



HAL
open science

Navigation de robot avec conscience sociale : entre l'évaluation des risques et celle des conventions sociales

Jorge Rios Martinez

► **To cite this version:**

Jorge Rios Martinez. Navigation de robot avec conscience sociale : entre l'évaluation des risques et celle des conventions sociales. Autre. Université de Grenoble, 2013. Français. NNT : 2013GRENM045 . tel-00957945

HAL Id: tel-00957945

<https://theses.hal.science/tel-00957945>

Submitted on 11 Mar 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ DE GRENOBLE

Spécialité : **Mathématiques-Informatique**

Arrêté ministériel : 7 août 2006

Présentée par

Jorge Alberto Rios Martinez

Thèse dirigée par **Christian Laugier**

et codirigée par **Anne Spalanzani**

préparée au sein du centre de recherche **INRIA Rhône-Alpes du
Laboratoire d'Informatique de Grenoble**

dans l'Ecole Doctorale **Mathématiques, Sciences et Technologies de
l'Information, Informatique**

**Socially-Aware Robot Navigation:
combining Risk Assessment and
Social Conventions**

Thèse soutenue publiquement le **8 Janvier 2013**,
devant le jury composé de :

James CROWLEY

Professeur Grenoble INP, France, Président

Rachid ALAMI

Directeur de Recherche, LAAS-CNRS, France, Rapporteur

Olivier SIMONIN

Maître de Conférences, HDR, Université de Lorraine, France, Rapporteur

David DANNEY

Chargé de Recherche, INRIA Sophia-Antipolis, France, Examineur

Christian LAUGIER

Directeur de Recherche, INRIA Rhône-Alpes, France, Directeur de thèse

Anne SPALANZANI

Maître de Conférences, UPMF, Grenoble, France, Co-Directeur de thèse



Acknowledgements

Finishing this thesis represents a big satisfaction both professional and personal to me. It has been a complete experience of life, a rich process that provided me great experiences and maturity in the field I like to work. I want to acknowledge to all the people that collaborated in the fulfillment of my objectives.

First, I want to thank my thesis directors Christian Laugier and Anne Spalanzani for their advice during all the thesis work and also for their oportune observations when I was stuck in the middle of the path.

Also, I would like to thank to the members of the jury who examine my thesis, David Daney and James Crowley for their pertinent questions and observations and specially to Rachid Alami and Olivier Simonin for accepting to review my thesis.

I want to thank to all members of e-motion team who support me and share with me not only their knowledge but also their friendship: Luis Martinez, Mao, Chiara Fulgenzi, Thiago Bellardi, Mathias Perrolaz, Alexandros Makris, Amaury Negre, Procopio Stein, Dizan Vazquez. Specially, thanks to the best officemates Chiara Troiani and Stephanie Lefevre, it was a pleasure to work with you. Thanks to Arturo Escobedo for the time invested and his help with the preparation of robotic wheelchair demonstrations.

Thanks to Alessandro Renzaglia for accepting to collaborate with me and to explain me the CAO method.

I must mention the great help that Myriam Etienne always provided with the administrative processes.

Thanks to CONACYT for the financial support granted for the realization of my studies (Reference. No. 250140).

Finally, I want to thank to my wife Karla for all the support she gave me, her patience and for motivating me all along this adventure. Definitely this achievement is also her achievement.

Abstract

This thesis proposes a risk-based navigation method including both the traditional notion of risk of collision and the notion of risk of disturbance. With the growing demand of personal assistance to mobility and mobile service robotics, robots and people must share the same physical spaces and follow the same social conventions. Robots must respect proximity constraints but also respect people interacting. For example, they may not break interaction between people talking, unless the robot task is to take part in the conversation. In this case, it must be able to join the group using a socially adapted behavior. The socially-aware navigation system proposed in this thesis integrates both an assessment of a risk of collision using predictive models of moving obstacles, and an assessment of accordance with social conventions. Human management of space (personal space, o-space, activity space...) inspired from sociology and social robotics literature is integrated, but also models of behavior that enable the robot to make medium-term prediction of the human positions. Simulation and experimental results on a robotic wheelchair validate the method by showing that our robot is able to navigate in a dynamic environment avoiding collisions with obstacles and people and, at the same time, minimizing discomfort in people by respecting spaces mentioned above.

Contents

| | |
|--|------------|
| Contents | iii |
| List of Figures | vii |
| Nomenclature | ix |
| 1 Introduction | 1 |
| 1.1 Motivation | 1 |
| 1.2 Problem description | 2 |
| 1.3 Contributions | 4 |
| 1.4 Overview | 4 |
| 1.5 Publications | 5 |
| 2 Social conventions: from comfort and proxemics to social robotics | 9 |
| 2.1 Introduction | 9 |
| 2.2 Social behavior and comfort | 11 |
| 2.3 Proxemics: human management of space | 14 |
| 2.3.1 Space related to one person | 14 |
| 2.3.1.1 Personal Space | 14 |
| 2.3.1.2 Information Process Space | 18 |
| 2.3.2 Space related to groups of people | 19 |
| 2.3.2.1 Interaction spaces: the O-Space concept | 20 |
| 2.3.2.2 Arrangements of groups. | 22 |
| 2.3.3 Space related to objects: Activity Space and Affordance Space . | 24 |
| 2.3.4 Space related to robots | 27 |
| 2.4 Social Robotics | 30 |
| 2.4.1 Social robot abilities | 30 |
| 2.4.2 Socially-aware robot navigation | 32 |

CONTENTS

| | | |
|----------|---|-----------|
| 2.4.3 | Socially-aware robot navigation related work | 32 |
| 2.4.3.1 | Unfocused interaction | 33 |
| 2.4.3.2 | Focused interaction | 35 |
| 2.4.3.3 | Focused and unfocused interaction | 37 |
| 2.5 | Conclusion and thesis positioning | 38 |
| 3 | Modeling social behaviors : The Social Filter | 41 |
| 3.1 | Introduction | 41 |
| 3.2 | Space related to one person | 42 |
| 3.2.1 | Model of Personal Space | 42 |
| 3.2.2 | Model of Information Process Space | 44 |
| 3.3 | Space related to groups of people | 45 |
| 3.3.1 | The shape of O-space: Preliminary psychophysics experiment | 45 |
| 3.3.2 | Model of O-Space in two-person formations | 48 |
| 3.3.3 | Approaching a two-person formation | 51 |
| 3.3.4 | Groups of more than two people | 52 |
| 3.4 | Space related to objects | 54 |
| 3.5 | Prediction of O-Space location in future conversations | 55 |
| 3.5.1 | Analysis of interaction scenarios | 56 |
| 3.5.2 | Model of interaction using Hidden Markov Models (HMM) | 64 |
| 3.5.2.1 | HMM definition and notation | 64 |
| 3.5.2.2 | Example of HMM | 65 |
| 3.5.2.3 | Model of interaction | 66 |
| 3.5.2.4 | Results of interaction vs no interaction classification | 68 |
| 3.6 | Conclusions | 69 |
| 4 | Including social conventions in robot navigation | 71 |
| 4.1 | Introduction | 71 |
| 4.2 | Navigation in dynamic environments: state of the art | 72 |
| 4.3 | Optimization-based navigation with Information Process Space and Personal Space conventions | 75 |
| 4.3.1 | Environment and robot model | 76 |
| 4.3.2 | Navigation strategy | 77 |
| 4.3.2.1 | Cognitive-based Adaptive Optimization formulation | 77 |
| 4.3.2.2 | Social conventions | 80 |
| 4.3.2.3 | Movement Prediction | 82 |

