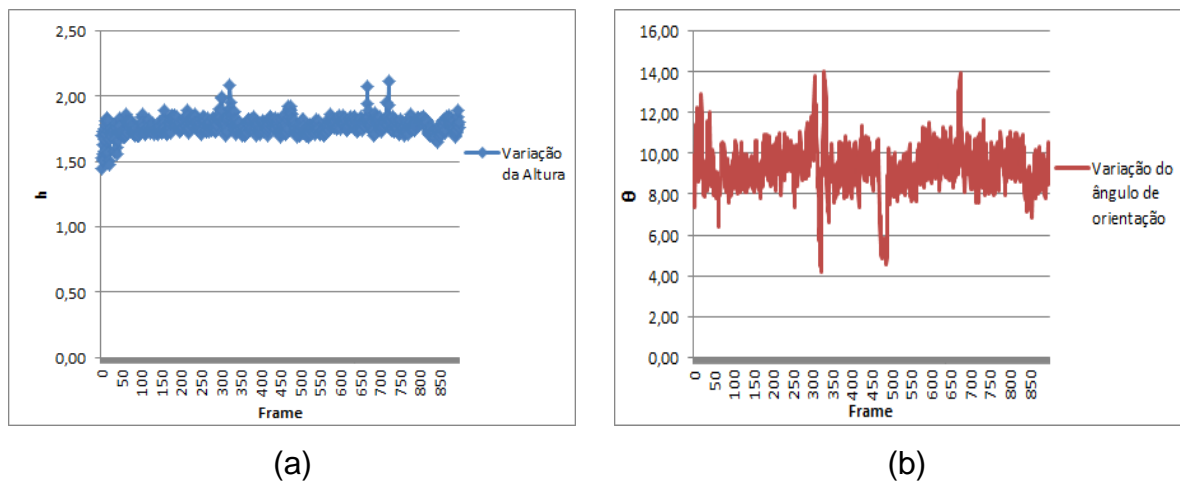


Figura 31 – Gráficos da variação dos graus de liberdade do estado da câmera. (a) Variação da altura h (em metros) em função do frame. (b) Variação do ângulo de orientação θ (em graus) em função do frame.

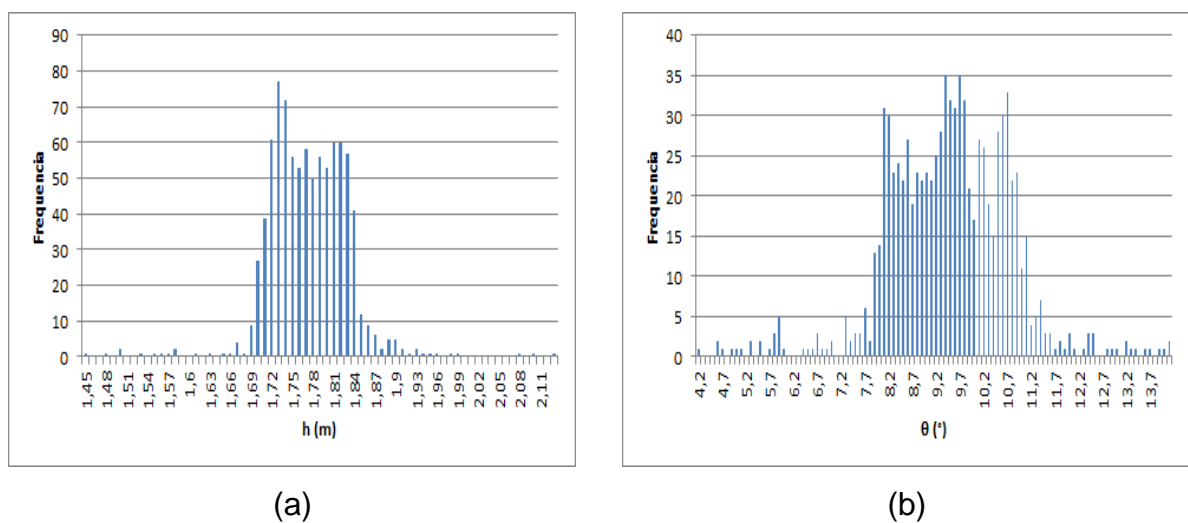


Figura 32 – Histogramas dos graus de liberdade do estado da câmera. (a) Histograma da altura h (em metros), com aproximação de 0.01 m. (b) Histograma do ângulo de orientação θ (em graus), com aproximação de 0.1° .

A partir da análise do gráfico apresentado na Figura 31(a) e do histograma apresentado na Figura 32(a) é possível observar que o grau de liberdade altura concentra seus valores estimados principalmente no intervalo de 1,69 m a 1,87 m, mas também apresenta valores estimados ao longo de todo o intervalo de 1,45 m a 2,11 m. A partir da análise do gráfico apresentado na Figura 31(b) e do histograma apresentado na Figura 32(b) é possível observar que o grau de liberdade ângulo de orientação concentra seus valores estimados principalmente no intervalo de 7,7° a 11,2°, mas também apresenta valores estimados ao longo de todo o intervalo de 4,2° a 13,7°. A alta dispersão dos valores estimados para os graus de liberdade permite evidenciar a natureza dinâmica do estado da câmera, reforçando a necessidade da utilização de mecanismos robustos para estimativa deste estado.

6.2 Resultados do Mapeamento para Imagens Estáticas

Os experimentos para avaliação do sistema de mapeamento foram realizados utilizando os conjuntos de treino e teste descritos na seção 5.1.1. Em todos os experimentos foram utilizadas coleções de misturas de gaussianas com 12 gaussianas e mapas de 50x40 m, representados como imagens de 500x400 pixels (resolução 0.1 m por pixel). Foi adotado um limiar de intensidade de pixel igual a 8 para binarização das imagens que representam o resultado da classificação, e um limiar de probabilidade igual a 0.5 como separador entre as localizações trafegáveis e as localizações não trafegáveis.

No conjunto de testes formado por imagens estáticas o desempenho do sistema de mapeamento da região trafegável foi avaliado em comparação com um classificador humano. Para isso, os mapas de *grid* de ocupação probabilísticos – resultados do sistema de mapeamento – são comparados com as projeções das imagens de referência que foram manualmente classificadas. É importante observar que são comparadas somente as localizações que estão no campo perceptual da imagem de referência. Esse detalhe é importante, pois no mapa de *grid* de ocupação probabilístico a região mapeada é mais abrangente – mapeia de fato toda a região de 50x40 m à frente do veículo – enquanto para o classificador humano a região mapeada é limitada pela HFOV de aproximadamente 66° da câmera que foi utilizada nos experimentos.

A Tabela 4 apresenta os resultados obtidos na comparação do sistema de mapeamento com o classificador humano. A Tabela 5 apresenta resultados da literatura obtidos por Thrun et al. (THRUN,2006) no mapeamento de regiões trafegáveis no deserto e nas montanhas. A Figura 33(a) e a Figura 33(b) apresentam graficamente os resultados da Tabela 4 e da Tabela 5 segundo as métricas de TPR e FAR.

Tabela 4 — Resultados da comparação do sistema de mapeamento com o classificador humano, divididos em diferentes alcances. Os números de 1, 2, 3 e 4 identificam as imagens estáticas utilizadas nos experimentos.

Alcance do mapeamento (m)	1	2	3	4
	TPR (%)			
Até 10	94,05	93,69	86,81	100,00
10-20	81,46	70,46	79,61	97,51
20-35	90,13	84,35	92,22	88,75
35-50	56,76	56,72	72,91	52,69
FAR (todos os alcances) (%)	2,29	3,57	2,94	20,69

Tabela 5 — Resultados da comparação do sistema de mapeamento com o classificador humano obtidos por Thrun et al. (THRUN,2006) no mapeamento de regiões trafegáveis em desertos e montanhas.

Alcance do mapeamento (m)	Deserto	Montanhas
	TPR (%)	
Até 10	-	-
10-20	90,46	88,32
20-35	91,18	86,65
35-50	87,97	80,11
50+	69,42	54,89
FAR (todos os alcances) (%)	3,70	2,60

Fonte: Adaptado de (THRUN,2006).

