UNIVERSIDADE ESTADUAL DE CAMPINAS

FACULDADE DE ENGENHARIA MECÂNICA

COMISSÃO DE PÓS-GRADUAÇÃO EM ENGENHARIA MECÂNICA

Arthur de Miranda Neto

Sistema de Percepção Visual Embarcado aplicado à Navegação Segura de Veículos

(Embedded Visual Perception System applied to Safe Navigation of Vehicles)

Campinas, 2011.

156/2011

Arthur de Miranda Neto

Sistema de Percepção Visual Embarcado aplicado à Navegação Segura de Veículos

(Embedded Visual Perception System applied to Safe Navigation of Vehicles)

Tese apresentada ao Curso de Doutorado da Faculdade de Engenharia Mecânica da Universidade Estadual de Campinas, como requisito para a obtenção do título de Doutor em Engenharia Mecânica.

Área de Concentração: Mecânica dos Sólidos e Projeto Mecânico (UNICAMP-FEM) e Tecnologia da Informação e de Sistemas (UTC-HEUDIASYC).

Orientador UNICAMP-FEM: Prof. Dr. Douglas Eduardo Zampieri
Orientadora: UTC-HEUDIASYC: Dra. Isabelle Fantoni Coichot

Campinas
2011

UNIVERSIDADE ESTADUAL DE CAMPINAS

FACULDADE DE ENGENHARIA MECÂNICA

COMISSÃO DE PÓS-GRADUAÇÃO EM ENGENHARIA MECÂNICA

DEPARTAMENTO DE MECÂNICA COMPUTACIONAL

TESE DE DOUTORADO
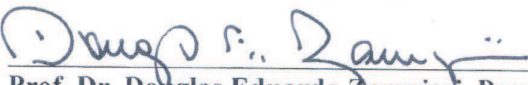
Sistema de Percepção Visual Embarcado aplicado à Navegação Segura de Veículos

Autor: Arthur de Miranda Neto
Orientador: Prof. Dr. Douglas Eduardo Zampieri – UNICAMP-FEM, Brasil.
Orientadora: Dra. Isabelle Fantoni Coichot – UTC-HEUDIASYC, França.
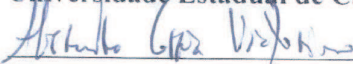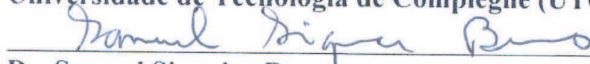
A Banca Examinadora composta pelos membros abaixo aprovou esta Tese:
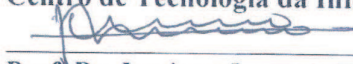
Prof. Dr. Douglas Eduardo Zampieri, Presidente
Universidade Estadual de Campinas (UNICAMP-FEM)
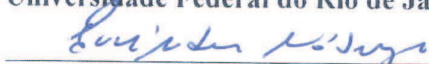
Prof. Dr. Alessandro Corrêa Victorino
Universidade de Tecnologia de Compiegne (UTC-HEUDIASYC), França
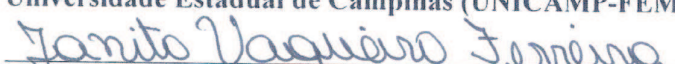
Dr. Samuel Siqueira Bueno
Centro de Tecnologia da Informação Renato Archer (CTI)

Prof. Dr. Luciano Luporini Menegaldo
Universidade Federal do Rio de Janeiro (COPPE/UFRJ)

Prof. Dr. Eurípedes Guilherme de Oliveira Nobrega
Universidade Estadual de Campinas (UNICAMP-FEM)

Prof. Dr. Janito Vaqueiro Ferreira
Universidade Estadual de Campinas (UNICAMP-FEM)

Campinas, 26 de agosto de 2011

Eu dedico esta tese,
a Deus,
a minha querida família, especialmente a minha esposa Débora e meus filhos Davi e Yasmin,
e para meus grandes amigos Douglas Eduardo Zampieri e Alessandro Corrêa Victorino.

Je dédie cette thèse,
à Dieu,
à ma chère famille, en particulier à mon épouse Débora et à mes enfants Davi et Yasmin,
et à mes grands amis Douglas Eduardo Zampieri et Alessandro Corrêa Victorino.

I dedicate this thesis,
to God,
to my dear family, especially to my wife Débora and my children Davi and Yasmin,
and to my great friends Douglas Eduardo Zampieri and Alessandro Corrêa Victorino.

Ensina-nos a contar os nossos dias, de tal maneira que alcancemos corações sábios.

Enseigne-nous à compter nos jours, tellement que nous puissions avoir un cœur sage.

So teach us to number our days, That we may get us a heart of wisdom.

Holy Bible: Psalm 90:12

Resumo

Esta tese aborda o problema de evitamento de obstáculos para plataformas terrestres semi- e autônomas em ambientes dinâmicos e desconhecidos. Baseado num sistema monocular, propõe-se um conjunto de ferramentas que monitoram continuamente a estrada a frente do veículo, provendo-o de informações adequadas em tempo real. A partir de um algoritmo robusto de detecção da linha do horizonte é possível investigar dinamicamente somente a porção da estrada a frente do veículo, a fim de determinar a área de navegação, e da deteção de obstáculos. Uma área de navegação livre de obstáculos é então representa a partir de uma imagem multimodal 2D. Esta representação permite que um nível de segurança possa ser selecionado de acordo com o ambiente e o contexto de operação. A fim de reduzir o custo computacional, um método automático para descarte de imagens é proposto. Levando-se em conta a coerência temporal entre consecutivas imagens, uma nova metodologia de gerenciamento de energia (*Dynamic Power Management*) é aplicada ao sistema de percepção visual a fim de otimizar o consumo de energia. Estas propostas foram testadas em diferentes tipos de ambientes, e incluem a deteção da área de navegação, navegação reativa e estimação do risco de colisão. Uma característica das metodologias apresentadas é a independência em relação ao sistema de aquisição de imagem e do próprio veículo. Este sistema de percepção em tempo real foi avaliado a partir de diferentes bancos de testes e também a partir de dados reais obtidos por diferentes plataformas inteligentes. Em tarefas realizadas com uma plataforma semi-autônoma, testes foram conduzidos em velocidades acima de 100 Km/h. A partir de um sistema em malha aberta, deslocamentos reativos autônomos foram realizados com sucesso..

Palavras-chave: Robôs Móveis, Veículos Autônomos, Navegação de Robôs Móveis, Visão por Computador, Tempo Real, Detecção de Obstáculos.

Abstract

This thesis addresses the problem of obstacle avoidance for semi- and autonomous terrestrial platforms in dynamic and unknown environments. Based on monocular vision, it proposes a set of tools that continuously monitors the way forward, proving appropriate road informations in real time. A horizon finding algorithm was developed to sky removal. This algorithm generates the region of interest from a dynamic threshold search method, allowing to dynamically investigate only a small portion of the image ahead of the vehicle, in order to road and obstacle detection. A free-navigable area is therefore represented from a multimodal 2D drivability road image. This multimodal result enables that a level of safety can be selected according to the environment and operational context. In order to reduce processing time, this thesis also proposes an automatic image discarding criteria. Taking into account the temporal coherence between consecutive frames, a new Dynamic Power Management methodology is proposed and applied to a robotic visual machine perception, which included a new environment observer method to optimize energy consumption used by a visual machine. This proposal was tested in different types of image texture (road surfaces), which includes free-area detection, reactive navigation and time-to-collision estimation. A remarkable characteristic of these methodologies is its independence of the image acquiring system and of the robot itself. This real-time perception system has been evaluated from different test-banks and also from real data obtained by two intelligent platforms. In semi-autonomous tasks, tests were conducted at speeds above 100 Km/h. Autonomous displacements were also carried out successfully. The algorithms presented here showed an interesting robustness.

Key Words: Mobile Robots, Autonomous Vehicles, Computer Vision, Real Time, Road Detection, Obstacle Avoidance.

List of Figures

List of Tables

List of Abbreviations

ADAS: Advanced Driver Assistance Systems.

ASER: Automatique, Systèmes Embarqués, Robotique.

API: Application Programming Interface.

CARMEN: An experimental vehicle at Heudiasyc Laboratory in Compiegne, France.

CTI: Center for Information Technology Renato Archer (or Renato Archer IT Center), Campinas, Brazil.

CRE: Collision Risk Estimation.

DARPA: The Defense Advanced Research Projects Agency (DARPA) is an agency of the United States Department of Defense responsible for the development of new technology for military.

DPM: Dynamic Power Management.

EATR: Energetically Autonomous Tactical Robot.

ITA: Interactive Thresholding Algorithm.

KF: Kalman filter.

KIT: Karlsruhe Institute of Technology, a German academic research and education institution.

LMA: Autonomous Mobility Laboratory at University of Campinas, Brazil.

MCT: Ministry of Science and Technology, Brazil.

OTM: Otsu Thresholding Method.

PCC: Pearson's Correlation Coefficient.

RTS: Real-Time Scheduling.

ROI: Regions-Of-Interest.

STRADA: An experimental vehicle at Heudiasyc Laboratory in Compiegne, France.

TIS: Technologies de l'Information et des Systèmes.

TTC: Time-To-Collision.

UTC: University of Technology of Compiègne, Compiègne, France.

UNICAMP: University of Campinas, Campinas, Brazil.

UAVs: Unmanned Aerial Vehicles.

VERO: An experimental platform, i.e. Robot Vehicle, at Center for Information Technology Renato Archer, Campinas, Brazil.

VBTP-MR: New Family of Armored Fighting Vehicles.

Contents

Chapter 3: Sky Removal and Navigable Area Detection

Chapter 4: Experimental Evaluation

Résumé

Système de Perception Visuel Embarqué appliqué à la Navigation Sûre de Véhicules

Arthur DE-MIRANDA-NETO

Directeur(s) de thèse : Isabelle FANTONI et Douglas ZAMPIERI.

Université de Technologie de Compiègne - UTC

Laboratoire HEUDIASYC UMR 6599 CNRS/UTC

Champ disciplinaire : Technologies de l'Information et des Systèmes - TIS

Domaine : Automatique, Systèmes Embarqués, Robotique - ASER

Les études développées dans ce projet doctoral ont concerné deux problématiques actuelles dans le domaine des systèmes robotiques pour la mobilité terrestre: premièrement, le problème associé à la navigation autonome et (semi)-autonome des véhicules terrestres dans un environnement inconnu ou partiellement connu. Cela constitue un enjeu qui prend de l'importance sur plusieurs fronts, notamment dans le domaine militaire. Récemment, l'agence DARPA[1] aux États-Unis a soutenu plusieurs challenges sur cette problématique robotique; deuxièmement, le développement de systèmes d'assistance à la conduite basés sur la vision par ordinateur. Les acteurs de l'industrie automobile s'intéressent de plus en plus au développement de tels systèmes afin de rendre leurs produits plus sûrs et plus confortables à toutes conditions climatiques ou de terrain. De plus, grâce à l'électronique embarquée et à l'utilisation des systèmes visuels, une interaction avec l'environnement est possible, rendant les routes et les villes plus sûres pour les conducteurs et les piétons.

---

[1] Defense Advanced Research Projects Agency (DARPA): est une agence du département de la Défense des États-Unis chargée de la recherche et développement des nouvelles technologies destinées à un usage militaire.

L'objectif principal de ce projet doctoral a été le développement de méthodologies qui permettent à des systèmes mobiles robotisés de naviguer de manière autonome dans un environnement inconnu ou partiellement connu, basées sur la perception visuelle fournie par un système de vision monoculaire embarqué. Un véhicule robotisé qui doit effectuer des tâches précises dans un environnement inconnu, doit avoir la faculté de percevoir son environnement proche et avoir un degré minimum d'interaction avec celui-ci.

Nous avons proposé un système de vision embarquée préliminaire, où le temps de traitement de l'information (point critique dans des systèmes de vision utilisés en temps-réel) est optimisé par une méthode d'identification et de rejet d'informations redondantes. Suite à ces résultats, on a proposé une étude innovante par rapport à l´état de l´art en ce qui concerne la gestion énergétique du système de vision embarqué, également pour le calcul du temps de collision à partir d'images monoculaires. Ainsi, nous proposons le développement des travaux en étudiant une méthodologie robuste et efficace (utile en temps-réel) pour la détection de la route et l'extraction de primitives d'intérêts appliquée à la navigation autonome des véhicules terrestres. Nous présentons des résultats dans un environnement réel, dynamique et inconnu. Afin d'évaluer la performance de l'algorithme proposé, nous avons utilisé un banc d'essai urbain et réel. Pour la détection de la route et afin d'éviter les obstacles, les résultats sont présents en utilisant un véhicule réel afin d'évaluer la performance de l'algorithme dans un déplacement autonome.

Cette Thèse de Doctorat a été réalisée à partir d'un accord de cotutelle entre l´ Université de Campinas (UNICAMP) et l'Université de Technologie de Compiègne (UTC), sous la direction du Professeur Docteur Douglas Eduardo ZAMPIERI, Faculté de Génie Mécanique, UNICAMP, Campinas, Brésil, et Docteur Isabelle FANTONI-COICHOT du Laboratoire HEUDIASYC UTC, Compiègne, France. Cette thèse a été soutenue le 26 août 2011 à la Faculté de Génie Mécanique, UNICAMP, devant un jury composé des Professeurs suivants:

- Professeur Docteur Douglas Eduardo ZAMPIERI - FEM/UNICAMP,
- Professeur Docteur Alessandro Corrêa VICTORINO - UTC/Compiègne,
- Docteur Samuel Siqueira BUENO - CTI/Campinas,
- Professeur Docteur Luciano Luporini MENEGALDO - COPPE/UFRJ/Rio de Janeiro,
- Professeur Docteur Eurípedes Guilherme de Oliveira NOBREGA - FEM/UNICAMP et
- Professeur Docteur Janito Vaqueiro FERREIRA - FEM/UNICAMP.

2

Publications :

- MIRANDA NETO, A., VICTORINO, A. C., FANTONI, I. AND ZAMPIERI, D. E., "Robust Horizon Finding Algorithm for Real Time Autonomous Navigation based on Monocular Vision", IEEE International Conference on Intelligent Transportation Systems (**ITSC 2011**), Washinton DC, US, October 5-7, 2011.

- MIRANDA NETO, A., VICTORINO, A. C., FANTONI, I., ZAMPIERI, D. E., "Automatic Regions-of-Interest Selection based on Pearson's Correlation Coefficient", IEEE/RSJ International Conference on Intelligent Robots and Systems (**IROS 2011**), IROS Workshop on Visual Control of Mobile Robots, San Francisco, California, US, September 25-30, 2011.

- MIRANDA NETO, A., VICTORINO, A. C., FANTONI, I. AND ZAMPIERI, D. E., "Real-Time Dynamic Power Management based on Pearson's Correlation Coefficient", IEEE International Conference On Advanced Robotics (**ICAR 2011**), Tallinn, Estonia, June 20-23, 2011.

- MIRANDA NETO, A., ZAMPIERI, D. E., VICTORINO, A. C. AND FANTONI-COICHOT, I., "Self-Organizing Maps for Environments and States Mapping of an Autonomous Navigation System Based on Monocular Vision", 20th International Congress of Mechanical Engineering (**COBEM 2009**), Gramado, Brazil, November 15-20, 2009.

- MIRANDA NETO, A., RITTNER, L., LEITE, N, ZAMPIERI, D. E. AND VICTORINO, A. C., "Nondeterministic Criteria to Discard Redundant Information in Real Time Autonomous Navigation Systems based on Monocular Vision", **ISIC Invited Paper**, IEEE Multi-conference on Systems and Control (**ISIC - MSC 2008**), San Antonio, Texas, US, September 3-5, 2008.

- MIRANDA NETO, A., RITTNER, L., LEITE, N, ZAMPIERI, D. E., LOTUFO R. AND MENDELECK, A., "Pearson's Correlation Coefficient for Discarding Redundant Information in Real Time Autonomous Navigation System", IEEE International Conference on Control Applications Part of IEEE Multi-conference on Systems and Control (**CCA - MSC 2007**), Singapura, October 1-3, 2007.

- MIRANDA NETO, A. AND RITTNER, L., "Simple and Efficient Road Detection Algorithm for Real Time Autonomous Navigation based on Monocular Vision", IEEE 3rd Latin American Robotics Symposium (**LARS 2006**), Santiago, Chile, October 26-27, 2006.

- MENDELECK, A., ZAMPIERI, D. E. AND MIRANDA NETO, A., "Classifying Systems Applied to Autonomous Robots´s Navigation", Brazilian Congress of Neural Networks (**CBRN 2005**), October, 2005.

Capítulo 1

Introdução

Conteúdo

1.1 Introdução

Historicamente, o desenvolvimento de novas tecnologias tem sido impulsionado pelas aplicações militares. Entretanto, tecnologias de emprego dual, i.e. militar e civil, são extrategicamente importantes. Robôs possuem muitos atributos valiosos que ajudam e complementam tarefas civis e militares. Mais importante, podem substituir humanos em tarefas perigosas. Robôs desempenharão um papel vital no futuro.

Nas últimas três décadas, a navegação visual para robôs móveis ou veículos não tripulados tornou-se uma fonte de incontáveis contribuições de pesquisa (Bonin-Fonte, et al., 2008). Várias aplicações para controle de plataformas semi- e autônomas estão em desenvolvimento. O desafio de construir métodos robustos, e, na maioria dos casos, sistemas de tempo real, está longe de ser alcançado. Isto pode ser observado pelo grande número de pesquisas publicadas nos últimos anos.

Para veículos terrestres em ambientes desconhecidos, a pesquisa tem sido conduzida em quatro frentes com objetivos diferentes. A primeira delas está relacionada à exploração do espaço. A segunda frente, com fins militares, tem apoio e incentivo da agência dos EUA conhecida como DARPA[2], cuja principal missão é prover as forças armadas com tecnologia inovadora o suficiente para manter a superioridade tecnológica. Uma subdivisão da área anterior diz respeito à aplicação da robótica móvel na agricultura. A quarta frente tem apoio da indústria automotiva, como exemplo, os Sistemas de Assistência ao Condutor, *Advanced Driver Assistance Systems* (ADAS).

Para fins militares ou civis, algumas destas aplicações incluem: o *Grand Challenge* (Thrun, et al., 2006) and *Urban Challenge* (Rojo, et al., 2007). O objetivo do *Grand Challenge* foi o desenvolvimento de um veículo autônomo capaz de atravessar um terreno fora de estrada não estruturado. Para o *Urban Challenge*, o objetivo do sistema foi conduzir um veículo, em modo autônomo, em um ambiente semi-urbano, que incluía cruzamentos e obstáculos estáticos e dinâmicos. Por outro lado, a robótica móvel na agricultura envolve áreas multidisciplinares, como a navegação inercial e reconhecimento de padrões a partir de sensores aéreos e terrestres (Bakker, et al., 2008), (Subramanian, et al., 2006). Finalmente, impulsionada pelo elevado número de veículos em todo o mundo, os sistemas ADAS surgem para ajudar motoristas na tarefa de condução (Gietelink, et al., 2006), (Rodríguez Flórez, 2010).

Em outra frente, os robôs aéreos também oferecem grandes perspectivas em muitas aplicações: busca e salvamento, fiscalização de obras, monitoramento em tempo real, em missões aéreas de alto risco, cartografia, deteção de incêndio ou em gravações de filmes no cinema (Kim, et al., 2003), (Bonin-fonte, et al., 2008). Da mesma forma, o desenvolvimento de veículos aéreos não tripulados (UAVs) tem sido alvo das aplicações militares. Estes veículos são utilizados principalmente para vigilância e reconhecimento. No entanto, uma limitação é o tempo máximo de vôo, uma vez que não podem transportar grandes volumes de combustível (Anton, et al., 2008). Futuras explorações em Marte, estabelecidas pela *Vision for Space Exploration*, exigem UAVs de longa resistência que utilizam recursos que são abundantes em Marte (Klesh, et al., 2009), (NASA, 2004). Incluindo inteligência reativa baseado em uma modalidade sentido-ação

---

[2] The Defense Advanced Research Projects Agency (DARPA) is an agency of the United States Department of Defense responsible for the development of new technology for military.

autônomos a partir do processamento óptico de imagem, o objetivo do projeto *Energetically Autonomous Tactical Robot* (EATR) ™ é desenvolver uma plataforma robótica autônoma capaz de realizar missões de longa distância e resistência sem a necessidade de reabastecimento manual ou convencional (DARPA, 2009), (Finkelstein, 2010).

Em todos estes casos anteriormente relatados, um fator importante é a variedade e a complexidade de ambientes e situações. Estes desenvolvimentos, que incluem plataformas inteligentes, tem um problema comum: proporcionar à plataforma a capacidade de perceber e interagir com seu ambiente vizinho. Essa capacidade, conhecida como percepção do ambiente, ainda está em evolução. Parte disso, a visão por computador aparece como uma importante ferramente devido à riqueza de informações que pode fornecer. Uma das aplicações possíveis para os sistemas de visão é a deteção da região navegável em imagens. Esta é uma tarefa fundamental no guiamento e evitamento de obstáculos para os veículos autônomos terrestres.

Desta forma, esta tese aborda o problema de desvio de obstáculos para plataformas semi- e autônomas em ambientes dinâmicos e desconhecidos. Baseado num sistema de visão monocular, propõe um conjunto de ferramentas que monitoram continuamente o caminho a frente do veículo, provendo informações apropriadas da via em tempo real. Este sistema foi testado com diferentes tipos de imagens e texturas, que inclui a deteção da área livre a frente do veículo, navegação reativa e estimação do risco de colisão. A fim de reduzir o tempo de processamento, este trabalho propõe também um método automático para descarte de imagens/informações redundantes, caracterizado posteriormente como um novo método para observação do ambiente, a fim de otimizar o consumo de energia utilizada por um sistema visual. Esta última proposta pode ser estendida a diferentes tipos de sistemas e sensores. Uma característica marcante das metodologias desenvolvidas é a independência ao sistema de aquisição de imagem e ao próprio robô. Este sistema de percepção em tempo real foi avaliado a partir de bases de testes e de dados reais obtidos por duas plataformas inteligentes.

1.1.1 Contexto Governamental, Institucional e da Tese

Robótica é um dos principais vetores de inovação. Especialmente com relação ao desenvolvimento de aplicações relacionadas aos Sistemas de Defesa, fazem parte de um contexto estratégico importante, e podem prover um salto revolucionário nas capacidades militares, aumentando a eficácia do combate. Particularmente, a navegação de robôs pode ser subdividida em duas classes: em ambiente previamente conhecido e em ambiente desconhecido. A última, tem sido de grande interesse para a comunidade militar.

Para veículos terrestres em ambientes desconhecidos, pesquisas tem sido conduzidas em algumas frentes, entre elas, com fins militares, e como mostrado na Figura 1.1, tem[3] apoio e incentivo do Departamento de Defesa dos EUA. Um dos objetivos expressos na visão dos EUA é a transformação de um terço da sua frota de veículos militares em veículos autônomos para uso geral, mesmo em áreas urbanas densamente povoadas.

Figura 1.1 – Carnegie Mellon University's Crusher: veículo terrestre de combate não-tripulado.

Em um contexto de renovação de exigências aplicados a interesses de curto e longo prazo, o Plano Estratégico Francês para Pesquisa e Tecnologia (PS R&T[4]) defini prioridades de pesquisa

---

[3] Crusher is a 6.5-ton autonomous off-road Unmanned Ground Combat Vehicle developed by researchers at the Carnegie Mellon University's National Robotics Engineering Center for Defense Advanced Research Projects Agency (DARPA).

[4] Strategic Plan for Research & Technology in defense and security, *Direction Générale de l'Armement* (DGA), 2009. The DGA is the French Government Defense procurement agency responsible for the program management, development and purchase of weapon systems for the French military.

no Plano de Prospecção de 30 anos (PP30). Este Plano de 30 anos prepara e recomenda as escolhas a serem feitas para adquirir e manter as capacidades operacionais das forças armadas francesas, agora e para o futuro. Para sistemas terrestres, o principal desafio continuará sendo o acesso a produtos inovadores (robótica) e também para equipamentos clássicos em desenvolvimento ou que exigem uma adaptação reativa. De maneira geral, podem ser classificados em três categorias: tecnologias de dupla utilização; tecnologias especificamente militares, e inovações oriundas do setor civil. O uso da robótica em operações terrestres deve permitir: a protecção dos soldados, substituí-los em situações perigosas, a produtividade e, consequentemente, a disponibilidade de soldados, a partir da realização de tarefas repetitivas, e, finalmente, a flexibilidade de uso. A Figura 1.2 apresenta uma família de plataformas robóticas em miniatura (MiniROC[5]) desenvolvidas na França para operar em terreno urbano.

Figura 1.2 – Tecnologia «*Mini-Robot de Choc*» (MiniROC).

Para o governo brasileiro, ciência, tecnologia e inovação são também questões de Estado, atravessam governos e devem ser tratados como compromissos que se movem de um período de governo para outro. A missão permanente de um Estado é manter a superioridade tecnológica,

---

[5] "Mini-Robot de Choc": A family of miniature robotic platforms developed in France is designed to operate in urban terrain.

desenvolvemendo tecnologias que as empresas privadas não podem alcançar, de forma rentável, a curto e médio prazo[6,7]. O Ministério da Ciência e Tecnologia (MCT) do Brasil, em consonância com o Plano de Aceleração do Crescimento do Governo Federal, desenvolveu um Plano de Acão 2007-2010. Desta forma, a estratégia nacional de desenvolvimento é inseparável da estratégia de defesa nacional, e o Estado Brasileiro tem procurado parcerias com outros países, a fim de desenvolver a capacidade tecnológica e fabricação de produtos de defesa.

Em acordo com as áreas de pesquisa básica de interesse do Ministério da Defesa do Brasil, e interessados em segurança veicular, a partir de 2008, um Laboratório de Mobilidade Autônoma (LMA) fora criado no Departamento de Mecânica Computacional da Universidade Estadual de Campinas. Este laboratório pretende estudar sistemas semi- e autônomos, essenciais para aplicações civis e militares. Atualmente, o LMA tem um grupo multidisciplinar que é constituído por membros especiais de diferentes instituições, como a Universidade Estadual de Campinas, Universidade de São Paulo e do Instituto Militar de Engenharia.

Com vista a compreender esforços conjuntos de pesquisa, desenvolvimento científico e tecnológico, e concernente a trabalhos de cooperação internacional, as primeiras atividades entre os membros do LMA e do Laboratório Heudiasyc UMR 6599 CNRS/UTC, França, começaram em meados de 2007, que inclui este programa de doutoramento em cooperação.

Durante os anos deste programa tese, este autor desenvolveu vários trabalhos para o estabelecimento do LMA, incluindo tarefas de relações institucionais. Além de co-fundador do LMA, este autor também foi mentor do projeto: Desenvolvimento de uma Plataforma Robótica para Estudos de Mobilidade Terrestre Semi- e Autônoma, programa Pró-Defesa[8] 2008, Ministério da Defesa do Brasil, CAPES[9].

Esta tese de doutorado foi realizada no Departamento de Mecânica Computacional, DMC/FEM, Universidade Estadual de Campinas, Brasil, de fevereiro de 2007 à julho de 2009, e no Laboratório Heudiasyc UMR 6599 CNRS/UTC, Universidade de Tecnologia de Compiègne,

---

[6] National Defense Strategy, Ministry of Defence, Brazil, 2nd edition.
[7] Action Plan 2007-2010, Ministry of Science and Technology, Brazil.
[8] Support Program for Education and Scientific Research and Technology in National Defense. It was a partnership between the Brazilian Ministry of Defense and CAPES. Its goal is to finance and stimulate the production of scientific and technological research and training of post-graduates in the area of Defense.
[9] Coordination for the Improvement of Higher Education Personnel, CAPES, Brazil.

França, de agosto de 2009 a agosto de 2011. Esta tese foi financiada em parte pelo CNPq-Brasil e em parte pelo Projeto ANR-France[10] / PERCOIVE[11].

1.1.2 Robôs Militares e Veículos em Área Urbana

De acordo com Chatila (2010), robôs entram no contexto militar especialmente quando é necessário reduzir a exposição humana a situações perigosas, principalmente em um ambiente hostil, onde um robô realiza missões especializadas. Muitas dessas missões robóticas podem ser observadas: desminagem; reconhecimento de campo, especialmente nuclear, radiológico, bacteriológico, químico, ou, como mostrado na Figura 1.3, em casos de IEDs[12].

Figura 1.3 – Uma explosão IED em uma estrada durante comboio militar.

Uma das maiores dificuldades em substituir humanos em qualquer tarefa é o fato de que ainda não é possível projetar um robô genérico adaptado a qualquer demanda, que inclui uma variedade muito grande de opções de projetos/design, tamanhos do robô, força motriz, instrumentação, desempenho esperado (incluindo a autonomia de decisão, "inteligência", etc).

---

[10] French National Research Agency, France.

[11] The problem addressed in PERCOIVE project involves increasing the safety on board and around of a road vehicles group, through observation and estimation of parameters of vehicle dynamics in interaction with the environment nearest neighbors and vehicles. Specifically, it aims to study and propose solutions to the environment perception (exteroceptive), for the estimation of vehicle dynamics and road attributes in a context estimation cooperative multi vehicle.

[12] Improvised Explosive Device (IED), also known as a roadside bomb, is constructed and deployed in ways other than in conventional military action.

Sejam em missões militares ou policiais, podem haver grande hostilidade do inimigo. Como exemplo, incursões terrestres em ambientes desconhecidos ou em área de predominância inimiga. Nestes casos, os sistemas semi- e autônomos podem contribuir para o sucesso destas missões, salvando vidas. Conflitos recentes mostram um uso emergente de veículos militares de grande porte em áreas urbanas. A complexidade das operações militares também está crescendo, pois em muitos casos, essas áreas têm dificuldade de acesso e são muito populosas. Consequentemente, aumentando significativamente a responsabilidade do condutor. A Figura 1.4 apresenta operações militares brasileiras em áreas urbanas.

Num contexto de veículos de grande porte, em 2007, nos desafios propostos pelo DARPA, a Equipe de Oshkosh adaptou um caminhão originalmente concebido para combate e utilizado pela Marinha dos EUA. O veículo conhecido como TerraMax ™ pode ser visto na Figura 1.5 (TEAM OSHKOSH, 2011).

Da mesma forma, a nova família de veículos blindados de combate é um projeto sob a direção do Exército Brasileiro. Conforme mostrado na Figura 1.6, o VBTP-MR é um veículo blindado desenvolvido em conjunto pela Fiat-IVECO e Exército Brasileiro. A principal característica desta nova família deve ser o seu design modular, permitindo a incorporação de diferentes torres, armas, sensores e sistemas de comunicação, incluindo uma versão de comunicações, uma versão ambulância e diferentes versões de apoio de fogo, armados com morteiros de grosso calibre e sistemas de armas.

Figura 1.4 – Operações militares brasileiras em áreas urbanas.

Figura 1.5 – TerraMax ™: 2007 DARPA Urban Challenge.

Figura 1.6 –VBTP-MR: um blindado 6x6 de transporte de pessoal desenvolvido em conjunto pela Fiat-IVECO e Exército Brasileiro (VBTP-MR Project, 2011).

No entanto, veículos robóticos são importantes não só para aplicações militares. Para o cidadão comum, os *Advanced Driver Assistance Systems* (ADAS) também tem grande importância. Impulsionados pelo elevado número de veículos em todo o mundo, a indústria e o fato de que, a cada ano, só na Europa, mais de 40.000 mortes e 1,4 milhões de lesões são causadas por acidentes com veículos, o ADAS surgiu para ajudar o motorista em sua tarefa de condução. Por exemplo, em função das mais complexas situações de tráfego, frações de segundo podem ser decisivas para se evitar um acidente de trânsito, onde um dos principais fatores é o tempo de reação humana, incluindo a fadiga, distrações e a possibilidade de adormecer ao volante (Gietelink, et al., 2006), (Rodríguez Flórez, 2010).

Figura 1.7 – *Advanced Driver Assistance Systems* (ADAS).

Incluindo também a robótica móvel na agricultura, veículos de mineração, etc, estes eixos de pesquisa, portanto, abrangem todos os aspectos da ciência e tecnologia relacionadas com a concepção de sistemas robóticos, sendo um esforço multidisciplinar que requer a entrada de engenheiros e cientistas de uma variedade de campos. Estes eixos podem ser classificados em quatro principais áreas (Chatila, 2010): percepção do ambiente; movimento e ação; antecipação e escolha; comunicação e interação. Esta tese explora aspectos da percepção do ambiente e pretende apresentar resultados comuns aos sistemas semi- e autônomos.

1.2 Problemática e Objetivos

Robôs autônomos têm motivado pesquisadores de diferentes grupos por causa do desafio que representa. Estes sistemas podem realizar tarefas em ambientes desestruturados sem a orientação humana contínua, e têm sido amplamente utilizados na indústria, em conflitos militares e em aplicações civis.

De uma maneira geral, um sistema autônomo deve ser capaz de gerenciar uma estrutura mecânica, utilizando de informações recebidas dos sensores. Como um exemplo bem-sucedido, a partir de algumas diretrizes do governo dos EUA, há alguns anos, o DARPA tem promovido excelentes pesquisas na área de navegação autônoma.
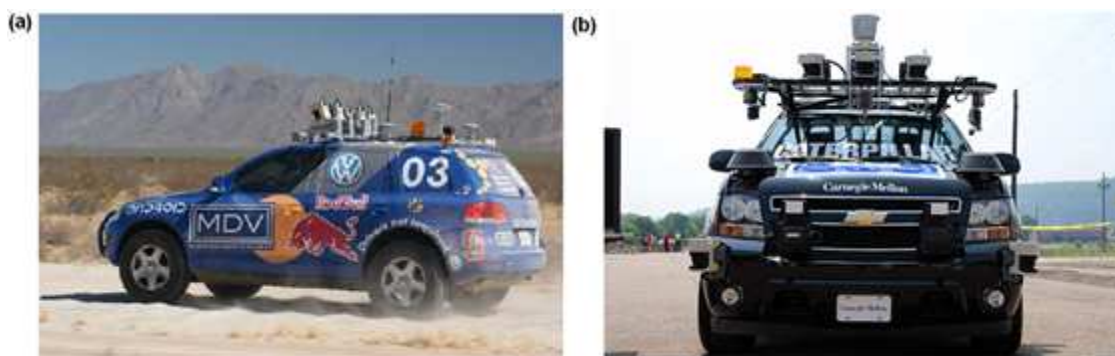
Figura 1.8 – (a) *Grand Challenge*: Stanford Racing. (b) *Urban Challenge*: Tartan Racing.

Desde 2002, o DARPA tem incentivado universidades, faculdades e empresas dentro e fora os EUA para desenvolverem veículos autônomos. Dois grandes desafios (*Grand Challenges*) foram organizados. Na primeira vez em 2004, nenhuma das equipes conseguiu concluir com sucesso o percurso criado para a competição. Já em 2005, cinco equipes completaram o desafio, e o primeiro lugar foi para a equipe da Universidade de Stanford. Stanford Racing cruzou o deserto de Mojave a uma velocidade média de 30.7 km/h. Mais recentemente, em Novembro de 2007, um desafio urbano (*Urban Challenge*) foi caracterizado por veículos autônomos que cumpriram missões em uma falsa área urbana. Entre 35 semi-finalistas, e onze finalistas, o primeiro lugar ficou com a equipe da Universidade Carnegie Mellon, Tartan Racing.

Além disso, nos últimos anos, houve também um aumento significativo em pesquisas envolvendo o desenvolvimento de veículos inteligentes, mas com foco em usuários humanos. Pesquisas sobre fatores humanos foram fundidas com a pesquisa na área de veículos inteligentes, com o objetivo de criar uma nova geração de ADAS. De acordo com Rodríguez Flórez (2010), estas pesquisas defendem a implantação de sistemas de segurança ativa do carro, e em contraste com os veículos autônomos, as funcionalidades dos sistemas ADAS são principalmente orientadas para a compreensão da percepção e visam apoiar e alertar o condutor em situações perigosas presumidas.

Com vista aos diferentes cenários e contextos de operações, é importante notar que em muitas ocasiões, o ser humano não pode ser deixado de lado, especialmente em determinadas operações militares. No entanto, entre outras tarefas específicas, esses sistemas devem monitorar

continuamente o caminho a frente do veículo, provendo informações apropriadas da estrada em tempo real e identificando obstáculos, a fim de evitar a colisão. Em ambientes externos, o cenário é dinâmico, com vários elementos em movimento, o que envolve a realização de operações complexas e não-deterministas. Desta forma, a exploração da área navegável em tempo real é uma forte ferramenta para avanços tecnológicos.

Em uma visão global, e como mostrado na Figura 1.9, os sistemas semi- e autônomos possuem camadas principais: Percepção, Navegação e Controle. De acordo com Thrun, et al. (2006), estes sistemas podem ser organizados em seis principais grupos funcionais: Interface com Sensores, Percepção, Controle, Interface com o Veículos, e Interface com o Usuário. De acordo com TEAM OSHKOSH (2011), este tipo de arquitetura facilita a separação de funções e desenvolvimento, permitindo a definição de interfaces entre os diferentes subsistemas.
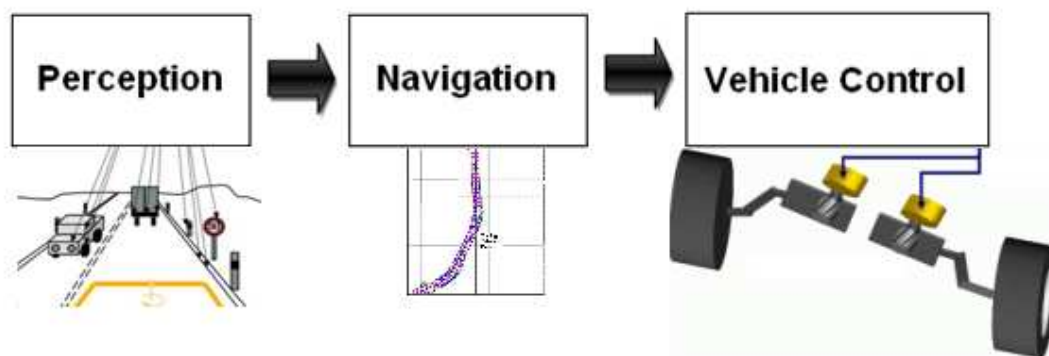
Figura 1.9 – Principais camadas dos sistemas semi- e autônomos.

A Percepção do ambiente é, então, uma questão importante para sistemas semi- e autônomos. Sistemas de percepção usam diferentes tipos de sensores, incluindo sensores ultra-sônicos, lasers, radar, câmeras, etc. Sistemas de visão tem sido projetados para investigar as informações estrada (Bonin-Fonte, et al., 2008). A Camada de Percepção pode receber dados provenientes de sistemas de visão. Obstáculos identificados pelos sistemas de visão são fundidos para melhorar a precisão da descrição do ambiente percebido. Posteriormente, as camadas de

Navegação e Controle utilizam as atualizações publicadas pela da camada de Percepção para determinar a próxima ação do veículo.

Câmeras são sensores passivos que são cada vez mais usados nos sistemas ADAS, por exemplo, no monitoramento de ponto cego, gerando alertas de mudança de faixa, nos sinais de limite de velocidade e no reconhecimento de pedestres. O amplo espectro funcional oferecido por este tipo de sensor o faz bastante atraente, pois fornece informações ricas e adicionais, substituindo sensores embarcados ou fornecendo redundância em aplicações de segurança (Rodríguez Flórez, 2010). Sistemas de visão monoculares são preferidos aos sistemas estéreo, visto que possuem vantagens em termos de redução de custos e a facilidade com que podem ser montados em veículos (Yamaguchi, et al., 2008). Além disso, o objetivo em termos de custo para preencher funções ADAS tem que ser muito menor do que os atuais, *Adaptive Cruise Control*, 500 Euros (Comissão Europeia, 2010).

Outro aspecto importante é determinar as velocidades relativas de um veículo com respeito a um outro, a fim de determinar o risco de colisão. Em geral, sensores como lidar[13], radar ou sonar, medem a distância e a direção de um obstáculo. De acordo com Rodríguez Flórez (2010), lasers constituem uma promissora tecnologia que concedem acesso a medições precisas e incluem a geometria da cena de maneira confiável inclusive em condições de mau tempo. A interação entre lidar e câmeras representa algo favorável para a melhoria dos existentes sistemas lidar baseados em rastreamento de múltiplos alvos.