| | | |
|----------|---|------------|
| 4.3.3 | Simulation results | 85 |
| 4.4 | Risk-based navigation with Personal Space, O-Space and Activity Space conventions | 91 |
| 4.4.1 | Environment and robot model | 92 |
| 4.4.2 | Navigation strategy | 93 |
| 4.4.2.1 | The Algorithm RiskRRT | 93 |
| 4.4.2.2 | Probabilistic Risk of Collision | 95 |
| 4.4.2.3 | Social conventions | 99 |
| 4.4.3 | Simulation results | 100 |
| 4.5 | Conclusions | 103 |
| 5 | Socially-aware navigation for assisted mobility | 107 |
| 5.1 | Introduction | 107 |
| 5.2 | Perception of human motion and activity | 109 |
| 5.2.1 | Detection techniques | 110 |
| 5.2.2 | Motion: models, tracking and prediction | 111 |
| 5.2.3 | Activity recognition | 113 |
| 5.2.4 | Measuring human activity | 114 |
| 5.3 | The assisted mobility system | 115 |
| 5.3.1 | Requirements | 115 |
| 5.3.2 | Structure of the proposed system | 117 |
| 5.3.2.1 | Risk-Based Socially-Aware Navigation | 118 |
| 5.3.2.2 | Pedestrians tracking and motion prediction | 118 |
| 5.4 | Experimental platform | 119 |
| 5.5 | Experimental Results | 121 |
| 5.5.1 | Perception of interactions with external camera | 121 |
| 5.5.2 | Perception of interactions with kinect sensor | 122 |
| 5.5.3 | Navigation results | 124 |
| 5.6 | Conclusions | 127 |
| 6 | Conclusions and perspectives | 129 |
| A | Assisted Mobility application | 133 |
| A.1 | Accessible Air Travel regulation | 133 |
| A.2 | Problem description and proposed application | 134 |

CONTENTS

List of Figures

| | | |
|------|--|----|
| 1.1 | Robots moving in dynamic environments populated by humans | 1 |
| 2.1 | The most important components of a socially-aware navigation system. | 10 |
| 2.2 | Typical arrangement exhibited by single humans | 15 |
| 2.3 | Different shapes of Personal Space | 16 |
| 2.4 | Information Process Space | 18 |
| 2.5 | Arrangement of people interacting in groups | 20 |
| 2.6 | Circular O-space in conversations for groups of more than two people. . | 24 |
| 2.7 | Activity and affordance spaces | 25 |
| 2.8 | Examples of unfocused interaction | 34 |
| 2.9 | Examples of focused interaction | 36 |
| 2.10 | Social spaces considered in this thesis | 38 |
| 3.1 | Flow of information related to the Social Filter | 41 |
| 3.2 | Personal space calculated by Social Filter Module | 43 |
| 3.3 | Information Process Space calculated by Social Filter Module | 45 |
| 3.4 | Psychophysic experiment | 46 |
| 3.5 | Scenario of the experiment at INRIA's Robotics Lab | 47 |
| 3.6 | Analysis of trajectories | 48 |
| 3.7 | Elements of O-space model | 50 |
| 3.8 | O-space calculated by Social Filter Module | 51 |
| 3.9 | Meeting points | 52 |
| 3.10 | Output of Social Filter, C-shape and Vis-a-vis formations detected . . . | 54 |
| 3.11 | Interaction of a human with an object of interest is detected | 55 |
| 3.12 | Two common situations appearing in human environments | 56 |
| 3.13 | Designed scenarios to study | 57 |
| 3.14 | Sequences recorded were labeled | 58 |

LIST OF FIGURES

| | | |
|------|---|-----|
| 3.15 | Behavior of the variable relative distance | 59 |
| 3.16 | Behavior of the variable relative speed | 60 |
| 3.17 | Behavior of the variable alignment | 60 |
| 3.18 | Behavior of the change in relative distance | 61 |
| 3.19 | Discretization of the space around people | 62 |
| 3.20 | Behavior of the formation variable | 63 |
| 3.21 | Graph used to represent an HMM model | 66 |
| 3.22 | HMM models for interaction and no-interaction | 67 |
| 3.23 | Average of the likelihood calculated for two models | 69 |
| | | |
| 4.1 | Idea of CAO approach | 81 |
| 4.2 | Models implemented to represent discomfort | 82 |
| 4.3 | An example of prediction | 83 |
| 4.4 | Simulation of the robot navigating | 85 |
| 4.5 | More simulations with different scenarios | 86 |
| 4.6 | CAO-based Navigation implemented in ROS | 87 |
| 4.7 | Trajectory generated by the CAO algorithm | 88 |
| 4.8 | Second scenario of CAO in ROS | 90 |
| 4.9 | Growth of the tree in RiskRRT | 92 |
| 4.10 | Example of execution of Risk-RRT algorithm. | 93 |
| 4.11 | Example of prediction for a pedestrian. | 97 |
| 4.12 | Predictive navigation example. | 100 |
| 4.13 | Qualitative comparison of predictive navigation | 101 |
| 4.14 | Example of execution of RiskRRT | 102 |
| 4.15 | Socially-Aware navigation in RiskRRT | 104 |
| 4.16 | Socially-Aware navigation in dynamic environments | 105 |
| | | |
| 5.1 | Transport and assistance for people with motor difficulties | 109 |
| 5.2 | Sensors offering big potential to be used in Social Signal Processing | 115 |
| 5.3 | Achitecture overview. | 117 |
| 5.4 | Experimental platform | 120 |
| 5.5 | Test environment | 120 |
| 5.6 | Interaction between two people detected by using markers | 121 |
| 5.7 | Interaction between three people detected by using markers | 121 |
| 5.8 | Model detects formations that favor the focused interaction | 122 |
| 5.9 | Interactions detected with Social Filter from Kinect | 123 |
| 5.10 | Experimental test | 124 |

LIST OF FIGURES

LIST OF FIGURES

| | |
|--|-----|
| 5.11 Social Filter activated | 125 |
| 5.12 Social Filter deactivated | 126 |
| A.1 Autonomous transport of people with reduced mobility | 135 |
| A.2 Topological map of the airport | 135 |

LIST OF FIGURES

Chapter 1

Introduction

1.1 Motivation

Some years ago it was predicted that a transformation in the scope and dimension of robotic applications would move robots from industrial structured environments to unstructured environments populated by humans. Others had predicted that, in the same way as computers, the emergent industry of robotics would place a robot in every home. Current world trends continue to point toward that direction. For example, researchers across Europe are creating new designs and tackling fundamental problems that eventually could lead to a world standard for domestic robots. In USA