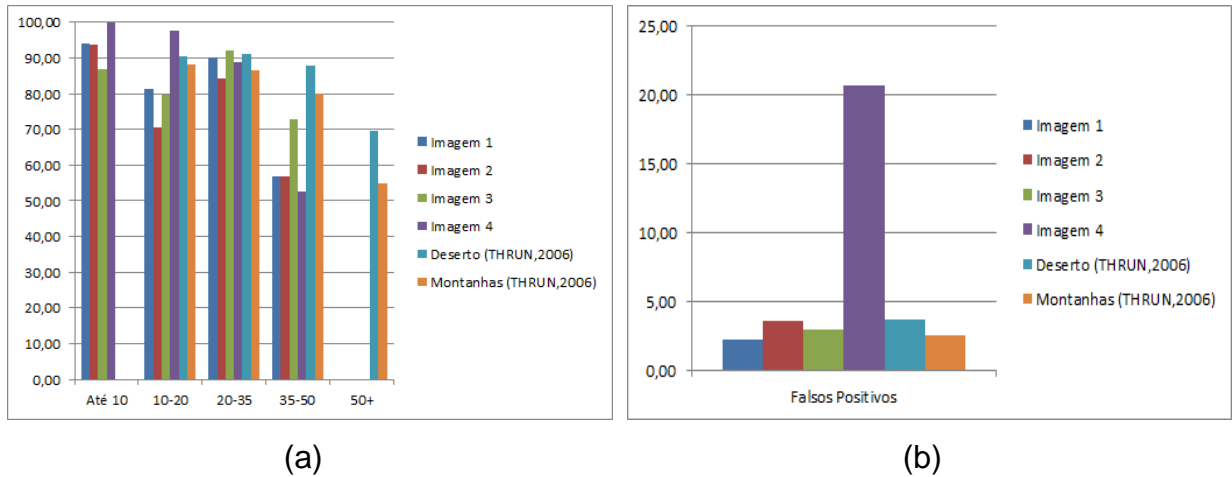


Figura 33 – Gráficos dos resultados da comparação do sistema de mapeamento com o classificador humano. (a) Gráfico em barras dos resultados segundo a métrica TPR. (b) Gráfico em barras dos resultados segundo a métrica FAR.

A Figura 34, Figura 35, Figura 36 e Figura 37 apresentam os resultados obtidos na comparação do sistema de mapeamento com o classificador humano para cada imagem do conjunto de testes. Nestas figuras, são apresentadas (a) a imagem de referência classificada manualmente, (b) o resultado do classificador humano após realizar o mapeamento inverso de perspectiva, (c) o resultado do mapeamento bayesiano, (d) a projeção para a imagem de referência da região trafegável obtida pelo sistema de mapeamento, (e) uma comparação entre (b) e (c), identificando a região trafegável corretamente mapeada – Verdadeiros Positivos - e a região não trafegável equivocadamente mapeada como trafegável – Falsos Positivos e (f) projeção para a imagem de referência da comparação dos resultados segundo as métricas de Verdadeiros Positivos e Falsos Positivos. Em (e) e (f) as regiões em verde identificam os verdadeiros positivos e as regiões em vermelho identificam os falsos positivos.

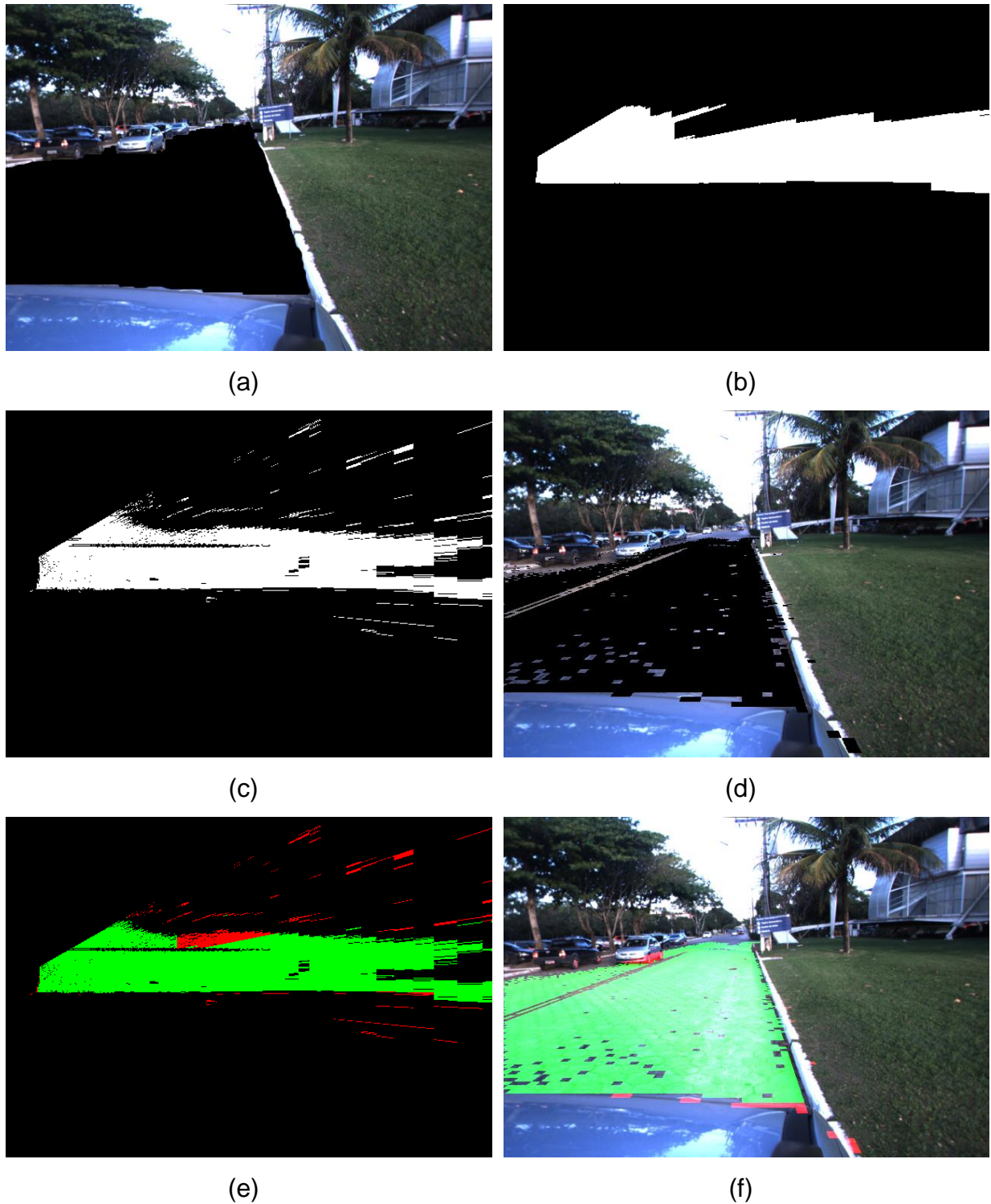


Figura 34 – Resultados da comparação do sistema de mapeamento com o classificador humano para a imagem estática 1. (a) Imagem de referência classificada manualmente. (b) Resultado do classificador humano - mapeamento inverso de perspectiva da imagem de referência classificada manualmente. (c) Resultado do mapeamento bayesiano – mapa de *grid* de ocupação probabilístico. (d) a projeção para a imagem de referência da região trafegável obtida pelo sistema de mapeamento. (e) Comparação do resultado do classificador humano (b) com o resultado do mapeamento bayesiano (c) – as regiões em verde identificam os verdadeiros positivos e as regiões em vermelho identificam os falsos positivos. (f) Projeção para a imagem de referência da comparação dos resultados segundo as métricas de Verdadeiros Positivos e Falsos Positivos.