Concernente ao trabalho desenvolvido nesta tese, uma abstração de uma camada de Percepção Visual baseada em Visão Monocular é apresentada na Figura 1.10. Esta proposta aborda o problema de desvio de obstáculos para veículos em ambientes dinâmicos e desconhecidos. Para entender melhor, após a aquisição da imagem por um dispositivo físico, tipo câmera, esta camada cumpre seu papel através das três subcamadas principais. Globalmente, as saídas das camadas (a) e (b) servem como entrada para um sistema de navegação semi- ou autônomo representado nas camadas (c) e (d). Por exemplo, pode ser usado em tarefas de estabilização dinâmica e controle de veículos. Esta abstração propõe um conjunto de ferramentas que monitoram continuamente o caminho a seguir, provendo as informaçes da estrada em tempo

---

[13] LIDAR (Light Detection And Ranging) is an optical remote sensing technology that can measure the distance to, or other properties of a target by illuminating the target with light, often using pulses from a laser.

real baseado num sistema de visão monocular. Detalhes sobre a troca de mensagens entre as camadas serão apresentados nos próximos capítulos. Como um resumo, temos:

- ✓ Camada Observador do Ambiente - *Environment observer* (a): A fim de reduzir o tempo de processamento, um critério automático da descarte de informações redundantes foi proposto em Miranda Neto, et al. (2007) e em Miranda Neto, et al. (2008). Este método melhora o desempenho de um sistema em tempo real, selecionando, de forma automática, quais imagens devem ser descartadas e quais devem ser tratadas pelo sistema de percepção visual. Portanto, levando-se em conta a coerência temporal entre consecutivas imagens, uma nova metodologia de gerenciamento dinâmico de energia (*Dynamic Power Management* - DPM) foi proposta por Miranda Neto, et al. (2011) e aplicado a um sistema de percepção visual robótico, que incluiu um método de Gerenciamento do Impacto Culmulativo de Dados (*Cumulative Impact Data Management*). Além disso, baseado no método de Pearson (Pearson, 1895), esta proposta também inclui um sistema de percepção em tempo real capaz de estimar o risco de colisão em ambientes dinâmicos e desconhecidos, usando um sistema de visão monocular.
- ✓ Camada de Deteção da Área Navegável - *Navigable Area Detection* (b): Um método robusto de deteção da região do horizonte proposto por Miranda Neto, et al. (2011) gera a região de interesse a partir de um método de limiarização dinâmico. Este método permite investigar dinamicamente apenas a porção da imagem a frente do veículo, a fim da deteção da estrada e evitamento de obstáculos.
- ✓ Camadas de Navegação Autônoma e Comandos - *Autonomous navigation* (c) *and Commands* (d): Não dizem respeito a esta tese.

1.2.1 Oranização do Manuscrito

Esta tese contém uma descrição detalhada de métodos, conceitos teóricos, observações experimentais, resultados e conclusões desta pesquisa. Cada capítulo constitui um passo para a implementação de um conjunto de ferramentas que monitoram continuamente o caminho a frente

do veículo baseado num sistea de visão monocular, provendo informações adequadas em tempo real. Os capítulos seguintes são:

- ✓ Capítulo 2: Coeficiente de Correlação de Pearson: Aplicações em Robótica Autônoma;
- ✓ Capítulo 3: Remoção do Horizonte e Deteção da Região Navegável;
- ✓ Capítulo 4: Avaliação Experimental;
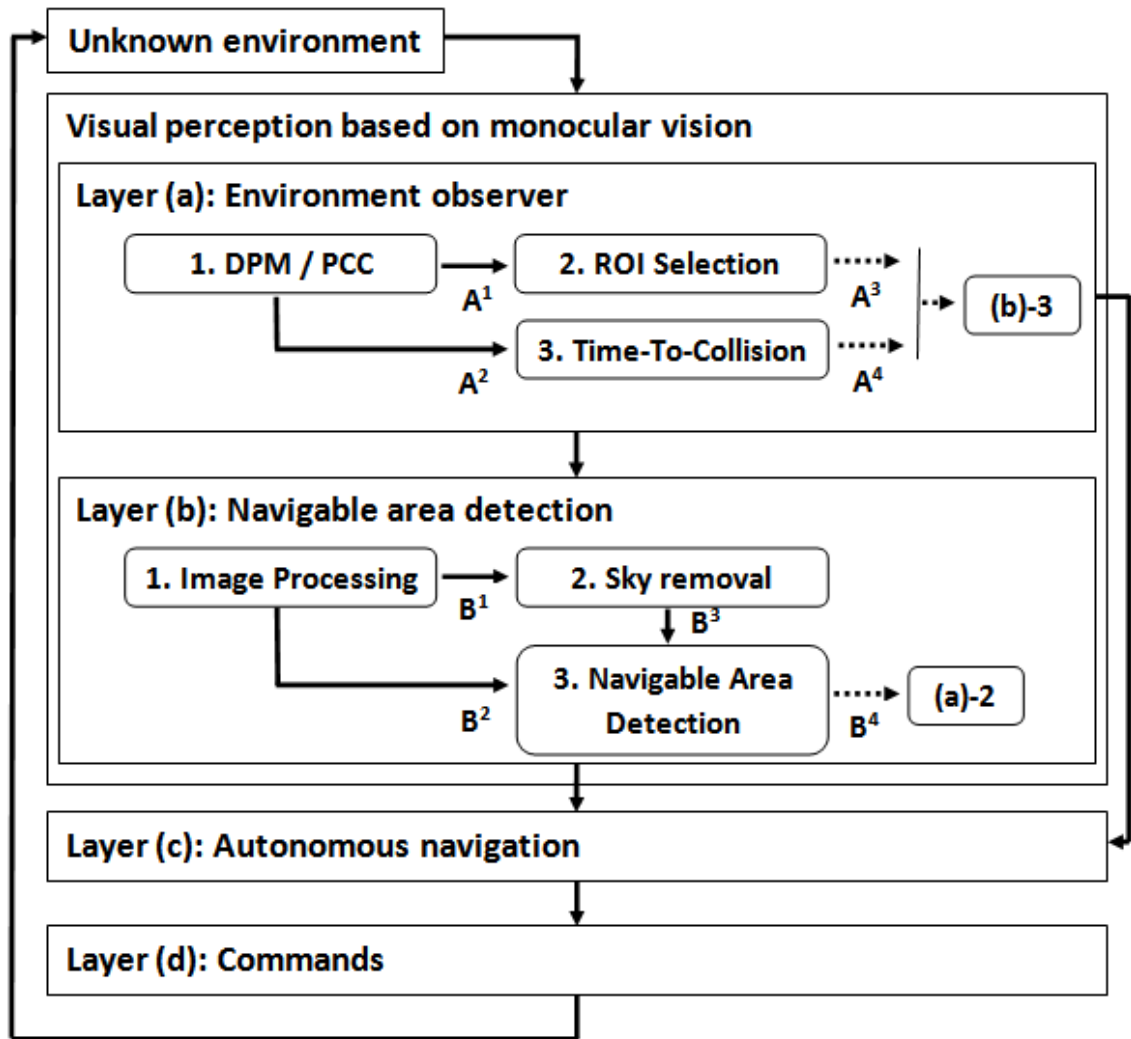
Figura 1.10 – Camada de Percepção Visual baseada em Visão Monocular.

Chapter 1

Introduction

Contents

1.1 Introduction

Historically, the development of new technologies has been driven by military applications. However, dual-use technology is a major issue as well. Robots have many valuable attributes that aid and complement civil and military tasks. Most importantly, they can replace humans in dangerous tasks, reducing casualties. They will play a vital role in the future.

In the last three decades, visual navigation for mobile robots or unmanned vehicles has become a source of countless research contributions (Bonin-Font, et al., 2008). Several applications for control of semi- and autonomous platforms are being developed. The challenge of constructing robust methods, and, in most cases, real-time systems, is far from being achieved. This can be observed by the great number of researches being published in the last few years.

For land vehicles in unknown environments, research has been driven in four fronts with different goals. The first of them is related to space exploration. The second front, with military

purposes, has had support and encouragement of the U.S. agency known as DARPA[14], whose main mission is to provide the armed forces with technology innovative enough to maintain its technological superiority. A subdivision of the previous area concerns the application of mobile robotics in agriculture. The fourth front has had support by the automotive industry, for example, the Advanced Driver Assistance Systems (ADAS).

For military or civil purposes, some of these applications include: the Grand Challenge (Thrun, et al., 2006) and Urban Challenge (Rojo, et al., 2007). The goal of the Grand Challenge was the development of an autonomous robot capable of traversing unrehearsed off-road terrain. For the Urban Challenge, the goal of the system was driving a car, autonomously, in a city environment, with way crossings and static and dynamic obstacles. On the other hand, mobile robotics in agriculture involves multidisciplinary areas, like inertial navigation and patterns recognition by aerial and terrestrial sensors (Bakker, et al., 2008), (Subramanian, et al., 2006). Finally, driven by the high number of vehicles in all the world, the ADAS emerged to help drivers in its driver task (Gietelink, et al., 2006), (Rodríguez Flórez, 2010).

On another front, the aerial robots also offer great perspectives in many applications: search and rescue, outdoor and indoor building inspection, real-time monitoring, high risk aerial missions, mapping, fire detection or cinema recording (Kim, et al., 2003), (Bonin-Font, et al., 2008). Similarly, the development of Unmanned Aerial Vehicles (UAVs) has been of interest for military applications. These vehicles are primarily used for surveillance and reconnaissance. However, one limitation is their maximum flight time, therefore they cannot carry large fuel payloads (Anton, et al., 2008). Future exploration of Mars, laid out by the Vision for Space Exploration, requires long-endurance UAVs that use resources that are plentiful on Mars (Klesh, et al., 2009), (NASA, 2004). Including reactive intelligence based on an autonomic sense-act modality and sensed by optical image processing, the purpose of the Energetically Autonomous Tactical Robot (EATR) ™ project is to develop an autonomous robotic platform able to perform long-range, long-endurance missions without the need for manual or conventional re-fueling (DARPA, 2009), (Finkelstein, 2010).

---

[14] The Defense Advanced Research Projects Agency (DARPA) is an agency of the United States Department of Defense responsible for the development of new technology for military.

In all these cases previously reported, an important factor is the variety and complexity of environments and situations. These intelligent platform developments have a common issue: providing to the platform the capability of perceiving and interacting with its neighbour environment. This capability, known as environment perception, is still in evolution. Part of this, machine vision appears as an important area due to the richness of information it deals with one of the possibilities is to use vision systems to identify the region that includes free-navigable area in images. It is one important task of autonomous terrestrial vehicle guidance.

In this way, this thesis addresses the problem of obstacle avoidance for semi- and autonomous platforms in dynamic and unknown environments. Based on a monocular vision, it proposes a set of tools that continuously monitors the way forward, proving appropriate road informations in real time. This system was tested in different types of image texture (road surfaces), which include free-area detection, reactive navigation and collision risk estimation. In order to reduce processing time, this work also proposes an automatic image discarding criteria, which includes a new environment observer method to optimize energy consumption used by a visual machine. This last proposal may be extended to other types of systems and sensors. A remarkable characteristic of this methodology is its independence of the image acquiring system and of the robot itself. This real-time perception system has been evaluated from real data obtained by two intelligent platforms.

1.1.1 Government, Institutional and Thesis Contexts

Robotics is one of the main vectors of innovation. Especially regarding the development of applications related to Defense systems, these applications are part of an important and strategic context. They can provide a revolutionary leap ahead in military capabilities, increasing combat effectiveness. Particularly, robot navigation can be subdivided into two classes: previously known environment and unknown environment. The later class has been of great interest to the military community.

For land vehicles in unknown environments it has been driven in some fronts, among them, with military purposes; as shown in Figure 1.1, it[15] has had support and encouragement of the U. S. Department of Defense. One of the goals expressed in its vision is the transformation of one third of the fleet of military vehicles in autonomous vehicles for general use, even in densely populated urban areas.

Figure 1.1 – Carnegie Mellon University's Crusher: an Unmanned Ground Combat Vehicle.

In a context of renewed requirements and sometimes contradiction between short and long-term interests, the French Strategic Plan for Research and Technology (PS R&T[16]) define the priorities of research in the *Plan prospectif à 30 ans* (PP30). This 30-Year Plan prepares and recommends the choices to be made to procure and maintain operational capabilities for the French forces now and in the future. For land systems, the main challenge will remain access to innovative products (robotics) and also for more classical equipment either under development or requiring reactive adaptation. It can be categorized into three categories: dual-use technologies; specific military technologies; and a number of under-employed innovations from the civilian sector. The use of robotics in land operations should improve: the protection of soldiers by replacing them in dangerous situations, the productivity and consequently the availability of soldiers, by carrying out repetitive tasks in the place of soldiers, and finally, flexibility of use.

---

[15] Crusher is a 6.5-ton autonomous off-road Unmanned Ground Combat Vehicle developed by researchers at the Carnegie Mellon University's National Robotics Engineering Center for Defense Advanced Research Projects Agency (DARPA).

[16] Strategic Plan for Research & Technology in defense and security, *Direction Générale de l'Armement* (DGA), 2009. The DGA is the French Government Defense procurement agency responsible for the program management, development and purchase of weapon systems for the French military.

Figure 1.2 presents a family of miniature robotic platforms (MiniROC[17]) developed in France to operate in urban terrain.

Figure 1.2 – «*Mini-Robot de Choc*» (MiniROC) technology.

For the Brazilian Government, science, technology and innovation are also matters of state, beyond, governments and should be treated as commitments that move from one period to another government. The permanent mission of a State is to maintain the technological superiority, developing technologies that private companies cannot reach, cost-effectively, in the short to medium term[18,19]. The Brazilian Ministry of Science and Technology (MCT), in line with the Federal Government Plan for Accelerated Growth, has made its Action Plain for 2007-2010. In this way, the national strategy of development is inseparable of the national defense strategy, and the Brazilian State has sought partnerships with other countries, in order to develop the technological capability and manufacturing of defense products.

In accordance with the basic research areas of interest of the Brazilian Ministry of Defense, and interested in vehicular security, from 2008 an Autonomous Mobility Laboratory

---

[17] "Mini-Robot de Choc": A family of miniature robotic platforms developed in France is designed to operate in urban terrain.
[18] National Defense Strategy, Ministry of Defence, Brazil, 2nd edition.
[19] Action Plan 2007-2010, Ministry of Science and Technology, Brazil.

(LMA) has been formed at Computational Mechanics Department of the University of Campinas. This laboratory intends to study the semi- and autonomous systems, essential for civil and military applications. Currently, the LMA has a multidisciplinary group and consists of special members from different institutions, as the University of Campinas, University of Sao Paulo and Military Institute of Engineering.

Aiming to understand the joint efforts of research, scientific and technological development, and concerning the work of international cooperation, the first activities between members of the LMA and the Heudiasyc laboratory UMR 6599 CNRS/UTC, France, began in mid-2007, which include this doctoral program in cooperation.

During the years of this thesis program, the author developed several works to the LMA establishment, including institutional relationships. Besides co-founder of LMA, the author is also the mentor of the project: Robot Platform Development for Studies of the (Semi)-Autonomous Terrestrial Mobility, 2008 Pro-Defense[20], Brazilian Ministry of Defense, CAPES[21].

This PhD thesis was carried out at the Computational Mechanics Department, DMC/FEM, University of Campinas, Brazil, from February 2007 to July 2009, and at the Heudiasyc laboratory UMR 6599 CNRS/UTC in the University of Technology of Compiegne, France, from August 2009 to August 2011. It was financed with a grant from the CNPq-Brazil[22] and from the ANR-France[23] / PERCOIVE Project[24].

---

[20] Support Program for Education and Scientific Research and Technology in National Defense. It was a partnership between the Brazilian Ministry of Defense and CAPES. Its goal is to finance and stimulate the production of scientific and technological research and training of post-graduates in the area of Defense.
[21] Coordination for the Improvement of Higher Education Personnel, CAPES, Brazil.
[22] National Council for Scientific and Technological Development (CNPq), Brazil.
[23] French National Research Agency, France.
[24] The problem addressed in PERCOIVE project involves increasing the safety on board and around of a road vehicles group, through observation and estimation of parameters of vehicle dynamics in interaction with the environment nearest neighbors and vehicles. Specifically, it aims to study and propose solutions to the environment perception (exteroceptive), for the estimation of vehicle dynamics and road attributes in a context estimation cooperative multi vehicle.

1.1.2 Military Robots and Vehicles in Urban Area

According to Chatila (2010), the robots enter in the military context especially when it is necessary to reduce human exposure to hazardous situations, mainly in a hostile environment, where a robot performs a specialized mission. Many of these robotic missions can be observed: mine clearance; field reconnaissance, especially nuclear, radiological, bacteriological, chemical, or, as shown in Figure 1.3, the IEDs[25].

Figure 1.3 – An IED explosion on a road during a military convoy.

A major difficulty in replacing human on any task is the fact that it is not yet possible to design a generic robot adapted to any demand, which includes a very large variety of design options, robot size, its driving force, instrumentation, expected performance (including decision-autonomy, "intelligence", etc).

Whether in military or police missions, there may be great hostility enemy. As an example, land incursions in unknown environments or in area of predominance enemy. In these cases, the semi- and autonomous systems can contribute to the success of these missions, saving lives. Recent conflicts show an emergent use of large military vehicles in urban areas. The complexity of military operations is also growing, because in many cases, these areas have difficult access and are very populous. Consequently, it increases significantly the driver responsibility. Figure 1.4 shows the Brazilian military operations in urban areas.

---

[25] Improvised Explosive Device (IED), also known as a roadside bomb, is constructed and deployed in ways other than in conventional military action.

In a context of large vehicles, in 2007 challenges of DARPA, the Team Oshkosh adapted a truck, originally designed for combat and used by the U.S. Marine Corps. The vehicle known as TerraMax ™ can be seen in Figure 1.5 (TEAM OSHKOSH, 2011).

In the same way, the New Family of Armored Fighting Vehicles is a project under the direction of the Brazilian Army. As shown in Figure 1.6, the VBTP-MR is an armored personnel carrier jointly developed by Fiat's IVECO and the Brazilian Army. The main feature of this new family should be its modular design, allowing the incorporation of different turrets, weapons, sensors and communications systems, including a communications version, an ambulance version and different fire support versions, armed with large caliber mortar and gun systems.

Figure 1.4 – The Brazilian Army operations.

Figure 1.5 – TerraMax ™: 2007 DARPA Urban Challenge.

Figure 1.6 – The VBTP-MR: a 6x6 armored personnel carrier jointly developed by Fiat's IVECO and the Brazilian Army (VBTP-MR Project, 2011).

But robotic vehicles are important not for military application only, for ordinary citizens, the Advanced Driver Assistance Systems (ADAS) has also great importance. Driven by the high number of vehicles in all the world, the industry and the fact that, every year in Europe alone, more than 40,000 casualties and 1.4 million injuries are caused by vehicle-related accidents, the ADAS emerged to help the driver in its driver task. For example, because the more complex traffic situations, fractions of a second can be decisive to avoid traffic accidents, where one of the main factors is the time for human reaction, in addition to fatigue, distractions and the ability to fall asleep at the wheel (Gietelink, et al., 2006), (Rodríguez Flórez, 2010).

Figure 1.7 – Advanced Driver Assistance Systems (ADAS).

Including also mobile robotics in agriculture, mining vehicles, etc, these axes of research, therefore, cover all aspects of science and technology related to the design of robotic systems, which is a multidisciplinary effort requiring input from engineers and scientist in a variety of

27

fields. They can be ranking into four main features (Chatila, 2010): environment perception; movement and action; anticipation and choice; communication and interaction. This thesis explores aspects of environment perception and intends to present common results to semi- and autonomous systems.

1.2 Issues and Objectives

Autonomous robots have motivated researchers from different groups because of the challenge that it represents. These systems can perform desired tasks in unstructured environments without continuous human guidance. They have been used extensively in industry, military conflicts and in civilian applications.

In a general way, autonomous systems must be capable to manage the mechanical structure, using the information received from sensors. As a successful example, from some guidelines of the U.S. government, a few years ago, the DARPA has been promoting excellent research in the area of autonomous navigation.
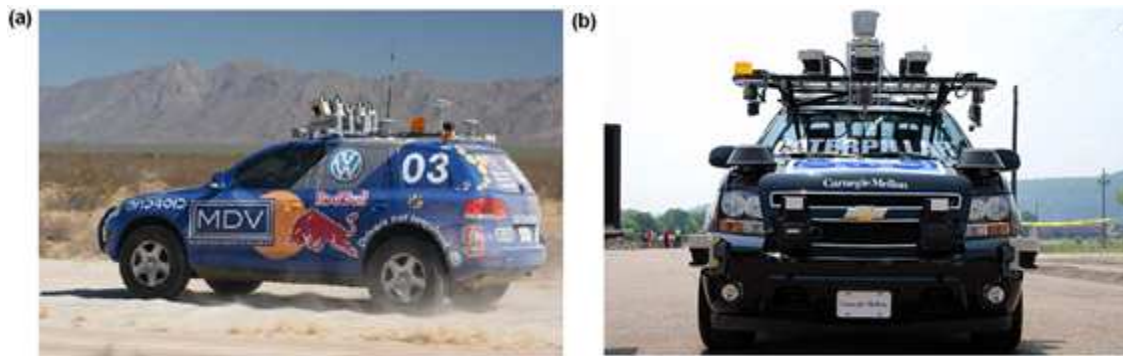
Figure 1.8 – (a) Grand Challenge: Stanford Racing. (b) Urban Challenge: Tartan Racing.

Since 2002, DARPA has been encouraging universities, colleges and businesses in and outside the U.S. to develop autonomous vehicles. Two Grand Challenges were organized. In the

first time in 2004, none of the teams managed to successfully completing the course set up for the competition. Already in 2005, five teams completed the challenge, and the first place was won by the Stanford University's team. Stanford Racing crossed the Mojave Desert with an average speed of 30.7 km/h. More recently, in November 2007, the last Urban Challenge was characterized by autonomous vehicles that managed their missions in a false urban area. Among thirty-five semi-finalists, eleven finalists, the first place was won by the Carnegie Mellon University's team, Tartan Racing.

Moreover, in recent years, there has also been a large increase research on the development of intelligent vehicles, but with a focus on human users. Research on human-factors has been merged with the investigation of intelligent vehicles, aiming to create a new generation of ADAS. According to Rodríguez Flórez (2010), it stands for the deployment of active car safety systems, and in contrast to autonomous vehicles, ADAS functionalities are mainly oriented towards proprio- and extero- perception understanding tasks which support and warn the driver by emitting alarms in presumed dangerous situations.

Regarding the different scenarios and contexts of operations, it's important to know that in many occasions the human being can not be cast aside, especially in some military operations. Nevertheless, among other specific tasks, these systems must continuously monitor the way forward, proving appropriate real-time road informations and identifying obstacles in order to avoid collision. In outdoor environments, the scenario is dynamic with several elements in motion, which involves carrying out complex, and non-deterministic operations. Then, the navigable area exploration in real time is a powerful driver for technological advances.

In a global view, as shown in Figure 1.9, the semi- and autonomous systems have major layers: Perception, Navigation and Vehicle Control. According to Thrun, et al. (2006), it can be divided into six main functional groups: Sensor Interface, Perception, Control, Vehicle Interface, and User Interface. According to TEAM OSHKOSH (2011), this type of architecture facilitates the separation of functions and development, and allows the definition of the interfaces between subsystems.

Environment perception is then a major issue in semi- and autonomous systems. The perception systems use many types of sensors, including ultrasonic sensors, laser rangefinders,

radar, cameras, etc. Machine vision has been designed to investigate the road informations (Bonin-Font, et al., 2008). Perception layer can receive data coming from vision systems. Obstacles identified by the vision systems are merged to improve the accuracy of the description of the perceived environment. Subsequently, the Navigation and Vehicle Control layers use the updates published by the Perception layer to determine the next action of the vehicle.
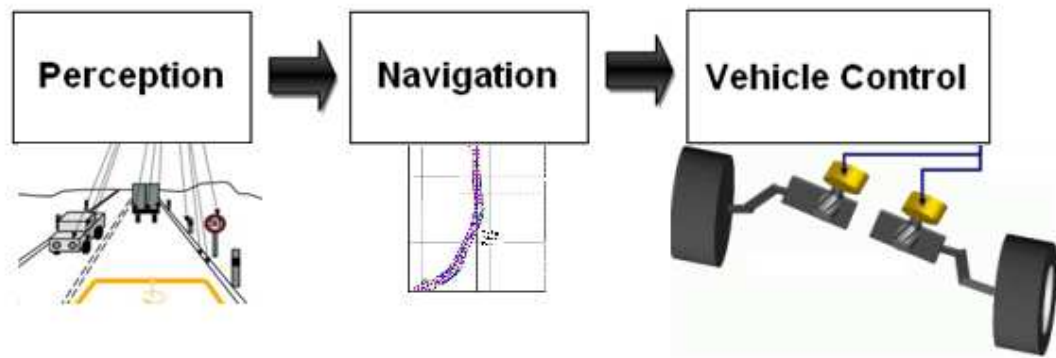
Figure 1.9 – Major layers of the semi- and autonomous systems.

Cameras are passive sensors which are more and more used in ADAS, like blind spot monitoring, lane departure warning, speed limit signals and pedestrian recognition. The large functional spectrum offered by this sensor makes it quite attractive, since it provides rich and additional information, replacing on-board sensors or provide redundancy in safety applications (Rodríguez Flórez, 2010). Monocular camera systems are preferred to stereo camera systems because monocular systems have advantages in terms of reduced costs and the ease with which they can be fitted to vehicles (Yamaguchi, et al., 2008). Moreover, the objective in terms of cost to fill ADAS functions has to be very lower than the current, Adaptive Cruise Control, 500 Euros (European Commission, 2010).

Another important aspect is to determine the relative speeds of one vehicle with respect to another one and to determine the risk of collision. In general, sensors like lidar[26], radar or sonar

---

[26] LIDAR (Light Detection And Ranging) is an optical remote sensing technology that can measure the distance to, or other properties of a target by illuminating the target with light, often using pulses from a laser.

measure the distance and direction to the obstacle. According to Rodríguez Flórez (2010), lasers constitute a promising sensor technology which grants access to accurate measurements of the scene geometry, reliable in poor weather conditions. The interaction between lidar and cameras seems to be favorable for the improvement of existing lidar-based multiple target tracking applications.

Concerning the work in this thesis, an abstraction of a Visual-Perception Layer based on Monocular Vision is presented in Figure 1.10. This proposal addresses the problem of obstacle avoidance for vehicles in dynamic and unknown environments. To better understand, after the image acquisition by a physical device, this layer fulfills its role through by the three main sublayers. Globally, the outputs of the layers (a) and (b) serve as input for a semi – or autonomous navigation systems represented by the layers (c) and (d). For example, it can be used in tasks as vehicle dynamics stabilization and control. This abstraction proposes a set of tools that continuously monitors the way forward, proving appropriate road informations in real time based on monocular vision. Details about the exchange of messages between the layers will be explained in the next chapters. However, as a resume:

- ✓ Layer (a) Environment observer: In order to reduce processing time, an automatic image discarding criteria was proposed by Miranda Neto, et al. (2007) and Miranda Neto, et al. (2008). It improves the performance of a real-time system by choosing, in an automatic way, which images should be discarded and which ones should be treated at the visual perception system. Therefore, taking into account the temporal coherence between consecutive frames, a new Dynamic Power Management (DPM) methodology was proposed by Miranda Neto, et al. (2011) and applied to a robotic visual machine perception, which includes a Cumulative Impact Data Management. Additionally, based on Pearson's method (Pearson, 1895), it also includes a real-time perception system capable of estimating the collision risk in dynamic and unknown environments by using a single monocular system.
- ✓ Layer (b) Navigable Area Detection: A robust horizon finding algorithm (Miranda Neto, et al., 2011) then generates the region of interest from a dynamic threshold

search method. It permits to investigate dynamically only a small portion of the image ahead of the vehicle, in order to road and obstacle detection.

✓ Layers (c) Autonomous navigation and (d) Commands: They are not concerning to this thesis.
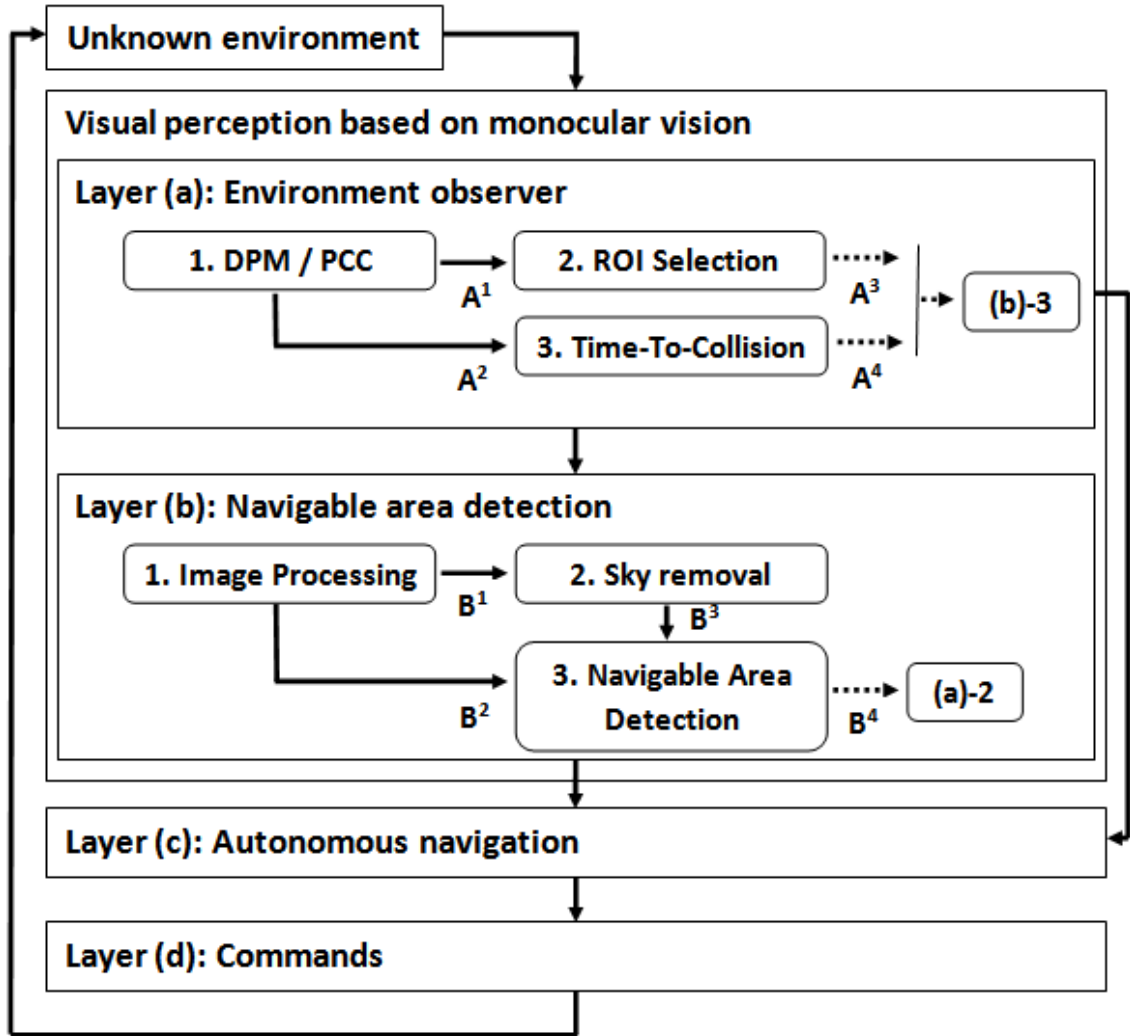
Figure 1.10 – The Visual-Perception Layer based on Monocular Vision.

1.2.1 Outline of the manuscript

This thesis contains a detailed description of the methods, theoretical concepts,

experimental observations, results and conclusions of this research. Every chapter constitutes a small step towards the implementation of a set of tools that continuously monitors the way forward based on monocular vision, proving appropriate road informations in real time. The following chapters are:

- ✓ Chapter 2: Pearson's Correlation Coefficient: Applications on Autonomous Robotics;
- ✓ Chapter 3: Sky Removal and Navigable Area Detection;
- ✓ Chapter 4: Experimental Evaluation;

Chapter 2

Pearson's Correlation Coefficient: Applications on Autonomous Robotics

Contents

This chapter contains a detailed description of the methods, theoretical concepts and experimental observations mainly based on two classical methods: Pearson Coefficient Correlation (Pearson, 1895) and Otsu Thresholding Method (Otsu, 1978). These methods are respectively presented in the Appendix A and B. More results and conclusions about this chapter will be presented in the Chapter 4.

2.1 Introduction

In the last three decades, visual navigation for mobile robots or unmanned vehicles has become a source of countless research contributions (Bonin-Font, et al., 2008). Several applications for control of semi- and autonomous platforms are being developed. The challenge to construct robust methods, and, in most cases, real-time systems, remains an open problem, and it is far from being achieved. This can be observed by the great number of researches being published in the last few years.

Some of these applications include: the aerial robots that offer great perspectives in many applications as search and rescue, real-time monitoring, high risk aerial missions, mapping, etc (Bonin-Font, et al., 2008), (Kim, et al., 2003). Similarly, the development of Unmanned Aerial Vehicles (UAVs) has been of interest for military applications, however, one limitation is their maximum flight time; therefore they cannot carry large fuel payloads (Anton, et al., 2008). Future exploration of Mars requires long-endurance UAVs that use resources that are plentiful on Mars (Klesh, et al., 2009), (NASA, 2004), (DARPA, 2009), (R. Finkelstein, 2010). Finally, for military or civil purposes, vehicular applications (Thrun, et al., 2006), (Berlin Team, 2007), (Gietelink, et al., 2006) have as objective the development of autonomous and semi-autonomous systems capable of traversing unrehearsed and off-road terrain, driving a car autonomously in an urban environment and also to help the driver in its driver task.

However, an important factor is the variety and complexity of environments and situations. These real-time intelligent platform developments have a common issue: providing to

the platform the capability of perceiving and interacting with its neighbour environment, managing power consumption, CPU usage, etc.

Primary interest in this work, the perceiving capability, known as environment perception, is still in evolution. Part of this, machine vision is an important tool that continuously monitors the way forward, proving appropriate road informations in real time. Although extremely complex and highly demanding, thanks to the great deal of information it can deliver, machine vision is a powerful means for sensing the environment and it has been widely employed to deal with a large number of tasks in the automotive field (Bertozzi, et al., 2000). However, it can lead to some losses due to the processing time. If the decision for a more complex machine vision system may increase the autonomy and "intelligence" degrees, especially in relation to navigation in unknown environments, on the other hand, it can lead to an excessively slow system for an independent real time application.

These problems of time-dependent and dynamic resource allocation have manifested themselves under different names in application domains such as embedded systems and digital circuits, which include energy and memory consumption for the embedded systems (Bouyer, 2010). It has been a topic of interest in the automotive industry (Obayashi, et al., 2004), (Morita and Shimamura, 2003), (Deshmukh, et al., 2010).

In 1885, an empirical and theoretical development that defined regression and correlation as statistical topics were presented by Sir Francis Galton (Rodgers and Nicewander, 1988). In 1895, Karl Pearson published the Pearson's Correlation Coefficient (PCC) (Pearson, 1895). The Pearson's method is widely used in statistical analysis, pattern recognition and image processing (Eugene and Johnston, 1996).

Based on Pearson's method, Miranda Neto, et al. (2007) proposed a visual-perception system based on an automatic image discarding method as a simple solution to improve the performance of a real-time navigation system by exploiting the temporal coherence between consecutive frames. Viewing problems of time-dependent and dynamic resource allocation, Miranda Neto, et al. (2011) also presents the PCC as an environment observer method to save processor energy (power) consumption.

An abstraction of a Visual-Perception Layer based on Monocular Vision was presented in Figure 1.10. Besides the works mentioned above, the system presented here also takes place in the obstacle avoidance context for vehicles in dynamic and unknown environments, proposing two others methods based on PCC: (1) a collision risk estimation; and (2) an extension of the environment observer method that selects automatically only the regions-of-interest (ROI) in order to perform a task. The main focus in this chapter is to present the Layer (a) Environment observer. As a resume of the Layer (a):

- ✓ Layer (a)-1 DPM / PCC: Dynamic Power Management (DPM) is a design methodology for dynamically reconfiguring systems. In this way, an environment observer method based on PCC observes if there are no significant changes in the environment, permitting that some logical components may be shut down to save processor energy consumption, and/or to make the CPU available for running concurrent processes.
- ✓ Layer (a)-2 ROI Selection: Taking into account the temporal coherence between consecutive frames, the regions-of-interest (ROI) are automatically selected in order to perform a task, in this case, navigable area detection and obstacle avoidance.
- ✓ Layer (a)-3 Collision Risk Estimation: Based on the PCC variation, this method estimates the collision risk in dynamic and unknown environments by using a single monocular system.

The Pearson's method is presented in Appendix A. The Discarding Criteria method is presented in section 2.2. Section 2.3 introduces the Visual-Perception Layer based on Monocular Vision. Based on Pearson's Correlation Coefficient (PCC), the following sections are:

- ✓ Section 2.4: Dynamic Power Management;
- ✓ Section 2.5: Automatic Regions-Of-Interest Selection;
- ✓ Section 2.6: Collision Risk Estimation.

2.2 Discarding Criteria

The discarding criteria was presented by Miranda Neto, et al. (2007) as a simple solution to improve the performance of a real-time navigation system by choosing, in an automatic way, which images should be discarded and which ones should be treated at the visual perception system. It was a new approach for using the Pearson's Correlation Coefficient (PCC).

In Figure 2.1, basically, if the PCC indicates that there is a high correlation between a reference frame and another new frame acquired, the new frame is discarded without being processed (for example, the system can repeat a last valid command). Otherwise, the frame is processed and it is set as the new reference frame for the subsequent frame.

The inclusion of an automatic image discarding method leads to a reduction of the processing time. Although the system spends some milliseconds computing the PCC, it gains much more time, in some cases, discarding more than 90% of the images (Miranda Neto, et al., 2008). However, it is important to notice that this percentage is not dependent on the video sequence or image size, but on the obstacles / objects influence.
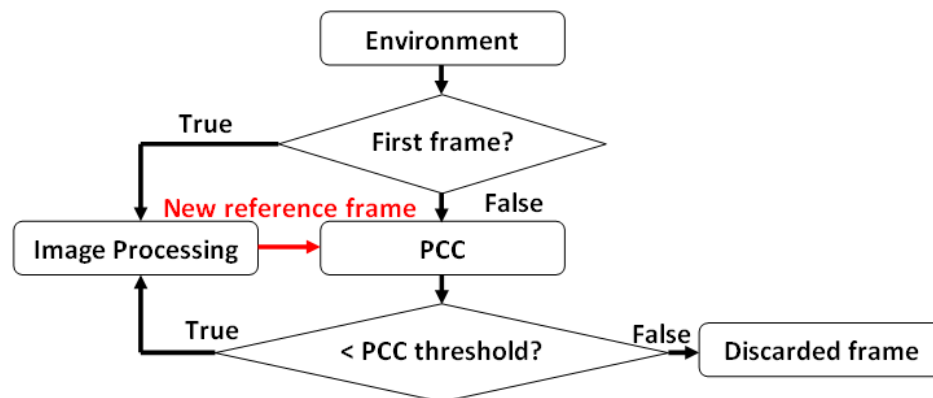
Figure 2.1 – Discarding criteria (Miranda Neto, et al., 2007).

2.3 Visual-Perception Layer Based on Monocular Vision

2.3.1 Sensor Perception Review

The perception of the environment is a major issue in autonomous vehicles. The perception layer uses many types of sensors (Thrun, et al., 2006), (Berlin Team, 2007), (Greco, 2008), including ultrasonic sensors, laser rangefinders, radar, cameras, etc, which in many cases may be limited in scope and subject to noise. These sensors are not perfect: ultrasonic sensors are cheap but suffer from specular reflections and are limited in range, and laser rangefinders and radar provide better resolution but are more complex and expensive (Ulrich, et al., 2000).

The vision-based sensors are defined as passive sensors. However, vision sensors are less robust than millimeter-wave radars in foggy, night, or direct sun-shine conditions (Bertozzi, et al., 2000). All range-based obstacle detection systems have difficulty for detecting small or flat objects on the ground, and range sensors are also unable to distinguish between different types of ground surfaces (Ulrich, et al., 2000). Notwithstanding, the main problem with the use of active sensors is represented by interference among sensors of the same type, hence, foreseeing a massive and widespread use of these sensing agents, the use of passive sensors obtains key advantages (Bertozzi, et al., 2000). For example, Figure 2.2 illustrates the monocular vision contribution to the DARPA Grand Challenge (Dahlkamp, et al., 2006). This diagram shows that the reach of lasers was approximately 22 meters (left), whereas the monocular vision module often looks 70 meters ahead (right).