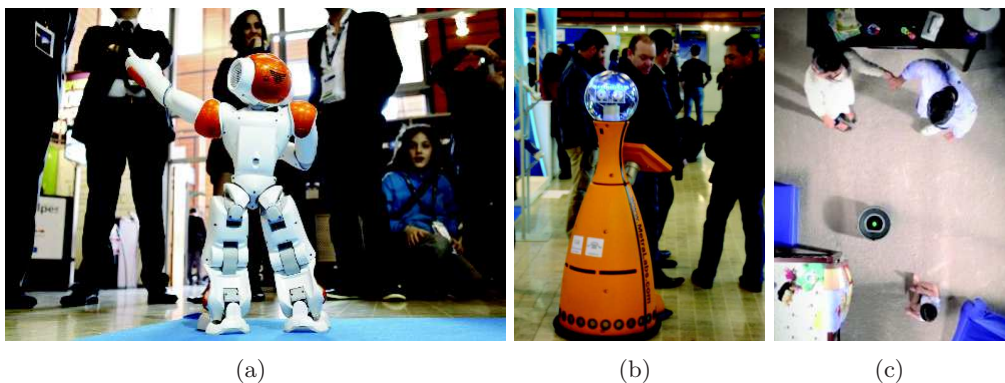


Figure 1.1: Examples of robots moving in dynamic environments populated by humans. In a) Nao by Aldebaran Robotics, in b) SCITOS A5 by Metralabs and in c) Roomba by iRobot.

robot vacuum cleaners, like Roomba, are example of successful non-factory robots. Some countries, like South-Korea, have established strategic plans involving companies,

1. INTRODUCTION

universities and research institutions with the objective of to seize the market of service robots.

Because of their multitude of forms and structures as well as application areas, service robots are not easy to define. In short, it can be said that service robots are those that work closely with humans, importantly the elderly and disabled, to help them with their lives ¹. In that context it can be observed that the physical space where the robots navigate will be shared with people who maintain a set of social conventions related to such space, Fig. 1.1 shows three examples of robots moving in the same space as humans. When moving, our robot may considerably disturb people around it, especially when its behavior is perceived as unsocial. Therefore to be integrated in a more natural way, a robot will have to understand the conventions mentioned before and to behave accordingly.

This thesis explores techniques about how to produce socially-aware motion in robots by including social conventions in their autonomous navigation strategy. It is expected that when the robot makes navigation decisions it can identify human behaviors (like meeting people to interact, talking with others, interacting with objects, etc.) in the environment, predict future states of the world, estimate risk associated to each possible solution and finally chooses the one with the lowest risk. Risk is not only related to collision but also to disturbance of social conventions.

As an example of scenario addressed in this work, imagine an airport where some people are walking alone at different speeds, small groups walking together, conversation groups, etc. A robotic wheelchair is providing a service to humans by transporting people with reduced mobility. For this system it is crucial to take into account the actual needs and characteristics of both its users and the people around them. From the robot side, the execution of an optimal planned trajectory that passes too close to humans talking will be considered as a violation to social behavior because it will interrupt a conversation. Even worse, the wheelchairs passenger may be held responsible for that behavior. In that cases a socially-aware navigation solution is needed.

1.2 Problem description

The general framework of the problem is that of autonomous robot navigation in dynamic environments which is still a challenging one where the main addressed aspect in the state of the art has been the safety of the robot. The assumption that environment

¹Taken from the foreword of the 6th International Conference on Field and Service Robotics, 2007. <http://www.inrialpes.fr/FSR07/index.html>

1. INTRODUCTION

is populated by human beings introduces a new important aspect: to include social awareness in the navigation solutions. Most of the previous approaches supposed that people are moving independently and considered people as dynamic obstacles or non-static objects. Recently, the social awareness issue has been addressed by considering humans not only as dynamic obstacles that respond to physical properties but also as social entities with intentions and feelings that follow social conventions according to psychological properties.

In this thesis the problem to be solved can be expressed as:

To autonomously and safely navigate in a dynamic environment populated by human beings while social awareness is included in the navigation decisions.

The main difficulties related to the presented problem are listed below:

1. In a dynamic environment there is uncertainty inherent to the future state of physical world. Models for the prediction of the most likely situation of the environment and the observed moving entities must be included. The representation of the environment must permit a fast updating with the new observations and be available when the navigation strategy requires information.
2. A second source of uncertainty is linked to the perception of the physical world. Each sensor has a particular model with specific advantages and drawbacks therefore a robust fusion technique of data coming from multiple sensors must be implemented considering uncertainty. Robust real-time techniques are needed in order to detect and track humans in the environment providing useful information to reach high level reasoning on scene situations. Many robust current techniques employed to register human motion rely on invasive sensors which are not suitable for common life settings.
3. Mathematical models capable to reproduce social behavior in practice must be developed. While a perfect model of human behavior could not be available, some techniques have already been used to replicate expected social behavior in robotic fields. Navigation techniques must rely on that kind of models both to understand the scene and to filter non socially acceptable behaviors. Social conventions are a creation of human societies, they are difficult to model because they include always a subjective component invisible to sensors, they are situation dependent and very adaptive. But their function as regulator of social behaviors make them essential to reach social awareness.

1. INTRODUCTION

4. The navigation strategy to use must produce collision free solutions based on the dynamic environment model, social conventions model and robot dynamic properties. The motion planning algorithm chosen must reach an equilibrium between the reactive properties desired in a dynamic environment with the stable solutions of a global planner.

In the present thesis the four points mentioned above have been addressed but the work has been focused on the last two. It is important to precise that even when social conventions models are inspired on psychological and sociological notions the main interest is on practical robot navigation, i.e., our models do not aim to fully explain human behavior but to promote social behavior in robot navigation.

1.3 Contributions

- Two new models based on social conventions of human management of space have been developed and implemented as part of a “Social Filter”. The first one represents the space of interaction between two people. The second one represents the space in front of a pedestrian related to his/her comfort. This models have been tested with fixed camera and mobile Kinect sensor.
- A new method for early detection of interactions between humans using HMM models has been designed and tested on data coming from real datasets.
- A socially-aware risk-based navigation method including both the traditional notion of risk of collision and a new formulation of risk of disturbance has been proposed. This method integrates mathematical models of social conventions followed in human interaction.
- A socially-aware optimization-based navigation strategy which includes comfort notion has been proposed. This method integrates mathematical models of social conventions related to comfort.

1.4 Overview

This thesis is organized in the following chapters:

Chapter 2 is a review of notions related to social conventions beginning from social sciences and finishing with robotics. Theories linking social behavior and comfort

1. INTRODUCTION

are first reviewed. Then the human management of space is addressed from the point of view of Proxemics field considering one person, groups, objects and robots. Aspects of social robotics and related work on socially-aware navigation are presented.

Chapter 3 presents the models that have been designed and implemented under the notion of Social Filter. Preliminary psychophysical experiments realized in order to validate the parameters of models are described. Also an analysis of a dynamic interaction situation is explained and the resulting model based on Hidden Markov Models theory is shown.

Chapter 4 explains the proposed approaches which have been designed to fulfill the requirements of robot autonomous navigation with social conventions. First the optimization-based approach is described and simulation results are shown. Then the Risk-based approach is introduced, its characteristics reviewed and the details of the inclusion of social convention models explained.

Chapter 5 presents a framework for Socially-aware navigation in the case of assisted mobility where the results of this thesis have been integrated. The problem of robust perception and posterior behavior understanding is emphasized and selected state of art techniques are reviewed. The structure of the system is described showing where the results of this thesis have been integrated. Different scenarios of perception and navigation with a real platform are presented.

Chapter 6 concludes this thesis and gives perspectives of future work.

1.5 Publications

J. Rios-Martinez, A. Escobedo, A. Spalanzani, and C. Laugier, *Intention driven human aware navigation for assisted mobility*, in Workshop on Assistance and Service robotics in a human environment, IROS 2012.