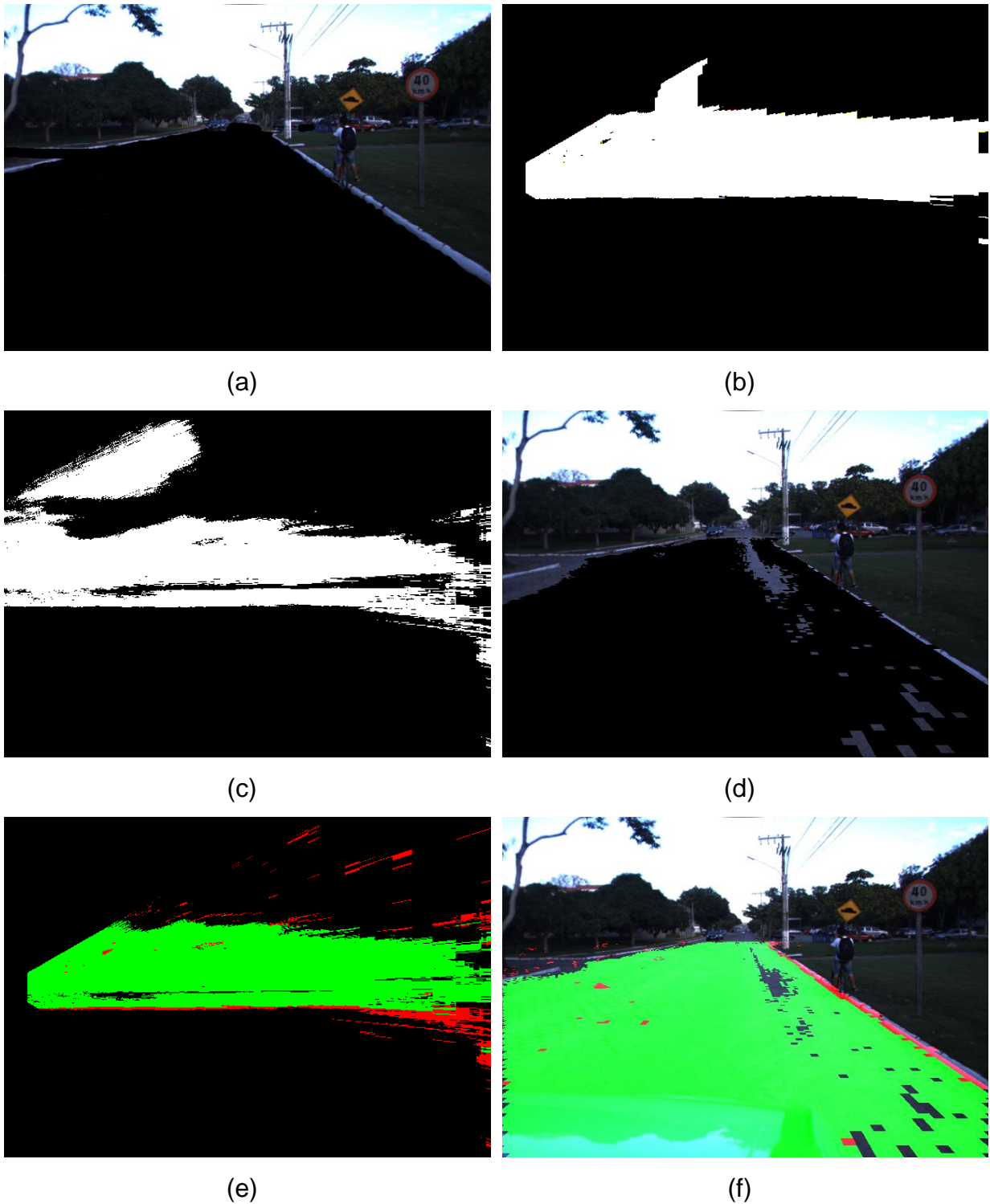


Figura 35 – Resultados da comparação do sistema de mapeamento com o classificador humano para a imagem estática 2. (a) Imagem de referência classificada manualmente. (b) Resultado do classificador humano - mapeamento inverso de perspectiva da imagem de referência classificada manualmente. (c) Resultado do mapeamento bayesiano – mapa de *grid* de ocupação probabilístico. (d) a projeção para a imagem de referência da região trafegável obtida pelo sistema de mapeamento. (e) Comparação do resultado do classificador humano (b) com o resultado do mapeamento bayesiano (c) – as regiões em verde identificam os verdadeiros positivos e as regiões em vermelho identificam os falsos positivos. (f) Projeção para a imagem de referência da comparação dos resultados segundo as métricas de Verdadeiros Positivos e Falsos Positivos.

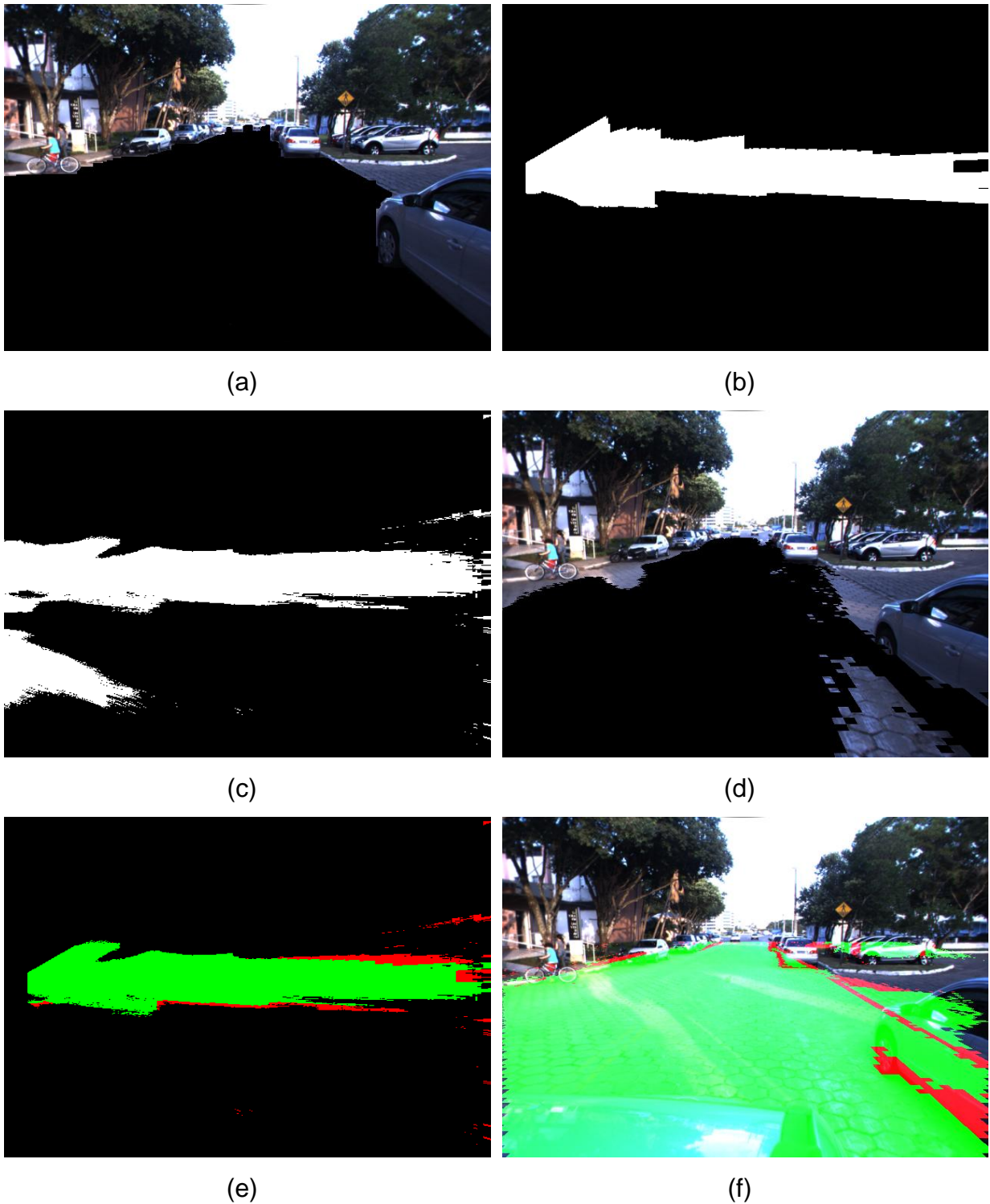


Figura 36 – Resultados da comparação do sistema de mapeamento com o classificador humano para a imagem estática 3. (a) Imagem de referência classificada manualmente. (b) Resultado do classificador humano - mapeamento inverso de perspectiva da imagem de referência classificada manualmente. (c) Resultado do mapeamento bayesiano – mapa de *grid* de ocupação probabilístico. (d) a projeção para a imagem de referência da região trafegável obtida pelo sistema de mapeamento. (e) Comparação do resultado do classificador humano (b) com o resultado do mapeamento bayesiano (c) – as regiões em verde identificam os verdadeiros positivos e as regiões em vermelho identificam os falsos positivos. (f) Projeção para a imagem de referência da comparação dos resultados segundo as métricas de Verdadeiros Positivos e Falsos Positivos.