On the safety front, the progressive safety systems will be developed through the manufacturing of an "intelligent bumper" peripheral to the vehicle in answering new features as: blind spot detection, frontal and lateral pre-crash, etc (European Commission, 2010).

Moreover, when incorporating several types of sensors, there is an increase of autonomy and "intelligence" degrees, especially in relation to navigation in unknown environments. In contrast, the type and quantity of sensors determine the volume of data for processing that requires, in most cases, a high computational cost. For unstructured environments, the scenario

for study is dynamic, with several elements in motion. Thus, running an autonomous or semi-autonomous system involves carrying out complex, and non-deterministic operations in real time.

Figure 2.2 – Monocular vision: comparison of the laser-based (left) and the image-based (right) mapper. Circles are spaced around the vehicle at 10 meters distance. It includes the sky removal process.

A real-time system must satisfy explicit response-time constraints, including failure. This system is one whose logical correctness is based on both the correctness of the outputs and their timeliness (Laplant, 2004). Moreover, there is a considerable complexity in the sense that correctness not only depends on the logical ordering of events of the systems, but also on the relative timing between them (Bouyer, 2010).

Aware that in the majority of the autonomous and semi-autonomous navigation systems, the machine-vision system is working together with other sensors, added to its low cost, this work uses a monocular vision-based sensor. Because it uses simple techniques and fast algorithms, the system is capable to achieve a good performance, where the compromise between processing time and images acquisition is fundamental.

2.3.2 Environment Observer Method

The machine vision (part of the embedded computers) is an important component of the set of sensors. Although extremely complex and highly demanding, thanks to the great deal of information it can deliver, it is a powerful means for sensing the environment and it has been widely employed to deal with a large number of tasks in the automotive field (Bertozzi, et al., 2000). However, complex machine vision systems can lead to some losses due to the processing time. Thinking about the existing relation between a real-time decision system and a camera acquisition system that operates in a specific acquisition rate, the decision for a more complex machine vision system possibly leads to an excessively slow system for an independent real time application. The great amount of information would not necessarily lead to better decisions and could also harm the performance of the system, overloading it.

Taking into account that it has been estimated that humans perceive visually about 90% of the environment information required for driving (Bertozzi, et al., 2000), it is not a bad idea to reduce information acquired by a vision system, in order to reduce processing time. But the definition of an automatic image discarding criteria, which leads to a minimum loss of information, may not be a trivial task for computational systems, specially real-time ones. Then, based on the idea to reduce information acquired and in order to reduce processing time, Miranda Neto, et al. (2007) propose an automatic image discarding criteria based on Pearson's Correlation Coefficient (PCC) (Pearson, 1895), a low complexity and easy implemented solution.

Furthermore, the fundamental premise for the applicability of Dynamic Power Management (DPM) is that systems experience nonuniform workloads during operation time. A second assumption is that it is possible to predict, with a certain degree of confidence, the fluctuations of workload (Benini, et al., 2000). In this case, a simple DPM method shuts down a component when it is idle and it is essentially a prediction problem (Yongguo Mei, et al., 2005). Additionally, in most real-world systems, there is little knowledge of future input events and DPM decisions have to be taken based on uncertain predictions. Thus, according to Benini, et al. (2000), the rationale in all predictive techniques is that of exploiting the correlation between the past history of the workload and its near future in order to make reliable predictions about future events. Moreover, workload observation and prediction should not consume significant energy.

41

2.4 Dynamic Power Management

2.4.1 Related Work

Autonomous robots can perform desired tasks in unstructured environments without continuous human guidance. These systems have some degree of self-sufficiency. Self-configuring, self-optimizing and self-protecting are still open questions. For advances in the energy autonomy, robots will need to extract energy from the environment. In many ways robots will face the same problems as animals (Deshmukh, et al., 2010). Beyond that an important variable is the system-state conditions in combination with environment events, because they may determine the system behavior.

In this way, a system must have knowledge of its available resources as well as its components, their desired performance characteristics and their current status. Dynamic Power Management (DPM) is a design methodology for dynamically reconfiguring systems to provide the requested services and performance levels with a minimum number of active components or a minimum load on such components. It encompasses a set of techniques that achieves energy-efficient computation by selectively turning off (or reducing the performance of) system components when they are idle (or partially unexploited) (Benini, et al., 2000). An autonomous robot planning task must be aware of power resources available (Deshmukh, et al., 2010). Furthermore, low-power consumption is required to achieve acceptable autonomy in battery-powered systems, as well as to reduce the environmental impact (e.g., heat dissipation, cooling-induced noise) and operation cost of stationary systems (Benini, et al., 2000).

Moreover, most electronic circuits and system designs are confronted with the problem of delivering high performance with a limited consumption of electric power, and achieving highly energy-efficient computation is a major challenge in electronic design (Benini, et al., 2000). In this context, a DPM and Real-Time Scheduling (RTS) techniques were presented by Yongguo Mei, et al. (2005). They were applied to reduce the power consumption of mobile robots. At the same time that scheduling is a key concept in computer multitasking and real-time operating system, the DPM dynamically adjusts power states of components adaptive to the task's need, reducing the power consumption without compromising system performance.

Mobile robots usually have multiple components, such as motors, microcontrollers and embedded computers. Whereas sensors collect data from environment, DC motors transform direct current into mechanical energy and are often used to drive the robots. In this model, the microcontrollers are used for low-level controls, whilst embedded computers are used for high-level computation, providing an application programming interface (API). Among several tasks, this high-level computation includes motion planning, image processing, and scheduling (Yongguo Mei, et al., 2005).

Finally, a case study of mobile robot's energy consumption and conservation is presented by Yongguo Mei, et al. (2005) and shown that motion accounts for less than 50% of the total power consumption. The power consumption of the microcontroller was very stable. This implies that other power consumers like computation have a big impact on power consumption (Deshmukh, et al., 2010). These values were estimated by dividing the battery capacity by the time the computer can run with a fully charged battery when running different programs (Yongguo Mei, et al., 2005).

2.4.2 Logical Dynamic Optimization

According to European Commission (2010), the road safety application can be important for other sectors such as aeronautics (for example guidance on the ground of planes), the railway, the sectors of security and defense, the property of equipment, etc. In this context, the fusion of sensors will allow to have a higher level of information and to get more various functions. However, it is important to consider all components to achieve better energy efficiency (Yongguo Mei, et al., 2005). Moreover, an important variable is the robot state in combination with environment events.

This section introduces a logical dynamic optimization methodology. Based on the PCC variation and by exploiting the temporal coherence between consecutive frames, it proposes a new environment observer method. As shown in Figure 2.3 (b), a monocular-vision system can observe if there are no significant changes in the environment, permitting that some logical components may be shut down to save processor energy consumption, and/or to make the CPU

available for running concurrent processes. To better understand, the Figure 2.3 (a) shows an autonomous displacement through the Mojave Desert (DARPA, 2005), where the robot Stanley has used an average speed of 30.7 km/h (Stanford, 2006). In Figure 2.3 (b), due to PCC nature, taking a reference frame, i.e. the first frame of the Figure 2.3 (a), a lower value of correlation is achieved when it is closer to the vehicle (vertical black line). In this case, a command obtained (go straight) to the first image can be repeated until when the PCC derivative approaches its maximum point, where there is the obstacle detection.
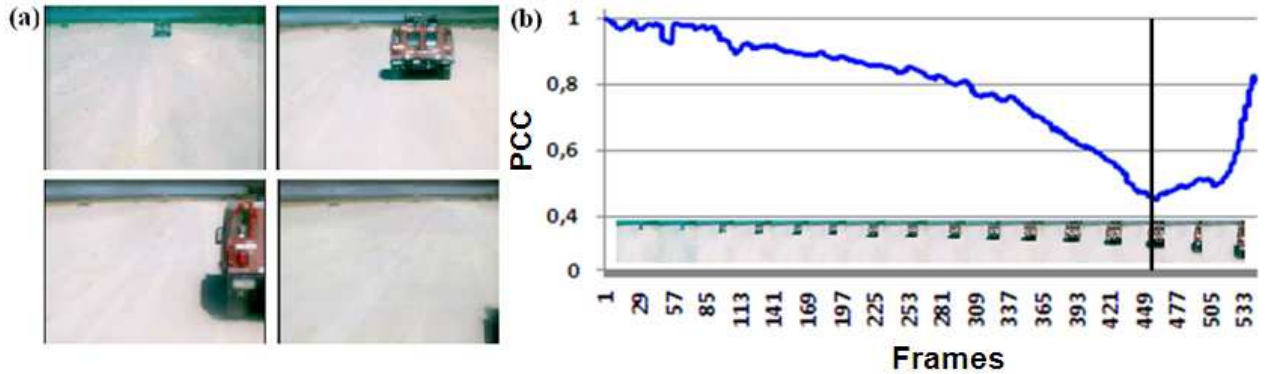
Figure 2.3 – DARPA Desert text-set 1 (DARPA, 2005): (a) original frames; (b) From a reference frame, its correlation with all others; In blue: the Pearson's correlation; The vertical black line: maximum point before the collision.

A robot can have many periodic tasks, such as motor and sensor control, sensing data reading, motion planning, and data processing. It may also have some aperiodic tasks, such as obstacle avoidance. Furthermore, for mobile robots, the tasks' deadlines are different at different traveling speeds. At a higher speed, the periodic tasks have shorter periods (Yongguo Mei, et al., 2005). In this way, it is important to notice that there is no diffeomorphism between the vehicle speed and the PCC variation, because if there are no changes between consecutive frames, the PCC threshold remains static. In this case, the isomorphism cannot be guaranteed and it ensures more efficiency for this proposal.

The Figure 2.4 (a) shows the same case of the Figure 2.3 from a different representation. From an empirical PCC threshold equal to 0.85 (green line), the reference frames (red points) are closer when it is near to an obstacle. Above of the green line all discarded images. Figure 2.5 also presents the discarding rate for an off-road context.

Whereas the main problem that has to be faced when real-time imaging is concerned and which is intrinsic to the processing of images is the large amount of data (Bertozzi, et al., 2000), the Figure 2.6 presents the accumulated time of a hypothetical image processing time (15ms) versus the gain obtained by using the discarding criteria (Miranda Neto, et al., 2007), which could allow significant savings in CPU time/power consumption. As shown in Figure 2.6 (a), in desert context, as shown in Figure 2.4 (b), were discarded 470 of 530 frames, whilst in Figure 2.6 (b), in off-road context, as shown in Figure 2.4 (b), were discarded 5595 of 6740 frames. For these two cases, the discarding rate remains over 80%.

Figure 2.4 – DARPA Desert text-set 1 (DARPA, 2005): (a) Green line: empirical PCC threshold equal to 0.85; Above of the green line it presents the discarded images; Red points: reference frames; (b) Red line: discarding rate; Blue line: vehicle speed; Green line: hypothetical image processing time (15ms).
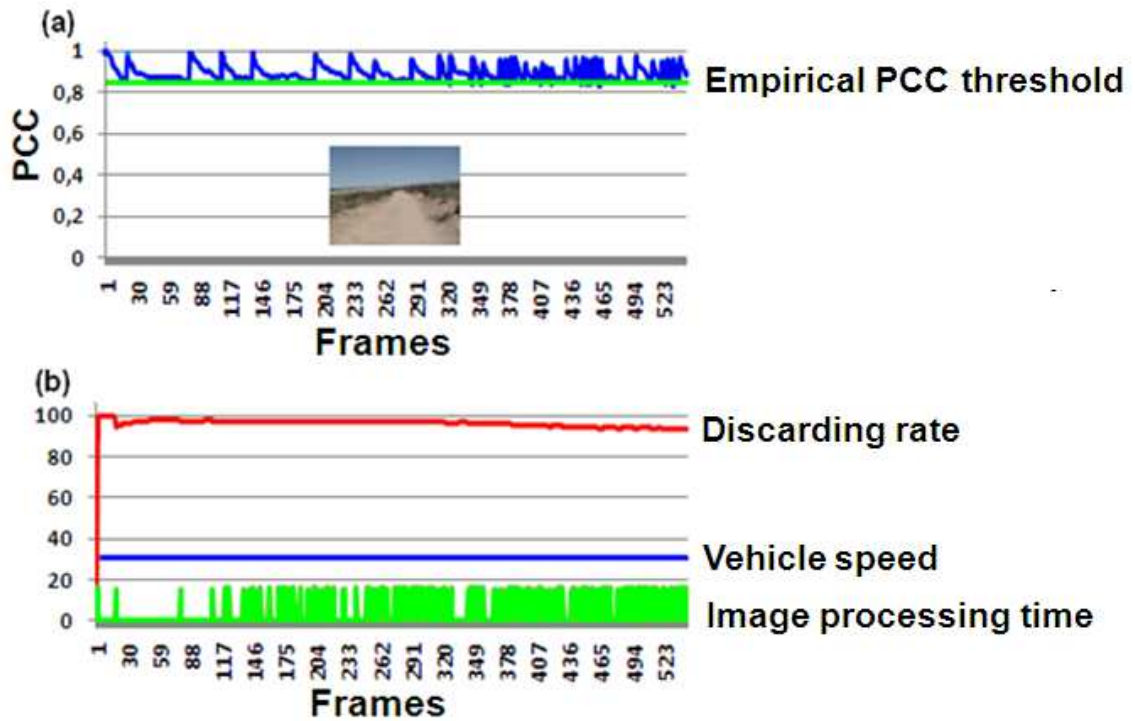
Figure 2.5 – DARPA Desert text-set 2 (DARPA, 2005): (a) Green line: empirical PCC threshold equal to 0.85; Above of the green line it presents the discarded images; (b) Red line: discarding rate; Blue line: vehicle speed; Green line: hypothetical image processing time (15ms).
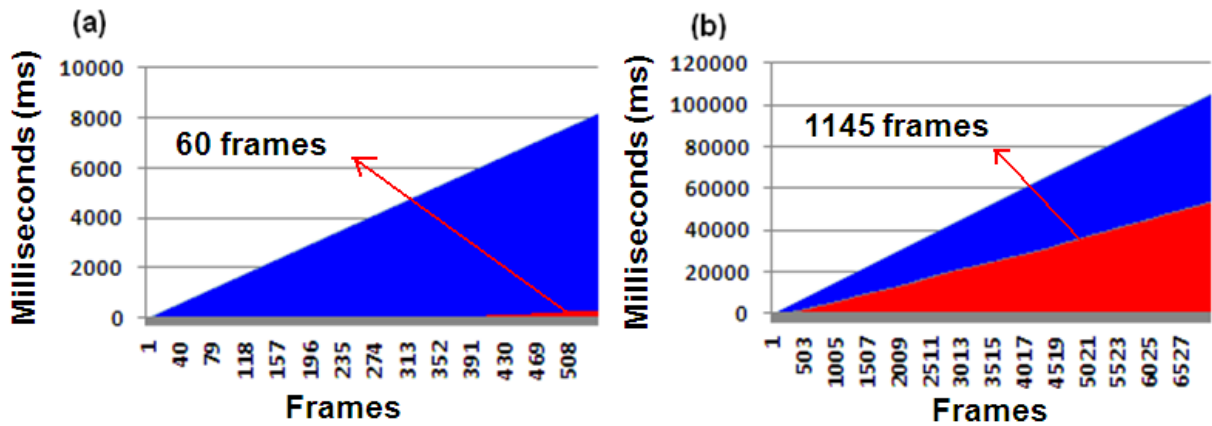
Figure 2.6 – DARPA Desert text-set (DARPA, 2005): (a) Desert video 1; (b) Off-road video 2; In blue: the cumulative impact computations (ms) without the discarding criteria; In red: the cumulative computations (ms) by using the discarding criteria.

2.5 Automatic Regions-Of-Interest (ROI) Selection

2.5.1 ROI Selection

Existing studies on energy reduction for robots focus on motion planning to reduce motion power. However, other components like sensing, control, communication and computation also consume significant amounts of power (Yongguo Mei, et al., 2005). According to the Pearson's correlation, in a certain analysis window (pair of frames), if the obstacle/object occupies a big portion of the scene, the PCC threshold tends to be low. Conversely, if obstacle/object occupies a small portion of the frame, it means that it is away from the vehicle and the system will have time enough to react. Nevertheless, where are these interest points/pixels? Or, which pixels (ROI) of the pair of images contribute most to the Pearson's coefficient computed? Which of them really need to be reprocessed?

Right after the Pearson's correlation obtained in Equation A.1, i.e. Figure 1.10 Layer (a): message A^1, it has x_m and y_m, respectively: the mean intensities of images 1 and 2. From these values, it begins again the process's correlation by applying:

$$r_2 = \frac{\sum_i (x_i - r_{1Xm})(y_i - r_{1Ym})}{\sqrt{\sum_i (x_i - r_{1Xm})^2}\sqrt{\sum_i (y_i - r_{1Ym})^2}} = \begin{cases} -1 \\ or \\ +1 \end{cases} \tag{2.1}$$

where x_i is the intensity of the ith pixel in image 1, y_i is the intensity of the ith pixel in image 2, r_{1Xm} and r_{1Ym} were respectively x_m and y_m obtained in Equation A.1.

ROI is a selected subset of samples within a dataset identified for a particular purpose. It is commonly used in medical imaging. In this work, for each pair of pixels analyzed in Equation 2.1, the only possible result is: [-1 or +1]. That is, all pixels with intensities below these means

obtained in Equation A.1 are candidates for interest points (ROI). Figure 2.7 (c) and (f) present this process, where the red pixels (interest points) represent $r_2 = -1$. A video containing all the images of this sequence is available in[27].

Taking as base an image resolution equal to image 96x72, in Figure 2.7 (g), by processing only when $r_2 = -1$, in desert video were processed about 205 thousand points, instead of 3.7 million points. In Figure 2.7 (h), in off-road context were processed about 10 million points, instead of 48 million points. Figures 2.7 (g) and (h) show, in blue, the cumulative impact data without the discarding criteria and, in red, the cumulative data by using the discarding criteria.
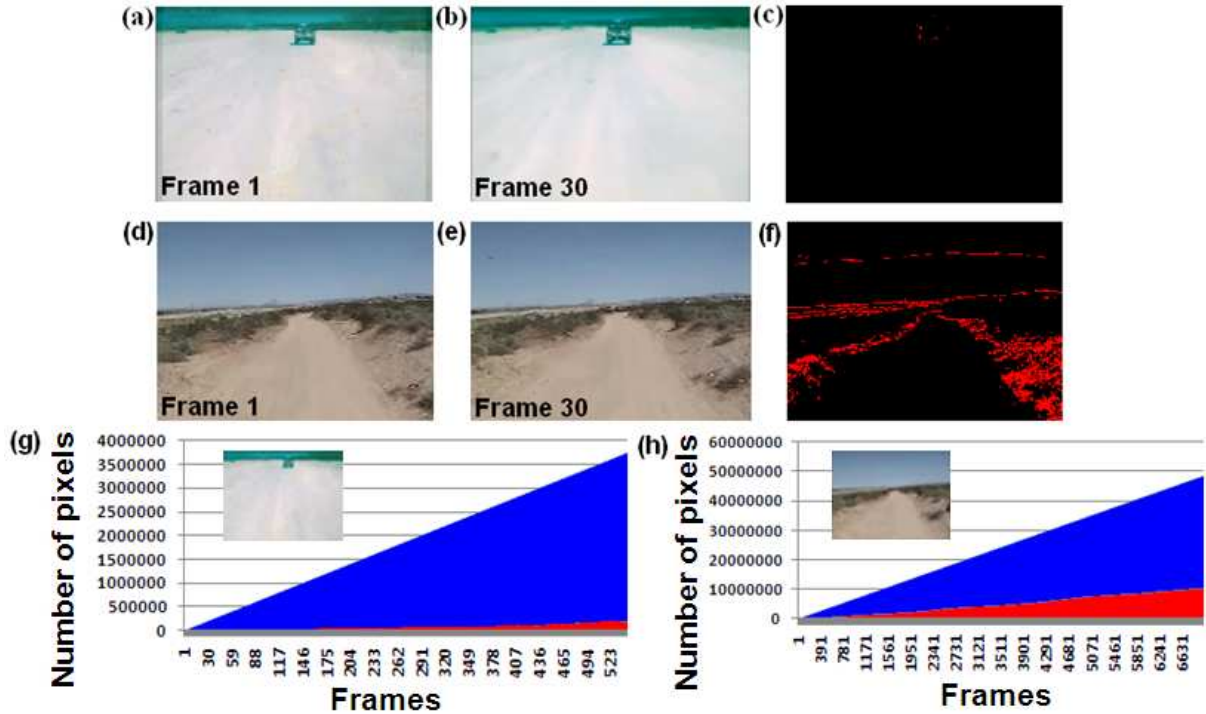
Figure 2.7 – DARPA Desert text-set (DARPA, 2005): (a), (b), (d) and (e) original frames; (c) and (f) present the process's correlation obtained in Equation 2.1, where the red pixels (interest points) represent $r_2 = -1$; (g) and (h) represent the cumulative impact data.

2.5.2 Navigable Area Detection after ROI Selection

With respect to navigable area detection tasks, Miranda Neto and Rittner (2006) have proposed a navigable area detection method based on Otsu thresholding algorithm (Otsu, 1978). More details about Otsu's method are provided in Appendix B.

As shown in Figure 1.10: Layer (b) Navigable Area Detection, this task is performed only when the PCC indicates that there is a low correlation (i.e. PCC <= 0.85) between the reference frame and the current frame. Otherwise, the current frame is discarded without being processed. That is, it keeps the previous segmentation result.

Consequently, in order to further improve the free-area detection, for each discarded image by the discarding criteria (Miranda Neto, et al., 2007), i.e. Figure 1.10 Layer (a): Environment observer: 1. DPM/PCC, and as shown in Figure 2.1, it classifies in Equation B.2 only the interest pixels (ROI) obtained in Equation 2.1 from the last Otsu threshold obtained by the navigable area detection algorithm (i.e. Figure 1.10 (b): Navigable Area Detection: B^4). As an example, Figure 2.8 (d) presents the interest pixels from the correlation between the Figure 2.8 (a) and (b), where the white pixels (ROI) represent $r_2 = -1$. In Figure 2.8 (e) these interest pixels are classified as navigable area (blue pixels).

A video containing all images of this sequence is available in[28].

2.5.3 Identifying the Limits (boundaries) of the Road after ROI Selection

In order to identify the limits of the road (which includes the obstacles), many research works were done on detect road boundary. As an example, the Canny edge detector (Canny, 1986) can be employed as input of Hough transform (Ballard, 1981) due to its robust performance and accurate edge localization. Nevertheless, to demonstrate another application for

---

[28] http://www.youtube.com/watch?v=VcUQVC1F8Xw – visited on October 1, 2011.

the proposed method in section 2.5.1, Figure 2.8 (f) presents the line detection using Hough transform from the Figure 2.8 (d), where the white pixels represent $r_2 = -1$.
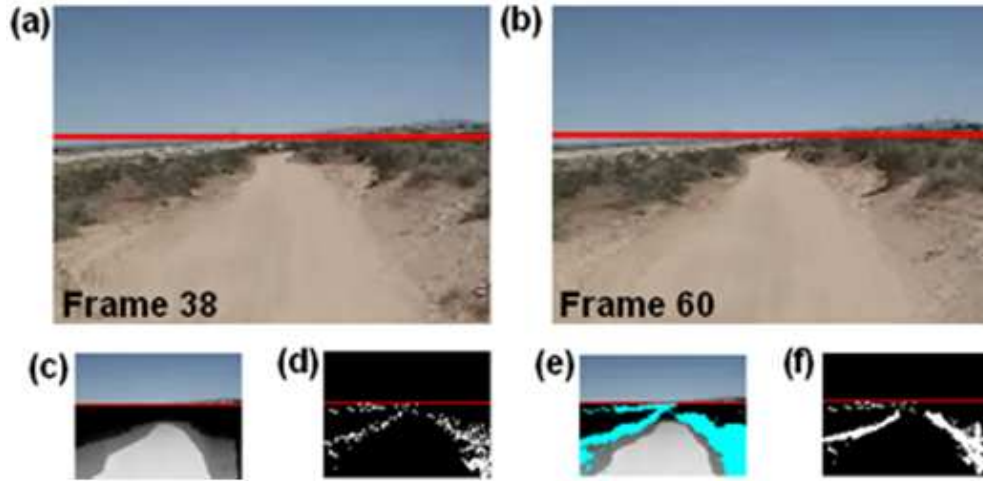
Figure 2.8 – DARPA Desert text-set (DARPA, 2005): (a) and (b) original frames; (c) the navigable area detection result; (d) the interest pixels (white pixels); (e) the interest pixels classified as navigable are represented in blue; (f) line detection using Hough transform from navigable area pixels; The red horizontal lines represent the horizon finding algorithm results (Miranda Neto, et al., 2011).

2.5.4 Obstacle Avoidance after ROI Selection

As presented earlier in chapter 1, the robots enter in the military context especially when it is necessary to reduce human exposure to hazardous situations (Chatila, 2010). Many of these robotic missions can be observed, among them, as shown in Figure 2.9 (g), an Improvised Explosive Device (IED) detonation. In this context, obstacle avoidance is a robotic discipline that includes reactive control in real time, i.e. reactive obstacle avoidance.

In order to make reliable predictions about future events, a predictive technique explores the correlation between the past history of the workload and its near future (Benini, et al., 2000). On the other hand, the Pearson's method explores regression and correlation aspects. Then, in

order to reduce the risk of collision, for the obstacle avoidance task, an empirical PCC threshold equal to 0.65 put in evidence the past history properties. This procedure also allows a greater level of security, especially when the camera does not "see" the navigable area (i.e. in front of a wall).

From a PCC threshold equal to 0.65 especially for the obstacle avoidance task, therefore, for each discarded image by the discarding criteria (Miranda Neto, et al., 2007) , i.e. Figure 1.10 Layer (a): Environment observer: 1. DPM/PCC, and as shown in Figure 2.1, it classifies in Equation B.2 only the interest pixels (ROI) obtained in Equation 2.1 from the last Otsu threshold obtained by the navigable area detection algorithm (i.e. Figure 1.10 (b): Navigable Area Detection: B^4). More details about Otsu's method are provided in Appendix B. Figure 2.9 (c), (f) and (i) present the interest pixels classified as obstacle, where the white pixels (ROI) represent $r_2 = -1$. As shown in Figure 2.9 (b), (e) and (h), the yellow pixels are the interest pixels classified as obstacle. A video containing all images of this sequence is available in[29].
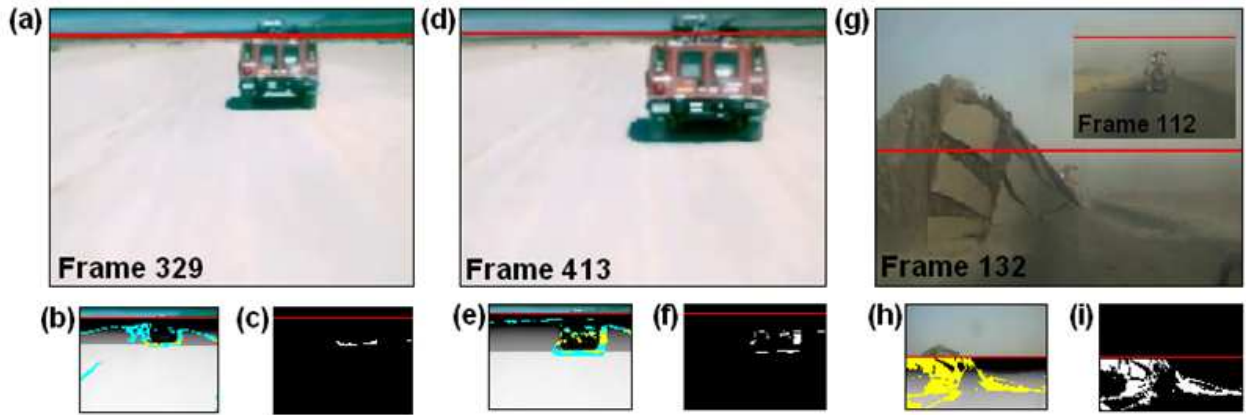
Figure 2.9 – DARPA Desert text-set (DARPA, 2005): (a) and (d) original frames; (g) Improvised Explosive Device detonation; (b), (e) and (h) the interest pixels classified as navigable are represented in blue; (b), (e) and (h) the interest pixels classified as obstacle are represented in yellow; (c), (f) and (i) represent the interest pixels classified as obstacle; The red horizontal lines represent the horizon finding algorithm results (Miranda Neto, et al., 2011).

---

[29] http://www.youtube.com/watch?v=VcUQVC1F8Xw – visited on October 1, 2011.

2.6 Collision Risk Estimation

2.6.1 Related Work

The importance of motion in visual processing cannot be understated (Beauchemin and Barron, 1995). The real nature of information used by humans to evaluate time-to-contact is still an open question. Humans adapt their motion to avoid collisions in order to preserve admissible time-to-contact. Velocity, distance, and time are intrinsically linked together (INRIA, 2011).

In the obstacle avoidance context, the collision warning algorithms typically issue a warning when the current range to an object is less than the critical warning distance, where the safety can be measured in terms of the minimum time-to-collision (TTC) (Gietelink, et al., 2009). To calculate the TTC several techniques are presented in the literature (Müller, et al., 2009), (Alenya, et al., 2009), (Dagan, et al., 2004), (Negre, et al., 2006), (Horn, 1986). For example, from the fusion of radar and vision, the results have demonstrated the advantages of both sensors to improve the collision-sensing accuracy (Wu, et al., 2009). The radar gives accurate range and range-rate measurements while vision solves the angular accuracy problem of radar; however this fusion solution is costly (Dagan, et al., 2004). Moreover, many authors focus on algorithms not suitable to perform under real-time requirements such as high and predictable response time as well as low computational costs (Müller, et al., 2009).

Measuring distances is a non-native task for a monocular camera system (Müller, et al., 2009). Nevertheless, TTC, or time-to-contact estimation is an approach to visual collision detection from an image sequence. It is a biologically inspired method that does not require scene reconstruction or 3D depth estimation (Alenya, et al., 2009). Actually, TTC is an interesting and well studied research topic (Müller, et al., 2009).

Optical flow may be used to TTC (Beauchemin and Barron, 1995), (Beyeler, et al., 2009), (Ruffier and Franceschini, 2005). However, to compute TTC from an optical flow has proven impractical for real applications in dynamic environment (Negre, et al., 2006). Additionally, gradient-based methods can be used with a certain degree of confidence in environments such as indoors were the lighting conditions can be controlled. It is computationally expensive (Mesbah,

1999). On the other hand, Inverse Perspective Mapping allows to transforms a front facing image to a top down bird's eye view (Mallot, et al., 1991), (Bertozzi and Broggi, 1998), (Bertozzi, et al., 1998). However, those equations have parameters that depend on the camera's position and its viewing angle (Tan, et al., 2006).

This work presents a novel approach to obtain the Collision Risk Estimation (CRE) estimates based on PCC from a monocular camera. Acting as a complement, a region-merging algorithm aims to represent homogeneous image regions. These image regions are equaled to reduce the Pearson's variation. From an adaptation to the Pearson's method it obtains the interest points. Finally, for finding the obstacle direction it proposes an interactive thresholding algorithm based on Otsu thresholding method (Otsu, 1978). The method proposed here does not take into account the relative acceleration between the host car and the subject "object", neither its distance nor its velocity are constant.

2.6.2 Region-Merging Algorithm

The region-merging algorithm mainly aims to represent homogeneous regions. In this context, distortions in the imaging system, pixel noise, slight variations in the object's position relative to the camera, and other factors produce a PCC threshold value less than 1, even if the object has not been moved or physically altered in any manner (Miranda Neto, et al., 2007). Whereas there are homogeneous regions in the image, and in order to obtain a Pearson's correlation closer to reality, the proposed method uses the Otsu Thresholding Method (OTM) and the Canny edge detector (Canny, 1986). More details about Otsu's method are provided in Appendix B. Additionally, only for the OTM, a Gaussian filter is applied. This filter acts as low-pass frequency filters (Gonzalez and Woods, 1991).

Figure 2.10 (a) and (e) represent an interval equivalent to 1s (a certain analysis window), namely the reference (first) frame and the current (last) frame of the series. This pair of frames was submitted to OTM in Equation B.2 and also to the Canny edge detector (Canny, 1986). The Canny edge detector results are presented in the Figure 2.10 (b) and (f). Then, the Figure 2.10 (c) and (g) present the Canny step 2: it initiates then analyzing from the slice closest to the vehicle,

going from the bottom edge of the image until a first detected edge; from this point, the image pixels are classified as foreground (black). Finally, the OTM results are presented in Figure 2.10 (d) and (h).
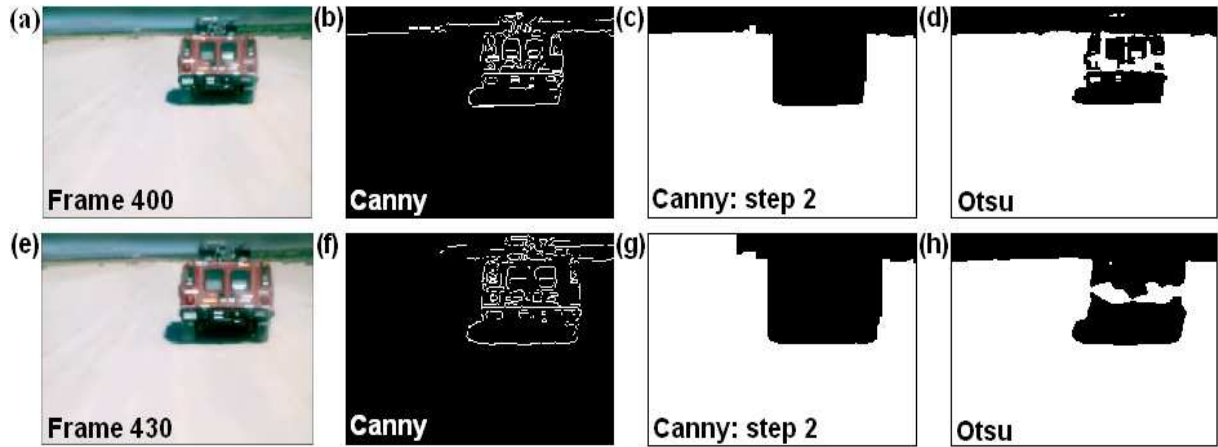
Figure 2.10 – DARPA Desert text-set (DARPA, 2005): (a) and (e) original frames; (b) and (f): the Canny edge detection (Canny, 1986); (c) and (g) are the binarized images by Canny step 2; (d) and (h) are the binarized images by Otsu thresholding method.

Right after these processes, the next step is shown in Figure 2.11 (c), where it has a yellow pixel for each pair of background pixels in both binarized images, whether in OTM or Canny. These pixels represent the homogeneous regions that will be equaled. Then, the last step is shown in Figure 2.11 (d): from this homogeneous region (yellow pixels) in Figure 2.11 (c), the pixels of the current frame will be copied to the reference frame, respectively: Figure 2.11 (b) to (a). In resume, the Figure 2.11 (d) is the Figure 2.11 (a) with the background of the Figure 2.11 (b). The Pearson's correlation (Equation A.1) between Figure 2.11 (a) and (b) is 0.790. After the region-merging algorithm, it was 0.800 between Figure 2.11 (d) and (b). Taken the DARPA Desert text-set (DARPA, 2005), the new Pearson's correlation (Equation A.1) results after the region-merging algorithm are presented in Figure 2.12 (b): red line.
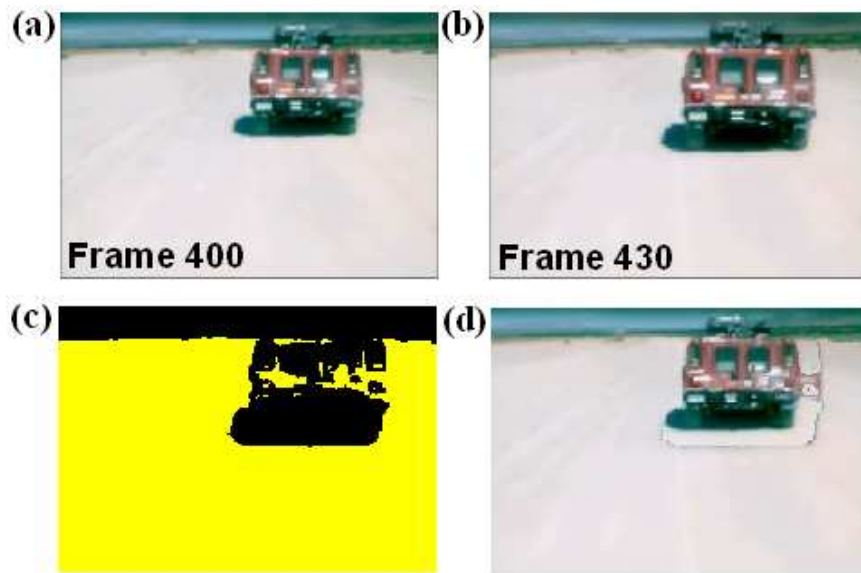
Figure 2.11 – DARPA Desert text-set (DARPA, 2005): (a) and (b) original frames; (c) Binarized image by Otsu thresholding method and Canny step 2; (d) the Frame 400 with the background of the Frame 430.
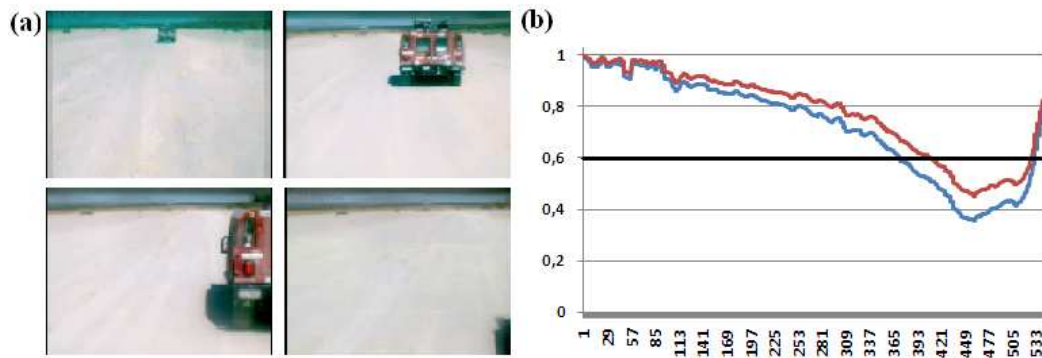
Figure 2.12 – DARPA Desert text-set 1 (DARPA, 2005): (a) original frames; (b) From a reference frame (Figure 2.12 (a): first frame), its correlation with all others; Blue line: the Pearson's correlation; Red line: the Pearson's correlation after the region-merging algorithm.

2.6.3 Collision Risk Estimation

This section presents a novel approach to Pearson's Correlation Coefficient (PCC). Based on the PCC variation, (i.e. Figure 1.10 Layer (a): A^2), and by exploiting the temporal coherence between consecutive frames, it estimates the Collision Risk Estimation (CRE) in dynamic and unknown environments by using a single monocular system.

The time-to-collision (TTC) was first defined as the distance to an obstacle divided by the relative velocity between them (Müller, et al., 2009). However, it is important to notice that there is no diffeomorphism between the vehicle speed and the PCC variation, because if there are no changes between consecutive frames captured, the PCC threshold remains static. Then, the isomorphism cannot be guaranteed. Moreover, the proposed method neglects relative acceleration between the host car and the subject obstacle.

The Figure 2.13 (a) shows an autonomous displacement through the Mojave Desert (DARPA, 2005), where the robot Stanley has used an average speed of 30.7 km/h (Stanford, 2006). In Figure 2.13 (b), taking a reference frame, i.e. the first frame of the Figure 2.13 (a), a lower value of correlation is achieved when it is closer to the vehicle, Figure 2.13 (b): vertical black line. That is, when the PCC derivative approaches its maximum point, there is the obstacle detection. In this work, this maximum point is presented in Figure 2.13 (b): vertical red line and it has named as Risk of Collision, $R_c = 1 - 0.6$ (i.e. 0.4). This empirical Risk of Collision factor can be changed. A video containing all images of this sequence is available in[30].

Taking into account R_c, the CRE is estimate in Equation 2.2:

$$CRE = \frac{R_c}{\left(1 - r_1\right)} \tag{2.2}$$

where 1 (one) represents the reference frame and r_1 was obtained in Equation A.1 and $R_c = 1 - 0.6$ (i.e. 0.4).