A. Spalanzani, J. Rios-Martinez, C. Laugier, and S. Lee, *Risk based navigation decisions*, in Handbook of Intelligent Vehicles, A. Eskandarian, Ed. Springer London, 2012, pp. 1459 – 1477.

D. Vasquez, P. Stein, J. Rios-Martinez, A. Escobedo, A. Spalanzani, and C. Laugier, *Human aware navigation for assistive robotics*, in 13th International Symposium on Experimental Robotics, 2012.

1. INTRODUCTION

J. Rios-Martinez, A. Renzaglia, A. Spalanzani, A. Martinelli, and C. Laugier, *Navigating between people: A stochastic optimization approach*, in 2012 IEEE International Conference on Robotics and Automation, 2012.

J. Rios-Martinez, A. Spalanzani, and C. Laugier, *Understanding human interaction for probabilistic autonomous navigation using Risk-RRT approach*, in International Conference on Intelligent Robots and Systems, Sep. 2011, pp. 2014 – 2019.

J. Rios-Martinez, A. Spalanzani, and C. Laugier, *Probabilistic autonomous navigation using risk-rrt approach and models of human interaction*, in Workshop on Perception and Navigation for Autonomous Vehicles in Human Environment, IROS 2011.

T. Bellardi, J. Rios-Martinez, D. Vasquez, and C. Laugier, *A novel approach for object extraction from video sequences based on continuous background/foreground classification*, in International Conference on Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on, oct. 2010.

1. INTRODUCTION

1. INTRODUCTION

Chapter 2

Social conventions: from comfort and proxemics to social robotics

2.1 Introduction

This chapter presents a review of the elements that guided us to a socially-aware robot navigation system starting from sociological concepts and finishing with applications in the new field of Social Robotics. Figure 2.1 outlines the socially-aware robot navigation concept that adjust better to our purposes, five main blocks have been identified:

1. a navigation in dynamic environments strategy,
2. a perception system,
3. a robot,
4. a set of social conventions and
5. an environment.

In the classic robot navigation framework the navigation algorithm receives an abstraction of the environment¹ via the robot in order to produce safe plans of navigation. The navigation solutions are returned to the robot which using its dynamic model and controllers sends commands to the action system in order to alter its position in the environment. In the new scheme social awareness is reached by the integration of both social conventions and a new set of techniques, dependent of the perception system, dealing with the processing of social behavior cues, semantics of space and prediction

¹It includes the case of an intelligent environment communicating with the robot.

2. SOCIAL CONVENTIONS: FROM COMFORT AND PROXEMICS TO SOCIAL ROBOTICS

of behaviors. The social conventions are dependent on the particular environment but also on the robot physical properties and task.

In this thesis the social conventions are focused in the shared environment and the humans around. Section 2.3.4 presents recent robotic experiments contributing to the knowledge of how the link between the blocks of Social Conventions and Robot must be constructed.

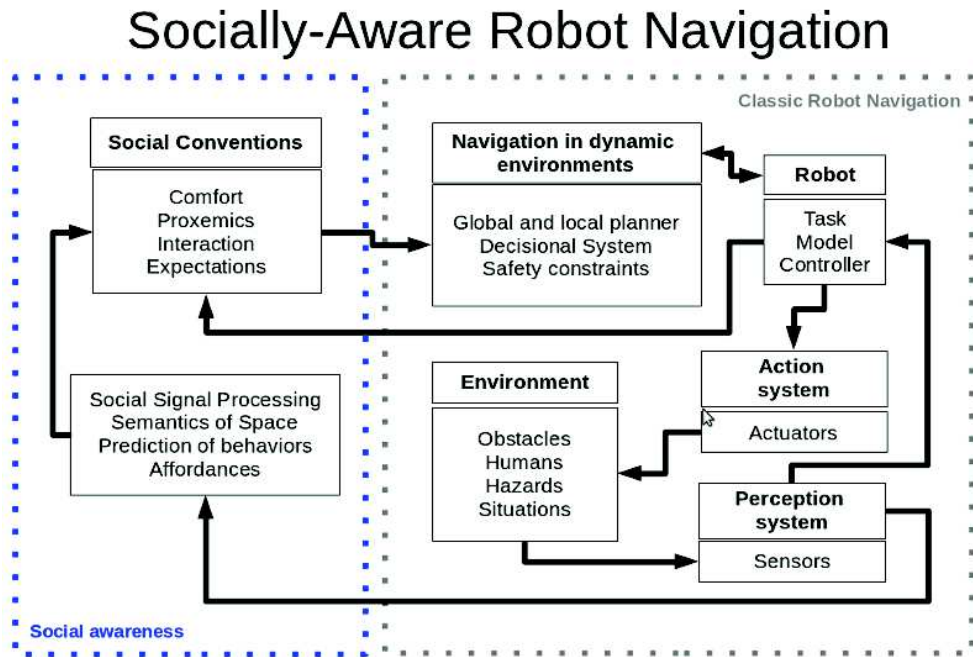


Figure 2.1: The most important components of a socially-aware navigation system.

Social conventions are these modes of behavior created and accepted by the society that help humans to understand intentions of others and facilitate the communication with them. For example, a social order is created as result of the conventions followed by pedestrians in order to respect mutual territories, such as not walking beside an unknown or not following someone too closely. We were concerned with social conventions that could be relevant to the robot navigation task because in that context robots must display not only safe but understandable behavior.

The basic idea is simple: *if a mobile robot can understand and follow social conventions then the humans will better understand robot intentions and will find the co-existence with robots more comfortable.*

2. SOCIAL CONVENTIONS: FROM COMFORT AND PROXEMICS TO SOCIAL ROBOTICS

Our research to understand what humans expectations are respecting to comfort-space relationship conducted us to the subject of human spatial behavior, in the sociology field, which is the topic of sections 2.2 and 2.3. Section 2.4 presents a state of the art in the emergent field of social robotics, focused on socially-aware robot navigation, one of the main desirable characteristics of a social robot. Chapter finishes with conclusions in section 2.5.

2.2 Social behavior and comfort

This section provides an overview of relevant theories of human spatial behavior in the context of this thesis that were found in sociology literature, the field is pretty vast and only selected references are cited.

In order to design an strategy for leading a robot safely in an environment populated by humans and at the same time preserving their comfort, it must be explored how humans manage their surrounding space when they are navigating and how their navigation decisions affect the comfort of others.

Let us begin by describing how human behavior can be perceived by mean of behavioral cues. No one seems to question the idea that organisms communicate via a number of modalities: auditory, visual, olfactory, and tactile; and that modalities can include a number of channels, for example, visual signals can be displayed in the face, body, and hands. According to many authors like (Hogan and Stubbs, 2003), more than sixty percent of the communication between two people or between one speaker and a group of listeners is nonverbal. Nonverbal communication is based on wordless cues that humans are sending and receiving constantly, mainly via visual modality. An emerging domain called Social Signal Processing (a survey can be found in (Vinciarelli et al., 2008)) aims at providing computers with the ability to sense and understand that kind of social messages. In that framework social messages are defined in terms of social signals and social cues.