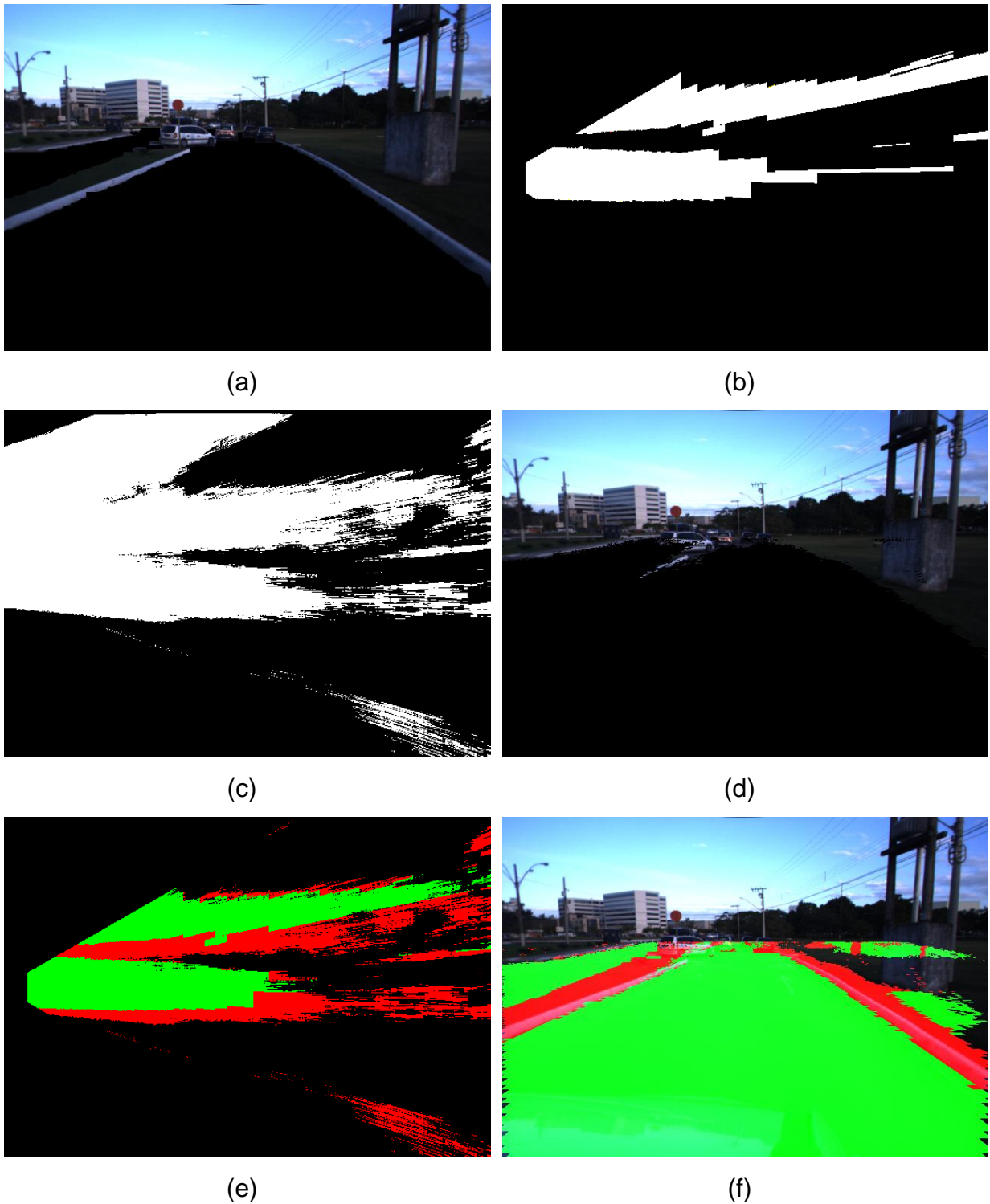


Figura 37 – Resultados da comparação do sistema de mapeamento com o classificador humano para a imagem estática 4. (a) Imagem de referência classificada manualmente. (b) Resultado do classificador humano - mapeamento inverso de perspectiva da imagem de referência classificada manualmente. (c) Resultado do mapeamento bayesiano – mapa de *grid* de ocupação probabilístico. (d) a projeção para a imagem de referência da região trafegável obtida pelo sistema de mapeamento. (e) Comparação do resultado do classificador humano (b) com o resultado do mapeamento bayesiano (c) – as regiões em verde identificam os verdadeiros positivos e as regiões em vermelho identificam os falsos positivos. (f) Projeção para a imagem de referência da comparação dos resultados segundo as métricas de Verdadeiros Positivos e Falsos Positivos.

Na avaliação do sistema de mapeamento os resultados obtidos para as imagens estáticas 1, 2 e 3 apresentaram uma TPR superior no mapeamento de regiões entre 20 e 35 m a frente do veículo do que no mapeamento de regiões mais próximas, entre 10 e 20 m a frente do veículo. Este resultado não se confirmou ao avaliar a imagem estática 4 e também não foi evidenciado em um trabalho similar desenvolvido por Thrun et al. (THRUN,2006) que obteve uma diminuição da TPR em função da distância. Assim, é provável que este resultado não esperado seja uma particularidade do conjunto de testes utilizado para a avaliação do sistema de mapeamento, embora não tenha sido possível identificar tal particularidade.

É possível observar que os resultados obtidos para todas as imagens estáticas apresentam regiões de falsos positivos adjacentes aos contornos da região trafegável. Esta característica pode ser explicada pelo fato de que mesmo pequenos erros na estimativa de estado da câmera causam distorções na imagem projetada a partir do mapeamento inverso de perspectiva, fazendo com que as mesmas regiões do mundo real se apresentem em localizações diferentes em imagens que foram capturadas em momentos sucessivos. A utilização de um mecanismo como o EKF para estimativa do estado da câmera é capaz de diminuir as distorções nas imagens obtidas a partir do mapeamento inverso de perspectiva, mas ainda assim essas distorções podem causar erros significativos no sistema de mapeamento. Para ilustrar um cenário no qual a qualidade do mapeamento inverso de perspectiva influencia negativamente o sistema de mapeamento, a Figura 38(a), Figura 38(b) e Figura 38(c) apresentam os resultados, com significativas distorções do mundo real, do mapeamento inverso de perspectiva realizado sobre imagens de referência que foram capturadas menos de 2 segundos antes da captura da imagem estática 4. A Figura 38(d) apresenta o resultado, sem significativas distorções do mundo real, do mapeamento inverso de perspectiva realizado sobre a imagem estática 4. A instabilidade da estimativa do estado da câmera, evidenciada na baixa qualidade dos resultados do mapeamento inverso de perspectiva para as imagens capturadas momentos antes da captura da imagem estática 4, contribui para a alta FAR observada nos resultados experimentais obtidos ao avaliar a imagem estática 4.

Os resultados apresentados neste capítulo demonstram que o sistema de mapeamento é bastante robusto para o mapeamento de regiões a até 35 m a frente do veículo, tendo desempenho muito bom para algo baseado primordialmente em visão computacional e permitindo estender o mapeamento baseado em outros sensores com alcance mais limitado. Além disso, a baixa

FAR observada permite que seja realizado planejamento de caminhos seguro com base no mapa de ocupação da região trafegável construído.