---

[30] http://www.youtube.com/watch?v=qMLG3icde9s – visited on October 1, 2011.
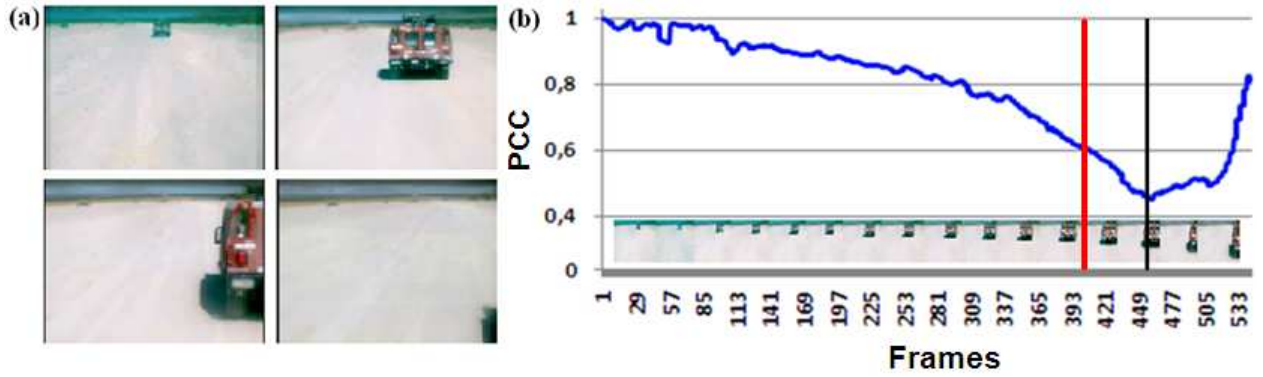
Figure 2.13 – DARPA Desert text-set 1 (DARPA, 2005): (a) original frames; (b) From a reference frame (Figure 2.12 (a): first frame), its correlation with all others; In blue: the Pearson's correlation; The vertical black line: maximum point before collision; The vertical red line: Risk of Collision.

2.6.4 Obstacle Direction: Interactive Thresholding Algorithm

From the interest points known in Equation 2.1, this section proposes an Interactive Thresholding Algorithm (ITA) that reclassifies the background and foreground pixels based on Otsu Thresholding Method (Otsu, 1978). The OTM was presented in Appendix B, where its goal is the partitioning of pixels in two classes: foreground and background.

In this context, Figure 2.14 (c), (g) and (k) present this partitioning from the process's correlation in Equation 2.1, where the red pixels (interest points) represent $r_2 = -1$. Taking as base an image resolution equal to image 96x72, the ITA process will be performed the Otsu's method (i.e. Equation B.2) N times until the result is invariably or until the red points (foreground) are less than 100. For example, from the first interaction result presented in Figure 2.14 (k), for each new interaction, the red points will be reclassified in Equation B.2. This process is shown in Figure 2.15 (a) to (f). Additionally, Figure 2.14 (d), (h) and (l) also present the ITA results. The blue line indicates the object direction based on the center of mass of the red mass points.

57
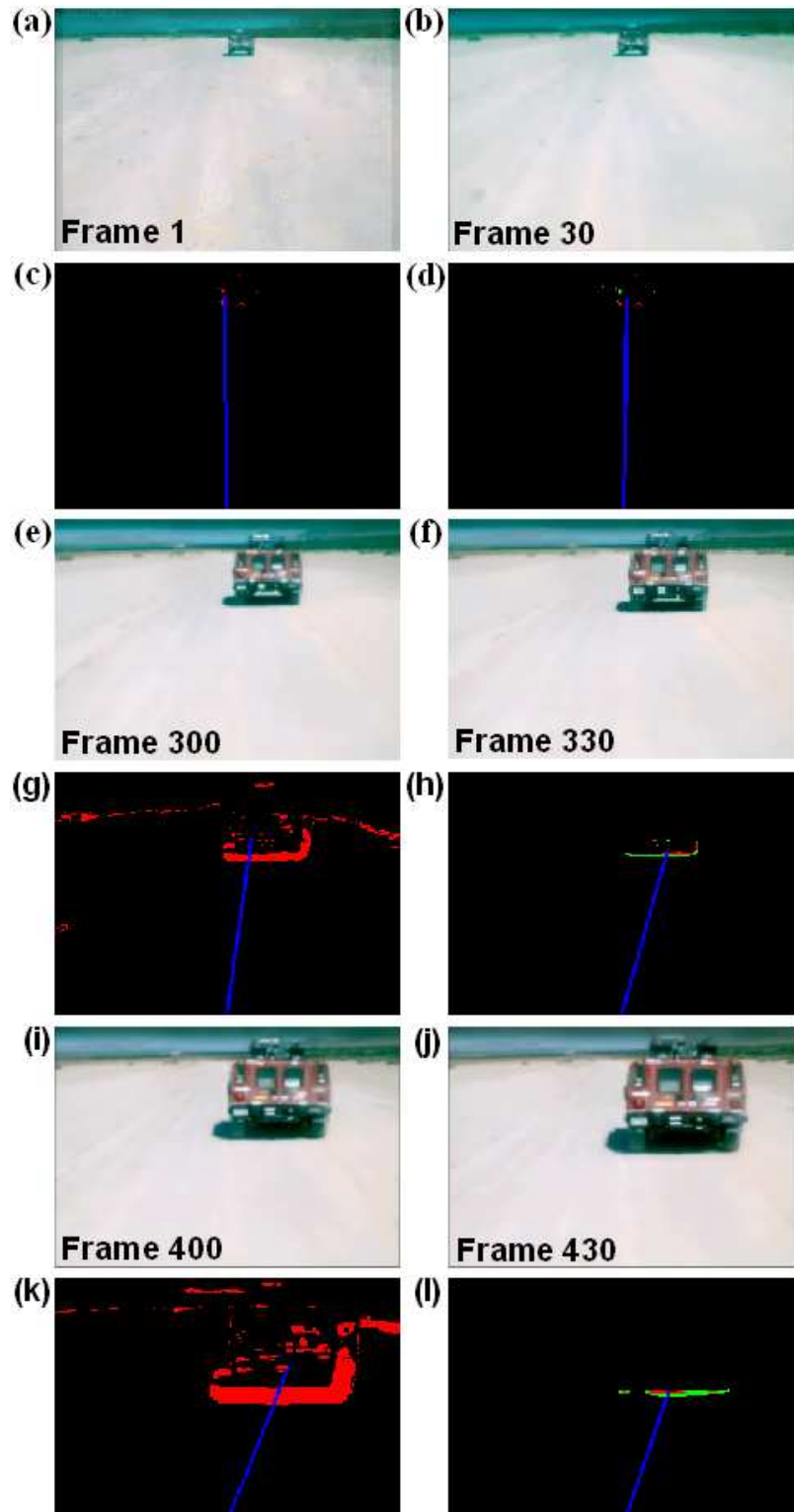
Figure 2.14 – DARPA Desert text-set (DARPA, 2005): (a), (b), (e), (f), (i) and (j) original frames; (c), (g) and (k) are the interest points from the process's correlation in Equation 2.1; (d), (h) and (l) are the Interactive Thresholding Algorithm results.
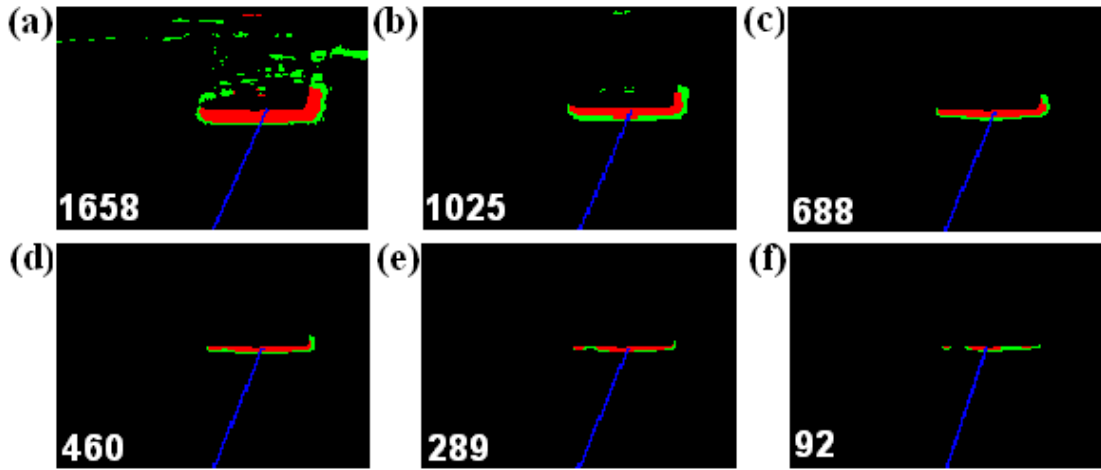
Figure 2.15 – The Interactive thresholding algorithm process.

2.6.5 Collision Risk Estimation: Case Study

Table 2.1 presents the Collision Risk Estimation (CRE) estimation from the Figure 2.14:

✓ Frames column: it represents the pairs of frames [1–30], [300–330] and [400–430], respectively: Figure 2.14: [(a), (b)], [(e), (f)] and [(i), (j)].

✓ $(1-r_1)$ column: it presents the Pearson's correlation obtained in Equation A.1 after the region-merging algorithm.

✓ Variation in the Range column: it presents the PCC variation between the first and last frames of the series.

✓ CRE Second column: it estimates the CRE in Equation 2.2.

✓ Distance Meters column: it presents an estimate in meters from the average speed of 30.7 Km/h (Stanford Team, 2006).

TABLE 2.1
RELATIONSHIP BETWEEN FRAMES OF THE FIGURE 2.14
AND COLLISION RISK ESTIMATION (CRE)
STANLEY AVERAGE SPEED: 30.7 KM/H (STANFORD TEAM, 2006)

Frames	$(1-r_1)$	Variation in the Range	Risk of Collision	CRE Second	Distance Meters
1–30	(1-0.968)	0.032	(R_c / 0.032)	12.43s	106m
300–330	(1-0.907)	0.093	(R_c / 0.093)	4.31s	36.75m
400–430	(1-0.800)	0.200	(R_c / 0.200)	2.00s	17.08m

Chapter 3

Sky Removal and Navigable Area Detection

Contents

This chapter contains a detailed description of the methods, theoretical concepts and experimental observations mainly based on a classical segmentation method known as Otsu Thresholding Method (Otsu, 1978). This method is presented in the Appendix B. More results and conclusions about this chapter will be presented in the Chapter 4.

3.1 Introduction

Navigation of an Autonomous Vehicle is based on its interaction with the environment, through information acquired by sensors. The perception of the environment is a major issue in semi- and autonomous systems. Perception for driverless vehicles area is still in evolution. These systems have been developed for improved safety and convenience, without the need to adapt the

environment. Machine Vision is an important tool to identify the region that includes the road in images. Navigable Area Detection is the major task of autonomous vehicle guidance.

For land vehicle navigation, monocular-vision systems have been designed to investigate the road informations, and in order to decrease the volume of data for processing, some systems have been designed to investigate only a small portion of the road ahead of the vehicle (Bertozzi, et al., 2000). Otherwise, the sky region is not a region of interest, and the horizon line threshold is applied to generate a road image (Lim, et al., 2009).

For navigable area detection, different techniques on automatic and semi-automatic road extraction methods are proposed in the literature (Ulrich, et al., 2000), (Aviña-Cervantes, et al., 2003), (Dahlkamp, et al., 2006), (Diego, et al., 2010), (Chetan, et al., 2010), (Yanqing, et al., 2010). The ability to identify the region that includes the road in images captured using vehicle mounted cameras is a useful function (Yamaguchi, et al., 2008). Determining the area of free road ahead is a key component of several driving assistance modules (Diego, et al., 2010).

An abstraction of a Visual-Perception Layer based on Monocular Vision was presented in Figure 1.10. This chapter considers a real-time perception problem applied to experimental vehicles (human operated or autonomous systems). It proposes an evolution of the segmentation method presented by Miranda Neto and Rittner (2006) and applied it to obstacle detection and reactive navigation, which takes place in obstacle avoidance context for vehicles in dynamic and unknown environments. It also includes a horizon finding algorithm that finds the horizon line and apply it to generate the navigable area. The main focus in this chapter is to present the Layer (a) Navigable Area Detection.

The Related Works are presented in section 3.2. Section 3.3 and 3.4 introduce the Horizon Finding and Navigable Area Detection algorithms. The main sections are:

✓ Section 3.3: Horizon Finding Algorithm;
✓ Section 3.4: Navigable Area Detection Algorithm;

3.2 Related Works

3.2.1 Machine Vision and Image Segmentation

According to Gonzalez and Woods (1991), the acquisition of values sampled by a camera can be compared with image processing performed by human vision. On the other hand, it is known that the human vision without assistance from the brain is not capable to allow people displacement (navigation) in an efficient way. This is also truth for computer systems, which, in order to navigate based on images, should contain software intelligent enough to manage mechanical structure through navigation.

What is trivial for the human system, that is, to construct three-dimensional scenes from two-dimensional images captured by the vision system and to use it on the decision process for navigation, is not necessarily trivial for the computer systems.

Different from the human system, complex computational vision systems can lead to some damages due to the processing time. Thinking about the existing relation between a real time decision system and an image reading system that operates in a specific acquiring/reading rate, that is, amount of images read per second, one can question: how many images acquired must be discarded by the image processing system to guarantee an acceptable real time navigation of an autonomous vehicle? A solution was proposed in Chapter 2.

Due to the general applicability of it, the problem of navigation of mobile robots is dealt with using more complex techniques (Bertozzi, et al., 2000). In this way, processes conducted by the machine vision have as goal, in addition to the visual information, to process image data for machine perception. Several methods can perform pattern recognition in images. When applied to image processing, generally are known as segmentation methods. According to Gonzalez and Woods (1991), segmentation methods can be considered as the partition of digital images in sets of pixels, considering application and previously defined criteria.

The purpose of segmentation is to distinguish objects in an image (Abutaleb, 1989); what can be extremely sophisticated and complex. Results can be very satisfactory with the use of well elaborated filters. However, these results (high quality segmentation) can generate a higher price, that is, normally robust segmentation algorithms present great complexity.

One way to perform segmentation of an image is to use thresholds. This type of

segmentation technique, called thresholding, is very simple and fast computationally, however the identification of the ideal threshold can be sufficiently complicated. The best thing to do in this case is to use techniques and algorithms that search the thresholds automatically.

Thresholding methods are divided in two groups: global and local. The global ones divide the image using only one threshold and the local ones are those that divide the image in sub-images and for each one of them a threshold is defined. In (Sahoo, et al., 1988) the local thresholds are defined as multilevel thresholds. Summarizing, from the definition of global thresholds and/or multilevel, the use of a global threshold in a two-dimensional image I (x, y) with levels of N = [0,255] intensity consists of determining an only threshold T that separates pixels in two distinct classes: object and background. Threshold T normally is applied to the histogram of an image h (N), which can be seen as a description of the distribution of pixels intensities for the image.

Multilevel thresholds are generally more laborious than the global threshold, because it is more complicate to find multiples thresholds that determine the interest regions effectively, especially when there are some groups of objects in the image.

On the other hand, according to Gonzalez and Woods (1991), the global thresholding only reaches good results when the illumination of the image is relatively uniform and the interest regions, represented by objects, possess a significant difference of intensity from the background (contrast).

In this context, region recognition can be handled by popular thresholding algorithm such as Maximum Entropy, Invariant Moment and Otsu Thresholding Method (OTM). For navigable area detection, because OTM supplies a more satisfactory performance in image segmentation, it was used to overcome the negative impacts caused by environmental variation (Yanqing, et al., 2010). Furthermore, some authors consider the OTM as one of the best choices for real-time applications in machine vision (Sahoo, et al., 1988), (Lee, et al., 1990). It still remains one of the most referenced thresholding methods (Sezgin and Sankur, 2004). The details of the Otsu Thresholding Algorithms are presented in Appendix B.

3.2.2 Image Processing

This work uses a color or gray-level image and smooth them using a Gaussian filter. The Gaussian smoothing operator is a 2-D convolution operator. It acts as low-pass frequency filters (Gonzalez and Woods, 1991).

If the image is colored, in order to utilize the most important information of the color image, the candidate color channel that was dominant in certain color space is selected to generate the histogram image (Yanqing, et al., 2010). For navigable area detection, this empirical space is a small portion of the road ahead of the vehicle, where the absence of other vehicles has been assumed (Chetan, et al., 2010). As an example, for find the drivable surfaces, Thrun, et al. (2006) projects drivable area from the laser scan analysis into the camera image. This quadrilateral area is between 10 and 20 meters ahead of the robot. As shown in Figure 3.6 (a): yellow area, the basic idea is to consider a given region in the actual image as drivable. It assumes that the bottom center of the image contains road pixels for a large majority of the time (Rauskolb, et al., 2008). This technique was first presented by Ulrich, et al. (2000).

For sky removal (Miranda Neto, et al., 2011), as shown in Figure 3.3, this space represents 60% of the image height. The candidate color channel is obtained in Equation B.1, as shown in Appendix B.

3.2.3 Sky Removal

Figure 2.2 illustrated the monocular vision contribution to the DARPA Grand Challenge (Dahlkamp, et al., 2006). This diagram showed that the reach of lasers was approximately 22 meters (left), whereas the monocular vision module often looks 70 meters ahead (right). It also shows the horizon detection for sky removal. Stanford Racing Team implemented the horizon finding algorithm originally proposed by Ettinger, et al. (2003) to eliminate all pixels above that horizon. Additionally, these techniques has been used for flight stability and control system to Micro Air Vehicles (Ettinger, et al., 2003), for the control of the airship (Rives, et al., 2004) and landing aircraft control (Entzinger, 2008). It also was employed as an absolute attitude sensor which is useful for low-level control of an unmanned aerial vehicle (Hwangbo, 2009).

The purpose of segmentation here is the line horizon detection (horizon finding) for sky

removal. Based on Otsu Thresholding Method (Otsu, 1978), this work proposes a global thresholding method, which seeks not the ideal threshold for the whole image, but an ideal threshold associated at each portion (local information), that contributes to the final decision.

In this sense, let's see an example for land vehicle navigation, where the main task is navigable area detection. Figure 3.1 illustrates that probably for images that possess the horizon (sky) in its composition; the algorithm may not have a satisfactory result. To deal with this problem, this work proposes a robust horizon finding algorithm. After sky removal, it is possible to investigate only a small portion of the image ahead of the vehicle, and the new result can be seen in Figure 3.2.

Figure 3.1 – Original image and Otsu segmentation.
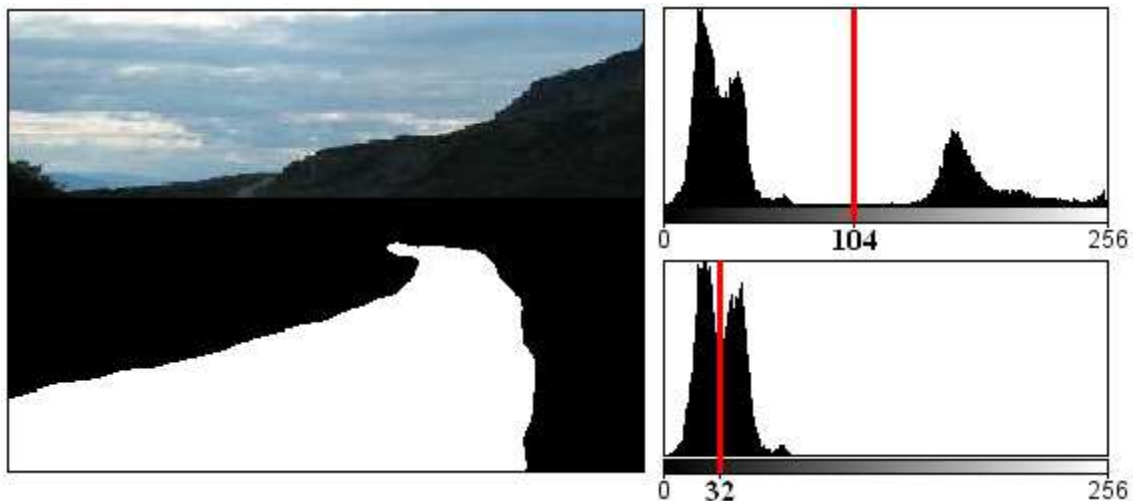
Figure 3.2 – Left: Otsu segmentation (navigable area detection) after sky removal; Right: Global Otsu threshold (104); new Local Otsu threshold (32) after sky removal.

3.2.4 Navigable Area Detection

Different techniques on automatic and semi-automatic road extraction methods are proposed in the literature. From a system trained by driving the robot through its environment, a vision-based obstacle detection allows that each individual image pixel is classified as belonging either to an obstacle or the ground based on its color appearance (Ulrich, et al., 2000).

For robot navigation in agricultural environments or hazardous related areas, a method for extracting and tracking man-made roads segments color images in small areas was proposed by Aviña-Cervantes, et al. (2003), where these small areas are characterized later by color and texture attributes, and features are classified using the K-NN rule or the Support Vector Machines method.

A method for identifying drivable surfaces in difficult unpaved and off-road terrain conditions was proposed based on computer vision and lidar. Computer vision then constructs appearance models to find drivable surface outward into the far range (Dahlkamp, et al., 2006).

From a previously extracted road segments (manually or semi-automatically) in a traffic-free reference video record on a first drive, a road-detection method to infer the areas of the image depicting road surfaces without performing any image segmentation was presented by (Diego, et al., 2010). It uses a dynamic background subtraction based on Otsu thresholding algorithm.

A partition-based algorithm for classification of outdoor terrains using monocular camera was proposed by Chetan, et al. (2010). A vision-based navigable area detection method was proposed by Yanqing, et al. (2010) to realize visual guiding navigation for autonomous land vehicles. In this case, the images are segmented into road and non-road region by using Otsu thresholding algorithm, which included the Monte Carlo method with the road boundary extraction based on Canny.

As seen earlier, some authors consider the Otsu thresholding algorithm as one of the best choices for real-time applications. This work therefore uses the Otsu's method due to its application in real time and its robustness to different types of terrain.

3.3 Horizon Finding Algorithm

According to Sezgin and Sankur (2004), cluster-Otsu (clustering thresholding) minimizing the weighted sum of intra-class variance of the foreground and background pixels to establish an optimum threshold, that minimization of intra-class variance is tantamount to the maximization of between-class scatters. The Otsu method gives satisfactory results when the numbers of pixels in each class are close to each other.

Using as reference Figure 3.3, the horizon finding results can be seen in Table 3.1. It presents an example of land vehicle navigation, where the range of horizon analysis was empirically limited to a maximum height of 60% of the image. This region is also the empirical space to select the candidate color channel as shown in Equation B.1.

The horizon finding algorithm creates empirically cuts that divide image in ten parts (sub-images) of equal heights. All these sub-images are then submitted to the segmentation algorithm in Equation B.2, where the output is a vector with Otsu thresholds values of each sub-image. Then, a threshold is calculated at each sub-image (Figure 3.4), which depends on local information and variance of the pixel neighborhood. Subsequently, in order to find the optimal threshold that is capable to get a more efficient segmentation and to distinguish with bigger precision, for each found local Otsu threshold, all sub-images are again submitted to the segmentation algorithm in Equation B.2. The purpose is to analyze how much each local sub-image contributes positively to increase or reduce horizon points. Then, the algorithm initiates by analyzing the sub-image closest to the vehicle, that is, the first more inferior slice of the image, going from the bottom edge of the image (red line, Figure 3.4) to the last more superior sub-image.
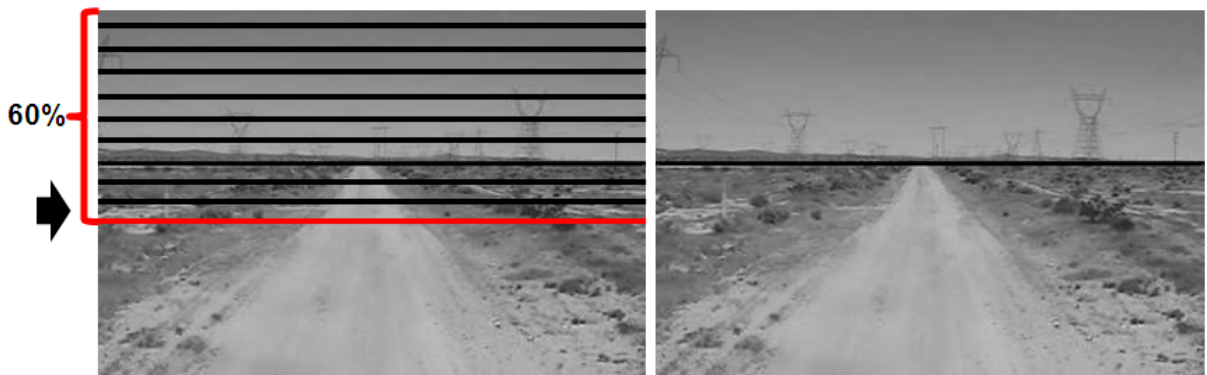
Figure 3.3 – Sub-images and the horizon line detection.
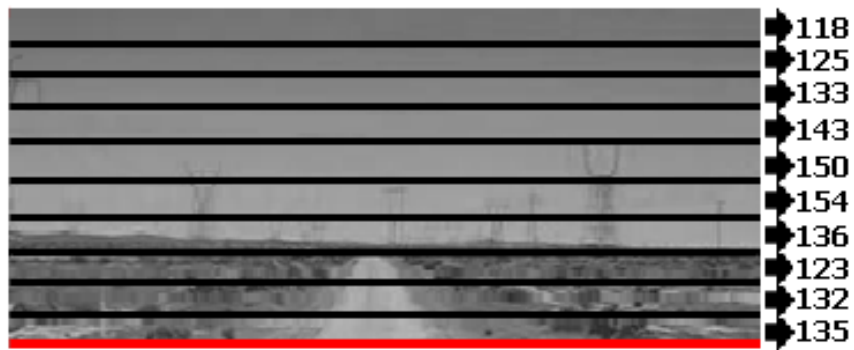
Figure 3.4 – The clusters-Otsu and its optimal local thresholds.

Using as reference Figure 3.4, it obtains the local Otsu thresholds and its percentage of foreground points (horizon). They are represented in blue by the diagonal matrix in Table 3.1, respectively, threshold (T) and percentage (P):

(a) T = 135 / P = 28% | (f) T = 150 / P = 57%
(b) T = 132 / P = 15% | (g) T = 143 / P = 48%
(c) T = 123 / P = 7% | (h) T = 133 / P = 53%
(d) T = 136 / P = 59% | (i) T = 125 / P = 53%
(e) T = 154 / P = 80% | (j) T = 118 / P = 52%

TABLE 3.1

LOCAL OTSU THRESHOLDS VALUES OF THE FIGURE 3.4

	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)	(j)	(k)	(l)
T	135	132	123	136	154	150	143	133	125	118	$\sum$	%
1	0	0	0	0	0	0	0	0	0	52	52	0
2	0	0	78	0	0	0	0	0	53	97	228	77
3	35	64	98	26	0	0	0	53	98	99	473	51
4	98	99	100	97	0	0	48	99	100	100	741	36
5	100	100	100	100	17	57	95	100	100	100	869	14
6	100	100	100	99	80	93	97	100	100	100	969	10
7	60	64	75	59	28	37	49	63	73	81	589	64
8	5	5	7	5	2	3	4	5	7	9	52	91
9	14	15	20	13	10	10	12	14	19	23	150	65
10	28	31	44	27	17	18	21	30	41	53	310	51

69

Another interaction is next necessary. For each local Otsu threshold found, all sub-images are again submitted to the segmentation algorithm in Equation B.2. Taking as reference the data in Table 3.1, from an Otsu threshold equal to T = 135, the result of the first interactions can be seen in column (a). For example, it has in row (a)-10 its percentage of foreground points equal to 28 (blue) which refers to the sub-image closest to the vehicle. In the following, using the same Otsu threshold, in row (a)-9, the result equal to 14 refers to the second sub-image analyzed, and so on for others rows in column (a).

As a second example, in column (j) the Otsu threshold is equal to 118. From this threshold (last more superior sub-image analyzed), in row (j)-10 the percentage of foreground points equal to 53 refers to the sub-image closest to the vehicle. In the following, in row (j)-9 the result equal to 23 refers to the second sub-image analyzed.

The columns (k) and (l) are, respectively, sum of rows and percentage difference between the value of a row and its previous one. Then, considering that the algorithm search the region of greatest difference in the image, the yellow filled row (8) is the one where it finds the value 91.

The horizon line separates the image into two regions that have different appearance (Ettinger, et al., 2003). The choice of this method was based on Miranda Neto and Rittner (2006), because not always the analysis of a bigger image portion (i.e. a global segmentation method) can contribute for a better result in the most critical region (Figure 3.1: in land navigation, the region closer to the vehicle), where obstacles should be detected and avoided as fast as possible. On the contrary, when separating the superior portion of the original image, it is capable to get a more efficient segmentation and to distinguish with bigger precision the obstacles from the navigation area.

Once the Otsu horizon line is found, the Hough transform (Ballard, 1981) is used, and a rapprochement between both results is triggered based on the weighted average. The Canny detector (Canny, 1986) was employed as input of Hough transform due to its robust performance and accurate edge localization. Finally, due to the fact that in small time intervals, there will be no abrupt changes in the positioning of new horizon line (NHL), it uses a simple decaying filter followed by a Kalman filter (KF) (Kalman, 1960). It is presented in Figure 3.5. The discrete Kalman filter tries to estimate the new horizon line position by the linear stochastic difference:

$$x_k = Ax_{k-1} + Bu_{k-1} + w_{k-1} \qquad (3.1)$$

where the measurement for the system is described by:

$$z_k = Hx_k + v_k \qquad (3.2)$$

where the measurement noise is also represented by a random variable v_k.

The random variables w_k and v_k represent the process and measurement noise, with normal probability distributions: $p(w) \approx N(0,Q)$ and $p(v) \approx N(0,R)$, where Q and R are constant in this model. In the difference Equation 3.1, the N x N matrix $A[110]$ relates the state at time step k to the state at step $k+1$. In Equation 3.2, the M x N matrix H relates the state to the measurement z_k. If x_k^- denotes a priori state estimate at step k provided the process prior to step k is known, and x_k denotes a posteriori state estimate at step k provided measurement z_k is known, then a priori and a posteriori estimate errors can be defined as $\begin{aligned}e_k^- &= x_k - x_k^- \\ e_k &= x_k - x_k\end{aligned}$. The a priori estimate error covariance and the a posteriori estimate error covariance are respectively $P_k^- = E\left[e_k^- e_k^{-T}\right]$ and $P_k = E\left[e_k e_k^{T}\right]$.
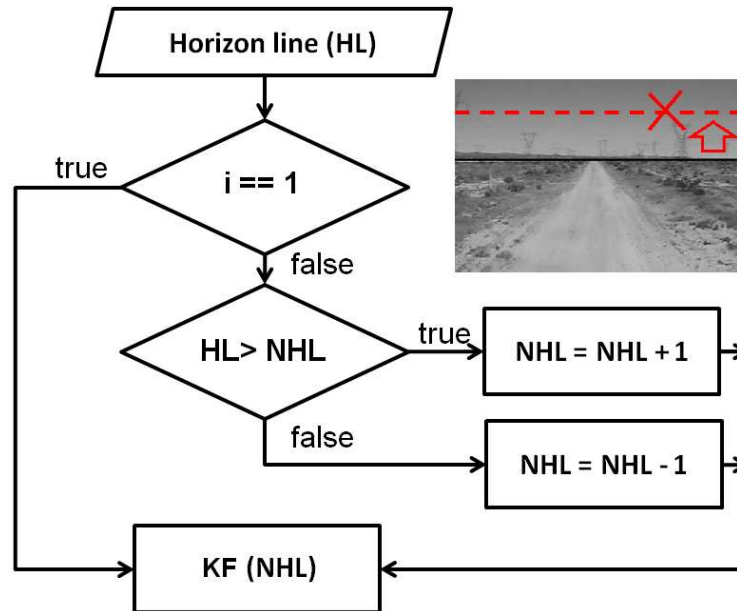
Figure 3.5 – Horizon Line (HL) and New Horizon Line (NHL).

3.4 Navigable Area Detection Algorithm

From the image processing and sky removal steps, in order to obtain a multimodal 2D drivability road image, RD_w, i.e. free-area detection, the algorithm performs the following tasks:

1. Task 1: Due to the different image textures on different roads, the original image, Figure 3.6 (a), and its negated, Figure 3.6 (c), are submitted to the Otsu's method in Equation B.2. The algorithm then selects an image with the highest percentage of navigable area (white points) in the bottom center of the image, as shown in Figure 3.6 (a): yellow area. The original image and its Otsu's result, Os^+, respectively, they can be seen in the Figure 3.6 (a) and (b). The negated image and its Otsu's result, Os^-, respectively, they can be seen in the Figure 3.6 (c) and (d).

2. Task 2: Whereas there are homogeneous regions in the image, and in order to identify the limits of the road (which includes the obstacles), the Canny edge detector (Canny, 1986) was employed as input of Hough transform (Ballard, 1981) due to its robust performance and accurate edge localization. Respectively, the results can be seen in the Figure 3.6 (e) and (f). Then, from the bottom center of the image, the algorithm concludes by finding a single image mass, Hc, Figure 3.6 (g). It may also help to identify the textureless regions classified as road region, specular surfaces, traffic markings, etc.

Figure 3.6 – (a) Original image and its Otsu's result in (b); (c) negated image and its Otsu's result in (d); (e) Canny edge detection result; (f) Hough transform result; (g) a single mass in the image (f).

A multimodal road image is then triggered based on the weighted average of the images intensities. It is described in Equation 3.3:

$$RD_{w(x,y)} = \frac{RD_{w(x,y)} + Os_{(x,y)} + Hc_{(x,y)}}{3}$$

(3.3)

where $RD_{w(x,y)}$ (left) is the intensity of the ith pixel, after update, in the new image RD_w, $RD_{w(x,y)}$ (right) is the intensity of the ith pixel in the old image RD_w, $Os_{(x,y)}$ is the intensity of the ith pixel in image obtained in tasks 1, $Hc_{(x,y)}$ is the intensity of the ith pixel in image obtained in tasks 2.

Due the small variance of shades between objects, false path-markings, false route-markings, shadows, etc, the next step is to re-project this 2D drivability road image inspired in a drag process. In fluid dynamics, drag forces act in a direction opposite to the oncoming flow velocity, i.e. forces (or resistance) that oppose the relative motion of an object through a fluid (White, 1986). Coming again to the domain of image, and as shown in Figure 3.9, the objective is to classify the road image in terms of fluidity and navigability, where the obstacles represent the forces that oppose the relative motion of a robot through a route. From the Equation 3.3, a new multimodal 2D drivability road image is described in Equation 3.4:

$$RD_{d(x,y_i)} = RD_{w(x,y_i)} - \left(RD_{w(x,y_i)} * \left(1 - \frac{Fv}{Df} \right) \right)$$

(3.4)

where for each image column x in $RD_{w(x,y)}$ obtained in Equation 3.3: (a) the algorithm initiates by analyzing each image row, y, from bottom to top; (b) the drag forces (resistance), Df, and flow velocities, Fv, are initialized with a value by default equal to $h*255$, where h is the image height and 255 represents the highest intensity level of a pixel. Then, for $y=h-1$ to $y=0$, the Df and Fv values are updated each line by:

$$Df = Df - 255 \tag{3.5}$$

$$Fv = (Fv - 255) - (255 - RD_{w(x,y_i)}) \tag{3.6}$$

where $RD_{d(x,y_i)}$ is the intensity of the ith pixel in the image RD_d. If $Fv \leq 0$ then $Fv = 0$. The others drag's attributes, as velocity, reference area and drag coefficient, they are equal to 1.

Right after applying the Equations 3.3 and 3.4, the final result to Figure 3.6 (a) is shown in Figure 3.7 (b). In Figure 3.8 (d) false-navigable areas (red) are eliminated by applying the drag process in Figure 3.8 (c).
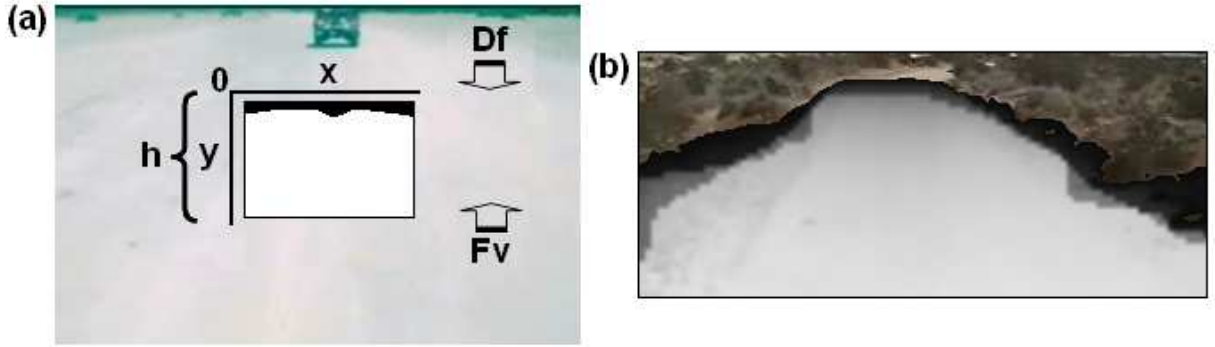
Figure 3.7 – (a) The drag process; (b) A new multimodal 2D drivability road image by considering the drag process: pixel intensities $RD_{d(x,y_i)} \geq 10$.

Figure 3.8 – (a) Original image after sky removal and its Otsu's result in (c); (b) Canny edge detection result; (d) A multimodal 2D drivability road image by considering the drag process: pixel intensities $RD_{d(x,y_i)} \geq 100$.

74
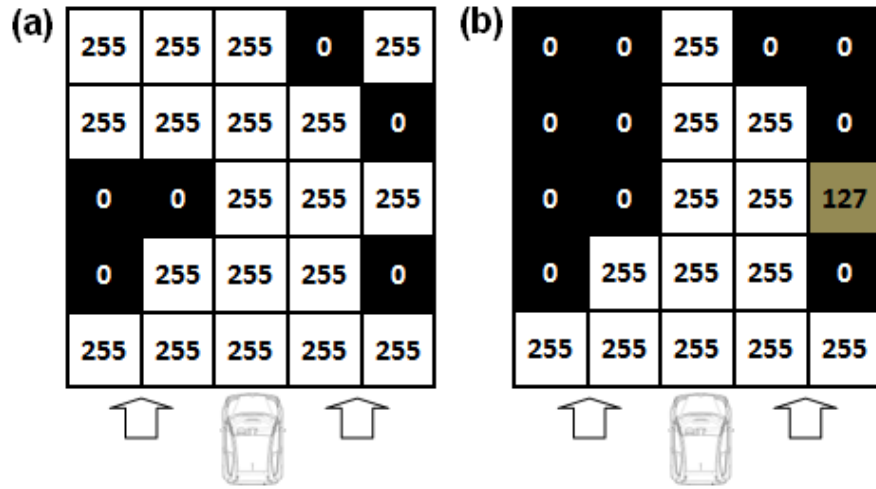
Figure 3.9 – (a) Hypothetical multimodal 2D drivability road image 5x5; (b) Right after apply the Equation 3.4, the new multimodal 2D drivability road image 5x5 by considering the drag process.

3.4.1 Reactive Navigation

Right after applying the Equations 3.3 and 3.4, a multimodal 2D drivability road image can be represented in two types of classes: c_1 navigable area and c_2 obstacles. In order to obstacle avoidance, the center of mass of the navigable area class is calculated by:

$$m_c = \sum_{x,y} h_c(x, y) \tag{3.7}$$

where $h_c(x, y) = \begin{cases} 1.se.S_{x,y} = c_1 \\ 0.se.S_{x,y} \neq c_1 \end{cases}$

The center of mass of the class c_1, denoted class o_{c1}, is given by the coordinates xo_{c1} and yo_{c1} in Equations 3.8 and 3.9:

$$xo_{c1} = \frac{\sum\limits_{x,y} x.h_c(x,y)}{m_c}$$ (3.8)

$$yo_{c1} = \frac{\sum\limits_{x,y} y.h_c(x,y)}{m_c}$$ (3.9)

Taking the origin point as: $x = \dfrac{image.width}{2}$ and $y = imageheight$; and getting the destination point in the Equations 3.8 and 3.9, the steering angle correction (new robot direction) can be calculated in Equation 3.10. As shown in Figure 3.10: red line, the main objective is to drive the robot to the center of the navigable area.

$$a = ((arc.\cos(\frac{|(yo_{c1} - y)|}{\sqrt{(xo_{c1} - x)^2 + (yo_{c1} - y)^2}}))*180)/PI$$ (3.10)
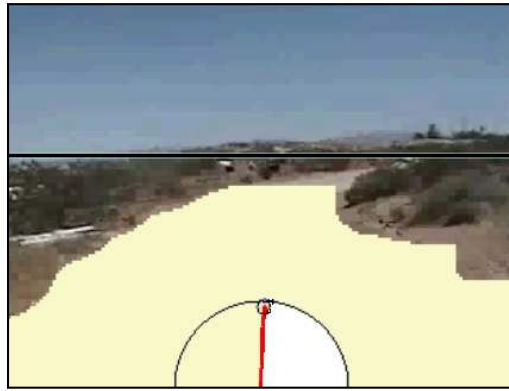
Fig. 3.10 – The center of mass of the navigable area: the new robot direction.