Definition 1 *The term **social cue** is typically used to describe a set of temporal changes in neuromuscular and physiological activity that last for short intervals of time (milliseconds to minutes).*

Definition 2 *Human **social signals** are acts or structures that influence the behavior or internal state of other individuals, that evolve because of that effect, and that are effective because the perceivers response has also evolved; signals may or may not convey conceptual information or meaning, Mehu and Scherer (2012).*

2. SOCIAL CONVENTIONS: FROM COMFORT AND PROXEMICS TO SOCIAL ROBOTICS

Examples of social signals are: attention, empathy, politeness, flirting and agreement. To recognize an specific social signal we need to detect and analyze multiple social cues which occurs in our brains regularly and most of time unconsciously. Table 2.1 shows the most important social cues and their relation with social signals. In that table,

| Social Cues | Social Signals | | | | | | |
|--------------------------------|----------------|-------------|--------|-----------|------------|------------|---------|
| | emotion | personality | status | dominance | persuasion | regulation | rapport |
| Physical appearance | | | | | | | |
| height | | | ✓ | ✓ | | | |
| attractiveness | | ✓ | ✓ | ✓ | ✓ | | ✓ |
| body shape | | ✓ | | ✓ | | | |
| Gesture and posture | | | | | | | |
| hand gestures | ✓ | | | | ✓ | ✓ | ✓ |
| posture | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| walking | | ✓ | ✓ | ✓ | | | |
| Face and eyes behaviour | | | | | | | |
| facial expressions | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| gaze behaviour | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| focus of attention | ✓ | ✓ | | | ✓ | ✓ | ✓ |
| Vocal behaviour | | | | | | | |
| prosody | ✓ | ✓ | | | ✓ | | |
| turn taking | | | ✓ | ✓ | | ✓ | ✓ |
| vocalizations | ✓ | ✓ | | ✓ | ✓ | ✓ | ✓ |
| silence | | | | | | | ✓ |
| Space and Environment | | | | | | | |
| distance | | ✓ | ✓ | | ✓ | | ✓ |
| seating arrangement | | | | ✓ | | | ✓ |

Table 2.1: Most important social cues and their relation to social signals. Taken from (Vinciarelli et al., 2008)

for example, it is presented that the social cue of height has a relation with the social signal of dominance, higher individuals are judged by others as more dominants. The importance of that proposal is that it permits to link perceivable traits to subjective concepts. Note also that in the vocal behavior cue it is not the meaning of the sound pronounced that matters but the way it is pronounced.

In the present thesis we use the behavioral cues of distance and posture of people to

2. SOCIAL CONVENTIONS: FROM COMFORT AND PROXEMICS TO SOCIAL ROBOTICS

understand the possible rapport among humans in the scene which is a high level cognitive concept. Observe in the table that posture has an influence on all the social signals described. As we will see in next section these cues can give us useful evidence to assess the concordance with social conventions during robot navigation.

Factors influencing the psychological comfort. Many psychological theories have been proposed to explain the relation between distance, visual behaviors and comfort in humans, see a review in (Aiello, 1977, 1987; Greenberg et al., 1980). People will become uncomfortable if they are approached at a distance that is judged to be too close: the greater invasion/intrusion the more discomfort or arousal is experienced by the person. The relation between intrusion and discomfort is linear, indicating that each increment of intrusion produces a comparable increment in discomfort (Hayduk, 1981a). Same reference explained that when threat is high or there is the potential for threat in a situation, distances tend to be larger, especially for females. The greater the control one feels, the greater the proclivity toward closer spaces.

In casual conversations, people claim an amount of space related to that activity. This space is respected by other people and only participants have permitted access to it, therefore intrusions cause discomfort (Kendon, 2010). In Thompson et al. (1979) subjects rated intermediate distances (between four and eight foot) as most comfortable, preferable and appropriate for interaction situations.

Related theories have postulated that exists an optimal range of distance preferred by interacting individuals. Deviations from optimal range result in discomfort, then forces are exerted to keep the system in equilibrium. In that sense theory presented by Argyle and Dean (1965) proposed that it exists an equilibrium of intimacy which involves components like level of physical proximity, eye contact, intimacy of topic and amount of smiling. If one of the components is disturbed compensatory changes may occur along the other dimensions. If this is not possible the subject will feel uncomfortable in one of two ways. If the disturbance is in the direction of too much intimacy, avoidance forces will predominate, and the subject will feel anxiety about rejection or revealing inner states; if in the direction of less intimacy, he will simply feel deprived of affiliative satisfactions. It seems that this theory is maintained even in virtual environments (Bailenson et al., 2001, 2003) where participants kept greater distance from virtual humans when approaching their fronts compared to their backs. In addition, participants gave more personal space to virtual agents who engaged them in mutual gaze.

Since comfort in the above context is a subjective notion it is clear that it cannot

2. SOCIAL CONVENTIONS: FROM COMFORT AND PROXEMICS TO SOCIAL ROBOTICS

be measured directly by any sensor. Studies made to explain how distance, posture and visual behavior affect comfort in humans can be used to develop useful models for robotics. Proxemics is one of those studies that will be presented in the next section.

2.3 Proxemics: human management of space

The concepts presented in this section have been borrowed from sociology literature and have been used in the present thesis as tools to complete the navigation framework with a social component.

Definition 3 *Proxemics is the study of the nature, degree, and effect of the spatial separation individuals naturally maintain (as in various social and interpersonal situations) and of how this separation relates to environmental and cultural factors. The term was first proposed by Hall (1966) to describe the human management of space.*

He observed the existence of some rules not written that conducted people to keep distances from others, and others to respect this space. His work is, perhaps, the most commonly known about proxemics.

When humans navigate they maintain spaces for themselves comparable to those they imagine the others would prefer, a notion that is supported by theory of mind¹.

It has been proposed that management of space done by a single person is different to the one realized by a group of people. According to Krueger (2011), social cognition is fundamentally an interactive form of space management, they use the term “we-space” which is the result of the coordinated engagement of interactants and is different from the perspective commonly studied, of an individual acting agent.

In this section, concepts related to Proxemics will be described. In order to gain clarity in the exposition of the human management of space, this section have been divided in space related to: one person, groups of people, objects and robots.

2.3.1 Space related to one person

2.3.1.1 Personal Space

Definition 4 *Personal Space is the area individual humans actively maintain around themselves into which others cannot intrude without arousing discomfort, (Hayduk, 1978).*

¹Analysis on this subject can be read in (Krueger, 2011)

2. SOCIAL CONVENTIONS: FROM COMFORT AND PROXEMICS TO SOCIAL ROBOTICS

The notion of Personal Space was introduced into the social psychological literature to describe the emotionally-tinged zone around the human body that people feel is “their space”. A very complete historical review of the notion can be found in (Sommer, 2002). In figure 2.2 it is shown an environment with people not interacting, it is possible to observe the natural arrangement of people motivated by the respect of individual Personal Spaces. Blue circles are used to give an idea of the extension of Personal Space.



Figure 2.2: Typical arrangement that humans exhibit caused by the respect of Personal Space, blue circles are marked as reference.

As the main interest is on navigation having a planar model of the environment, the works have been classified according to the resulting shape that Personal Space would display in the ground plane. In the fig. 2.3 a comparison of Personal Space shape according to four different works is shown.