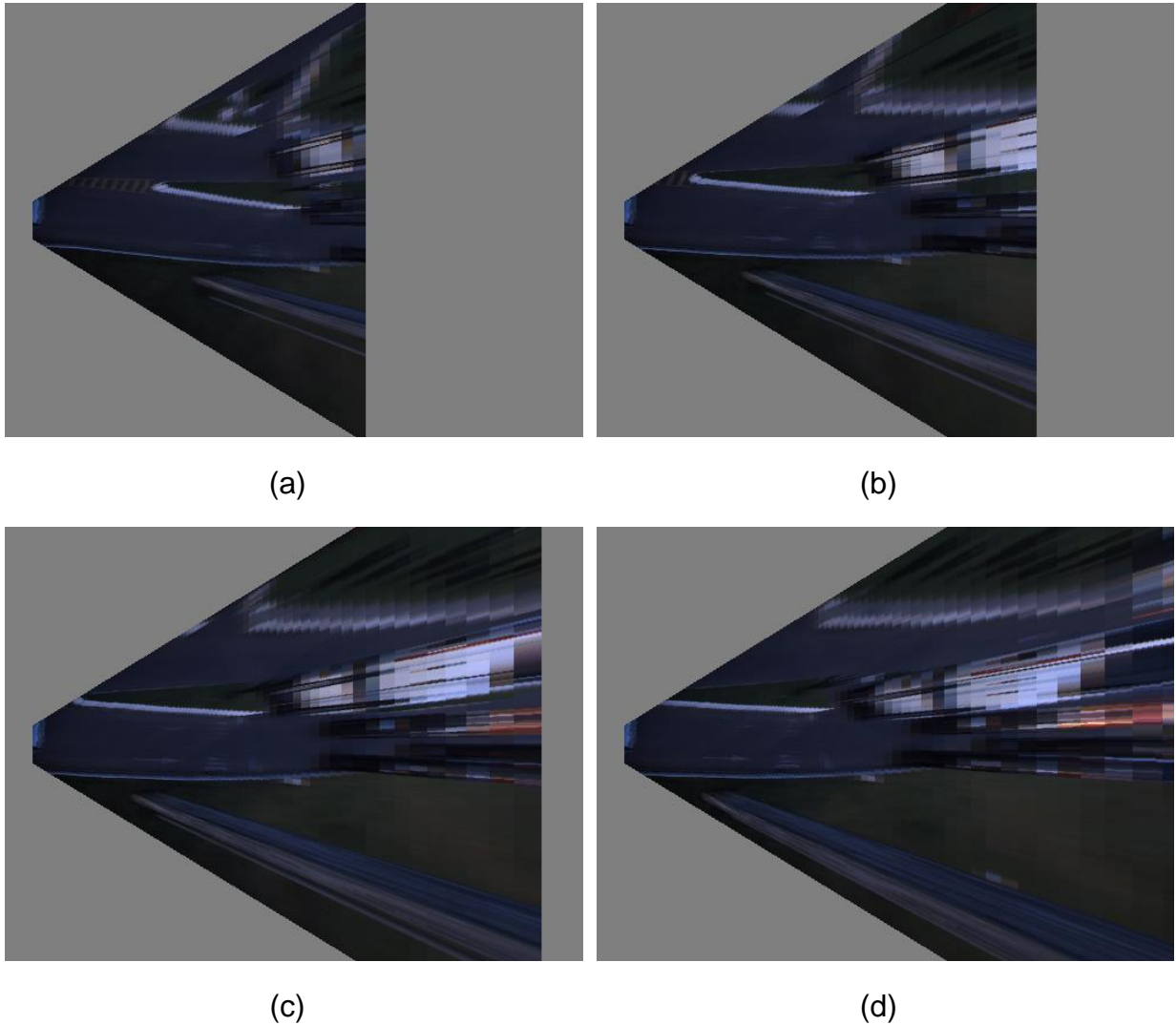


Figura 38 – Mapeamento inverso de perspectiva de imagens capturadas pouco antes da captura da imagem estática 4. (a), (b) e (c) Resultados com significativas distorções do mapeamento inverso de perspectiva realizado sobre imagens de referência capturadas menos de 2 segundos antes da captura da imagem estática 4. (d) Resultado do mapeamento inverso de perspectiva realizado sobre a imagem estática 4.

6.2.1 Limitações do classificador

Uma vez que o classificador das regiões trafegáveis é baseado em uma coleção de mistura de gaussianas que modela as propriedades de cor da região trafegável, regiões trafegáveis que possuem propriedades de cor que ainda não foram aprendidas pelo classificador não podem ser devidamente classificadas como trafegáveis. Essas novas regiões se apresentam em um

cenário real de várias formas: presença de sombras, mudança de condições climáticas, mudança de tipo de pavimentação e presença de faixas e sinalização desenhadas na região trafegável. A maioria dessas situações é devidamente tratada através do aprendizado contínuo, uma vez que assim que o classificador incorpora uma gaussiana que represente apropriadamente novas propriedades de cor da região trafegável, as novas regiões trafegáveis podem ser devidamente classificadas. No entanto, a presença de faixas e sinalizações desenhadas na região trafegável representa um caso difícil de ser tratado, pois dificilmente uma janela de segurança é capaz de obter uma gaussiana que modele adequadamente suas propriedades de cor. A Figura 39 ilustra a limitação do classificador na identificação de uma faixa desenhada sobre a região trafegável. A Figura 39(a) apresenta a imagem de referência, capturada pela câmera direita, contendo uma faixa de pedestres desenhada na cor branca. A Figura 39(b) apresenta o resultado da classificação realizada na imagem de referência após realizar o mapeamento inverso de perspectiva, evidenciando que a parte da faixa de pedestres que está desenhada com maior intensidade de branco foi classificada como não trafegável.



Figura 39 – Limitações do classificador para identificação de faixas e sinalizações desenhadas sobre a região trafegável. (a) Imagem de referência (capturada pela câmera direita). (b) Resultado da classificação realizada sobre a imagem de referência após realizar o mapeamento inverso de perspectiva.

6.3 Avaliação de Desempenho Computacional

Uma vez que funcionamento em tempo real é um requisito essencial para que um sistema de mapeamento possa ser utilizado em situações de navegação autônoma, este trabalho desen-

volveu uma versão paralelizada do classificador baseado em coleção de mistura de gaussianas utilizando programação paralela com a tecnologia CUDA. A decisão de paralelizar a etapa de classificação da região trafegável foi tomada após avaliar o desempenho computacional de 10 iterações do sistema de mapeamento. A Figura 40 apresenta um gráfico de pizza que representa o percentual médio de tempo gasto por cada etapa do sistema de mapeamento, considerando a versão sequencial do classificador, que permite ao sistema de mapeamento processar na média 3,6 FPS. A análise do gráfico de pizza da Figura 40 permite identificar que a maior parte do tempo de uma iteração do sistema de mapeamento, aproximadamente 79,5%, é gasto pela etapa de classificação da região trafegável, indicando que melhorias de desempenho nesta etapa podem contribuir significativamente para a melhoria de desempenho computacional do sistema de mapeamento. A Figura 41 apresenta um gráfico de pizza que representa o percentual médio de tempo gasto por cada etapa do sistema de mapeamento, considerando a versão paralelizada do classificador, que permite ao sistema de mapeamento processar na média 17,55 FPS. Esta média pode ser considerada satisfatória, visto que a câmera estéreo utiliza nos experimentos possui um *frame rate* máximo de apenas 16 FPS.