Chapter 4

Experimental Evaluation

Contents

This chapter contains experimental observations based on the last two chapters.

4.1 Introduction

Theoretical, computational and experimental simulations were considered in the previous chapters, mainly using an experimental DARPA Desert text-set (DARPA, 2005). In order to use an urban text-set, this chapter include the Karlsruhe Dataset (Geiger, 2011) provided by the Karlsruhe Institute of Technology. Besides, others results are also presented on real, dynamic and unknown environments by using two experimental vehicles. The computational time was measured on a 2.5GHz Intel Core 2 Quad processor, 3.48 GB RAM, Microsoft Windows XP Professional SP3, Visual Studio C++ and OpenCV 2.1.0.

Figure 4.1 (a) and (b) present the Carmen and Strada platforms, which are equipped with GPS receivers, stereo and monocular vision systems, a multi-layer lidar and a CAN-bus gateway which grants access to embedded sensors used by built-in ABS and ESP functions, Wheel Speed Sensors (WSS) and yaw rate gyroscope. The Carmen platform is also equipped with Velodyne lidar scanner. In Figure 4.1 (c), the VERO platform is equipped with SICK LMS and Hokuyo UTM30 outdoor laser scanners, GPS receiver, a monocular camera, and a CAN-bus interface which grants access to encoder data for the four wheels and steering, and allows commands to be sent to two independent motors driving the rear wheels and to control the steering angle.

The experimental works reported here was carried out at the Heudiasyc[31] Laboratory UMR 6599 CNRS/UTC, Compiegne, France, from August 2009 to August 2011, and at the Center for Information Technology Renato Archer[32] (CTI), Campinas, Brazil, from October, 28 2010 to November, 07 2010. Results for different types of image texture (road surfaces) were selected and its results are presented bellow from the same structure of the last chapters:

- ✓ Section 4.2: Pearson's Correlation Coefficient (PCC): Applications on Autonomous Robotics;
- ✓ Section 4.3: Sky Removal and Navigable Area Detection;
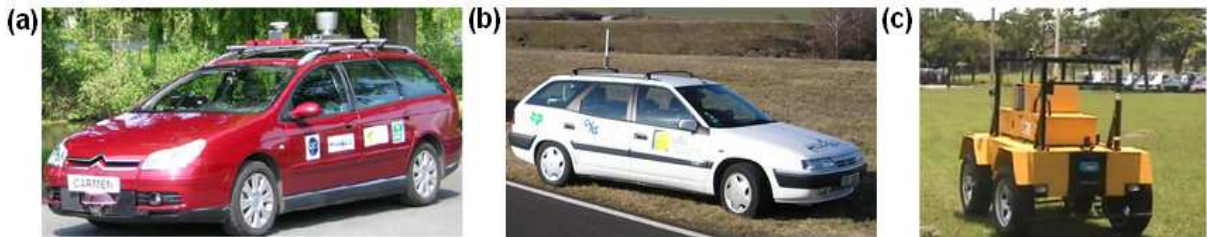
Figure 4.1 – The experimental vehicles: (a) Carmen and (b) Strada vehicles at Heudiasyc Laboratory in Compiegne, France; (c) VERO platform at Renato Archer IT Center (CTI) in Campinas, Brazil.

---

[31] http://www.hds.utc.fr/ : Created in 1980/81, the laboratory Heudiasyc, which means Heuristic and Diagnostic Methods for Complex Systems (Heudiasyc), is a joint research unit. It is attached to the University of Technology of Compiègne (UTC) and the French National Center for Scientific Research (CNRS) – visited on October 1, 2011.
[32] http://www.cti.gov.br/english/ : As unit of the Ministry of Science and Technology, the Center for Information Technology Renato Archer (CTI) conducts research and development in many strategic areas, such as micro and nano electronics, systems, software and IT applications, e.g. robotics, decision support systems and 3D technologies for industry and medicine. With strong interaction with universities, research centers and industry, CTI has been able to act as a key agent in the Brazilian R&D system – visited on October 1, 2011.

4.2 Pearson's Correlation Coefficient: Applications on Autonomous Robotics

Based on Pearson's method, the chapter 2 proposed a visual-perception system based on an automatic image discarding method as a simple solution to improve the performance of a real-time navigation system by exploiting the temporal coherence between consecutive frames. It also presented the Pearson's Correlation Coefficient (PCC) as an environment observer method to save processor energy (power) consumption. This section includes experimental results in dynamic and unknown environments. It is organized as follows:

- ✓ Section 4.2.1: Dynamic Power Management;
- ✓ Section 4.2.2: Automatic Regions-Of-Interest Selection;
- ✓ Section 4.2.3: Collision Risk Estimation.

4.2.1 Dynamic Power Management

The autonomous robots have some degree of self-sufficiency. For advances in the energy autonomy, robots will need to extract energy from the environment. The state conditions in combination with environment events are an important variable of the system, which must have knowledge of its available resources as well as its components, their desired performance characteristics and their current status. In this way, Dynamic Power Management (DPM) encompasses a set of techniques that achieves energy-efficient computation by selectively turning off (or reducing the performance of) system components when they are idle (or partially unexploited) (Benini, et al., 2000).

As presented earlier in chapter 2, in automobiles, the stop-start systems automatically shut down and restart the internal combustion engine improving fuel economy and reducing emissions. On the other hand, thanks to the great deal of information that it can deliver, the new architectures based on multiple sensors (fusion technology) can increase the perception-safety level. When incorporating several types of sensors, there is an increase of autonomy and "intelligence" degrees, especially in relation to navigation in unknown environments. However, it

also increases energy consumption. In contrast, the great amount of information would not necessarily lead to better decisions and could also harm the performance of the system, overloading it. This section then presents results of a system that observes the environment and enables energy savings based on monocular vision. This visual-observer methodology may be extended to other sensors and components.

The section 2.4.2 introduced a logical dynamic optimization methodology based on PCC. This method observes if there are no significant changes in the environment, permitting that some logical components could be shut down to save processor energy consumption, and/or to make the CPU available for running concurrent processes.
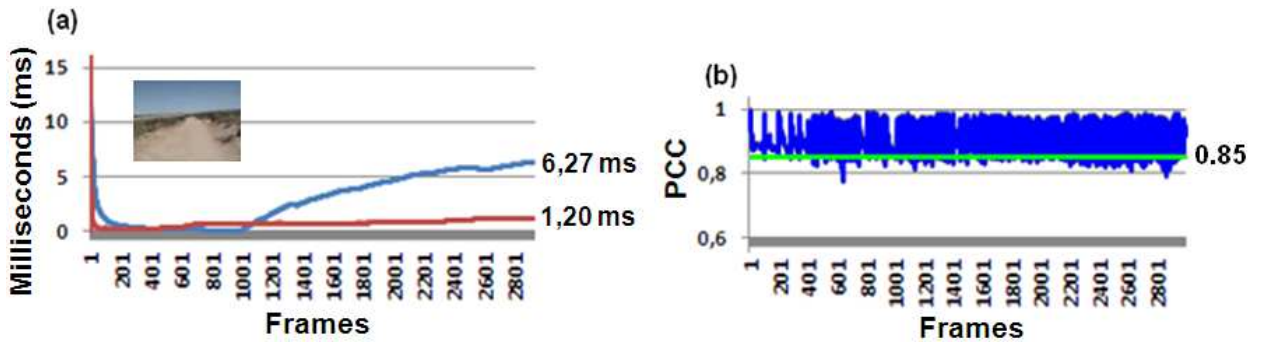
Figure 4.2 – DARPA Desert text-set (DARPA, 2005): The computational mean time of a horizon finding algorithm in unknown and urban environment (Miranda Neto, et al., 2011); (a) The red line: the computational mean time was 1.20 ms with the discarding criteria; (a) The blue line: the computational mean time was 6.27 ms without the discarding criteria; (b) The green line: an empirical PCC threshold equal to 0.85; (b) In blue: DPM performance based on discarding criteria: above the green line, it presents all discarded images.

Based on DARPA Desert text-set (DARPA, 2005), as an example of the potential of this proposal, the Figure 4.2 shows the computational gain by using the discarding criteria. From an empirical PCC threshold equal to 0.85 (Miranda Neto, et al., 2011), the red line shows that the computational mean time was 1.20 ms, against 6.27 ms without the discarding criteria. In Figure

4.2 (b), above the green line, it presents all discarded images. A video containing all images of this sequence is available in[33].

To better understand, Figures 4.3 and 4.4 present the performance in real, dynamic and unknown environments. From displacements on the outskirts of the Karlsruhe Institute of Technology (KIT) in Germany, these data were obtained in (Geiger, 2011). It contains high-quality image sequences (1344x391 pixels) recorded at 10 frames per second from a moving vehicle using a Pointgrey Flea2 firewire camera. For these cases, in order to reduce the number of data, it includes a resolution reduction of image to 250x120 pixels.

Figure 4.3 (a) presents the cumulative impact computations (ms) versus the performance gain obtained by using the discarding criteria. In the same way, Figure 4.3 (b) shows the number of frames versus the number of discarded frames by using the discarding criteria. In Figure 4.3 (c) the percentage of discarded frames is presented in blue. The black vertical lines present an analysis window with the performance evaluation of the discarding criteria at red lights. This analysis window represents a time interval of about 3 minutes, which comprises the interval between frames 913 and 2375, where it was possible to reject 1454 images of a total of 1463 (99%). A video containing all images of this sequence is available in[34].

Figure 4.4 (a) presents the cumulative impact computations (ms) versus the performance gain obtained by using the discarding criteria. In the same way, Figure 4.4 (b) shows the number of frames versus the number of discarded frames by using the discarding criteria. In Figure 4.4 (c) the percentage of discarded frames is presented in blue. The black vertical lines present an analysis window with the performance evaluation of the discarding criteria at red lights. This analysis window represents a time interval of about 40 seconds, which comprises the interval between frames 105 and 480, where it was possible to reject 374 images of a total of 376 (99%). A video containing all images of this sequence is available in[35].

---

[33] http://www.youtube.com/watch?v=XaZndmMaieE – visited on October 1, 2011.
[34] http://www.youtube.com/watch?v=pLI1c7kLKk0 – visited on October 1, 2011.
[35] http://www.youtube.com/watch?v=A99WtcvOs9U – visited on October 1, 2011.
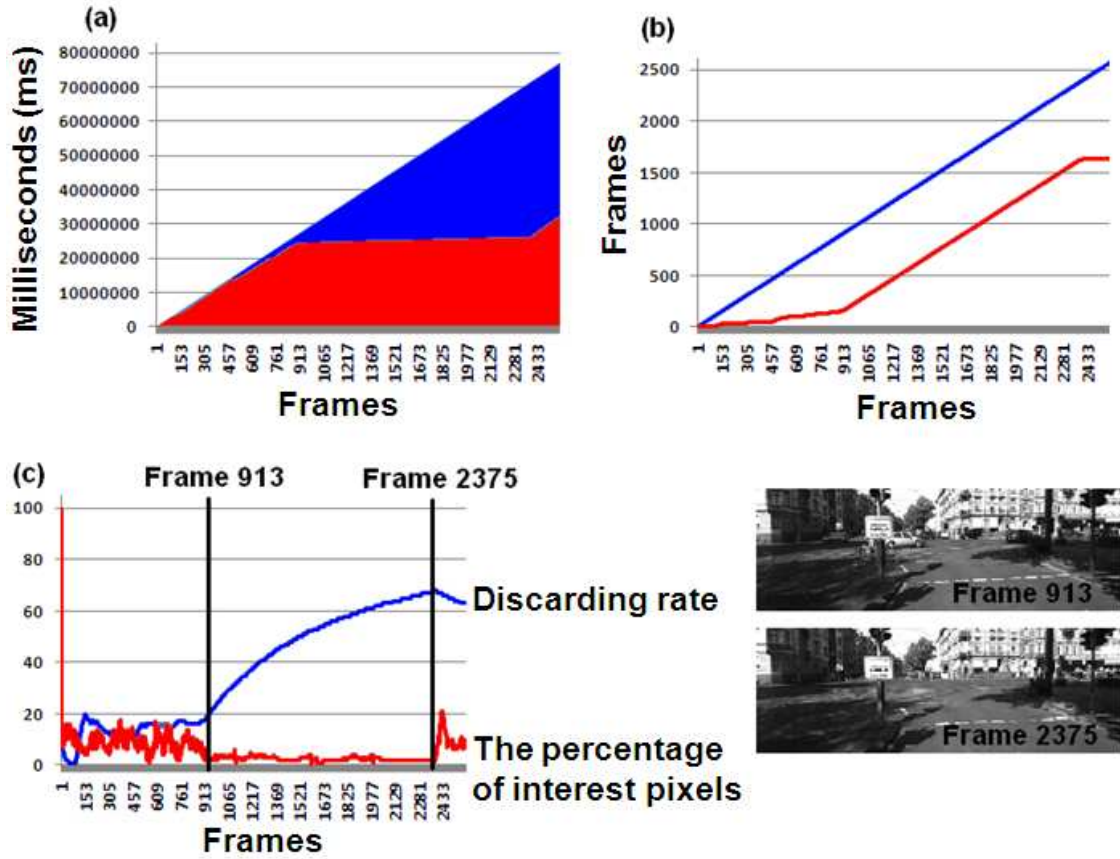
Figure 4.3 – Karlsruhe Dataset on September, 8th 2009 – drive 0012 (Geiger, 2011): (a) In blue: the cumulative impact computations (ms); In red: the cumulative computations (ms) by using the discarding criteria. (b) In blue: the number of frames; In red: the number of discarded frames by using the discarding criteria. (c) In blue: discarding rate; In red: the percentage of interest pixels; In the analysis window, represented by two black vertical lines, the performance evaluation of the discarding criteria at red lights.
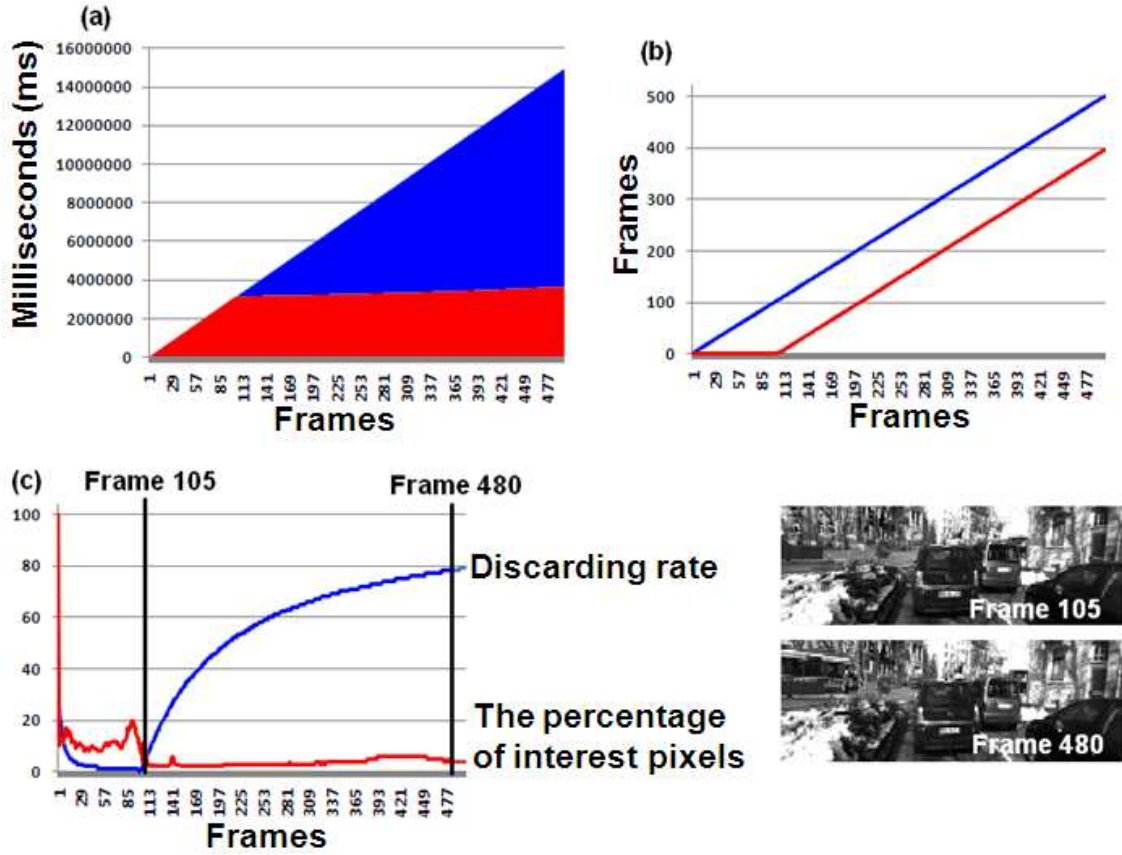
Figure 4.4 – Karlsruhe Dataset on March, 9th 2010 – drive 0082 (Geiger, 2011): (a) In blue: the cumulative impact computations (ms); In red: the cumulative computations (ms) by using the discarding criteria. (b) In blue: the number of frames; In red: the number of discarded frames by using the discarding criteria. (c) In blue: discarding rate; In red: the percentage of interest pixels; In the analysis window, represented by two black vertical lines, the performance evaluation of the discarding criteria at red lights.

Taken displacements on the outskirts of the Heudiasyc Laboratory in France, the next text-set contains image sequences (320x240 pixels) recorded at 20 frames per second from a moving vehicle using a Sony DFW-VL500 camera. This stage of testing also evaluates the proposed algorithm at low and high speeds on real-time conditions using the Carmen and Strada vehicles shown in Figures 4.1 (a) and (b). For these cases, in order to reduce the number of data, an image resolution reduction to 160x120 pixels was introduced.

In this way, taken displacements performed in the downtown Compiegne, Figure 4.5 (a) presents the cumulative impact computations (ms) versus the performance gain obtained by using the discarding criteria. Figure 4.5 (b) shows the number of frames versus the number of discarded frames by using the discarding criteria. In Figure 4.5 (c) the percentage of discarded frames is presented in blue. The black vertical lines present an analysis window with the performance evaluation of the discarding criteria in a narrow street. This analysis window represents a time interval of about 0.25 seconds, which comprises the interval between frames 2227 and 2232, where it was possible to reject 4 images of a total of 5 (80%). A video containing all images of this sequence is available in[36].
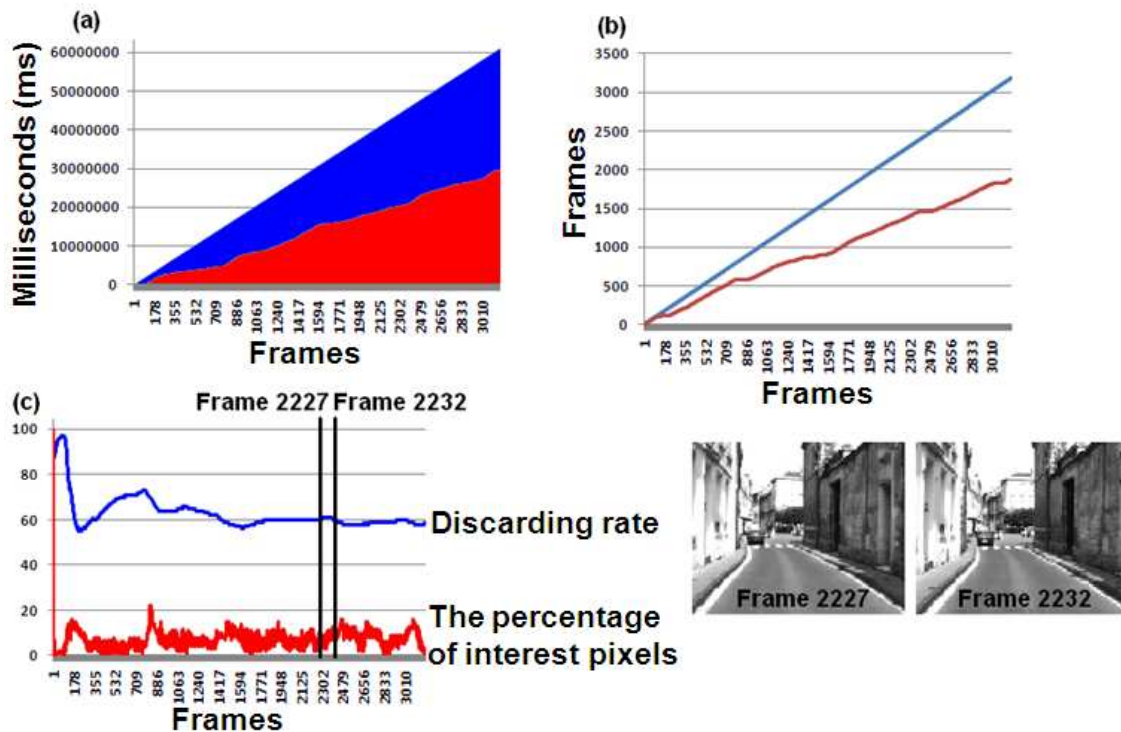
Figure 4.5 – Heudiasyc Laboratory in France, 2005, Strada: (a) In blue: the cumulative impact computations (ms); In red: the cumulative computations (ms) by using the discarding criteria. (b) In blue: the number of frames; In red: the number of discarded frames by using the discarding criteria. (c) In blue: discarding rate; In red: the percentage of interest pixels; In the analysis window, represented by two black vertical lines, the performance evaluation of the discarding criteria in a narrow street.

Especially in the urban context, a real-time system must satisfy explicit response-time constraints. The compromise between processing time and images acquisition is therefore fundamental. In this way, Figure 4.6 (a) presents the computational mean time of a horizon finding algorithm in unknown and urban environment (Miranda Neto, et al., 2011). This algorithm was also presented in Chapter 3. In this way, from an empirical PCC threshold equal to 0.85, the red line shows that the computational mean time was 5.09 ms, against 15.62 ms without the discarding criteria. In Figure 4.6 (b), above the green line, it presents all discarded images.
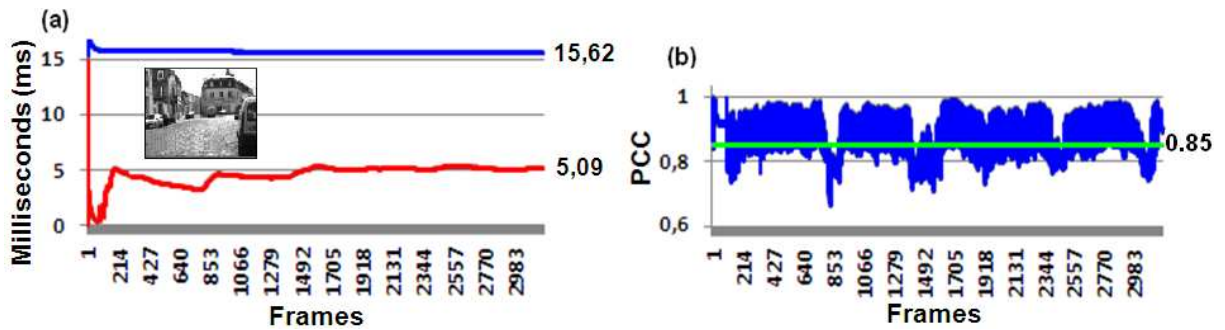
Figure 4.6 – Heudiasyc Laboratory in France, 2005, Strada: The computational mean time of a horizon finding algorithm in unknown and urban environment (Miranda Neto, et al., 2011); (a) The red line: the computational mean time was 5.09 ms with the discarding criteria; (a) The blue line: the computational mean time was 15.62 ms without the discarding criteria; (b) The green line: an empirical PCC threshold equal to 0.85; (b) In blue: DPM performance based on discarding criteria: above the green line, it presents all discarded images.

In the following two cases, the speed informations were obtained in real time. As shown in Figure 4.7 (c), values greater than 100 Km/h can be found (green line). In this context, for mobile robots, the tasks' deadlines are different at different traveling speeds. It is therefore important to notice that there is no diffeomorphism between the vehicle speed and the PCC variation, because if there are no changes between consecutive frames, the PCC threshold remains static. In this case, the isomorphism cannot be guaranteed and it ensures more efficiency for the discarding criteria.

In the first case, taken displacements performed on a highway next to Compiegne, Figure 4.7 (a) presents the cumulative impact computations (ms) versus the performance gain obtained by using the discarding criteria. Figure 4.7 (b) shows the number of frames versus the number of discarded frames by using the discarding criteria. In Figure 4.7 (c) the percentage of discarded frames is presented in blue. The black vertical lines present an analysis window with the performance evaluation of the discarding criteria in acceleration from 37 Km/h to 86 Km/h. This analysis window represents a time interval of about 30 seconds, which comprises the interval between frames 1800 and 2370, where it was possible to reject 430 images of a total of 571 (75%). A video containing all images of this sequence is available in[37].
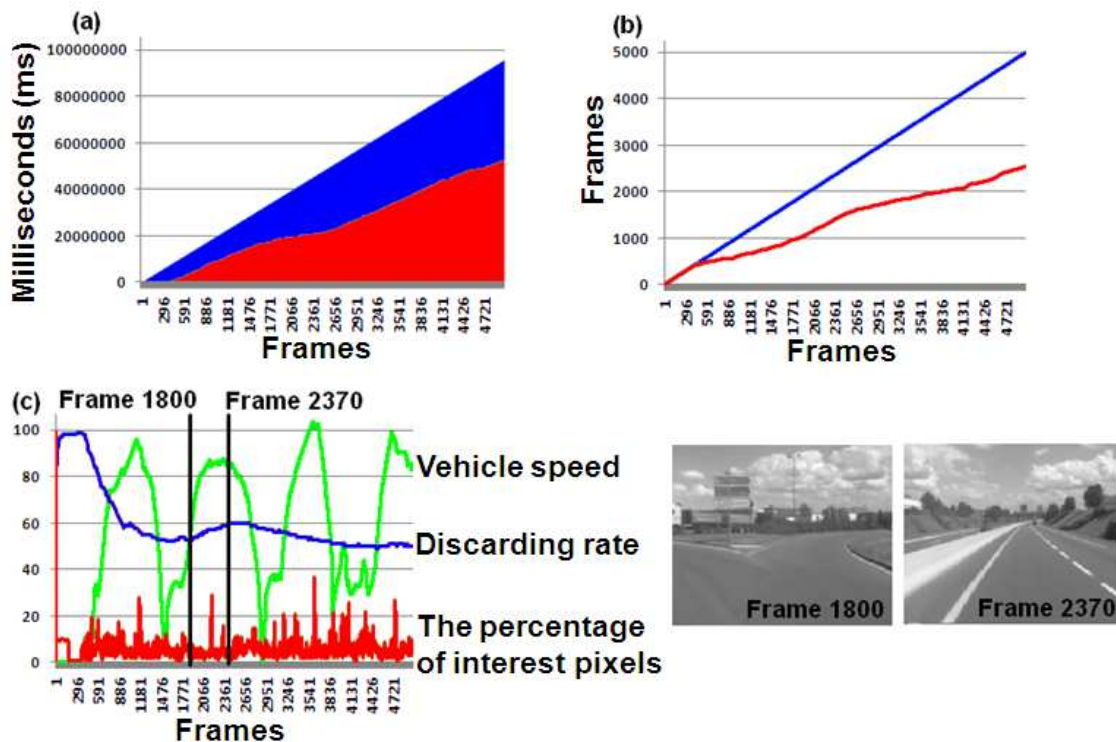
Figure 4.7 – Heudiasyc Laboratory in France, 2010, Carmen: (a) In blue: the cumulative impact computations (ms); In red: the cumulative computations (ms) by using the discarding criteria. (b) In blue: the number of frames; In red: the number of discarded frames by using the discarding criteria. (c) In blue: discarding rate; In red: the percentage of interest pixels; In green: The vehicle speed; In the analysis window, represented by two black vertical lines, the performance evaluation of the discarding criteria in acceleration from 37 Km/h to 86 Km/h.

---

[37] http://www.youtube.com/watch?v=WpViME5XaeQ – visited on October 1, 2011.

In the second case, taken displacements performed in the Forest of Compiegne, Figure 4.8 (a) presents the cumulative impact computations (ms) versus the performance gain obtained by using the discarding criteria. Figure 4.8 (b) shows the number of frames versus the number of discarded frames by using the discarding criteria. In Figure 4.8 (c) the percentage of discarded frames is presented in blue. The black vertical lines present an analysis window with the performance evaluation of the discarding criteria in overtaking task. This analysis window represents a time interval of about 3 seconds, which comprises the interval between frames 1700 and 1755, where it was possible to reject 7 images of a total of 55 (12%). A video containing all images of this sequence is available in[38].
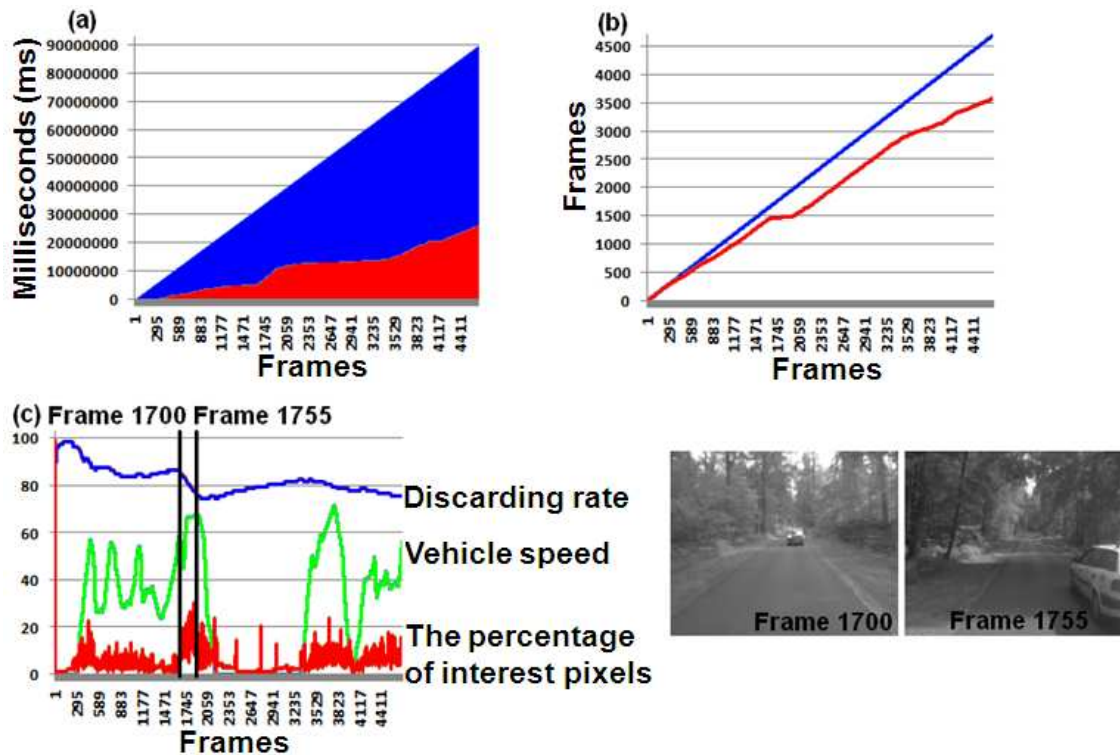
Figure 4.8 – Heudiasyc Laboratory in France, 2010, Carmen: (a) In blue: the cumulative impact computations (ms); In red: the cumulative computations (ms) by using the discarding criteria. (b) In blue: the number of frames; In red: the number of discarded frames by using the discarding criteria. (c) In blue: discarding rate; In red: the percentage of interest pixels; In green: The vehicle speed; In the analysis window, represented by two black vertical lines, the performance evaluation of the discarding criteria in overtaking task.

---

[38] http://www.youtube.com/watch?v=YPFuzAH8P7c – visited on October 1, 2011.

The following three cases present autonomous displacements at Center for Information Technology Renato Archer (CTI) in Brazil. An open-loop reactive scheme was used to navigate the vehicle. These stages of testing evaluate the proposed algorithm at low speed on real-time conditions using the VERO platform shown in Figure 4.1 (c). These text-sets contain image sequences (640x480 pixels) recorded at a frame rate of 25 frames per second from a moving vehicle using a Sony DFW-VL500 camera. For these cases, in order to reduce the number of data, an image resolution reduction to 160x120 pixels was introduced.

The first case is presented in Figure 4.9 (a), which includes the cumulative impact computations (ms) versus the performance gain obtained by using the discarding criteria in off-road context. Figure 4.9 (b) shows the number of frames versus the number of discarded frames by using the discarding criteria. In Figure 4.9 (c) the percentage of discarded frames is presented in blue. The black vertical lines present an analysis window with the performance evaluation of the discarding criteria in area (road image) with high noise level. It is part of the task execution to go through a gate. This analysis window represents a time interval of about 8 seconds, which comprises the interval between frames 303 and 490, where it was possible to reject 144 images of a total of 188 (76%). A video containing all images of this sequence is available in[39].

In Figure 4.10 (a), the second case also presents the cumulative impact computations (ms) versus the performance gain obtained by using the discarding criteria in off-road context. In the same way, Figure 4.10 (b) shows the number of frames versus the number of discarded frames by using the discarding criteria. In Figure 4.10 (c) the percentage of discarded frames is presented in blue. The black vertical lines present an analysis window with the performance evaluation of the discarding criteria in real-time obstacle avoidance and reactive navigation. This analysis window represents a time interval of about 4 seconds, which comprises the interval between frames 700 and 777, where it was possible to reject 17 images of a total of 77 (22%). A video containing all images of this sequence is available in[40].

As expected and shown in Figure 4.10 (c), in grass areas, and even on the parallelepiped streets or where an excessive noise is observed, the efficiency of the method decreases. However,

---

[39] http://www.youtube.com/watch?v=F7g9GkY4LEQ – visited on October 1, 2011.
[40] http://www.youtube.com/watch?v=lfamoWAtiJs – visited on October 1, 2011.

in this case the discarding rate remains over 60% most of the time. On the other hand, in Figure 4.9 (c) the discarding rate remains over 85% most of the time.
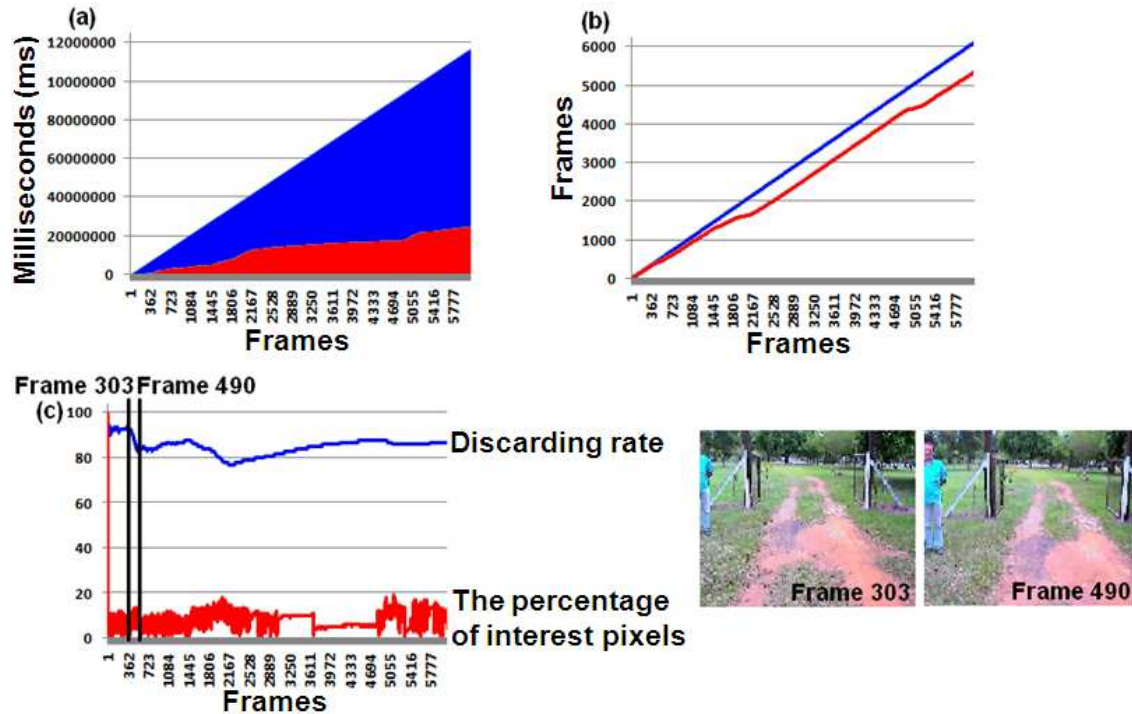
Figure 4.9 – Renato Archer IT Center (CTI) on November, 4th 2010 – VERO platform in off-road context: (a) In blue: the cumulative impact computations (ms); In red: the cumulative computations (ms) by using the discarding criteria. (b) In blue: the number of frames; In red: the number of discarded frames by using the discarding criteria. (c) In blue: discarding rate; In red: the percentage of interest pixels; The black vertical lines present an analysis window with the performance evaluation of the discarding criteria in area (road image) with high noise level.
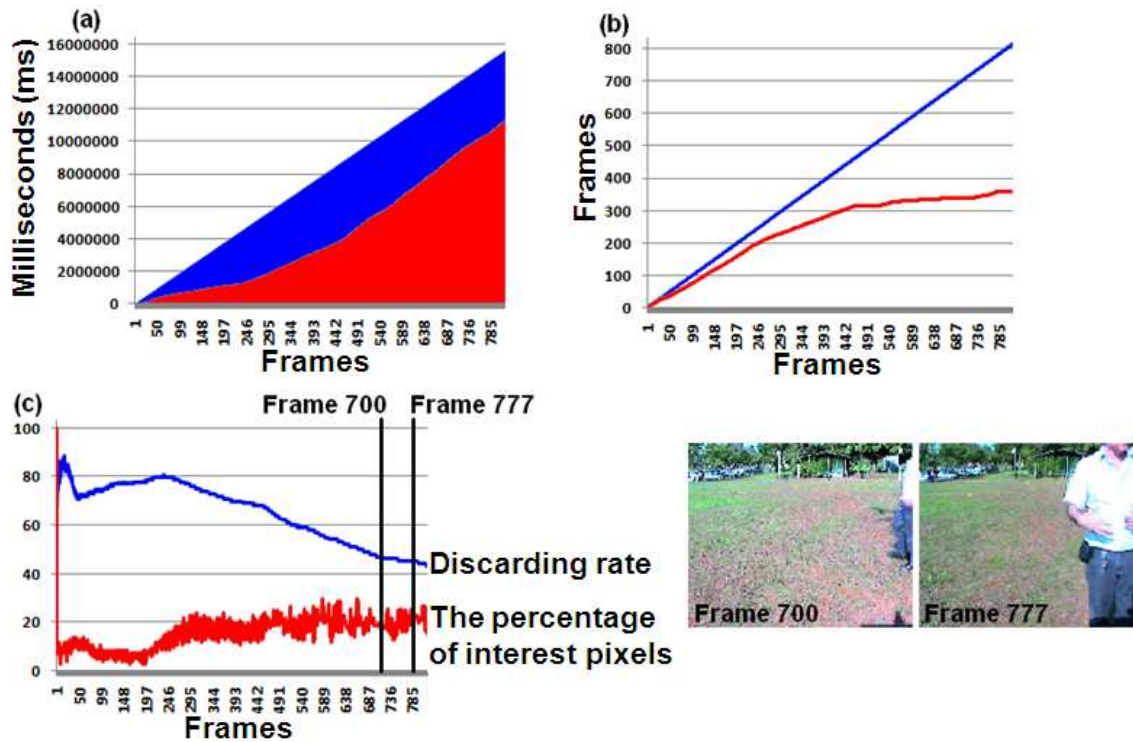
Figure 4.10 – Renato Archer IT Center (CTI) on November, 4th 2010 – VERO platform in off-road context: (a) In blue: the cumulative impact computations (ms); In red: the cumulative computations (ms) by using the discarding criteria. (b) In blue: the number of frames; In red: the number of discarded frames by using the discarding criteria. (c) In blue: discarding rate; In red: the percentage of interest pixels; The black vertical lines present an analysis window with the performance evaluation of the discarding criteria in real-time obstacle avoidance and reactive navigation.