Concentric circles. According to Hall (1966) it is possible to classify the space around a person with respect to social interaction in four specific zones whose distances from human body are listed below:

- the public zone $> 3.6\text{m}$,
- the social zone $> 1.2\text{m}$
- the personal zone $> 0.45\text{m}$
- the intimate zone $< 0.45\text{m}$

2. SOCIAL CONVENTIONS: FROM COMFORT AND PROXEMICS TO SOCIAL ROBOTICS

The categories proposed explain very well reactions like the uncomfortable sense of somebody invading your intimate zone or the perception of a person aiming interact with you by entering to your social zone. It is well known that the measures proposed by Hall are not strict because they vary depending on age, culture, type of relationship and context. Studies in [Bar-Haim et al. \(2002\)](#) suggested that tolerance to intrusion in childrens personal space is related to infancy attachment security/insecurity with the mother and professional caregivers.

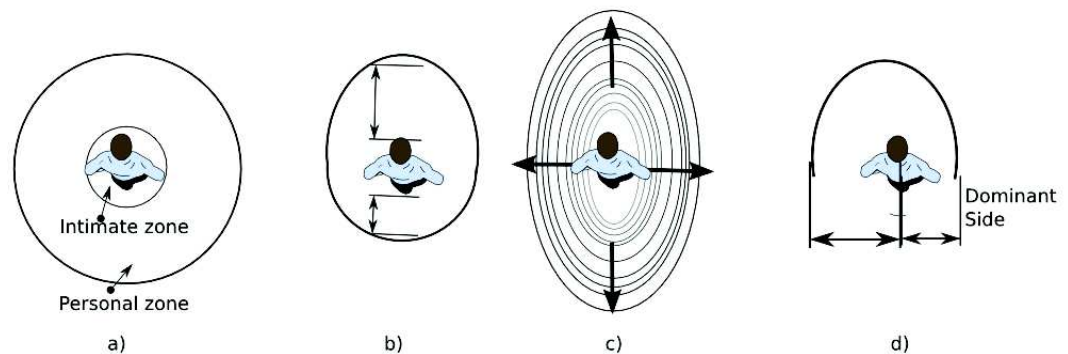


Figure 2.3: Different shapes of Personal Space: a) Concentric circles [Hall \(1966\)](#) b) Egg shape, bigger in the front [Hayduk \(1981b\)](#) c) Ellipse [Helbing and Molnar \(1995\)](#) d) Smaller in the dominant side [Grin-Lajoie et al. \(2008\)](#).

Egg shape. [Hayduk \(1981b\)](#) considered that in general people are more strict regarding their frontal space therefore frontal invasions are more uncomfortable.

Concentric Ellipses. Personal Space was referred as “the private sphere” in the *Social Force Model* proposed by [Helbing and Molnar \(1995\)](#) which stated that the motion of pedestrians is influenced by other pedestrians by mean of repulsive forces. Interestingly the potential repulsive, according to this model, is a monotonic decreasing function with equipotential lines having the form of an ellipse that is directed into the direction of motion. The Social Force Model has been very used to represent human behavior in agent simulation and recently has attracted the attention of robotics community.

2. SOCIAL CONVENTIONS: FROM COMFORT AND PROXEMICS TO SOCIAL ROBOTICS

Asymmetric shape. More recent work, [Grin-Lajoie et al. \(2008\)](#), showed that personal space is not modulated according to walking speed during circumvention of obstacles and that personal space is asymmetrical, i.e., it is smaller in the dominant side of pedestrian. Their findings suggested that personal space is used to control navigation in a cluttered environment. In that sense, [Higuchi et al. \(2006\)](#) explored the fact that to walk through a narrow passage, the information about the size of the passage relative to the width of the body is required. That work supported the idea that the representation of space accurately reflects the action capabilities of humans, they classified space around human in two regions, the first which is in reach of the hand and a second one which is out of reach of the hand.

Other relevant aspects Experiments presented in [Hayduk \(1994\)](#) supported the idea that personal space is dynamic and situation dependent, i.e., it is more correct to consider personal space as a momentary spatial preference. The spatio-temporal model of personal space presented in ([Park and Trivedi, 2007](#)) can be adjusted according to a velocity parameter representing low and high velocity profiles. There is a lot of work already done but in our opinion more studies are needed, for example, to have a better idea of the three-dimensional shape of personal space and how it evolves on time. The Personal Space is not only a psychological concept. Recent work has provided biological evidence that the amygdala may be required to trigger the strong emotional reactions normally following Personal Space violations, thus regulating interpersonal distance in humans [Kennedy et al. \(2009\)](#).

Considerations taken into account in this thesis. In this thesis Personal Space is used as synonymous of the personal zone plus the intimate zone defined by Hall, their measures were employed as reference for the model of Personal Space proposed in the next chapter (see section [3.2.1](#)) . By contrast the discomfort is not considered as constant in all the personal space but as being higher in front of the human than behind. Some assumptions made in the model follow. Personal Space has shape of ellipse similar to the one proposed in the Social Force model. The size of personal space will not be modulated by pedestrian speed and, for simplicity reasons, will remain constant during all navigation situations although its evolution on time is considered as perspective. Personal space will be centered in human position and its front must be bigger.

2. SOCIAL CONVENTIONS: FROM COMFORT AND PROXEMICS TO SOCIAL ROBOTICS

2.3.1.2 Information Process Space

Definition 5 *Information Process Space (IPS) is the space in which pedestrians take into account obstacles and other pedestrians in order to calculate next moves and where psychological comfort is evaluated, (Kitazawa and Fujiyama, 2010).*

Kitazawa and Fujiyama (2010) presented a work where the size and the shape of IPS was explored. They explained that many current models of pedestrian movements share the IPS notion as common element. As result of their experiments the IPS would have a cone shape area more than semicircular as proposed in similar works. Pedestrians are more interested in the exact front having a small relative lateral distance. Moreover the subjects in the study did not pay attention in the area with an angle more than 45 degrees or more from the walking direction. Fig. 2.4 shows an schema showing IPS characteristics. IPS is strongly related to visual behavior and it is important to under-

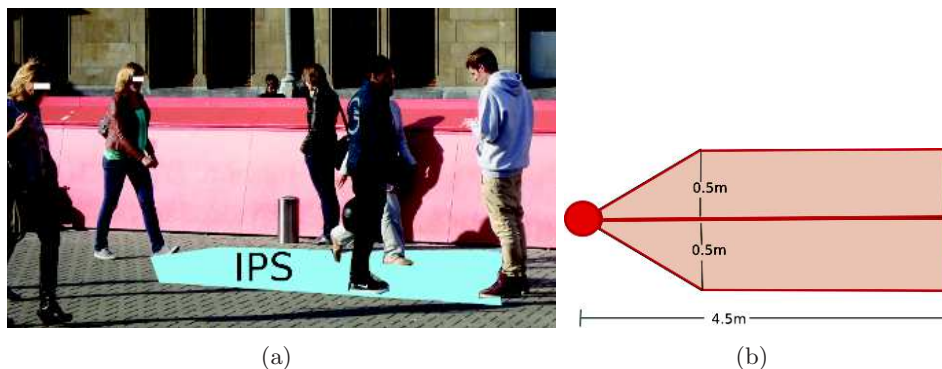


Figure 2.4: Information Process Space shape according Kitazawa and Fujiyama (2010). a) Pedestrian is more interested in the exact front to detect obstacles and other pedestrians in order to calculate next moves. b) Shape and measures obtained for the IPS, pedestrian is represented by the circle.

stand how it can be used to estimate proxemics behavior. Work done by Argyle and Dean (1965) showed that reducing eye-contact makes greater proximity possible and that greater proximity reduces eye-contact. According to Goffman (1963), approaching pedestrians are comfortable glancing at one another until they reach a separating distance of approximately 2.5 meters, at this distance people typically look down.