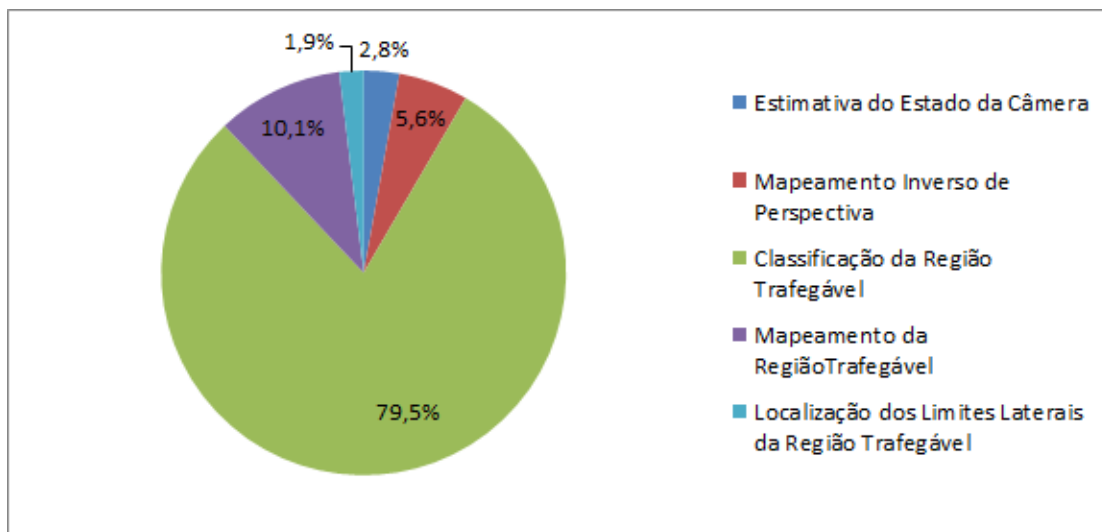


Figura 40 – Gráfico de pizza com o percentual médio de tempo gasto por cada etapa do sistema de mapeamento, considerando a versão sequencial do classificador.

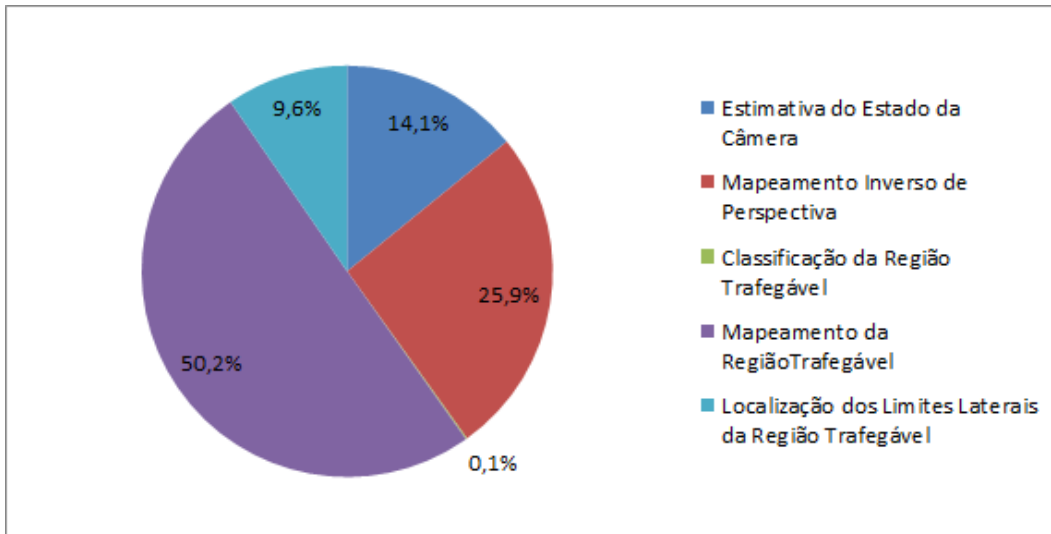


Figura 41 – Gráfico de pizza com o percentual médio de tempo gasto por cada etapa do sistema de mapeamento, considerando a versão paralelizada do classificador.

7 DISCUSSÃO

Neste capítulo são discutidos os principais trabalhos correlatos, bem como as principais limitações deste trabalho.

Este trabalho apresenta um sistema para mapeamento da região trafegável que se baseia principalmente em visão computacional. Seu principal diferencial é utilizar unicamente uma câmera estéreo frontal para a captura de informações do ambiente e para construção de mapas de ocupação, dispensando a necessidade de utilização de outros sensores de alcance.

7.1 Trabalhos Correlatos

A utilização de câmeras digitais para a classificação de regiões trafegáveis tem sido objeto de estudos recentes. Katramados (KATRAMADOS,2009) emprega uma câmera monocular localizada sobre um robô autônomo para mapeamento de ambientes não estruturados, utilizando um classificador baseado em análise de histogramas. Este classificador utiliza na fase de aprendizado uma pequena janela de segurança fixa nas imagens capturadas pela câmera monocular, extraíndo informações de cor e textura de vários canais dos sistemas de cores *Hue/Saturation/Lightness* (HSL), *YCbYr* e *LAB*. Na fase de classificação, o classificador determina a partir da análise de histograma os intervalos de valores de intensidade que dominam a janela de segurança, e classifica como trafegáveis apenas os pixels da imagem que estejam contidos em um desses intervalos dominantes.

Sales (SALES,2010) emprega uma câmera digital monocular localizada sobre um robô autônomo para mapeamento de ambientes estruturados, utilizando redes neurais artificiais para a classificação de regiões trafegáveis e não trafegáveis. No trabalho de Sales, a fase de treinamento ocorre ao classificar manualmente o primeiro frame capturado pela câmera monocular. Após o treinamento, o método funciona de maneira autônoma, empregando quatro redes neurais que analisam propriedades dos modelos de cores *Hue/Saturation/Value* (HSV) e *Red/Green/Blue* (RGB) e classificam blocos de pixel das imagens em trafegável com valor 1,0 e não trafegável com valor 0,0. O resultado do classificador para cada bloco de pixels da imagem é dado pela média aritmética dos resultados das quatro redes neurais, e varia no in-

tervalor entre 0,0 e 1,0. Este resultado é simplificado para três possíveis estados: não trafegável, com valor 0,0, se o resultado do classificador é menor ou igual a 0,3; trafegável, com valor 1,0, se o resultado do classificador é maior ou igual a 0,7; e não determinado, com valor 0,5, se o resultado do classificador está presente no intervalo de 0,3 a 0,7. Os mapas obtidos após a classificação das imagens são utilizados para a tomada de decisões durante a navegação do robô autônomo.

Embora ainda seja objeto de estudo de trabalhos recentes, a utilização de visão computacional e algoritmos de aprendizado de máquina para construção de mapas de ocupação da região trafegável em situações de navegação autônoma de veículos em ambientes não estruturados vem sendo objeto de pesquisa da comunidade científica pelo menos desde 1993, quando o sistema *Supervised Classification Applied to Road Following* – SCARF foi descrito por Crisman e Thorpe (CRISMAN,1993). O sistema SCARF empregava aprendizado de máquina baseado em coleções de mistura de gaussianas, com as gaussianas definidas sobre o sistema de cores RGB, mantendo uma coleção de gaussianas para reconhecimento da região trafegável e outra coleção para reconhecimento de regiões não trafegáveis. A maior limitação desse sistema é justamente o aprendizado realizado *offline*, visto que esta abordagem limita a utilização do classificador. Como o SCARF aprende a classificar as regiões previamente - antes da navegação do veículo - o classificador tende a não ser robusto o suficiente na classificação de regiões diferentes que não foram aprendidas na fase de treinamento.

Com intuito de efetivar mapeamento de longo alcance (mais de 70 m) para navegação autônoma de um veículo terrestre, Thrun et al. (THRUN,2006) apresentam uma evolução do sistema SCARF que permite aprendizagem continuada durante a navegação do veículo Stanley, vencedor do desafio DARPA *Grand Challenge 2005*. Para isso, é utilizado um conjunto de LIDARs fixados no topo de Stanley que permite validar uma pequena região trafegável – conhecida como janela de segurança - a frente do veículo. A janela de segurança determinada pelo conjunto de LIDARs é localizada nas imagens obtidas a partir de uma câmera monocular que também se localizava no topo de Stanley e os pixels que pertencem à janela de segurança são utilizados como amostras para evoluir uma coleção de mistura de gaussianas que modela as propriedades de cor da região trafegável. Esta abordagem permite que o classificador de Stanley consiga adaptar-se a novos tipos de pavimentação em menos de um segundo (THRUN,2006).

Além de Stanley, outro veículo de destaque no desafio DARPA *Grand Challenge* 2005 e de especial interesse para esta dissertação foi o Desert Buckeyes, que conseguiu superar a etapa de Qualificação e dirigir autonomamente em terreno completamente desconhecido por aproximadamente 50 km na corrida final (HUMMEL,2006). O Desert Buckeyes foi equipado com um sistema de mapeamento em tempo real baseado em visão computacional, utilizando tanto visão estéreo quanto visão monocular, e seu planejamento de caminho era realizado diretamente sobre o mapa obtido a partir dos algoritmos de visão computacional.