The last case in open-loop reactive way is presented in Figure 4.11. In this case, the main task of the system was to perform displacements toward the middle of a paved avenue (road). Figure 4.11 (a) presents the cumulative impact computations (ms) versus the performance gain obtained by using the discarding criteria. In the same way, Figure 4.11 (b) shows the number of frames versus the number of discarded frames by using the discarding criteria. In Figure 4.11 (c) the percentage of discarded frames is presented in blue. The black vertical lines present an analysis window with the performance evaluation of the discarding criteria while the system performs the displacements toward the middle of a road. This analysis window represents a time

interval of about 40 seconds, which comprises the interval between frames 1 and 930, where it was possible to reject 844 images of a total of 930 (90%). A video containing all images of this sequence is available in[41].
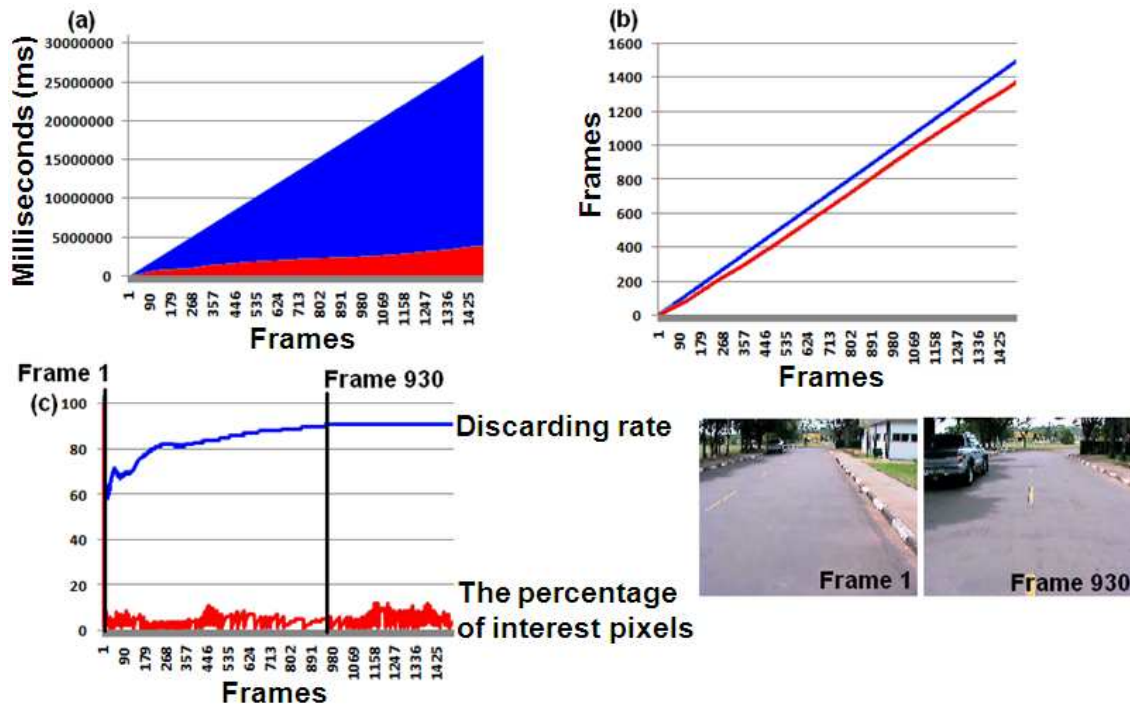

Figure 4.11 – Renato Archer IT Center (CTI) on November, 4th 2010 – VERO platform in off-road context: (a) In blue: the cumulative impact computations (ms); In red: the cumulative computations (ms) by using the discarding criteria. (b) In blue: the number of frames; In red: the number of discarded frames by using the discarding criteria. (c) In blue: discarding rate; In red: the percentage of interest points; The black vertical lines present an analysis window with the performance evaluation of the discarding criteria in displacements toward the middle of a road.

[41] http://www.youtube.com/watch?v=HYfqgTUlK7s – visited on October 1, 2011.

4.2.2 Automatic Regions-Of-Interest Selection

As has been shown in Chapter 2, in a certain analysis window (pair of frames), if the obstacle/object occupies a big portion of the scene, the PCC threshold tends to be low. Conversely, if obstacle/object occupies a small portion of the frame, it means that it is away from the vehicle and the system will have time enough to react. In this context, the question was: where are these interest points/pixels? Or, which pixels (ROI) of the pair of images contributed most to the computed Pearson's coefficient? Which of them really need to be reprocessed (or resent to a server)? Currently, many projects use remote processing and this proposal could also optimize the data transmission over the network.

The section 2.5.1 presented that for each pair of pixels analyzed in Equation 2.1, the only possible result is: [-1 or +1]. That is, all pixels with intensities below the means obtained in Equation A.1 are candidates for interest points (ROI). It was presented in the Figure 2.9 (c), (f) and (i). In red, the percentage of interest pixels for each frame was also presented in: Figure 4.3 (c), Figure 4.4 (c), Figure 4.5 (c), Figure 4.7 (c), Figure 4.8 (c), Figure 4.9 (c), Figure 4.10 (c) and Figure 4.11 (c).

4.2.2.1 Navigable Area Detection

Chapter 3 presented implementation details of a method for navigable area detection based on Otsu thresholding. As shown in Figure 1.10: Layer (b) Navigable Area Detection, this task is performed only when the PCC indicates that there is a low correlation between the reference frame and the current frame. Otherwise, the current frame is discarded without being processed. When an image is discarded, the system keeps the previous segmentation result, which is linked to an Otsu threshold. This stored threshold represents the navigable area. Consequently, in order to further improve the navigable area detection, for each discarded image by the discarding criteria in Figure 2.1, it classifies only the interest pixels (ROI) obtained in Equation 2.1 from the last Otsu threshold obtained by the navigable area detection algorithm, i.e. Figure 1.10 (b): message B[4].

For the cases presented here, the interest pixels obtained in Equation 2.1 represent $r_2 = -1$. As a first example, from a stored Otsu threshold, in Figure 4.12 (b), (e) and (h) the interest pixels classified as navigable are represented in blue; the interest pixels classified as obstacle are represented in yellow. A video containing all images of this sequence is available in[42].
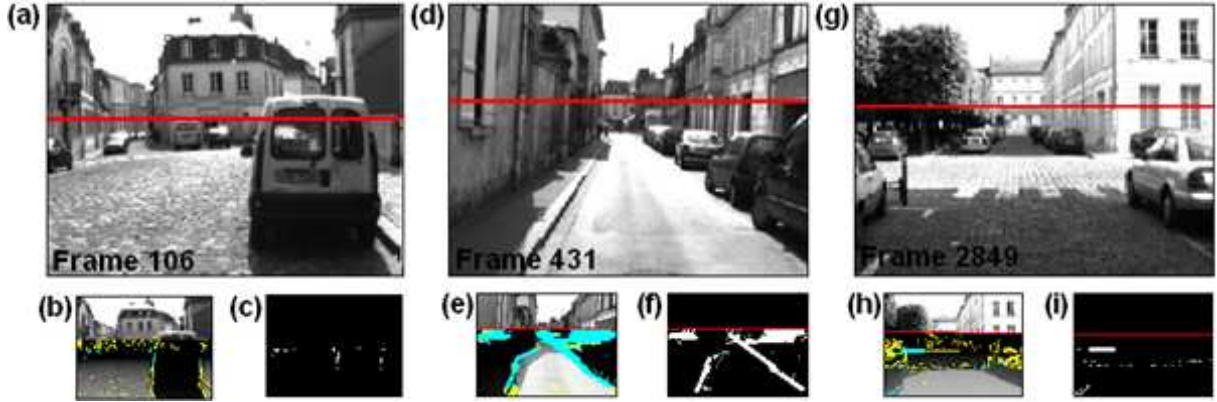
Figure 4.12 – Heudiasyc Laboratory in France, 2005, Strada: (a), (d) and (g) are frames in real environments; (b), (e) and (h) the interest pixels classified as navigable are represented in blue; the interest pixels classified as obstacle are represented in yellow; (f) and (i): line detection using Hough transform from navigable area pixels; The red horizontal lines represent the horizon finding algorithm results (Miranda Neto, et al., 2011).

At first stage of testing, based on displacements on the outskirts of the Karlsruhe Institute of Technology in Germany, the images of the Figure 4.13 were obtained in (Geiger, 2011). It contains high-quality image sequences (1344x391 pixels) recorded at 10 frames per second from a moving vehicle using a Pointgrey Flea2 firewire camera. For these cases, in order to reduce the number of data, an image resolution reduction to 250x120 pixels was introduced. As following, in the top images of the Figure 4.13 (a) to (f), the interest pixels classified as navigable are represented in blue; and the interest pixels classified as obstacle are represented in yellow:

---

[42] http://www.youtube.com/watch?v=czR-SGPeG7w – visited on October 1, 2011.

✓ Figure 4.13 (a): Frame 446 of the drive 0010; The interest pixels represent 12% of the image;

✓ Figure 4.13 (b): Frame 560 of the drive 0010; The interest pixels represent 6% of the image;

✓ Figure 4.13 (c): Frame 45 of the drive 0012; The interest pixels represent 10% of the image;

✓ Figure 4.13 (d): Frame 451 of the drive 0012; The interest pixels represent 13% of the image;

✓ Figure 4.13 (e): Frame 925 of the drive 0012; The interest pixels represent 4% of the image;

✓ Figure 4.13 (f): Frame 750 of the drive 0016; The interest pixels represent 7% of the image.

Videos containing all images sequence of the Figure 4.13 are presented in[43,44,45].

---

[43] http://www.youtube.com/watch?v=Mcc15y1VBoQ – visited on October 1, 2011.
[44] http://www.youtube.com/watch?v=pLI1c7kLKk0 – visited on October 1, 2011.
[45] http://www.youtube.com/watch?v=5ao_pX9Xa_s – visited on October 1, 2011.
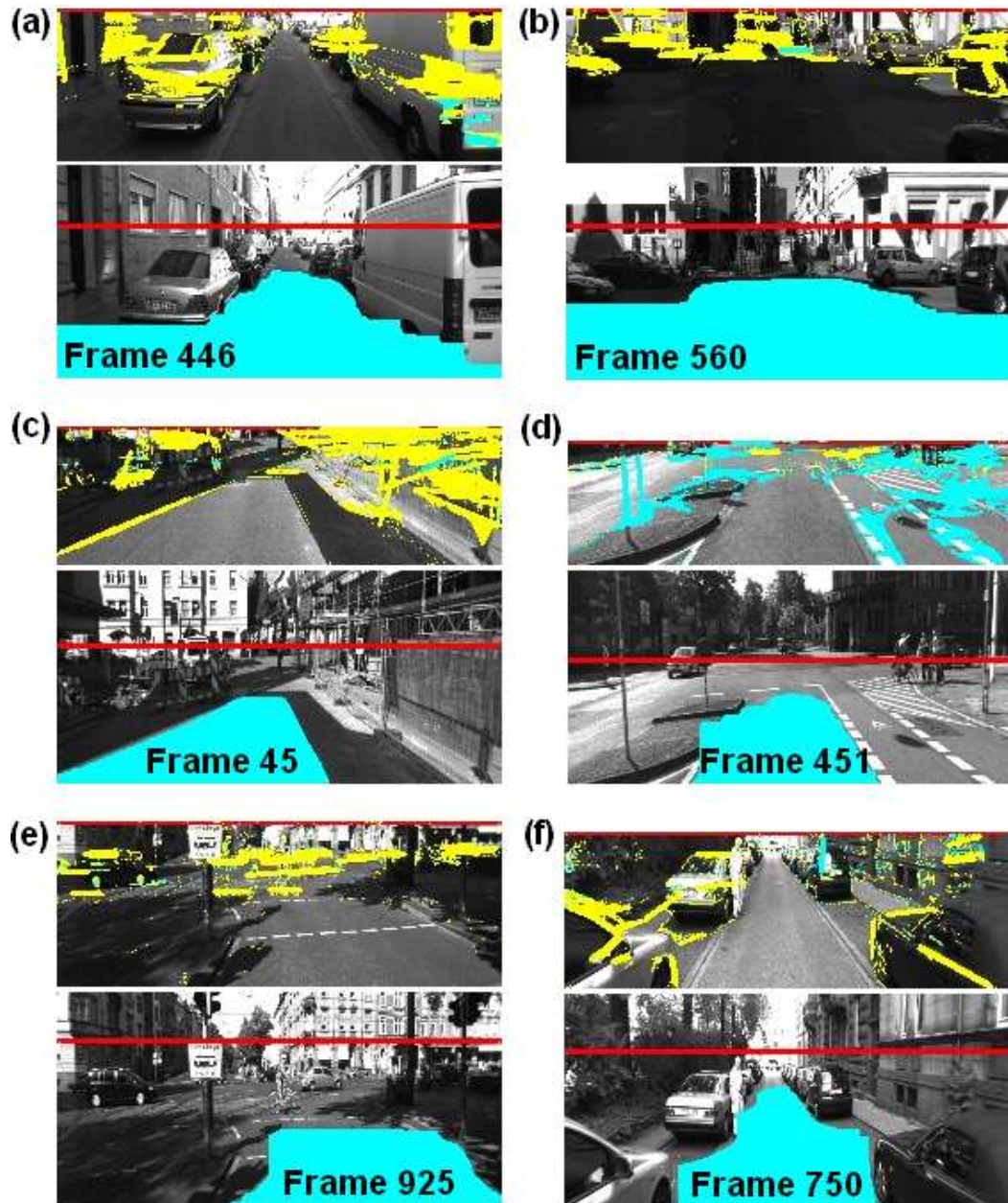
Figure 4.13 – Karlsruhe Dataset on September, 8th 2009 (Geiger, 2011): From a stored Otsu threshold, in the top images, the interest pixels classified as navigable area are represented in blue; and the interest pixels classified as obstacle are represented in yellow; after line detection using Hough transform; In the bottom images, the stored Otsu threshold is represented in blue; The red horizontal lines represent the horizon finding algorithm results (Miranda Neto, et al., 2011).

The second stage of testing evaluates the proposed algorithm at high speed on real-time conditions using the Carmen vehicle shown in Figures 4.1 (a). The speed informations were obtained in real time on a highway next to Compiegne. As shown in Figure 4.7 (c), values greater than 100 Km/h can be found. This text-set contains image sequences (320x240 pixels) recorded at 20 frames per second from a moving vehicle using a Sony DFW-VL500 camera. For these cases, in order to reduce the number of data, an image resolution reduction to 160x120 pixels was introduced. As following, in the top images of the Figure 4.14 (a) to (f), the interest pixels classified as navigable area are represented in blue; and the interest pixels classified as obstacle are represented in yellow:

- ✓ Figure 4.14 (a): Frame 1106; Speed: 93.8 Km/h; The interest pixels represent 3% of the image;
- ✓ Figure 4.14 (b): Frame 1133; Speed: 96 Km/h; The interest pixels represent 4% of the image;
- ✓ Figure 4.14 (c): Frame 3527; Speed: 97.01 Km/h; The interest pixels represent 5% of the image;
- ✓ Figure 4.14 (d): Frame 3563; Speed: 100.2 Km/h; The interest pixels represent 2% of the image;
- ✓ Figure 4.14 (e): Frame 4740; Speed: 96.35 Km/h; The interest pixels represent 6% of the image;
- ✓ Figure 4.14 (f): Frame 4760; Speed: 94.66 Km/h; The interest pixels represent 6% of the image.

A video containing all images of this sequence is available in[46].

---

[46] http://www.youtube.com/watch?v=WpViME5XaeQ – visited on October 1, 2011.
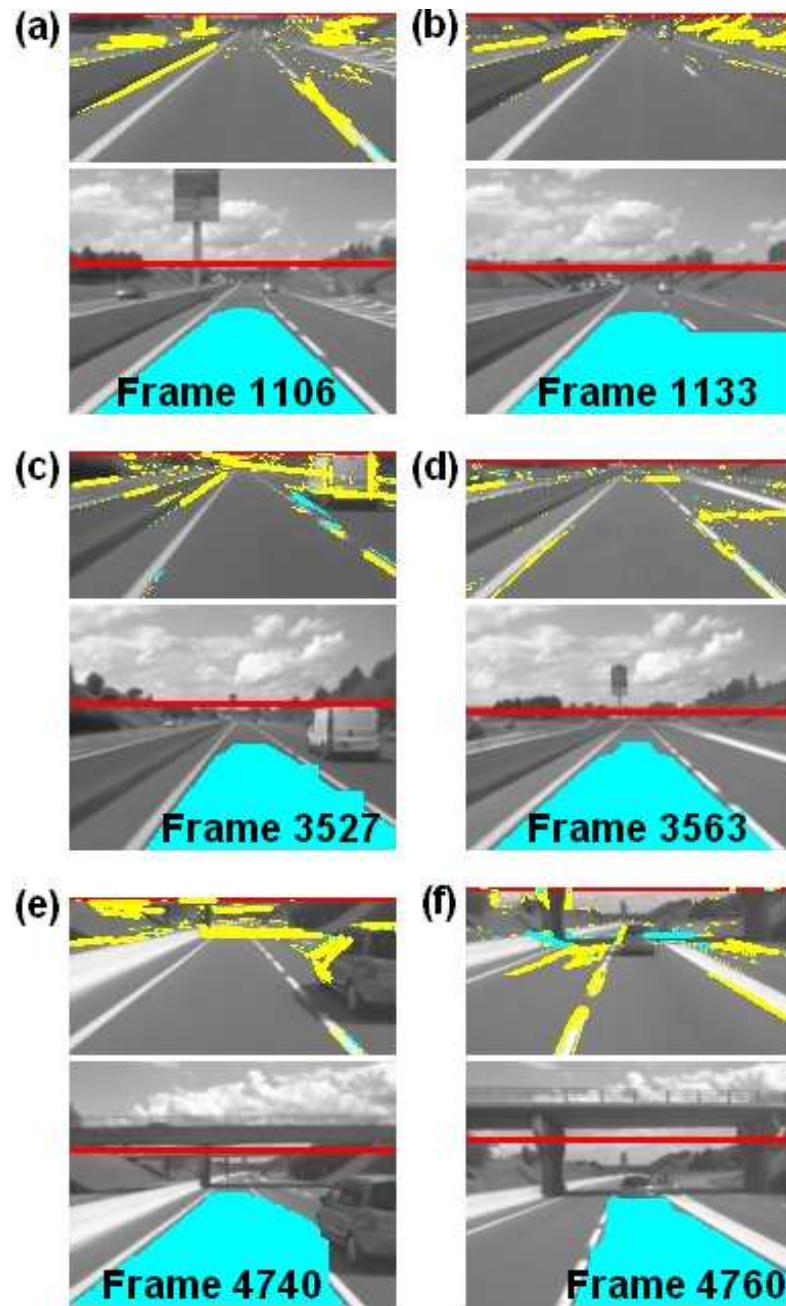
Figure 4.14 – Heudiasyc Laboratory in France, 2010, Carmen: From a stored Otsu threshold, in the top images, the interest pixels classified as navigable area are represented in blue; and the interest pixels classified as obstacle are represented in yellow; after line detection using Hough transform; In the bottom images, the stored Otsu threshold is represented in blue; The red horizontal lines represent the horizon finding algorithm results (Miranda Neto, et al., 2011).

4.2.2.2 Obstacle Avoidance

As presented earlier in section 2.5.4, the Pearson's method explores regression and correlation aspects. Consequently, from an empirical PCC threshold equal to 0.65, it puts in evidence the past history properties. As shown in Figure 1.10: Layer (b) Navigable Area Detection, the navigable area detection task is performed only when the PCC indicates that there is a low correlation between the reference frame and the current frame. In a computational process running in parallel, for each frame processed, evidencing the past history properties from a PCC threshold equal to 0.65, another old Otsu threshold is also stored. Besides the results presented in the Figures 4.12, 4.13 and 4.14 (i.e. the interest pixels classified as obstacle represented in yellow), this procedure also allows a greater level of security, especially when the camera does not "see" the navigable area (i.e. in front of a wall), as shown in the Figure 4.15 (a), (c) and (d).

Therefore, in order to further improve the obstacle detection, for each discarded image by the discarding criteria in Figure 2.1, it classifies only the interest pixels obtained in Equation 2.1 from an old Otsu threshold obtained by the navigable area detection algorithm. It happens when the PCC threshold is equal or less to 0.65, i.e. Figure 1.10 (b): message B[4]. As following, in the top images of the Figure 4.15 (a) to (g), the interest pixels classified as navigable area are represented in blue; and the interest pixels classified as obstacle are represented in yellow:

✓ Figure 4.15 (a): Heudiasyc Laboratory in France, 2005, Strada: Frame 834; The interest pixels represent 22% of the image;

✓ Figure 4.15 (b): Heudiasyc Laboratory in France, 2005, Strada: Frame 925; The interest pixels represent 15% of the image;

✓ Figure 4.15 (c): Heudiasyc Laboratory in France, 2005, Strada: Frame 1500; The interest pixels represent 10% of the image;

✓ Figure 4.15 (d): Heudiasyc Laboratory in France, 2005, Strada: Frame 1525; The interest pixels represent 12% of the image;

✓ Figure 4.15 (e): Karlsruhe Dataset on March, 9th 2010 – drive 0081 (Geiger, 2011): Frame 50; The interest pixels represent 13% of the image;

✓ Figure 4.15 (f): Karlsruhe Dataset on March, 9th 2010 – drive 0081 (Geiger, 2011): Frame 118; The interest pixels represent 20% of the image;

✓ Figure 4.15 (g): Karlsruhe Dataset on March, 9th 2010 – drive 0081 (Geiger, 2011): Frame 163; The interest pixels represent 22% of the image.
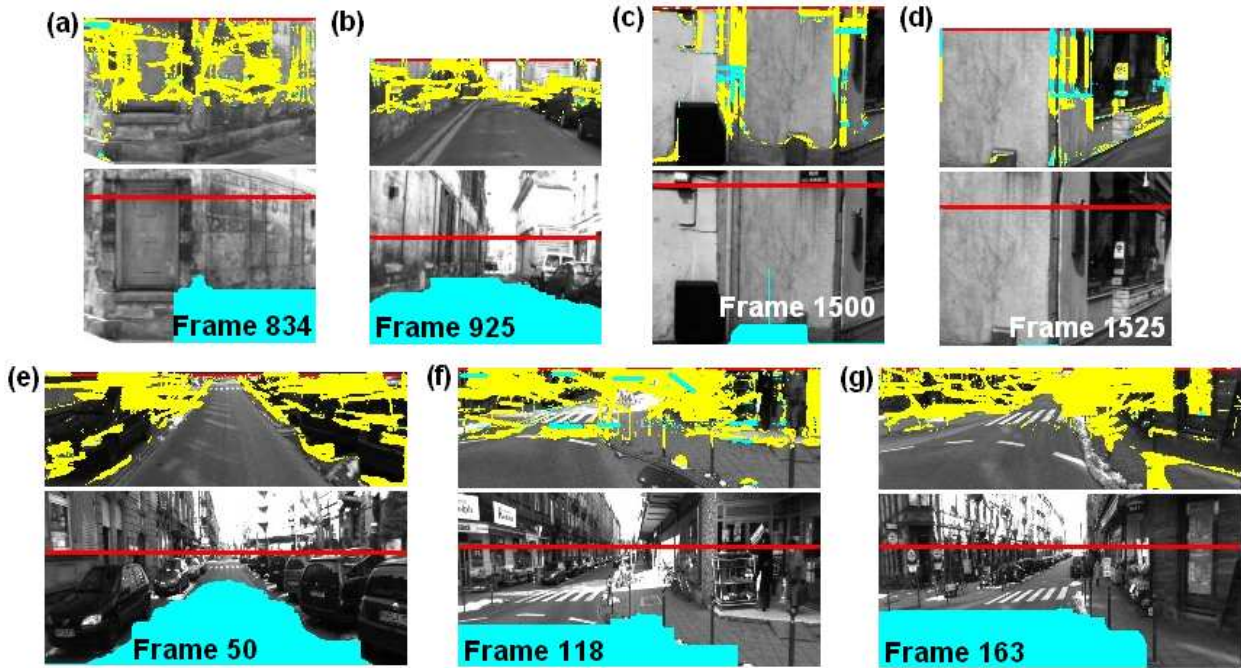
Figure 4.15 – Obstacle avoidance from an old stored Otsu threshold (urban context and in front of a wall): In the top images, the interest pixels classified as navigable area are represented in blue; and the interest pixels classified as obstacle are represented in yellow; after line detection using Hough transform; In the bottom images, the stored Otsu threshold is represented in blue; The red horizontal lines represent the horizon finding algorithm results (Miranda Neto, et al., 2011).

As shown in Figure 4.16, in grass areas: (a), (b) and (c); on the parallelepiped streets: (d), (e) and (f); or where an excessive noise is observed, the efficiency of the method may decrease considerably, what can be improved with the application of a smoothing filter, and/or from region-merging algorithm that mainly aims to represent homogeneous regions i.e. as presented in section 2.6.2. As following, in the top images of the Figure 4.16 (a) to (f), the interest pixels

classified as navigable area are represented in blue; and the interest pixels classified as obstacle are represented in yellow:

- ✓ Figure 4.16 (a): Renato Archer IT Center (CTI) on November, 4th 2010 – VERO platform: Frame 93; The interest pixels represent 9% of the image;
- ✓ Figure 4.16 (b): Renato Archer IT Center (CTI) on November, 4th 2010 – VERO platform: Frame 97; The interest pixels represent 11% of the image;
- ✓ Figure 4.16 (c): Renato Archer IT Center (CTI) on November, 4th 2010 – VERO platform: Frame 730; The interest pixels represent 26% of the image;
- ✓ Figure 4.16 (d): Heudiasyc Laboratory in France, 2005, Strada: Frame 34; The interest pixels represent 1% of the image;
- ✓ Figure 4.16 (e): Heudiasyc Laboratory in France, 2005, Strada: Frame 166; The interest pixels represent 15% of the image;
- ✓ Figure 4.16 (f): Heudiasyc Laboratory in France, 2005, Strada: Frame 317; The interest pixels represent 5% of the image.

Videos containing all images sequence for this section are presented in[47, 48, 49].

---

[47] http://www.youtube.com/watch?v=czR-SGPeG7w – visited on October 1, 2011.
[48] http://www.youtube.com/watch?v=DhexaG489Sc – visited on October 1, 2011.
[49] http://www.youtube.com/watch?v=lfamoWAtiJs – visited on October 1, 2011.
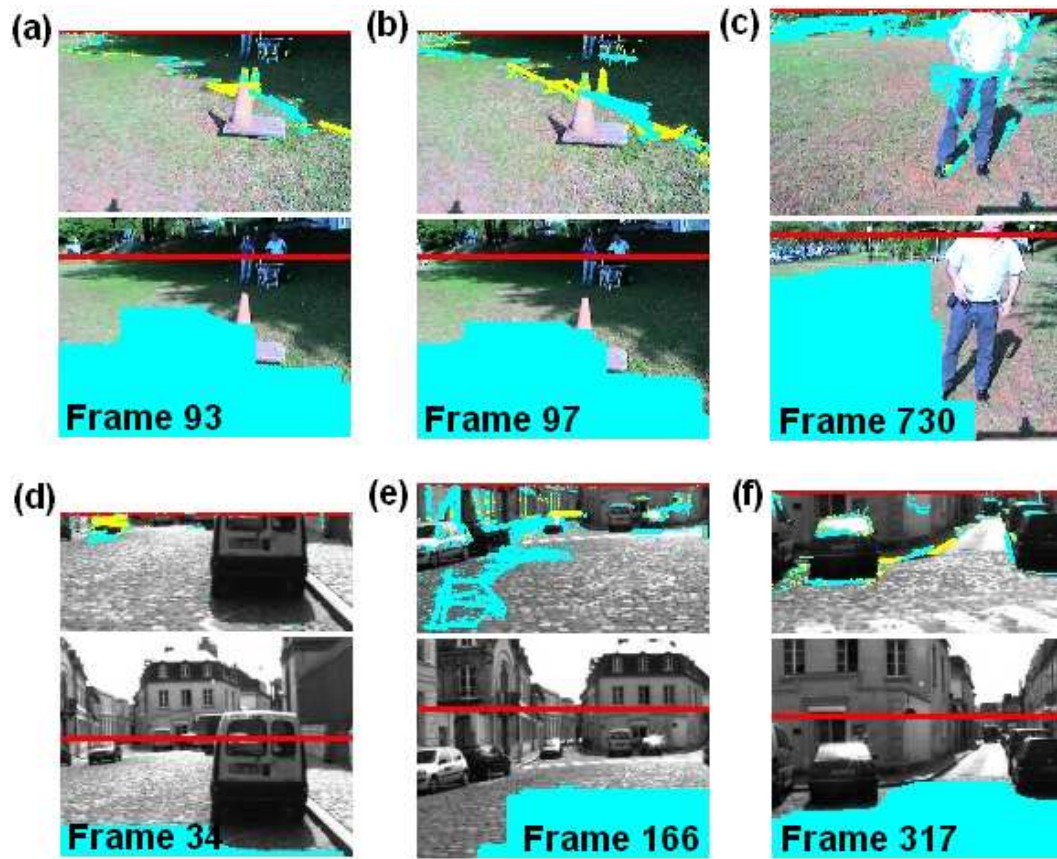
Figure 4.16 – Obstacle avoidance from an old stored Otsu threshold (grass areas and parallelepiped streets): In the top images, the interest pixels classified as navigable area are represented in blue; and the interest pixels classified as obstacle are represented in yellow; after line detection using Hough transform; In the bottom images, the stored Otsu threshold is represented in blue; The red horizontal lines represent the horizon finding algorithm results (Miranda Neto, et al., 2011).

4.2.3 Collision Risk Estimation

Section 2.6.3 presented a novel approach to obtain Collision Risk Estimation (CRE), or time-to-collision (TTC), estimates based on Pearson's Correlation Coefficient (PCC) in dynamic and unknown environment from a monocular camera. Acting as a complement, a region-merging algorithm (i.e. section 2.6.2) aims to represent homogeneous image regions. These image regions are equaled to reduce the Pearson's variation. From an adaptation to the Pearson's method it obtains the interest pixels (i.e. section 2.5.1). Finally, for to find the obstacle direction, an interactive thresholding algorithm is proposed in section 2.6.4 based on Otsu thresholding method (OTM). This method does not take into account the relative acceleration between the host car and the subject "object", neither its distance nor its velocity are constant. This section therefore evaluates the proposed algorithm at high speed on real-time conditions using the vehicle shown in Figures 4.1 (a). The speed informations were obtained in real time on a highway and in the Forest of Compiegne. This text-set contains image sequences (320x240 pixels) recorded at 20 frames per second from a moving vehicle using a Sony DFW-VL500 camera. For these cases, in order to reduce the number of data, an image resolution reduction to 96x72 pixels was introduced. A video containing all images sequence for this section are presented in[50].

In real conditions this monocular-vision system has been designed to investigate only a small portion of the road ahead of the vehicle, where the absence of other vehicles has been assumed (Bertozzi, et al., 2000). The Figure 4.17 (d) presents this fix analysis region (yellow line). In the Figures 4.17 to 4.21, the left side column images (a) to (f) are such that:

- ✓ (a): The first interaction of the interactive thresholding algorithm (ITA) presented in section 2.6.4; and its direction based on mass center of the red area;
- ✓ (b): The last interaction and final result of the interactive thresholding algorithm (ITA) presented in section 2.6.4; and its direction based on mass of center of the green area;
- ✓ (c): Original image after reduction scale; and the obstacle direction;
- ✓ (d): Result of the region-merging algorithm presented in section 2.6.2;
- ✓ (e): Binarized image by OTM and Canny presented in section 2.6.2;
- ✓ (f): Binarized image by Canny presented in section 2.6.2;

---

[50] http://www.youtube.com/watch?v=qMLG3icde9s – visited on October 1, 2011.
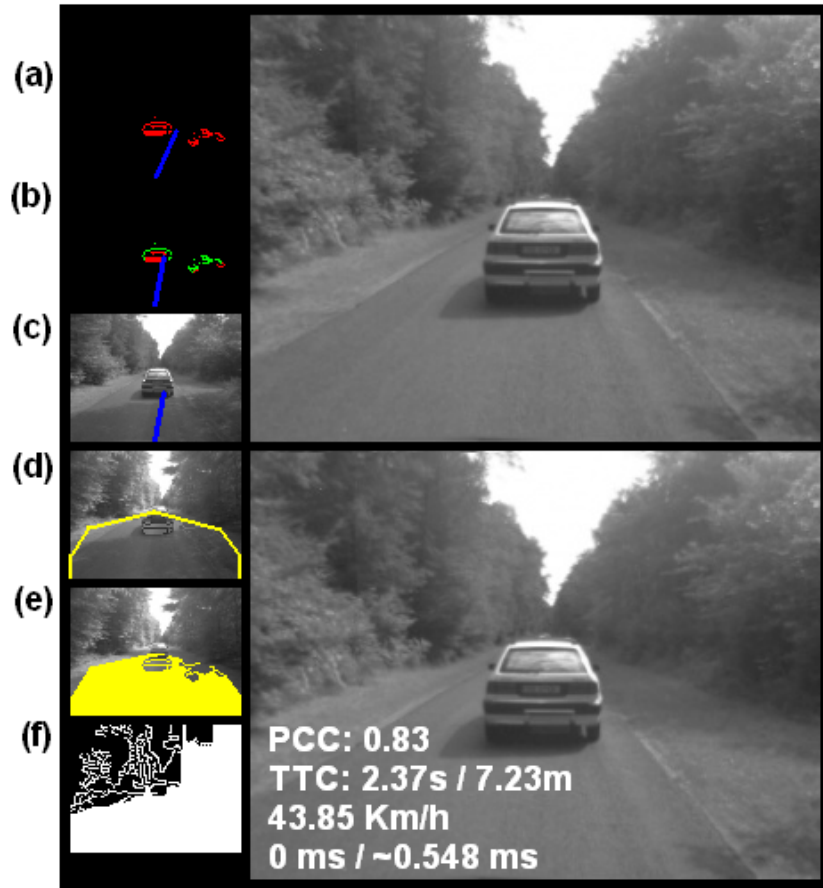
Figure 4.17 – Heudiasyc Laboratory in France, 2010, Carmen: Speed 43.85 Km/h: (a): The first interaction of the interactive thresholding algorithm (ITA) presented in section 2.6.4; and its direction based on mass of center of the red area; (b): The last interaction and final result of the interactive thresholding algorithm (ITA) presented in section 2.6.4; and its direction based on mass of center of the green area; (c): Original image after reduction scale; and the obstacle direction; (d): Result of the region-merging algorithm presented in section 2.6.2; (e): Binarized image by OTM and Canny presented in section 2.6.2; (f): Binarized image by Canny presented in section 2.6.2.

TABLE 4.1: COLLISION RISK ESTIMATION (CRE): FIGURE 4.17
CARMEN AVERAGE SPEED: 43.85 KM/H

PCC r_1	$(1-r_1)$	Variation in the Range	Risk of Collision	CRE Second	Distance Meters
0.8315	(1-0.8315)	0.1685	(R_c / 0.1685)	2.37s	7.23m
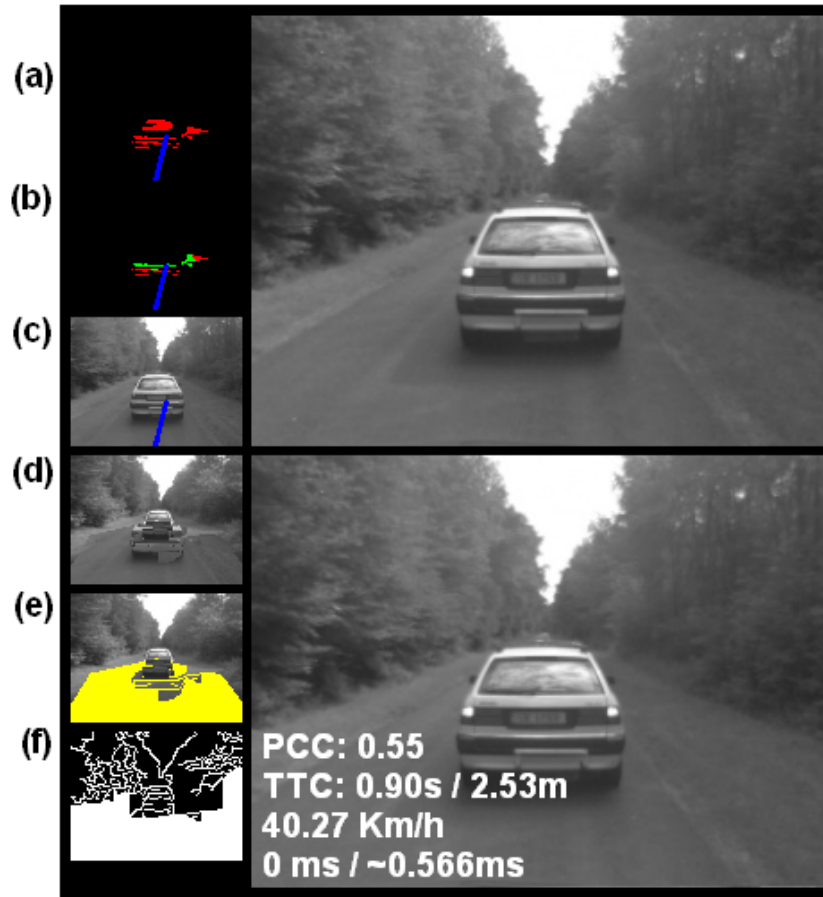
Figure 4.18 – Heudiasyc Laboratory in France, 2010, Carmen: Speed 40.27 Km/h: (a): The first interaction of the interactive thresholding algorithm (ITA) presented in section 2.6.4; and its direction based on mass of center of the red area; (b): The last interaction and final result of the interactive thresholding algorithm (ITA) presented in section 2.6.4; and its direction based on mass of center of the green area; (c): Original image after reduction scale; and the obstacle direction; (d): Result of the region-merging algorithm presented in section 2.6.2; (e): Binarized image by OTM and Canny presented in section 2.6.2; (f): Binarized image by Canny presented in section 2.6.2.

TABLE 4.2: COLLISION RISK ESTIMATION (CRE): FIGURE 4.18
CARMEN AVERAGE SPEED: 40.27 KM/H

PCC r_1	$(1-r_1)$	Variation in the Range	Risk of Collision	CRE Second	Distance Meters
0.5584	(1-0.5584)	0.4416	(R_c / 0.4416)	0.90s	2.53m
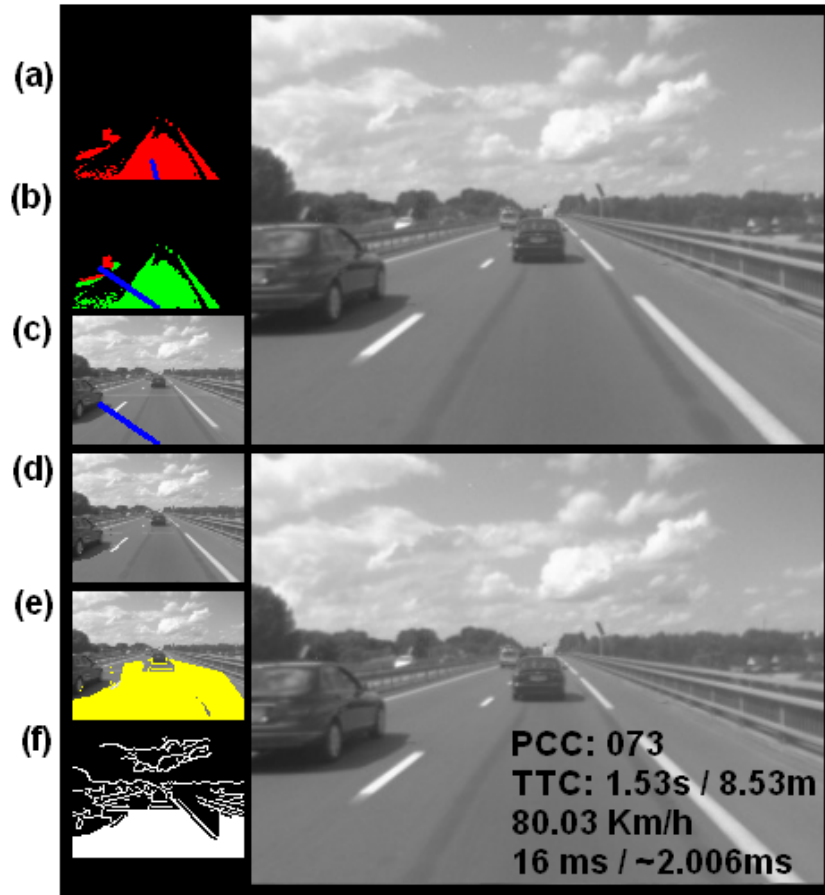
Figure 4.19 – Heudiasyc Laboratory in France, 2010, Carmen: Speed 80.03 Km/h: (a): The first interaction of the interactive thresholding algorithm (ITA) presented in section 2.6.4; and its direction based on mass of center of the red area; (b): The last interaction and final result of the interactive thresholding algorithm (ITA) presented in section 2.6.4; and its direction based on mass of center of the green area; (c): Original image after reduction scale; and the obstacle direction; (d): Result of the region-merging algorithm presented in section 2.6.2; (e): Binarized image by OTM and Canny presented in section 2.6.2; (f): Binarized image by Canny presented in section 2.6.2.