It seems that pedestrian visual behavior is also conditioned by culture and gender, for example, experimental results in Patterson et al. (2007) supported the hypothesis that Japanese pedestrians are less responsive as they pass non acquaintances on the sidewalk than Americans and that female confederates would receive more glances from

2. SOCIAL CONVENTIONS: FROM COMFORT AND PROXEMICS TO SOCIAL ROBOTICS

pedestrians in the same scenario. Experiments on obstacles avoiding in Higuchi et al. (2006) suggested that vision is used to obtain information on space out of hand reach and to control locomotor action in a feed-forward manner. Turner and Penn (2002) proposed the concept of exosomatic visual architecture which permits agents guided by visual affordances reproduce a natural movement which correlate well with observed human behavior. Moussad et al. (2011) suggested that, guided by visual information, namely the distance of obstructions in candidate lines of sight, pedestrians apply two simple cognitive procedures to adapt their walking speeds and directions. The first one is concerned with trade-off between avoiding obstacles and minimizing detours from the most direct route. The second one is based on the idea that a time period is required for the pedestrian to stop in the case of an unexpected obstacle then pedestrians should compensate for this delay by keeping a safe distance.

Considerations taken into account in this thesis. The IPS concept was compatible to the purposes of this thesis. Justified by the enlisted related works it was proposed that a human will be disturbed if the robot invades his/her IPS. The model developed for IPS is presented in section 3.2.2 and the navigation approach that integrates this model is described in section 4.3.

2.3.2 Space related to groups of people

This thesis is mainly concerned with the interactions observed in standing people in conversation forming groups. It exists evidence to posit that people keep more space around a group than the mere addition of single Personal Spaces (Efran and Cheyne, 1973; Knowles et al., 1976; Krueger, 2011), it is, therefore, important to study groups separately.

According to Goffman (1963), humans react to societal regulation through the concepts of focused and unfocused interactions.

Definition 6 *Focused interaction occurs when individuals agree to sustain a single focus of cognitive and visual attention.*

Definition 7 *Unfocused interactions are interpersonal communications resulting solely by virtue of an individual being in another's presence.*

Conversations are focused interactions because people share a common focus of attention which can be traduced to a shared common space. In unfocused interactions people negotiate their position with others by means of nonverbal behaviors (like group

2. SOCIAL CONVENTIONS: FROM COMFORT AND PROXEMICS TO SOCIAL ROBOTICS

arrangements) which improve comfort and predictability of human actions.

In this thesis it is proposed that a robot with a socially-aware navigation must exhibit an unfocused interaction by respecting focused interactions, specifically conversations, of people in the environment. To fulfill that proposal it is important to detect conversations. The approach chosen is based on the observation that, in environments where several individuals are co-present, the manner which they orient and place themselves in relation to one another directly reflects how they may be involved with one another. The concepts of O-Space and F-formations are at the center of our solution to detect conversations, both of them will be presented in the following section.

2.3.2.1 Interaction spaces: the O-Space concept

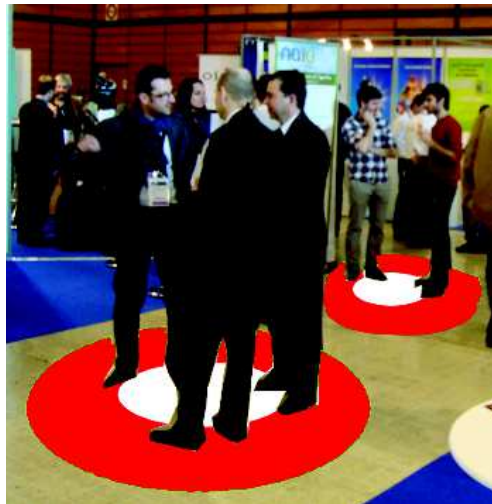


Figure 2.5: People interacting in groups follow some spatial patterns of arrangement. The situation of O-Space is marked with a white circle for the two groups and it is surrounded by the p-Space in red.

Definition 8 *The O-Space is the joint or shared area reserved for the main activity that is established by groups in focused interaction, only participants have permitted access to it, they protect it and others tend to respect it, (Kendon, 2010). O-Space varies depending on body size, posture, position and orientation of participants during the activity.*

Definition 9 *The p-Space is the space surrounding the O-Space which is used for the placement of the participant's bodies and also personal belongings, (Kendon, 2010).*

2. SOCIAL CONVENTIONS: FROM COMFORT AND PROXEMICS TO SOCIAL ROBOTICS

People in conversation follow spatial patterns of arrangement which function as social signals to inform the kind of activity that is taking place, in fig. 2.5 it is observed how the position and orientation of people can help us to decide what groups are in conversation and where it would be located the O-Space.

Definition 10 *The term F-formation is used to designate the system of spatial-orientation arrangement and postural behaviors that people create and maintain in order to sustain their O-space. F-formations are characteristic of people who come together to accomplish joint activity, they are typically associated with the occurrence of small conversational gatherings, (Ciolek and Kendon, 1980).*

An F-formation seems to ensure that equality of social participation among members is preserved. Another important aspect is that an F-formation system can be viewed as an adaptive mechanism that enhances people's chances for smooth and undisturbed engagement in their social encounter, whatever external circumstances may prevail. The shape of the F-formation strongly depends on the number of people involved, the interpersonal relationship among them, the group attentional focus and on environmental constraints like furniture. In the case of two people some F-formations have been identified as the most frequent, see table 2.2 for a description of each one. In the present context, Marshall et al. (2011) studied the support that physical spaces give to social interaction by using F-formations, it is stated that physical structures in the space can encourage and discourage particular kinds of interactions. In general F-formations have been less studied than Personal Space. To the best of our knowledge few works in robotics (especially in social robotics) have been interested on them and only one work (Cristani et al., 2011) have described a way to estimate the location of O-space related to the F-formation by using position and head orientation but they assumed a circular shape. The work presented in (Mead et al., 2011a) was not conclusive considering F-formation metrics.

Considerations taken into account in this thesis. Very few quantitative information was found in references of social sciences, but it was enough to propose suitable models to integrate in robot navigation, our model of O-Space is presented in 3.3.2. In the case of more than two people O-space is commonly accepted as a circle (that is the reason for the O), we think that, in the same way that personal space, the O-space has more sensible zones and for this reason a circle it is not appropriated in all the cases. We have used the torso direction and the body position to detect F-formations. A different O-Space has been designed for selected F-formation between

2. SOCIAL CONVENTIONS: FROM COMFORT AND PROXEMICS TO SOCIAL ROBOTICS

two people: Vis-a-vis, V-shape, L-shape and C-shape, described in table 2.2.

In the present document, from now on, the term formation will be used in place of F-formation.

2.3.2.2 Arrangements of groups.

Groups of two people. According to (Ciolek and Kendon, 1980; Kendon, 2010), for two people in conversation, six formations or arrangements are the most frequent: N-shape, Vis-a-vis, V-shape, L-shape, C-shape and side-by-side. A description of each one is presented in table 2.2. The arrangements classification follows the criteria of body position and orientation, for example a Vis-a-vis formation is identified because each body is parallel and oriented to the other body, forming a letter H viewed from above. Other information provided by the table is the type of environment where the formation were more frequent. N-Shape, Vis-a-vis and V-Shape were more frequents in open spaces not heavily used by pedestrians.