Além da capacidade de mapeamento, veículos autônomos precisam, em muitos casos, determinar os limites da região trafegável com o intuito de obedecer às leis de trânsito vigentes e minimizar as chances de se encontrar obstáculos à frente do veículo. Para determinar os limites laterais da região trafegável, Broggi (BROGGI,1999) emprega processamento de imagens obtidas a partir de uma câmera monocular localizada sobre um veículo para detecção das faixas de sinalização que determinam os limites laterais da via. O método de Broggi primeiro projeta a imagem original obtida a partir da câmera monocular aplicando o mapeamento inverso de perspectiva e em seguida filtra a imagem projetada para destacar as faixas de sinalização baseando-se no fato de que essas faixas são desenhadas sobre as vias em cores de maior intensidade (nos resultados experimentais apresentados por Broggi as vias possuíam asfalto cinza escuro enquanto as faixas eram brancas). Por fim, o método reconhece cadeias de pixels que correspondem às faixas de sinalização e reúne cadeias próximas em polígonos que correspondem aos limites laterais das vias.

A principal limitação do trabalho de Broggi (BROGGI,1999) para a localização dos limites laterais da região trafegável é a necessidade de existência das faixas de sinalização que nem sempre estão presentes nas vias. Para navegação autônoma no deserto e nas montanhas durante o desafio DARPA *Grand Challenge* 2005, onde tipicamente não existiam faixas de sinalização desenhadas nas vias, Trun et al. (THRUN,2006) utilizaram o mapa construído a partir de um conjunto de LIDARs e informações de trajeto previamente fornecidas - *Route Description Data File* (RDDF) - para determinar os limites da região trafegável. Para determinar os limites laterais de ambos os lados das vias, o método proposto por Thrun et al. espera que esses limites sejam paralelos às posições fornecidas pelo RDDF, embora a exata distância lateral do RDDF para os limites laterais seja desconhecida e varie com o tempo. Thrun et al. utilizaram um filtro de Kalman para estimar a distância lateral de cada limite da região trafegável ao centro do RDDF. Para fornecer uma observação ao filtro de Kalman, o método pes-

quisa no mapa construído a partir do conjunto de LIDARs por possíveis obstáculos em localizações ortogonais ao RDDF, e a observação fornecida ao filtro de Kalman corresponde ao maior vão sem obstáculos.

7.2 Análise Crítica Deste Trabalho de Pesquisa

Com desenvolvimento do sistema de mapeamento baseado em visão computacional e os resultados das avaliações efetuadas, foram identificadas algumas limitações neste trabalho, que podem ser tratadas em trabalhos futuros.

a) O mapeamento de objetos móveis, tais como animais, pedestres, ciclistas e outros veículos que trafegam nas vias não é preciso, uma vez que a integração bayesiana das informações dos mapas instantâneos não é apropriada para o mapeamento de objetos móveis. Esta limitação é evidenciada na avaliação do sistema de mapeamento, pois tipicamente os outros veículos que trafegam nas vias representam regiões de Falsos Positivos nos mapas de ocupação probabilísticos.

b) A aplicação e avaliação dos mapas de *grid* de ocupação probabilísticos construídos pelo sistema de mapeamento para planejamento das ações em situações reais de navegação autônoma ainda não foi realizada. Esta pendência torna este trabalho menos completo que outros trabalhos correlatos, tais como Thrun et al. (THRUN,2006) e Hummel (HUMMEL,2006) que empregaram suas técnicas de mapeamento para planejamento de ações com sucesso no desafio DARPA *Grand Challenge* 2005.

c) O método de localização dos limites laterais da região trafegável pode apresentar limitações para modelagem da região trafegável na presença de curvas. Além disso, o desempenho do método não foi avaliado segundo métricas padronizadas.

d) A avaliação do sistema de mapeamento não possui uma base de comparação única e padronizada. Em virtude desta limitação, a comparação direta dos resultados deste trabalho com outros trabalhos correlatos tais como o sistema de mapeamento proposto por Thrun et al.

(THRUN,2006) não é conclusiva sobre qual trabalho apresentou os resultados mais significativos.

8 CONCLUSÃO

Neste capítulo é apresentado um breve sumário deste trabalho, suas principais conclusões e direções para trabalhos futuros.

8.1 Síntese

Navegação de veículos terrestres autônomos atraiu um nível de atenção sem precedentes quando a DARPA anunciou seus desafios em 2004 e 2005 (HUMMEL,2006). Navegação de veículos autônomos envolve ao menos três tarefas principais: percepção do ambiente (mapeamento), planejamento de caminhos e controle do veículo (HUMMEL,2006). Enquanto esta última tarefa (controle do veículo) provou-se um problema já resolvido, as duas primeiras constituem, excetuando falhas de hardware, as principais fontes de erros nos desafios DARPA *Grand Challenge* de 2004 e 2005 (HUMMEL,2006).

Neste trabalho, é investigada uma solução computacional para a tarefa de percepção do ambiente (mapeamento) motivada na habilidade do ser humano de construir mapas sofisticados utilizando seu sistema visual e suas habilidades cognitivas. Para isso, é empregada uma câmera estéreo frontal localizada sobre o veículo para a extração de informações visuais, com a finalidade de permitir a construção de mapas da região trafegável robustos que possam ser utilizados para planejamento das ações em aplicações reais de navegação autônoma de veículos. A avaliação do sistema de mapeamento proposto evidencia que a abordagem é apropriada, permitindo identificar com precisão a região trafegável, o que torna o sistema apto para ser empregado, em conjunto com outros módulos de detecção de obstáculos e objetos móveis, em uma solução para navegação de veículos autônomos.

8.2 Conclusões

Este trabalho teve como objetivo verificar a viabilidade de resolução do difícil problema de mapeamento de ambientes não estruturados em situações de navegação autônoma de veículos terrestres, empregando nesta tarefa apenas uma câmera estéreo frontal como sensor de per-

cepção do ambiente. Para isso foi desenvolvido um sistema de mapeamento baseado essencialmente em visão computacional, aprendizado de máquina e métodos probabilísticos que permite transformar as informações visuais para o sistema de coordenadas do mundo real, classificar as regiões em trafegáveis ou não trafegáveis, construir um mapa de ocupação de alta qualidade da região trafegável e localizar os limites laterais da região trafegável.

Para a tarefa de transformação das imagens da câmera de referência para o sistema de coordenadas do mundo real, foi empregada a técnica de mapeamento inverso de perspectiva (BERTOZZI, 1998). Para fazer uso desta técnica, foi necessário estimar o estado da câmera estéreo - posição e orientação da câmera em relação ao plano da região trafegável – empregando uma implementação do filtro de Kalman para a obtenção de estimativas de alta qualidade. No passo de predição do filtro foram utilizadas informações de localizações estimadas a partir de fusão de sensores, enquanto no passo de correção foram utilizadas apenas informações derivadas a partir de visão estéreo, computadas a partir da análise de mapas em v-disparidades (LABAYRADE,2002; LAMOSA,2005).

Na tarefa de classificação das regiões trafegáveis, foi empregada uma técnica de aprendizado de máquina baseada em coleção de mistura de gaussianas que permite modelar as propriedades de cor da região trafegável. Com base no modelo de cores é possível classificar regiões mais afastadas do veículo em função dos valores das funções de densidade de probabilidade aplicadas aos pixels das imagens. A abordagem proposta neste trabalho constitui uma extensão do método proposto por Thrun et al. (THRUN,2006), pois como a identificação e validação das janelas de segurança – regiões seguras utilizadas para evolução da coleção de mistura de gaussianas - são realizadas utilizando apenas técnicas de visão computacional, o sistema de mapeamento emprega apenas uma câmera estéreo frontal, dispensando a utilização de outros sensores de alcance.

Na tarefa de construção do mapa de ocupação de alta qualidade da região trafegável, é realizada a integração bayesiana das observações locais (mapas instantâneos). Para isso, é utilizado o algoritmo de mapeamento com localizações conhecidas (THRUN,2005), sendo as localizações do veículo obtidas a partir de fusão de sensores.

Por fim, o mapa de ocupação da região trafegável é analisado para localização dos limites laterais da região trafegável.