TABLE 4.3: COLLISION RISK ESTIMATION (CRE): FIGURE 4.19
CARMEN AVERAGE SPEED: 80.03 KM/H

PCC r_1	$(1-r_1)$	Variation in the Range	Risk of Collision	CRE Second	Distance Meters
0.7394	(1-0.7394)	0.2606	(R_c / 0.2606)	1.53s	8.53m

105
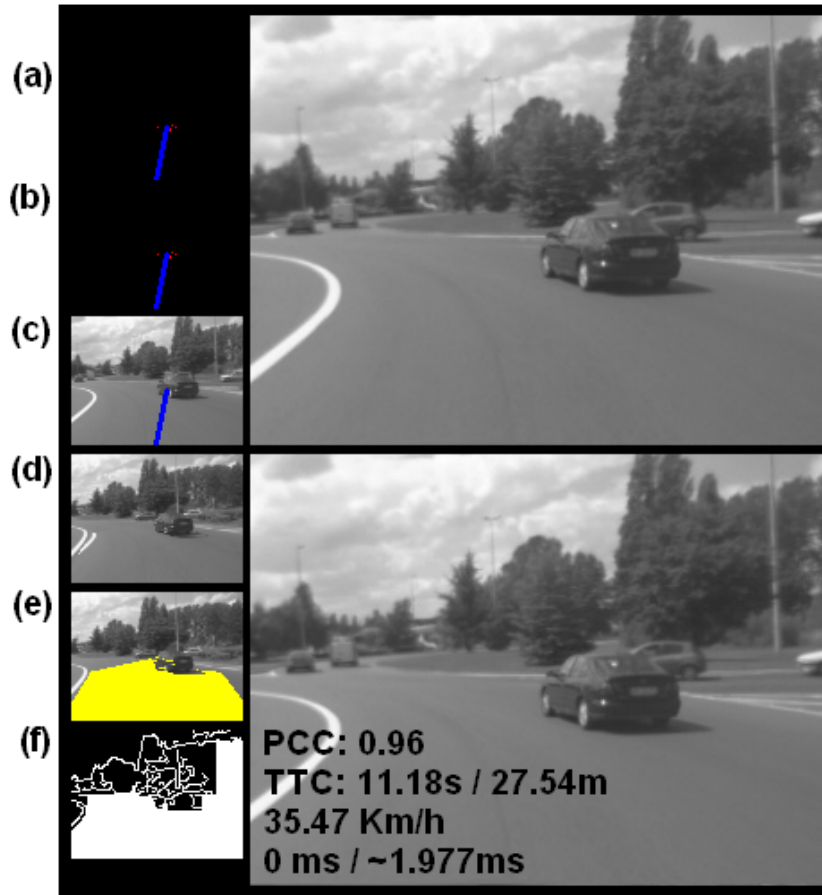
Figure 4.20 – Heudiasyc Laboratory in France, 2010, Carmen: Speed 35.47 Km/h: (a): The first interaction of the interactive thresholding algorithm (ITA) presented in section 2.6.4; and its direction based on mass of center of the red area; (b): The last interaction and final result of the interactive thresholding algorithm (ITA) presented in section 2.6.4; and its direction based on mass of center of the green area; (c): Original image after reduction scale; and the obstacle direction; (d): Result of the region-merging algorithm presented in section 2.6.2; (e): Binarized image by OTM and Canny presented in section 2.6.2; (f): Binarized image by Canny presented in section 2.6.2.

TABLE 4.4: COLLISION RISK ESTIMATION (CRE): FIGURE 4.20
CARMEN AVERAGE SPEED: 35.47 KM/H

| PCC r_1	$(1-r_1)$	Variation in the Range	Risk of Collision	CRE Second	Distance Meters
0.9642	(1-0.9642)	0.0358	(R_c / 0.0358)	11.18s	27.54m
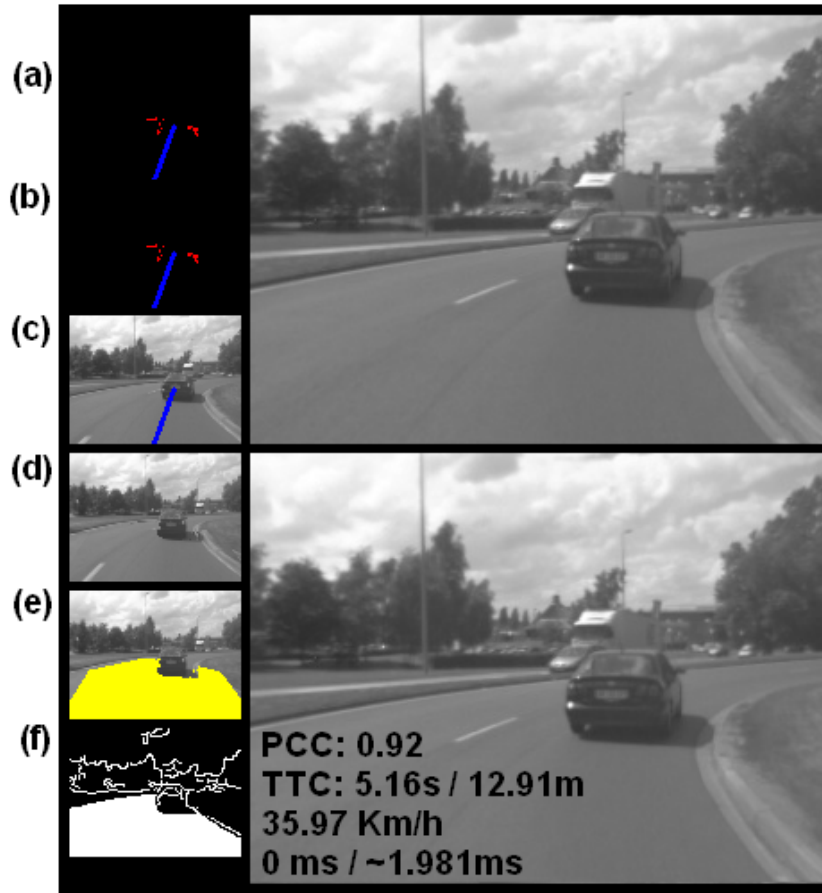
Figure 4.21 – Heudiasyc Laboratory in France, 2010, Carmen: Speed 35.97 Km/h: (a): The first interaction of the interactive thresholding algorithm (ITA) presented in section 2.6.4; and its direction based on mass of center of the red area; (b): The last interaction and final result of the interactive thresholding algorithm (ITA) presented in section 2.6.4; and its direction based on mass of center of the green area; (c): Original image after reduction scale; and the obstacle direction; (d): Result of the region-merging algorithm presented in section 2.6.2; (e): Binarized image by OTM and Canny presented in section 2.6.2; (f): Binarized image by Canny presented in section 2.6.2.

TABLE 4.5: COLLISION RISK ESTIMATION (CRE): FIGURE 4.21
CARMEN AVERAGE SPEED: 35.97 KM/H

PCC r_1	$(1-r_1)$	Variation in the Range	Risk of Collision	CRE Second	Distance Meters
0.9226	(1-0.9226)	0.0774	(R_c/ 0.0774)	5.16s	12.91m

107

4.3 Sky Removal and Navigable Area Detection

Chapter 3 presented a real-time perception problem applied to experimental vehicles. It proposed an evolution of the segmentation method presented in Miranda Neto and Rittner (2006) and applied it to obstacle detection and reactive navigation, which takes place in obstacle avoidance context for vehicles in dynamic and unknown environments. It also included a horizon finding algorithm that finds the horizon line and apply it to generate the navigable area. The results are presented below as follows:

- ✓ Section 4.3.1: Horizon Finding Algorithm;
- ✓ Section 4.3.2: Navigable Area Detection Algorithm;

4.3.1 Horizon Finding Algorithm

As showed in section 4.2.1, a real-time system must satisfy explicit response-time constraints and the compromise between processing time and images acquisition is fundamental. The sky removal process may also allow a significant reduction of data, since the main task is navigable area detection. Moreover, as presented in section 3.2.3, the analysis of a bigger image portion not always can contribute for a better result in the most critical region (Figure 3.1: in land navigation, the region closer to the vehicle is the most important one), where obstacles should be detected and avoided as fast as possible. As a sample of the potential of this horizon finding algorithm execution in real time, the Figure 4.6 (a) presented the computational mean time in unknown and urban environment. Additional results of this sky removal can also be found in Figures 2.9, 2.10, 4.12, 4.13, 4.14, 4.15, 4.16, 4.26, 4.27, 4.28, 4.29, 4.30 and 4.31.

A video containing more images of the next sequences is available in[51].

For each image a real horizon line (black) was registered manually.

---

[51] http://www.youtube.com/watch?v=8KbZ1J0txUE – visited on October 1, 2011.

Based on two DARPA Desert text-set (DARPA, 2005), Figure 4.22 and 4.23 present the performance of the horizon line detection process. These text-sets contain image sequences of 320x240 pixels. In order to reduce the number of data, an image resolution reduction to 128x96 pixels was introduced.
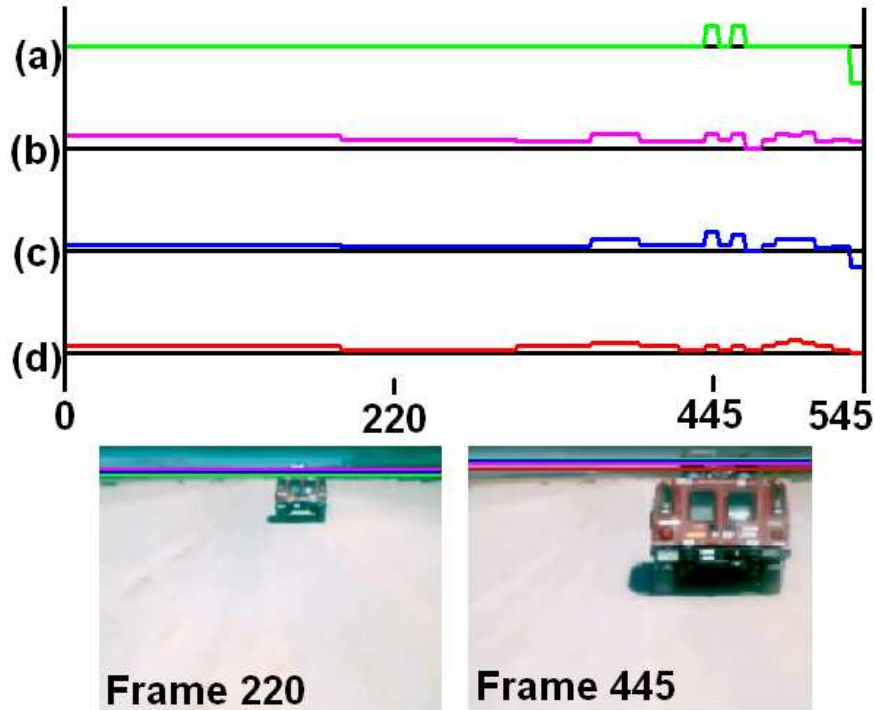
Figure 4.22 – DARPA Desert text-set 1 (DARPA, 2005): For each image a real horizon line (black) was registered manually; Horizon Finding Algorithm: (a) the green line represents the Otsu horizon line detection; (b) the magenta line represents the weighted average of the Hough transformation lines; (c) the blue line represents the weighted average between the Otsu horizon line detection and the Hough transformation result; finally, (d) the red line represents the robust horizon finding algorithm based on Otsu segmentation, Hough transformation and filtering.
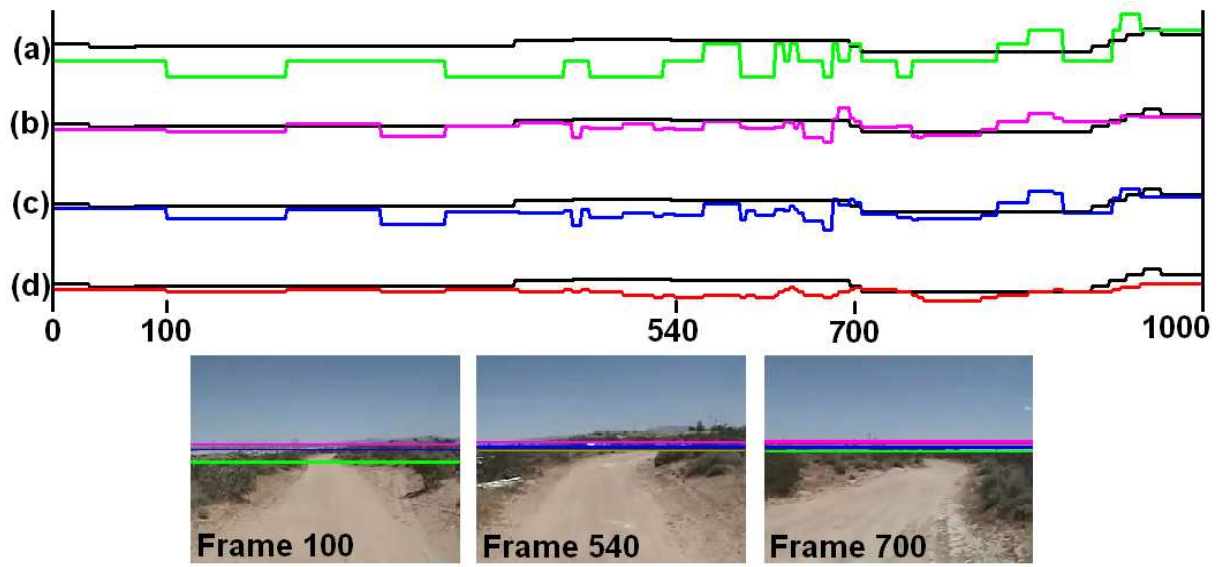
Figure 4.23 – DARPA Desert text-set 2 (DARPA, 2005): For each image a real horizon line (black) was registered manually; Horizon Finding Algorithm: (a) the green line represents the Otsu horizon line detection; (b) the magenta line represents the weighted average of the Hough transformation lines; (c) the blue line represents the weighted average between the Otsu horizon line detection and the Hough transformation result; finally, (d) the red line represents the robust horizon finding algorithm based on Otsu segmentation, Hough transformation and filtering.

Taken displacements on the outskirts of the Heudiasyc Laboratory in France, the Figure 4.24 contains images (320x240 pixels) recorded at 20 frames per second from a moving vehicle using a Sony DFW-VL500 camera. This stage of testing also evaluates the performance of the horizon line detection process on real-time conditions using the Strada vehicle shown in Figure 4.1 (b). In order to reduce the number of data, an image resolution reduction to 128x96 pixels was introduced.

As shown in Figure 4.25, the last case presents the performance of the horizon line detection process in infrared road video made by FLIR (2010). This text-set contains image sequences of 128x96 pixels.

In all cases presented here, for each image, a real horizon black line was registered manually. In this way, (a) the green line represents the Otsu horizon line detection; (b) the

110

magenta line represents the weighted average of the Hough transformation lines; (c) the blue line represents the weighted average between the Otsu horizon line detection and the Hough transformation result; finally, (d) the red line represents the robust horizon finding algorithm based on Otsu segmentation, Hough transformation and filtering presented in section 3.3.
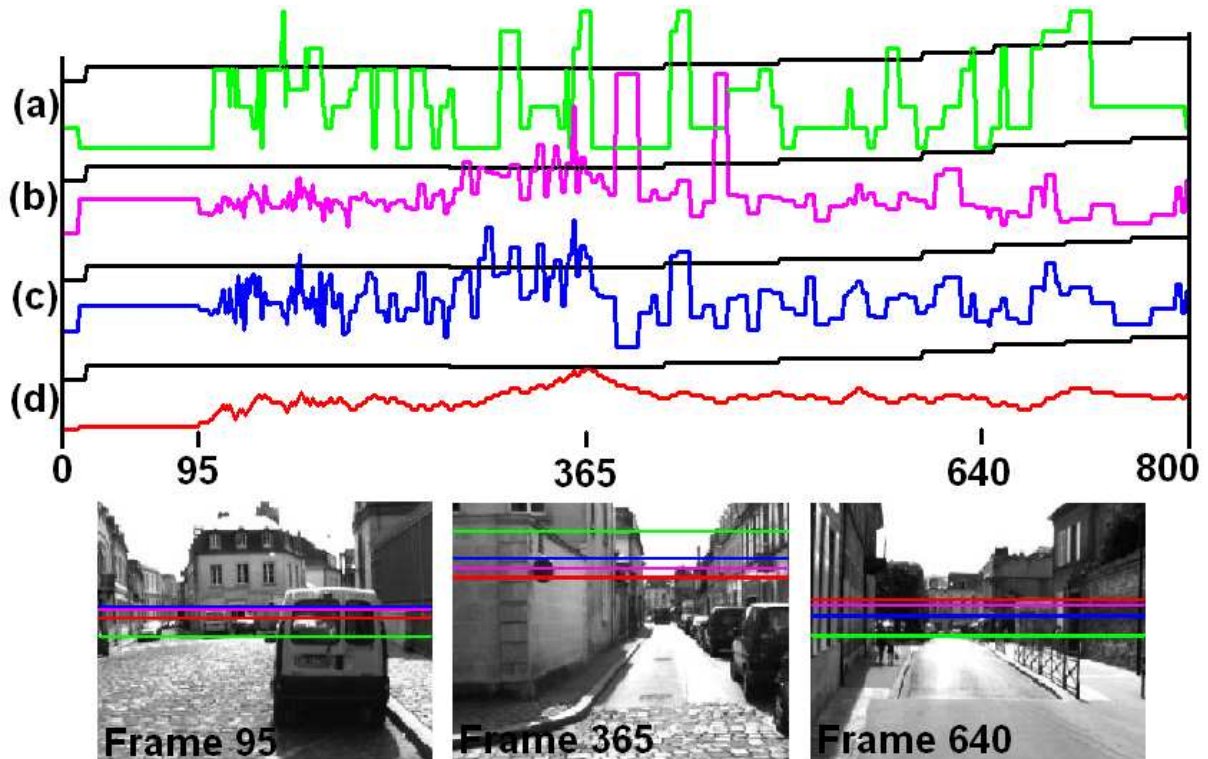
Figure 4.24 – Heudiasyc Laboratory in France, 2005, Strada: For each image a real horizon line (black) was registered manually; Horizon Finding Algorithm: (a) the green line represents the Otsu horizon line detection; (b) the magenta line represents the weighted average of the Hough transformation lines; (c) the blue line represents the weighted average between the Otsu horizon line detection and the Hough transformation result; finally, (d) the red line represents the robust horizon finding algorithm based on Otsu segmentation, Hough transformation and filtering.
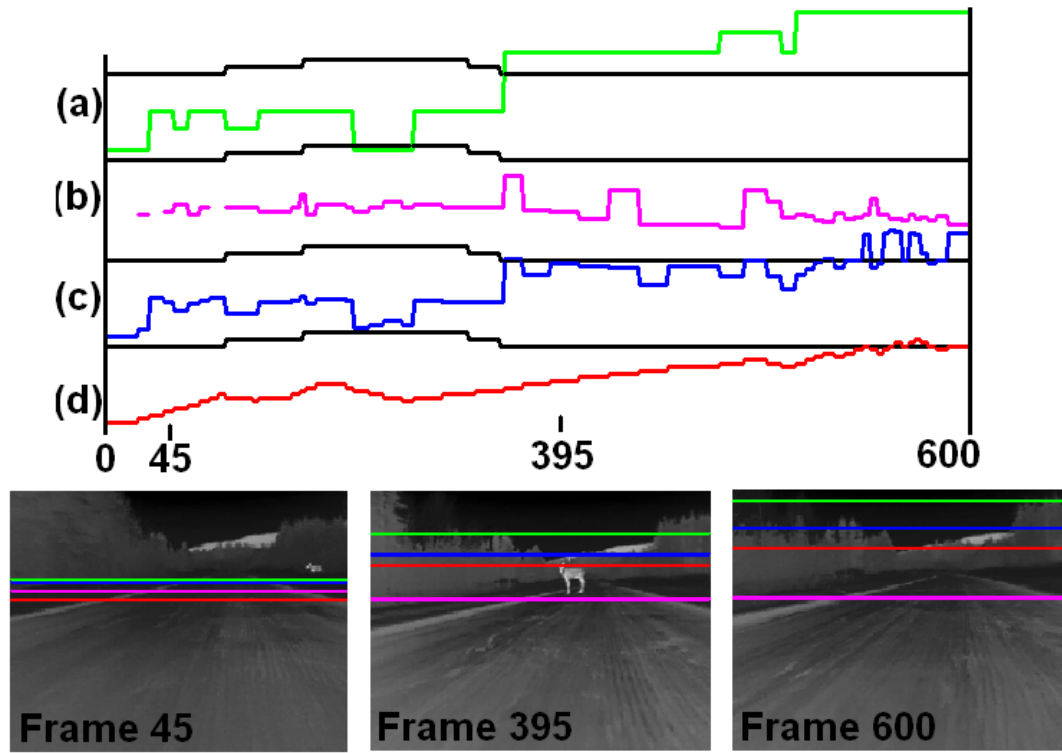
Figure 4.25 – FLIR: infrared road video (FLIR, 2010): For each image a real horizon line (black) was registered manually; Horizon Finding Algorithm: (a) the green line represents the Otsu horizon line detection; (b) the magenta line represents the weighted average of the Hough transformation lines; (c) the blue line represents the weighted average between the Otsu horizon line detection and the Hough transformation result; finally, (d) the red line represents the robust horizon finding algorithm based on Otsu segmentation, Hough transformation and filtering.

4.3.2 Navigable Area Detection Algorithm

Section 3.4 presented an evolution of the segmentation method proposed by Miranda Neto and Rittner (2006). From the sky removal process, this section presents results of a navigable area detection algorithm, which takes place in free-area detection and obstacle avoidance context for vehicles in dynamic and unknown environments. In this way, for each result presented below, a free-navigable area is therefore represented from a multimodal 2D drivability road image

obtained in Equation 3.3 and 3.4. This multimodal result enables that a level of safety may be selected according to the environment and operational context. Examples of this proposal can also be found in the bottom images of the Figures 4.13, 4.14, 4.15 and 4.16, where the blue pixels represent $RD_{d(x,y)} >= 70$.

Based on two DARPA Desert text-set (DARPA, 2005), Figure 4.26 presents the performance of the navigable area detection process. This text-set contains image sequences of 320x240 pixels. In order to reduce the number of data, an image resolution reduction to 128x96 pixels was introduced. Some results are presented below:

- ✓ (b), (e), (h) and (k): free-area detection obtained in Equation 3.3;
- ✓ (c), (f), (i) and (l): free-area detection obtained in Equation 3.3, where the blue pixels represent $RD_{d(x,y)} >= 70$;
- ✓ (e) and (f): an obstacle avoidance example;
- ✓ (h), (i), (k) and (i): the line change examples;

Taken displacements on the outskirts of the Karlsruhe Institute of Technology in Germany, the images of the Figure 4.27 were obtained in (Geiger, 2011). It contains high-quality image sequences (1344x391 pixels) recorded at 10 frames per second from a moving vehicle with a Pointgrey Flea2 firewire camera. For these cases, in order to reduce the number of data, an image resolution reduction to 250x120 pixels was introduced. Some results are presented below:

- ✓ (b), (e), (h), (k) and (n): free-area detection obtained in Equation 3.3;
- ✓ (c), (f), (i), (l) and (o): free-area detection obtained in Equation 3.3, where the blue pixels represent $RD_{d(x,y)} >= 70$;
- ✓ (b) and (c); (e) and (f): shadow examples;
- ✓ (h) and (i); (k) and (l): obstacle avoidance examples;
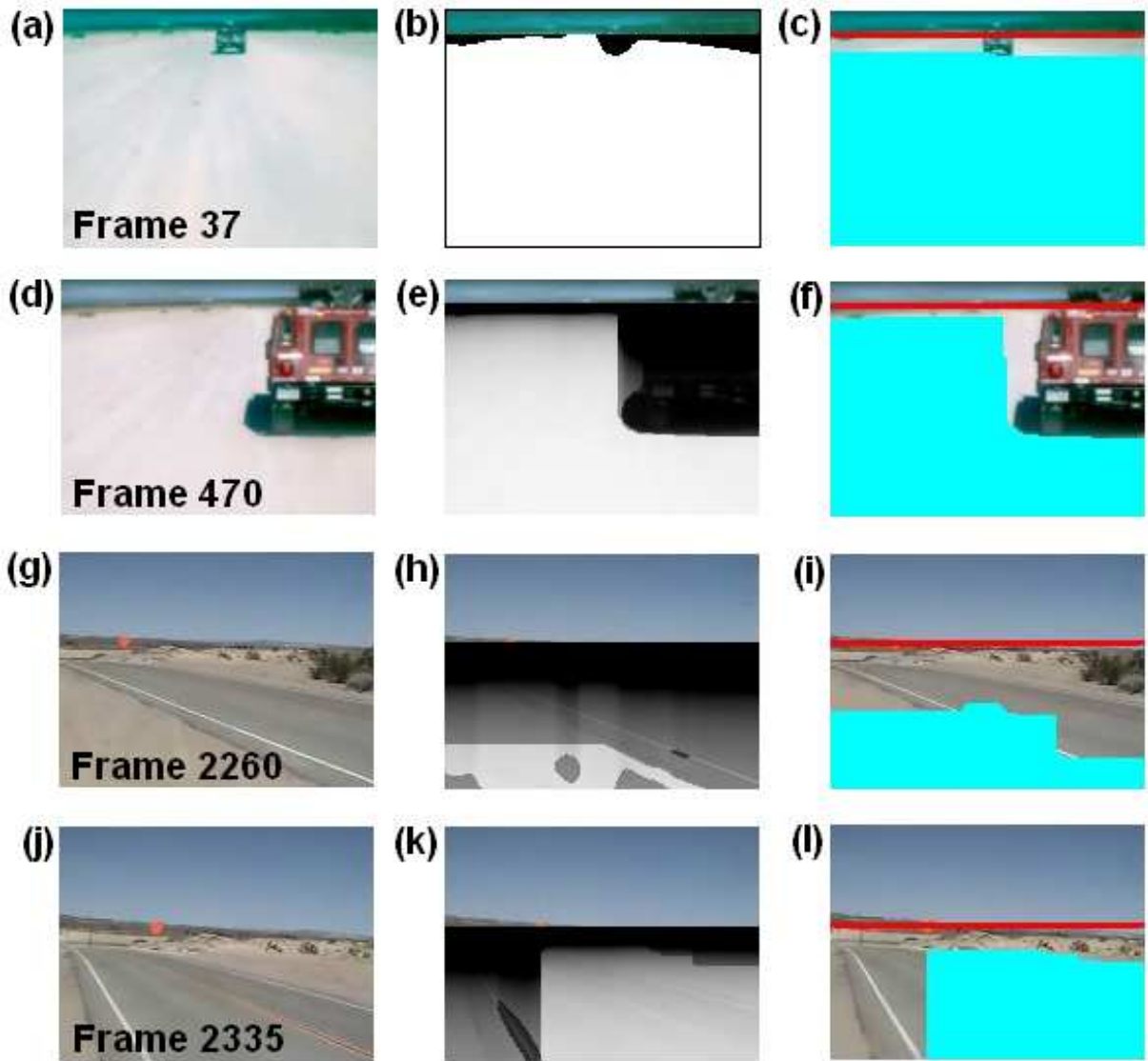
Figure 4.26 – DARPA Desert text-set (DARPA, 2005): Navigable Area Detection Algorithm: (b), (e), (h) and (k): free-area detection obtained in Equation 3.3; (c), (f), (i) and (l): free-area detection obtained in Equation 3.3, where the blue pixels represent $RD_{d(x,y)} >= 70$; (e) and (f): obstacle avoidance; (h), (i), (k) and (i): line change; The red horizontal lines represent the horizon finding algorithm results (Miranda Neto, et al., 2011).

Figure 4.27 – Karlsruhe Dataset on September, 8th 2009 (Geiger, 2011): Navigable Area Detection Algorithm: (b), (e), (h), (k) and (n): free-area detection obtained in Equation 3.3; (c), (f), (i), (l) and (o): free-area detection obtained in Equation 3.3, where the blue pixels represent $RD_{d(x,y)} >= 70$; (b) and (c); (e) and (f): shadow examples; (h) and (i); (k) and (l): obstacle avoidance examples; The red horizontal lines represent the horizon finding algorithm results (Miranda Neto, et al., 2011).

Taken displacements on the outskirts of the Heudiasyc Laboratory in France, the Figure 4.28 contains images (320x240 pixels) recorded at 20 frames per second from a moving vehicle using a Sony DFW-VL500 camera. In order to reduce the number of data, an image resolution reduction to 128x96 pixels was introduced. Some results are presented below:

✓ (b), (e) and (h): free-area detection obtained in Equation 3.3;

✓ (c), (f) and (i): free-area detection obtained in Equation 3.3, where the blue pixels represent $RD_{d(x,y)} >= 70$;

✓ (h) and (i): narrow passage example;

Figure 4.28 – Heudiasyc Laboratory in France, 2005, Strada: Navigable Area Detection Algorithm: (b), (e) and (h): free-area detection obtained in Equation 3.3; (c), (f) and (i): free-area detection obtained in Equation 3.3, where the blue pixels represent $RD_{d(x,y)} >= 70$; (h) and (i): narrow passage example; The red horizontal lines represent the horizon finding algorithm results (Miranda Neto, et al., 2011).

116

Taken displacements on the outskirts of the Heudiasyc Laboratory in France (Forest of Compiegne), Figure 4.29 contains images (320x240 pixels) recorded at 20 frames per second from a moving vehicle using a Sony DFW-VL500 camera. In order to reduce the number of data, an image resolution reduction to 128x96 pixels was introduced. Some results are presented below:

✓ (b), (e) and (h): free-area detection obtained in Equation 3.3;

✓ (c), (f) and (i): free-area detection obtained in Equation 3.3, where the blue pixels represent $RD_{d(x,y)} >= 70$;

✓ (a), (d) and (g); (b), (e) and (h); (c), (f) and (i): overtaking task example;
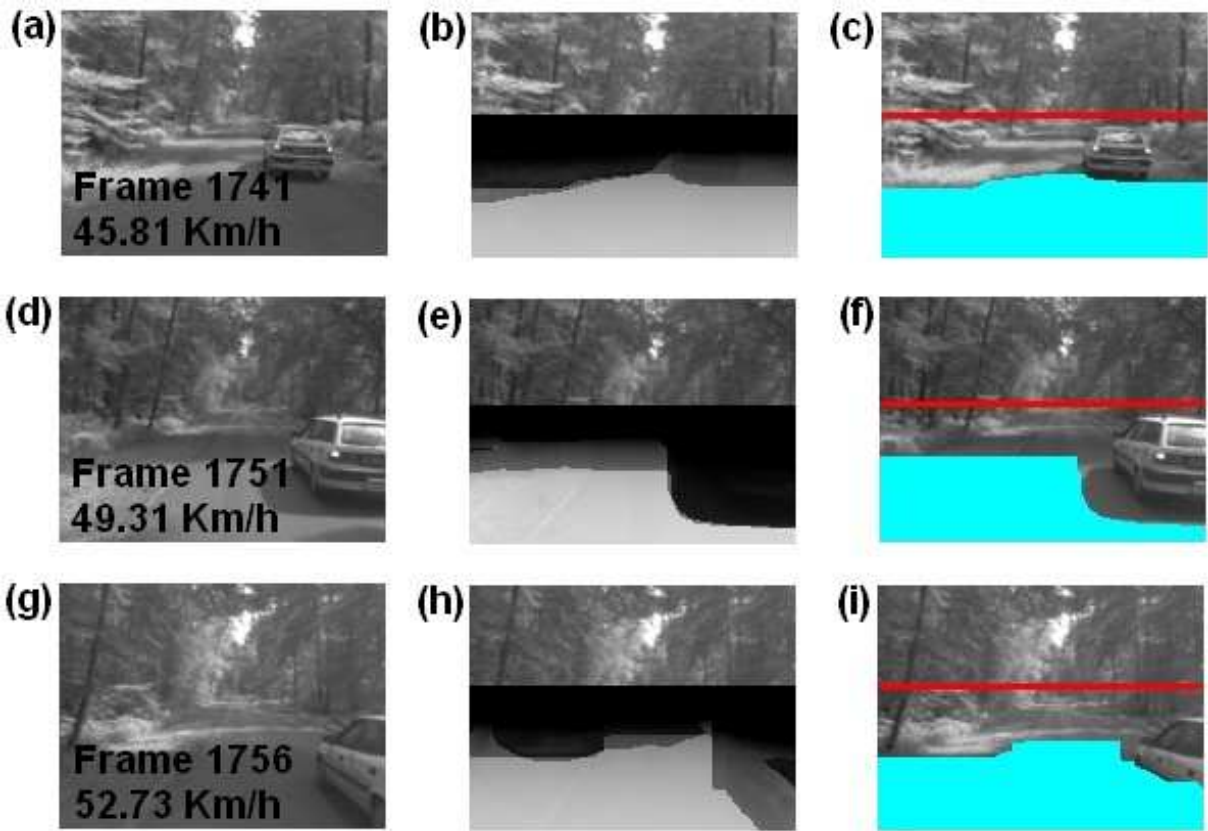
Figure 4.29 – Heudiasyc Laboratory in France, 2010, Carmen: Navigable Area Detection Algorithm in Forest of Compiegne: (b), (e), (h) and (k): free-area detection obtained in Equation 3.3; (c), (f), (i) and (l): free-area detection obtained in Equation 3.3, where the blue pixels represent $RD_{d(x,y)} >= 70$; (e) and (f): obstacle avoidance; (h), (i), (k) and (i): line change; The red horizontal lines represent the horizon finding algorithm results (Miranda Neto, et al., 2011).

117

Taken displacements on the outskirts of the Heudiasyc Laboratory in France (highway), the Figure 4.30 contains images (320x240 pixels) recorded at 20 frames per second from a moving vehicle using a Sony DFW-VL500 camera. In order to reduce the number of data, an image resolution reduction to 128x96 pixels was introduced. As shown in Figure 4.7 (c), values greater than 100 Km/h can be found. Some results are presented below:

- ✓ (b), (e), (h) and (k): free-area detection obtained in Equation 3.3;
- ✓ (c), (f), (i) and (l): free-area detection obtained in Equation 3.3, where the blue pixels represent $RD_{d(x,y)} >= 70$;
- ✓ (b) and (c): an obstacle avoidance example;

The last case presents autonomous displacements at Renato Archer IT Center (CTI) in Brazil. Figure 4.31 evaluates the proposed algorithm at low speed on real-time conditions using the VERO platform shown in Figure 4.1 (c). These text-sets contain image sequences (640x480 pixels) recorded at a frame rate of 25 frames per second from a moving vehicle using a Sony DFW-VL500 camera. For these cases, in order to reduce the number of data, an image resolution reduction to 160x120 pixels was introduced. Some results are presented below:

- ✓ (b), (e), (h), (k) and (n): free-area detection obtained in Equation 3.3;
- ✓ (c), (f), (i), (l) and (o): free-area detection obtained in Equation 3.3, where the blue pixels represent $RD_{d(x,y)} >= 70$;
- ✓ (h) and (i); (k) and (l); (n) and (o): area with high noise level;
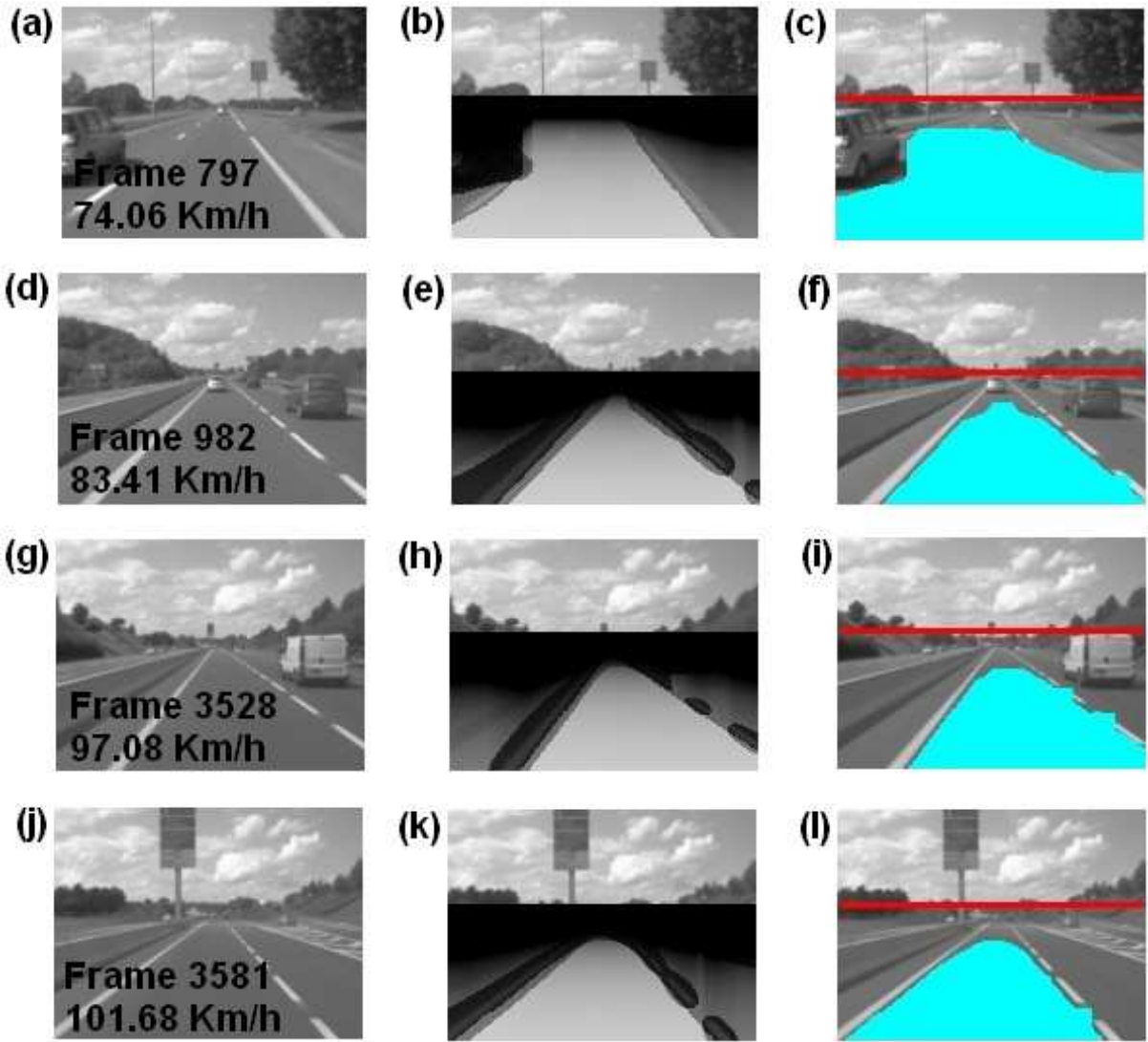
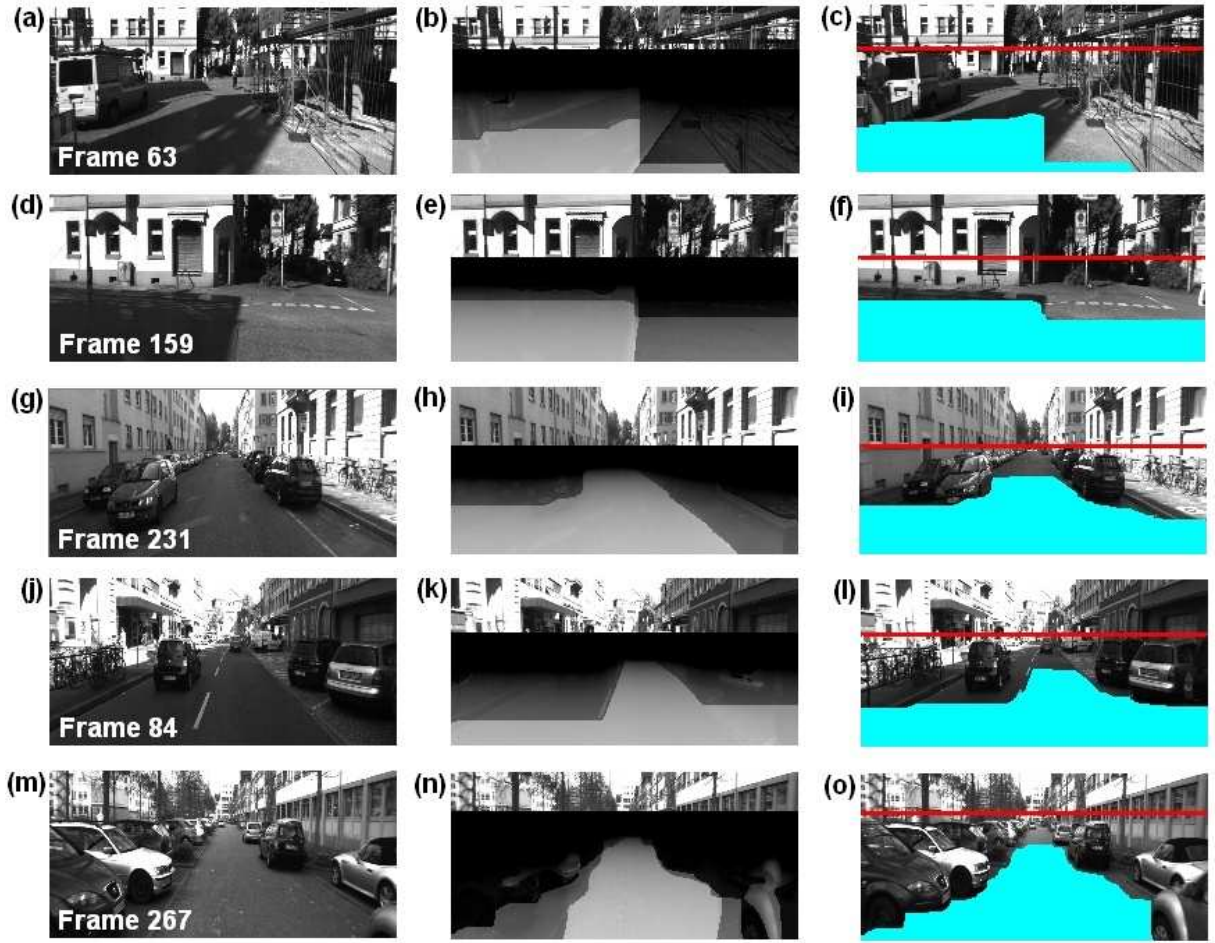Figure 4.30 – Heudiasyc Laboratory in France, 2010, Carmen: Navigable Area Detection Algorithm in highway: (b), (e), (h) and (k): free-area detection obtained in Equation 3.3; (c), (f), (i) and (l): free-area detection obtained in Equation 3.3, where the blue pixels represent $RD_{d(x,y)} >= 70$; (e) and (f): obstacle avoidance; (h), (i), (k) and (i): line change; The red horizontal lines represent the horizon finding algorithm results (Miranda Neto, et al., 2011).
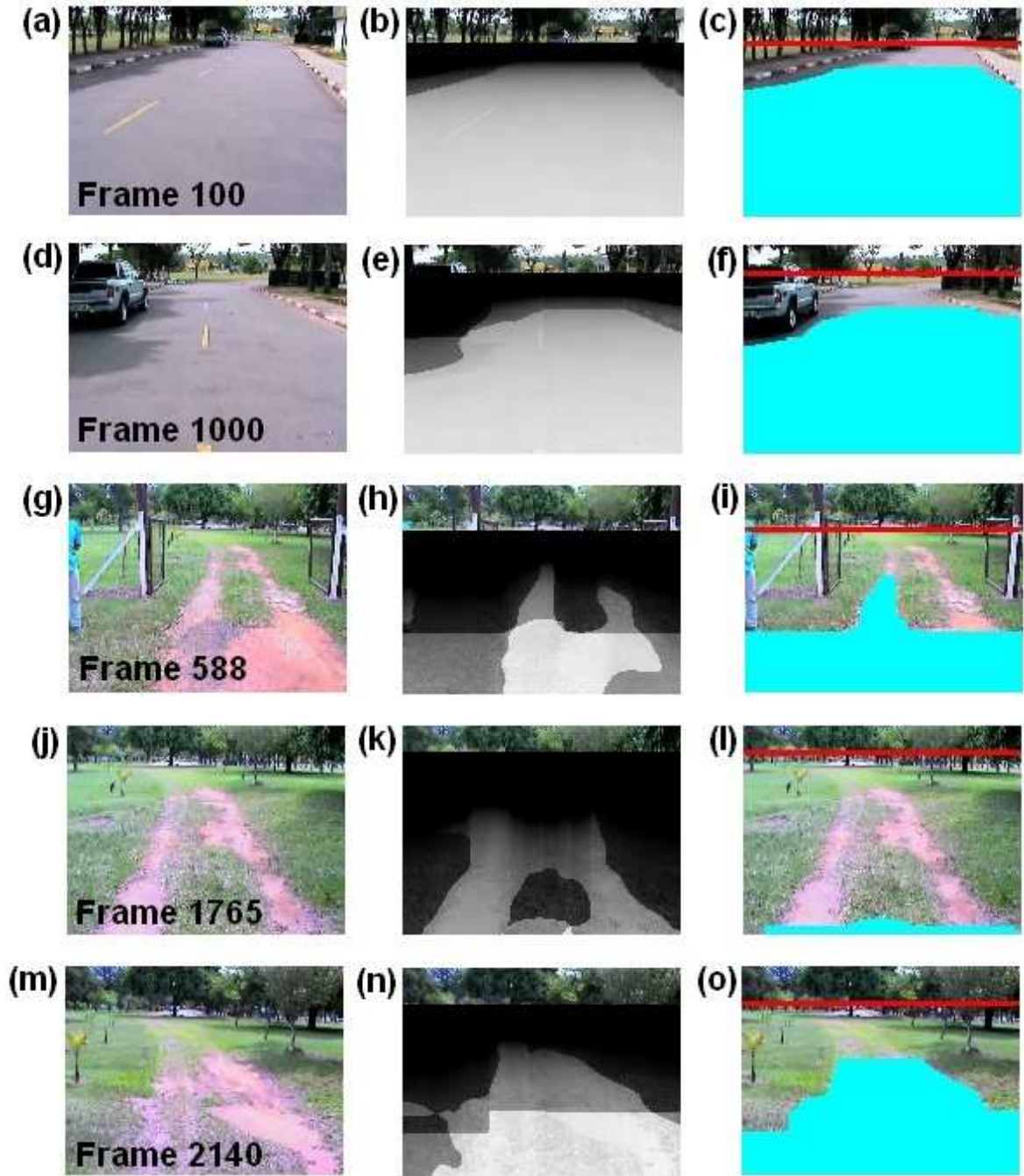
Figure 4.31 – Renato Archer IT Center (CTI) on November, 4th 2010 – VERO platform in off-road context: Navigable Area Detection Algorithm in highway: (b), (e), (h) and (k): free-area detection obtained in Equation 3.3; (c), (f), (i) and (l): free-area detection obtained in Equation 3.3, where the blue pixels represent $RD_{d(x,y)} >= 70$; (e) and (f): obstacle avoidance; (h), (i)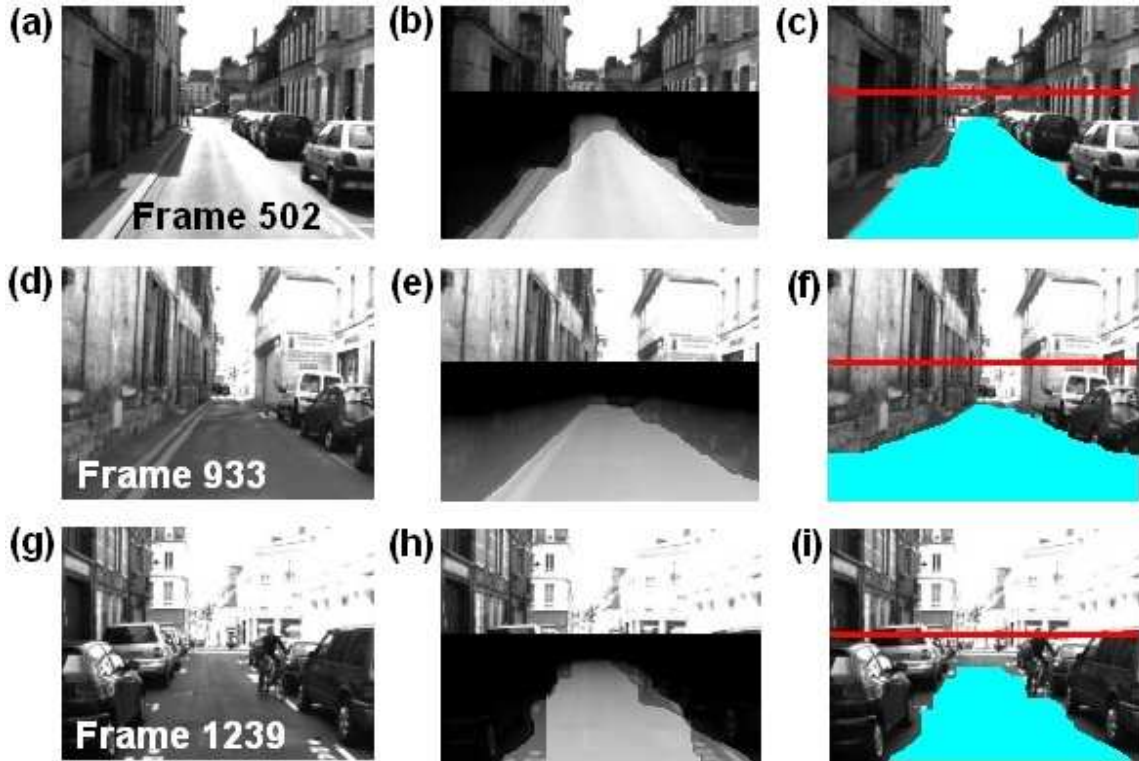, (k) and (i): line change; The red horizontal lines represent the horizon finding algorithm results (Miranda Neto, et al., 2011).

Videos containing more images sequence presented in this section are available in[52].

---
[52] http://www.youtube.com/watch?v=4cnU3rS41Cs – visited on October 1, 2011.

Conclusion

This thesis addressed the problem of obstacle avoidance for semi- and autonomous terrestrial platforms in dynamic and unknown environments. Based on an onboard monocular vision system, a set of tools was proposed to continuously monitor the way forward in the space ahead of the vehicle, proving appropriate road informations in real time. These tools were tested in different types of image texture (road surfaces). They also included free-area detection, reactive navigation and collision risk estimation.

A horizon finding algorithm was developed to sky removal (Miranda Neto, et al., 2011). This algorithm generates the region of interest from a dynamic threshold search method, allowing to dynamically investigate only a small portion of the image ahead of the vehicle, in order to road and obstacle detection. An evolution of the segmentation method presented by Miranda Neto & Rittner (2006) was proposed. A free-navigable area is therefore represented from a multimodal 2D drivability road image. This multimodal result enables the selection of a safety index according to the environment and operational context.

In order to reduce processing time, an automatic image discarding criteria was proposed by Miranda Neto, et al. (2007) and Miranda Neto, et al. (2008). Taking into account the temporal coherence between consecutive frames, a new Dynamic Power Management methodology was also proposed and applied to a robotic visual perception scheme (Miranda Neto, et al., 2011), which includes a new environment observer method to optimize energy consumption used by a visual machine. Additionally, based on Pearson's method (Pearson, 1895), it also includes a real-time perception system capable of estimating the collision risk in dynamic and unknown environments by using a single monocular system. It is noted the originality of these works.

A remarkable characteristic of the presented methodologies is the independence of the image acquisition system and of the robot itself. This real-time perception system has been evaluated by using different text-sets and also from real data obtained by two intelligent platforms. The presented methods were submitted to more than 47.5 thousand images. In semi-autonomous tasks, some tests were conducted at speeds above 100 Km/h. Open-loop reactive

autonomous displacements were also carried out successfully. It is important to notice that these algorithms were not based on previous knowledge of the environment (lane shape, geometric inference, etc) neither camera calibration.

Future work will focus the evaluation of the proposed methodologies with respect to ground truth. Another research effort will consider a multi-sensors system based mainly on vision, radar and laser.

Conclusão

Esta tese abordou o problema de evitamento de obstáculos para plataformas terrestres semi- e autônomas em ambientes dinâmicos e desconhecidos. Baseado num sistema de visão monocular embarcado, um conjunto de ferramentas foi proposto para monitorar continuamente o espaço à frente do veículo, provendo informações adequadas em tempo real. Estas ferramentas foram testadas com diferentes tipos imagens e texturas (tipos de superfícies). Resultados incluem a deteção da área navegável, navegação reativa e estimatição do risco de colisão.

Um algoritmo de deteção da linha do horizonte foi desenvolvido para a remoção da região do horizonte (Miranda Neto, et al., 2011). Este algoritmo gera a região de interesse a partir de um método dinâmico de limiarização, permitindo investigar dinamicamente apenas a porção da imagem à frente do veículo, a fim da deteção da área de navegação e evitamento de obstáculos. Uma evolução do método de segmentação apresentado por Miranda Neto & Rittner (2006) foi proposto. Uma área navegável é, portanto, representada a partir de uma imagem multimodal 2D (*a multimodal 2D drivability road image*). Este modelo multimodal permite a escolha de um índice de segurança de acordo com o ambiente e contexto operacional.

A fim de reduzir o tempo de processamento, um critério automático da descarte de informações redundantes foi proposto por Miranda Neto, et al. (2007) e Miranda Neto, et al. (2008). Este método melhorou o desempenho de um sistema em tempo real, selecionando, de forma automática, quais imagens devem ser descartadas e quais devem ser tratadas pelo sistema de percepção visual. Por conseguinte, levando-se em conta a coerência temporal entre consecutivas imagens, uma nova metodologia de gerenciamento dinâmico de energia (*Dynamic Power Management* - DPM) foi proposta por Miranda Neto, et al. (2011) e aplicado a um sistema de percepção visual robótico, que incluiu um método de Gerenciamento do Impacto Culmulativo de Dados (*Cumulative Impact Data Management*). Além disso, esta proposta também apresenta um sistema de percepção em tempo real capaz de estimar o risco de colisão baseado no método de Pearson (Pearson, 1895) em ambientes dinâmicos e desconhecidos. Nota-se a originalidade destes trabalhos.

Uma característica marcante das metodologias apresentadas é a independência ao sistema de aquisição de imagem e do próprio robô. Estes sistemas de percepção em tempo real foram avaliados usando diferentes conjuntos de dados e também a partir de dados reais obtidos por duas plataformas inteligentes. Os métodos apresentados nesta tese foram submetidos a mais de 47.500 imagens. Em tarefas semi-autônomas, testes foram conduzidos em velocidades acima de 100 Km/h. A partir de um sistema em malha aberta, deslocamentos reativos autônomos foram realizados com sucesso. É importante notar que esses algoritmos não são baseados em conhecimento prévio do ambiente (forma da pista, geométria, etc) nem necessitam de calibração da câmera.

Trabalhos futuros terão foco na avaliação das metodologias propostas em relação à bases de testes (*ground truth*). Esforços de pesquisa também incluirão sistemas multi-sensores baseados principalmente em radar, visão e laser.

Conclusion

Cette thèse a abordé la problématique d'évitement d'obstacles pour les plates-formes terrestres semi- et autonomes dans des environnements dynamiques et inconnus. Basé sur un système de vision monoculaire embarqué, un ensemble d'outils a été proposé pour surveiller en permanence l'espace avant du véhicule, prouvant le système d'informations routières appropriées en temps réel. Ces outils ont été testés dans différents types d'images et de textures (revêtements routiers), y compris également la détection de la région navigable, la navigation réactive et l'estimation du risque de collision.

Un algorithme pour la détection de la ligne d'horizon a été développé pour enlever la région de l'horizon (Miranda Neto, et al., 2011). Cet algorithme génère la région d'intérêt à partir d'une méthode de recherche de seuil dynamique, permettant d'examiner dynamiquement seulement une petite portion de l'image avant du véhicule, pour détecter la région navigable et éviter les obstacles. Une évolution de la méthode de segmentation présentée par Miranda Neto & Rittner (2006) a été proposée. Une zone navigable, libre d'obstacles, est donc représentée par une image multimodale 2D. Ce modèle multimodal permet la sélection d'un indice de sécurité selon l'environnement et le contexte opérationnel.

Afin de réduire les temps de traitement de l'information, une méthode d'identification et de rejet d'informations redondantes a été proposée par Miranda Neto, et al. (2007) et Miranda Neto, et al. (2008). Tenant compte de la cohérence temporelle entre des images consécutives, une étude innovante par rapport à l´état de l´art en ce qui concerne la gestion énergétique du système de vision embarqué a également été proposée et appliquée à un système de perception visuelle robotique (Miranda Neto, et al., 2011). Le développement d'une estimation du risque de collision basé sur le Coefficient de Corrélation de Pearson (Pearson, 1895) est également à noter dans le cadre de l'originalité de ce travail.

Une caractéristique remarquable des méthodologies présentées dans cette thèse est l'indépendance du système d'acquisition d'image et du robot lui-même. Ce système de perception en temps réel a été évalué en utilisant différents ensembles de base de données et aussi à partir

des données réelles obtenues par deux plates-formes intelligentes. Les méthodes présentées ici ont été soumis à plus de 47 500 images. En mode semi-autonome, certains tests ont été effectués à des vitesses supérieures à 100 Km/h. Des déplacements réactifs autonomes en boucle ouverte ont également été effectués avec succès. Il est important de noter que ces algorithmes ne sont pas fondés sur les connaissances antérieures de l'environnement (forme de voie, d'inférence géométrique, etc), ni de calibration de la caméra.

Les travaux futurs porteront sur l'évaluation des méthodologies proposées par rapport à la réalité de terrain. Un autre effort de recherche étudiera un système multi-capteurs basé principalement sur la vision, radar et laser.

References

Abutaleb, H. A. (1989). Automatic Thresholding of Gray-Level Pictures Using Two-Dimensional Entropy. *Computer Vision, Graphics, and Image Processing* .

Alenya, G. N. (2009). A Comparison of Three Methods for Measure of Time to Contact. *IEEE Proceedings of the International Conference on Intelligent Transportation Systems (ITSC 2009)* .

Anton, S. R., & Inman, D. J. (2008). *Energy Harvesting for Unmanned Aerial Vehicles.* Retrieved 07 17, 2011

Aviña-Cervantes, G., Devy, M., & Marín, A. (2003). Lane Extraction and Tracking for Robot Navigation in Agricultural Applications. *Proceedings of the International Conference on Advanced Robotics, ICAR 2003* .

Bakker, T., Wouters, H., Asselt, K. v., Bontsema, J., Tang, L., Muller, J., et al. (2008). A vision based row detection system for sugar beet. *Computers and Electronics in Agriculture, Elsevier* , 87–95.

Ballard, D. (1981). Generalized Hough transform to detect arbitrary shapes. *IEEE Trans. Pattern Anal. Machine Intell. , 2*, pp. 111–122.

Beauchemin, S. a. (1995). The computation of optical flow. *ACM Computing Surveys , 27*, pp. 433–467.

Benini, L. B. (2000). A Survey of Design Techniques for System-Level Dynamic Power Management. *IEEE Transactions on Very Large Scale Integration Systems , 3*, pp. 299-316.

Berlin, T. (2007). *Spirit of Berlin: An Autonomous Car for the DARPA Urban Challenge Hardware and Software Architecture.* Retrieved 12 02, 2010, from http://www.darpa.mil/grandchallenge/ TechPapers/Team_Berlin.pdf

Bertozzi, M. a. (1998). GOLD: a parallel real-time stereo vision system for generic obstacle and lane detection. *IEEE Transactions on Image Processing* , pp. 62–81.

Bertozzi, M. B. (1998). Stereo inverse perspective mapping: theory and applications. *Image and Vision Computing , 16*, pp. 585–590.

Bertozzi, M. B. (2000). Vision-based intelligent vehicles: state of the art and perspectives. *Robotics and Autonomous systems* , pp. 1–16.

Beyeler, A. Z. (2009). Vision-based control of near-obstacle flight. *Autonomous Robots , 27* (3), pp. 201–219.

Bonin-Font, F., Ortiz, A., & Oliver, G. (2008). Visual Navigation for Mobile Robots: A Survey. *Journal of Intelligent and Robotic Systems* , 263-296.

Bouyer, P. e. (2010). Quantitative analysis of real-time systems. *Journal Communications of the ACM* .

Canny, J. F. (1986). A computational approach to edge detection. *IEEE Trans. Pattern Anal. Machine Intell. , 6*, pp. 679–698.

Chatila, R. (2010). Dual research: Issues in robotics. *Oral presentation, Mission for Research and Innovation Scientific 2010. French Armament Procurement Agency (DGA), Paris, France* .

Chetan, J., Madhava, K., & Jawahar, C. V. (2010). An Adaptive Outdoor Terrain Classification Methodology using Monocular Camera. *Proceedings of the International Conference on Intelligent Robots and Systems, IROS 2010* .

Dagan, E. M. (2004). Forward collision warning with a single camera. *IEEE Intelligent Vehicles Symposium* , pp. 37–42.

Dahlkamp, H., Kaehler, A., Stavens, D., Thrun, S., & Bradski, G. (2006). Self-Supervised Monocular Road Detection in Desert Terrain. *Proceedings of the Robotics Science and Systems Conference* .

DARPA. (2005). *DARPA Grand Challenge.* Retrieved 02 10, 2008, from http://www.darpa.mil/grandchallenge05/

DARPA. (2009). *Energetically Autonomous Tactical Robot, DARPA Contract W31P4Q-08-C-0292.* Retrieved 05 27, 2011, from EATR at Robotic Technology: http://www.robotictechnologyinc.com/index.php/EATR

Deshmukh, A. V. (2010). *Towards Socially Constrained Power Management for Long-Term Operation of Mobile Robots.* Retrieved 07 17, 2011, from http://www.macs.hw.ac.uk/~ad166/papers/2010/TAROS.pdf

Diego, F., Álvarez, J. M., Serrat, J., & López, A. M. (2010). Vision based Road Detection via On line Video Registration. *Proceedings of the IEEE Int. Conf. on Intelligent Transportation Systems, ITSC 2010* .

Entzinger, J. O. (2008). Modeling of the visual approach to landing using neural networks and fuzzy supervisory control. *Proceedings of the 26th International Congress of the Aeronautical Sciences (ICAS2008)* .

Ettinger, S., Nechyba, M., Ifju, P., & Waszak, M. (2003). Vision-guided flight stability and control for micro air vehicles. *Advanced Robotics , 17*, pp. 617-640.

Eugene, Y. K. (1996). *The Ineffectiveness of the Correlation Coefficient for Image Comparisons.* Retrieved 07 17, 2011, from http://jps.anl.gov/vol.2/3-Correlation.pdf

European Commission. (2010, 07 05). *Radio Spectrum Committee, RSCOM10-35.* Retrieved 05 28, 2011, from Radio Spectrum Committee, European Commission, Public Document, Brussels: http://ec.europa.eu/information_society/policy/ecomm/radio_spectrum/_document_storage/rsc/rsc32_public_docs/rscom10_35.pdf

Finkelstein, R. a. (2010). *Patent No. 12/612,489 / US 2010/0155156 A1.* US.

FLIR. (2010). *FLIR Thermal Imaging, Night Vision and Infrared Camera Systems*. Retrieved 07 10, 2010, from Thermal imaging and the PathFindIR: An automotive night vision camera: http://www.flir.com/

Geiger, A. (2011). *Institute of Measurement and Control Systems, Karlsruhe Institute of Technology.* Retrieved 07 01, 2011, from http://www.rainsoft.de/software/datasets.html.

Gietelink, O. J. (2009). Development of a driver information and warning system with vehicle hardware-in-theloop simulations. *Mechatronics , 19*, pp. 1091–1104.

Gietelink, O. J. (2006). Development of advanced driver assistance systems with vehicle hardware-in-the-loop simulations. (V. S. Dynamics, Ed.) *Vehicle System Dynamics ,* 569-590.

Gonzalez, C. R. (1991). Digital Image Processing. *Addison-Wesley Publishing Company .*

Greco, C. R. (2008). *Real-Time Forward Urban Environment Perception for an Autonomous Ground Vehicle Using Computer Vision And Lidar.* Retrieved 07 17, 2011

Horn, B. K. (1986). Robot Vision. *The MIT Press .*

Hwangbo, M. (2009). Robust monocular vision-based navigation for a miniature fixed-wing aircraft. *Ph.D. proposal, Robotics institute, Carnegie Mellon University.*

INRIA, R. (2011). *State of the Art Report and Requirement Specifications.* Retrieved 02 03, 2011, from Hybrid Intelligent Virtual Actors, Integrating Research in Interactive Storytelling: http://iris.scm.tees.ac.uk/public-wiki/images/a/a6/IRIS-D4.1.pdf

Kalman, R. (1960). A new approach to linear filtering and prediction problems. *Journal of Basic Engineering ,* pp. 35–45.

Kim, B., Hubbard, P., & Necsulescu, D. (2003). *Swarming Unmanned Aerial Vehicles: Concept Development and Experimentation, A State of the Art Review on Flight and Mission Control.* Retrieved from http://pubs.drdc.gc.ca/PDFS/unc35/p520549.pdf

Klesh, A. T., & Kabamba, P. T. (2009). Solar-Powered Aircraft: Energy-Optimal Path Planning and Perpetual Endurance. *Journal of Guidance, Control, and Dynamics ,* 1320-1329.

Laplant, P. (2004). *Real-Time System Design and Analysis.* Retrieved 07 17, 2011, from http://www.ee.ui.ac.id/muis/course_file/Real-Time_Systems_Design_and_Analysis.pdf

Lee, U. C. (1990). A Comparative Performance Study of Several Global Thresholding Techniques for Segmentation. *Computer Vision, Graphics, and Image Processing , 52* (2), pp. 171-190.

Lim, K. H., & al., e. (2009). Vision-based Lane-Vehicle Detection and Tracking. *American Institute of Physics , 3 Special Edition*, pp. 157-171.

Mallot, H. A. (1991). Inverse perspective mapping simplifies optical flow computation and obstacle detection. *Biological Cybernetics ,* pp. 177–185.

Mesbah, M. (1999). Gradient-based optical flow: a critical review. *Proc. of the Fifth Int. Symp. on Signal Processing and Its Applications , 1*, pp. 467–470.

Miranda Neto, A., & Rittner, L. (2006). A Simple and Efficient Road Detection Algorithm for Real Time Autonomous Navigation based on Monocular Vision. *Proceedings of the 2006 IEEE 3rd Latin American Robotics Symposium* .

Miranda Neto, A., Rittner, L., Leite, N., Zampieri, D. E., & Victorino, A. C. (2008). Nondeterministic Criteria to Discard Redundant Information in Real Time Autonomous Navigation Systems based on Monocular Vision. *Proceedings of the IEEE Multi-conference on Systems and Control, MSC 2008, ISIC Invited Paper* , 420-425.

Miranda Neto, A., Rittner, L., Leite, N., Zampieri, D. E., Lotufo, R., & Mendeleck, A. (2007). Pearson's Correlation Coefficient for Discarding Redundant Information in Real Time Autonomous Navigation Systems. *Proceedings of the IEEE Multi-conference on Systems and Control, MSC 2007* , 426-431.

Miranda Neto, A., Victorino, A. C., Fantoni, I., & Zampieri, D. E. (2011). Real-Time Dynamic Power Management based on Pearson's Correlation Coefficient. *Proceedings of the IEEE International Conference on Advanced Robotics, ICAR 2011* .

Miranda Neto, A., Victorino, A. C., Fantoni, I., & Zampieri, D. E. (2011). Robust Horizon Finding Algorithm for Real-Time Autonomous Navigation based on Monocular Vision. *Proceedings of the IEEE Int. Conf. on Intelligent Transportation Systems, ITSC 2011* .

Morita, K. a. (2003). Examination of the Evaluation Methods for Hybrid Electric Vehicles. *Japan Automobile Research Institute Research Journal , 25* (11), pp. 437-440.

Müller, D. P.-S. (2009). Time To Contact Estimation Using Interest Points. *IEEE Proceedings of the International Conference on Intelligent Transportation Systems (ITSC 2009)* .

NASA. (2004). *The Vision for Space Exploration.* Retrieved 05 27, 2011, from National Aeronautics and Space Administration: 202.41.85.234:8000/InfoUSA/tech/space/space_exp.pdf

Negre, A. B. (2006). Real-time Time-To-Collision from variation of Intrinsic Scale. *INRIA base, Proc. of the Int. Symp. on Experimental Robotics* .

Obayashi, K. a. (2004). Concept of Vehicle Electric Power Flow Management System (VEF). *SAE 2004 World Congress & Exhibition No.2004-01-0361* .

Otsu, N. (1978). A threshold selection method from gray-level histogram. *IEEE Transactions on Systems, Man, and Cybernetics* , pp. 62-66.

Pearson, K. (1895). *Royal Society Proceedings* , p. 241.

R. Finkelstein, R. T. (2010). *Patent No. 20100155156.*

Rauskolb, F. W., & al., e. (2008). Caroline: An autonomously driving vehicle for urban environments. *Journal of Field Robotics , 25 (9)*, pp. 674–724.

Rives, P., & Azinheira, J. R. (2004). Linear Structures Following by an Airship using Vanishing Point and Horizon Line in Visual Servoing Scheme. *ICRA'04 , 1*, pp. 255-260.

Rodgers, J. L. (1988). *Thirteen Ways to Look at the Correlation Coefficient.* Retrieved 07 17, 2011, from http://data.psych.udel.edu/laurenceau/PSY861Psychological%20Statistics%20II%20Spring%202010/REA DINGS/rodgers-nicewander-1988-r-13-ways.pdf

Rodríguez Flórez, S. A. (2010). *Contributions by Vision Systems to Multi-sensor Object Localization and Tracking for Intelligent Vehicles.* University of Technology of Compiegne. Heudiasyc Laboratory, UMR UTC/CNRS 6599.

Rojo, J., Rojas, R., Gunnarsson, K., Simon, M., Wiesel, F., Ruff, F., et al. (2007). *Spirit of Berlin: An Autonomous Car for the DARPA Urban Challenge Hardware and Software Architecture.* Team Berlin, DARPA.

Ruffier, F. a. (2005). Optic flow regulation: the key to aircraft automatic guidance. *Robotics Autonomous Systems , 50*, pp. 177–194.

Sahoo, P. K. (1988). A survey of thresholding techniques. *Comput. Vision Graphics Image Processing , 41* (2), pp. 233-260.

Sezgin, M. a. (2004). Survey over image thresholding techniques and quantitative performance evaluation. *Journal of Electronic Imaging , 13*, pp. 146-165.

Stanford. (2006). *Stanford Racing Team's Entry In The 2005 DARPA Grand Challenge.* Retrieved 07 17, 2011, from http://ai.stanford.edu/~dstavens/darpa/Stanford.pdf

Subramanian, V., Burks, T. F., & Arroyo, A. (2006). Development of machine vision and laser radar based autonomous vehicle guidance systems for citrus grove navigation. *Computers and Electronics in Agriculture, Elsevier ,* 130–143.

Tan, S. D. (2006). Inverse perspective mapping and optic flow: A calibration method and a quantitative analysis. *Image and Vision Computing , 24* (2), pp. 153–165.

TEAM OSHKOSH. (2011). *Terramax - Darpa Urban Challenge.* Retrieved 05 28, 2011, from Terramax: http://www.terramax.com/technology/the_brawn.cfm

Thrun, S., Montemerlo, M., Dahlkamp, H., Stavens, D., Aron, A., Diebel, J., et al. (2006). Stanley, the robot that won the DARPA Grand Challenge. *Journal of Robotic Systems, Volume 23, Issue 9, ISSN:0741-2223 ,* 661-692.

Ulrich, I., & Nourbakhsh, I. (2000). Appearance-Based Obstacle Detection with Monocular Color Vision. *Proceedings of the AAAI National Conference on Artificial Intelligence .*

VBTP-MR Project. (2011). *Department of Science and Technology (DCT), Brazilian Army.* Retrieved 05 28, 2011, from DCT, Brazilian Army: http://www.dct.eb.mil.br/

White, F. M. (1986). Fluid Mechanics. *2nd Ed. McGraw Hill .*

Wu, S. D. (2009). Collision sensing by stereo vision and radar sensor fusion. *IEEE Transactions on Intelligent Transportation Systems , 10* (4).

Yamaguchi, K., Watanabe, A., Naito, T., & Ninomiya, Y. (2008). Road region estimation using a sequence of monocular images. *Pattern Recognition, 2008. ICPR 2008* .

Yanqing, W. D. (2010). Vision-based Road Detection by Monte Carlo Method. *Information Technology Journal , 9*, pp. 481-487.

Yongguo Mei, H. Y.-H. (2005). A case study of mobile robot's energy consumption and conservation techniques. *Proceedings of the IEEE International Conference on Advanced Robotics, ICAR 2005* , pp. 492-497.

Appendix A

Pearson´s Correlation Coefficient

The Pearson's Correlation Coefficient (PCC) is widely used in statistical analysis, pattern recognition and image processing. Applications on the latter include comparing two images for image registration purposes, disparity measurement, etc [K. Pearson, (1895) [53]; J. L. Rodgers and W. A. Nicewander, (1988)[54]; Eugene and Johnston, (1996)[55]]. It is described as:

$$r_1 = \frac{\sum_i (x_i - x_m)(y_i - y_m)}{\sqrt{\sum_i (x_i - x_m)^2} \sqrt{\sum_i (y_i - y_m)^2}}$$

(A.1)

where x_i is the intensity of the ith pixel in image 1, y_i is the intensity of the ith pixel in image 2, x_m is the mean intensity of image 1, and y_m is the mean intensity of image 2. The PCC threshold, r_1, has value 1 if the two images are identical, 0 if they are completely uncorrelated, and −1 if they are completely anti-correlated.
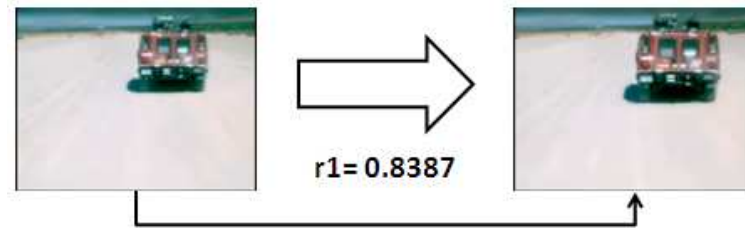

Figure A.1 – Pearson's Correlation Coefficient.

[53] K. Pearson, (1895), Royal Society Proceedings, 58, 241.
[54] J. L. Rodgers and W. A. Nicewander, (1988), "Thirteen Ways to Look at the Correlation Coefficient", The American Statistician, 42, 59-66.
[55] Eugene and R.G. Johnston, "The Ineffectiveness of the Correlation Coefficient for Image Comparisons", Technical Report LA-UR-96-2474, Los Alamos, 1996.

Appendix B

Otsu Thresholding Method

Region recognition can be handled by popular thresholding algorithm such as Maximum Entropy, Invariant Moment and Otsu Thresholding Method (OTM) (Otsu, 1978)[56]. Especially, for navigable area detection, because OTM supplies a more satisfactory performance in image segmentation, it was used to overcome the negative impacts caused by environmental variation (Yanqing, et al., 2010)[57]. Furthermore, some authors consider the OTM as one of the best choices for real-time applications in machine vision (Sahoo, et al., 1988)[58], (Lee, et al., 1990)[59]. It still remains one of the most referenced thresholding methods (Sezgin and Sankur, 2004)[60].

This work considers a color or gray-level image. If the image is colored, in order to utilize the most important information of the color image, the candidate color channel that was dominant in certain color space is selected to generate the histogram image (Yanqing, et al., 2010). This empirical space is a small portion of the road ahead of the vehicle, where the absence of other vehicles has been assumed (Bertozzi, et al., 2000)[61]. For this thesis, an empirical space is presented in section 3.4 and shown in Figure 3.8 (a): yellow area. From this empirical space, the candidate color channel is then obtained in Equation B.1:

$$Cc = \underset{\{R,G,B\}}{\mathrm{argmax}}(N_R, N_G, N_B)$$

(B.1)

[56] Otsu, N. (1978). A threshold selection method from gray-level histogram. IEEE Transactions on Systems, Man, and Cybernetics , pp. 62-66.

[57] Yanqing, W., Deyun, C., Chaoxia, S. and Peidong, W. "Vision-based Road Detection by Monte Carlo Method." Information Technology Journal, 2010: 481-487.

[58] Sahoo, P. K., Soltani, S. and Wong, A. K. C. "A survey of thresholding techniques." Comput. Vision Graphics Image Processing, 1988: 233-260.

[59] Lee, U.S., Chung, Y.S. and Park, H.R. "A Comparative Performance Study of Several Global Thresholding Techniques for Segmentation." Computer Vision, Graphics, and Image Processing, 1990: 171-190.

[60] Sezgin, M. and Sankur, B. "Survey over image thresholding techniques and quantitative performance evaluation." Journal of Electronic Imaging, 2004: 146-165.

[61] Bertozzi, M., Broggi, A. and Fascioli, A. "Vision-based intelligent vehicles: state of the art and perspectives." Robotics and Autonomous systems, 2000, 32 ed.: 1–16.

where Cc means the color channel and N_R means the number of the dominant color channel in certain referenced region.

The OTM exhaustively searches for the threshold that minimizes the intra-class variance, defined as a weighted sum of variances of the two classes. The thresholding process is seen as the partitioning of pixels of an image in two classes (foreground and background): C1={0, 1,…, T} and C2={T+1, T+2,…, N−1}, where T is the chosen threshold and N the number of intensity levels of the image. It is described by:

$$\sigma_W^2(T) = q_1(T)\sigma_1^2(T) + q_2(T)\sigma_2^2(T) \tag{B.2}$$

where $\sigma_W^2(T)$ is the intra-class variance, $q_1(T)$ is the proportion of background pixels, $\sigma_1^2(T)$ is the intensity variance of background pixels, $q_2(T)$ is the proportion of foreground pixels and $\sigma_2^2(T)$ is the intensity variance of foreground pixels.

The class probabilities are estimated as:

$$q_1(T) = \sum_{i=1}^{T} H(i) \text{ and } q_2(T) = \sum_{i=T+1}^{N} H(i) \tag{B.3}$$

The class means are given by:

$$\mu_1(T) = \sum_{i=1}^{T} \frac{iH(i)}{q_1(T)} \text{ and } \mu_2(T) = \sum_{i=T+1}^{N} \frac{iH(i)}{q_2(T)} \tag{B.4}$$

Finally, the individual class variances are:

$$\sigma_1^2(T) = \sum_{i=1}^{T} [i - \mu_1(T)]^2 \frac{H(i)}{q_1(T)} \tag{B.5}$$

$$\sigma_2^2(T) = \sum_{i=T+1}^{N} [i - \mu_2(T)]^2 \frac{H(i)}{q_2(T)} \tag{B.6}$$
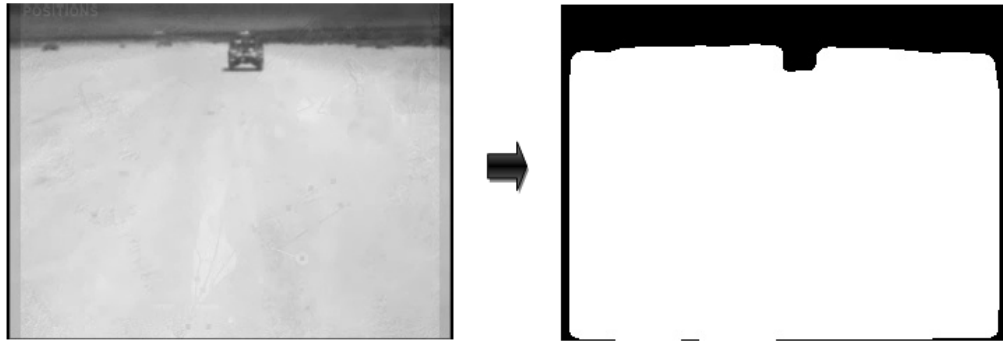
where H is the histogram of the selected channel in Equation B.1.

Figure B.1 – Otsu Thresholding Method.