O desempenho do sistema de mapeamento proposto neste trabalho foi avaliado, apresentando resultados muito significativos. Nos experimentos realizados foi possível identificar corretamente até 92,22% das localizações trafegáveis a até 35m de distância do veículo com FAR não superior a 3,57% no mapeamento do pavimento de paralelepípedos. É importante destacar que estes bons resultados foram obtidos apesar do sistema de mapeamento não modelar explicitamente obstáculos móveis que estavam constantemente presentes no conjunto de testes utilizado para avaliação. Caso obstáculos móveis fossem apropriadamente modelados, o desempenho do sistema tenderia a melhorar.

A interpretação dos resultados obtidos nos experimentos permite concluir que o sistema de mapeamento proporciona informações de qualidade para o planejamento de caminhos. Isto pode ser constatado pela alta TPR, que não representa limitações consideráveis para o planejamento em um horizonte de até 35m e pela baixa FAR, que minimiza os riscos de segurança decorrentes de planejamento de ações que resultem em colisão com obstáculos.

8.3 Trabalhos Futuros

Uma direção para trabalhos futuros é o desenvolvimento de técnicas para detecção de objetos móveis (outros veículos, pessoas, ciclistas e animais) nos mapas de ocupação. Outra direção para pesquisas futuras consiste na avaliação dos mapas de ocupação da região trafegável que são resultados deste trabalho para realização do planejamento de ações em uma situação real de navegação autônoma, inclusive comparando-os com mapas construídos por outras abordagens, tais como mapeamento a partir de sensores de alcance e mapeamento pela fusão de sensores de alcance e visão computacional. Outra oportunidade para trabalhos futuros consiste na avaliação de diferentes abordagens para classificação de regiões trafegáveis, como classificação em blocos de pixels ao invés de classificação pixel a pixel. Por fim, pretende-se aperfeiçoar o método de localização dos limites laterais da região trafegável para modelar regiões curvas e ainda avaliar seu desempenho utilizando métricas padronizadas.

9 REFERÊNCIAS BIBLIOGRÁFICAS

BERTOZZI, M.; BROGGI, A.; FASCIOLI, A.; FASCIOLI, R. Stereo Inverse Perspective Mapping: Theory and Applications. **Image and Vision Computing Journal**, v. 16, n. 8, p. 585–590, jun. 1998.

BRADSKI, G.; KAEHLER, A. **Learning OpenCV Computer Vision with OpenCV Library**. O'Reilly Media, 2008.

BROGGI, A.; BERTOZZI, M.; FASCIOLI, A.; GUARINO, C.; BIANCO, L.; PIAZZI, A. The ARGO Autonomous Vehicle's Vision and Control Systems. **International Journal of Intelligent Control and Systems**, p. 409–441, 1999.

BUEHLER, M.; IAGNEMMA, K.; SINGH, S. **The 2005 DARPA Grand Challenge: The Great Robot Race**. Springer, 2005.

BUEHLER, M.; IAGNEMMA, K.; SINGH, S. Special Issue on the 2007 DARPA Urban Challenge, Part I. **Journal of Field Robotics**, v. 25, n. 8, p.423-566, ago. 2008.

BUEHLER, M.; IAGNEMMA, K.; SINGH, S. Special Issue on the 2007 DARPA Urban Challenge, Part II. **Journal of Field Robotics**, v. 25, n. 9, p.567-724, set. 2008.

CRISMAN, J.; THORPE, C. SCARF: A Color Vision System that Tracks Roads and Intersections. **IEEE Transactions on Robotics and Automation**, v. 9, n. 1, p. 49–58, fev. 1993.

DUDA, R. O., HART, P. E. Use of the Hough Transformation to Detect Lines and Curves in Pictures. **Communications of the Association for Computing Machinery**, v. 15, n. 1, p. 11–15, jan. 1972.

ELFES, A. Sonar-based real-world mapping and navigation. **Journal of Robotics and Au-**

tomation, v. 3, n. 3, p. 249–265, jun. 1987.

FAWCETT, T. An Introduction to ROC Analysis. **Pattern Recognition Letters**, v. 27, n. 8, p. 861-874, jun. 2006.

GALLUP, D.; FRAHM J.-M.; MORDOHAI, P.; POLLEFEYS, M. Variable Baseline/Resolution Stereo. **Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition**, p. 1-8, jun. 2008.

GEIGER, A.; ROSER, M.; URTASUN, R. Efficient Large-Scale Stereo Matching. **Proceedings of the Asian Conference on Computer Vision**, Queenstown, Nova Zelândia, v. 1, p. 25-38, nov. 2010.

GEIGER, A.; ZIEGLER, J.; STILLER, C. StereoScan: Dense 3D Reconstruction in Real-time. **Proceedings of the IEEE Intelligent Vehicles Symposium**, Baden-Baden, Alemanha, p. 963-968 jun. 2011.

HUMMEL, B.; KAMMEL, S.; DANG, T.; DUCHOW, C.; STILLER, C. Vision-based Path-Planning in Unstructured Environments. **Proceedings of the IEEE Intelligent Vehicles Symposium**, Tokyo, Japão, p. 176-181, jun. 2006.

KALMAN, R. E. A New Approach to Linear Filtering and Prediction Problems. **Journal of Basic Engineering**, v. 82, p. 35–45. 1960.

KATRAMADOS, I.; CRUMPLER, S.; BRECKON, T.P. Real-time Traversable Surface Detection by Colour Space Fusion and Temporal Analysis. **Proceedings of the International Conference on Computer Vision Systems: Computer Vision Systems**, Liège, Bélgica, p. 265–274, out. 2009.

LABAYRADE, R.; AUBERT, D.; TAREL, J. Real Time Obstacle Detection in Stereovision on Non Flat Road Geometry through V-Disparity Representation. **Proceedings of the IEEE**

Intelligent Vehicle Symposium, v. 2, p. 646–651, jun. 2002.

LAMOSA, F.; UCHIMURA, K. A Complete U-V-Disparity Study for Stereovision Based 3D Driving Environment Analysis. **Proceeding of the International Conference on 3-D Digital Imaging and Modeling**, p. 204–211, jun. 2005.

LIMA, D.A; PEREIRA, G.A.S. Um Sistema de Visão Estéreo para Navegação de um Carro Autônomo em Ambientes com Obstáculos. **Anais do XVIII Congresso Brasileiro de Automática**, Bonito, MS, Brasil, set. 2010.

LYRIO, L. J. **Mapeamento 3D para um Veículo Autônomo Ackermann**. 2012. 69 p. Monografia de conclusão de Curso (Graduação em Ciência da Computação) – Departamento de Informática, Universidade Federal do Espírito Santo, Vitória, Espírito Santo.

MONTEMERLO, M.; ROY, N.; THRUN, S. Perspectives on Standardization in Mobile Robot Programming: The Carnegie Mellon Navigation (CARMEN) Toolkit. **Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems**, p. 2436–2441. 2003.

MORAVEC, H. P. Sensor fusion in certainty grids for mobile robots. **AI Magazine**, v. 9, n. 2, p. 61-74, jul. 1988

SALES, D.O.; SHINZATO, P.; PESSIN, G.; WOLF, D.F.; OSORIO, F.S. Vision-Based Autonomous Navigation System Using ANN and FSM Control. **Proceedings of the Robotics Symposium and Intelligent Robotic Meeting (LARS)**, p. 85-90, out. 2010.

SE, S.; BRADY, M. Ground Plane Estimation, Error Analysis and Applications. **Robotics and Autonomous Systems**, v. 39, n. 2, p. 59–71, mai. 2002.

THRUN, S.; BURGARD, W., FOX, D. **Probabilistic Robotics**. The MIT Press, 2005.

THRUN, S.; MONTEMERLO, M.; DAHLKAMP, H.; STAVENS, D.; ARON, A.; DIEBEL, J.; FONG, P.; and et al. Stanley: The Robot that Won the DARPA Grand Challenge. **Journal of Field Robotics**, v. 23, n. 9, p. 661–692, jun. 2006.

ZHANG, Z. A Flexible New Technique for Camera Calibration. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 22, n. 11, p. 1330–1334, nov. 2000.