The vis-a-vis formation (also referred as face-to-face) is the basic mode of human sociality (See analysis done by Krueger (2011)). Related to this, it has been shown that the perception of aperture between two people is constrained by psychosocial factors, for example, experiments in (Morgado et al., 2011) showed that the closer, affectively, participants felt to their classmates, the more they felt able to pass between them.

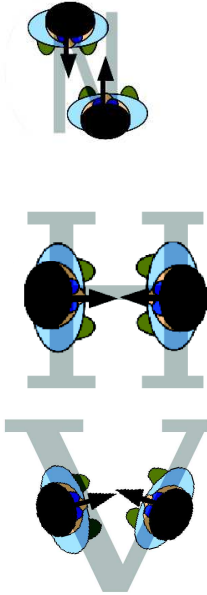
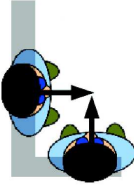
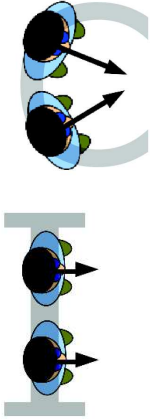
Groups of more than two people. When more than two people are in conversation, in absence of furniture, they commonly exhibit an arrangement with circular shape with the focus in the center (see fig. 2.6) therefore, in this case, the O-space could be taken as a circle whose center coincides with that of the inner space (Kendon, 2010). In the framework of navigation it would be more interesting to estimate that an agglomeration of humans is not a group because in that case the robot would have more space to pass.

A related study presented in (Groh et al., 2010) proposes a geometric based approach to detect co-located real world social interactions by looking interpersonal distance and torso orientation. They modeled social situations as the probability that a social situation among n persons exists at some point in space and time, given m social signals from these persons.

Regarding the management of space done by a group of people interacting, it was proposed by Lloyd (2009) that interpersonal spatial behavior is modulated not only by the distance between the interactants but also by the nature of the interaction, for example,

2. SOCIAL CONVENTIONS: FROM COMFORT AND PROXEMICS TO SOCIAL ROBOTICS

Table 2.2: Taxonomies of arrangements for a two-person formation defined by Ciolek and Kendon (1980). Models for O-Space in Vis-a-vis, V-shape, L-shape and C-shape cases are presented in section 3.3.2.

| Environment type | Arrangement | Description |
|---|---|--|
| Spaces open, spacious and not heavily used by pedestrians |  <p>The diagrams show three types of two-person formations. The top diagram, labeled 'N-shape', shows two figures facing each other with their bodies slightly offset, forming an 'N' shape. The middle diagram, labeled 'Vis-a-vis', shows two figures facing each other directly, forming an 'H' shape. The bottom diagram, labeled 'V-shape', shows two figures facing each other with their bodies intersecting at an angle, forming a 'V' shape.</p> | <p>N-shape. Individuals face each other and have their body planes parallel while they stand or sit slightly displaced by approximately half a body width.</p> <p>Vis-a-vis. People face each other directly, when seen from above, they appear to form with their body contours and nose and array resembling a letter H.</p> <p>V-shape. The participants body planes intersect outside the formation at an angle of approximately 45 degrees</p> |
| Spaces that are semi-open and heavily trafficked by pedestrians |  <p>The diagram shows two figures standing at right angles to each other, forming an 'L' shape.</p> | <p>L-shape. Participants are standing at right angles to each other, with their body planes intersecting outside the gathering.</p> |
| Areas delineated by the presence of a large, solid and impenetrable object and in places where there was little or no pedestrian movement |  <p>The diagrams show two types of two-person formations. The top diagram, labeled 'C-shape', shows two figures standing at an obtuse angle, forming a 'C' shape. The bottom diagram, labeled 'Side-by-side', shows two figures standing close together, facing the same direction, forming an 'I' shape.</p> | <p>C-shape. Participants are standing at an obtuse(open) angle of approximately 135 degrees, and, when seen from above, they form a figure resembling the letter C.</p> <p>Side-by-side. Two individuals face in the same direction but stand close enough to still have full access to each other's transactional segment.</p> |

2. SOCIAL CONVENTIONS: FROM COMFORT AND PROXEMICS TO SOCIAL ROBOTICS



Figure 2.6: Circular O-space in conversations for groups of more than two people.

threatening or not-threatening social interactions.

It is important to mention the existence of a more general, theory of spatial organization and classification proposed by [Shefflen \(1976\)](#). Schefflen argues that, when people cluster together, they define a territory which organizes the space around them in a particular structured fashion. The structured space influences and is sustained by a class of specific behaviors called territorial behaviors. The proposed classification of human territories ranges from simple ones like f-formations to more complex structures as gatherings and hubs. Future extensions of the current thesis consider the use of this theory.

In ([Ge et al., 2009](#)) groups traveling together are discovered using a bottom-up hierarchical clustering approach that compares sets of individuals based on a pairwise distance that combines proximity and velocity cues. The bottom-up approach proposed seems also applicable to standing groups (with velocity zero) in conversation, for this reason we chose detecting group interaction by using formations in the Social Filter (chapter 3). Moreover this framework will enable us to include proxemic behavior of walking people in conversation, like the one studied in ([Costa, 2010](#)), in our approach.

Until now it has been discussed the management of space made by humans related to comfort for other humans and their interactions. Next section is focused on the management of space around objects related to its function.

2.3.3 Space related to objects: Activity Space and Affordance Space

When a human is performing an activity he/she claims an amount of space, related to that activity, which is respected by other people. Concepts describing the kind of spaces we are interested in are activity space and affordance space.

2. SOCIAL CONVENTIONS: FROM COMFORT AND PROXEMICS TO SOCIAL ROBOTICS

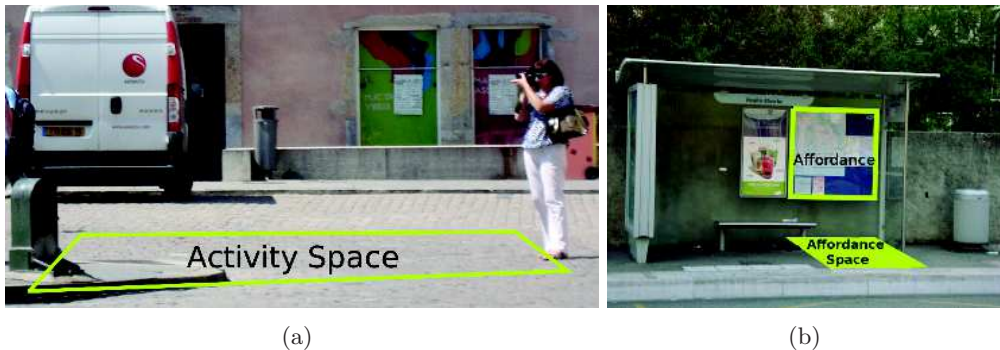


Figure 2.7: Activity and affordance spaces. In a) a woman is taking a picture, the space between her and her objective becomes an activity space. In b) the bus schedule represents an affordance for humans then the space in front of this information becomes an affordance space.

Definition 11 *Activity Space is a social space which is constituted by means of actions performed by agents. The notion implies a geometric space but does not give an explicit definition for the shape, it is assumed that it can take multiple shapes depending of specific actions, (Lindner and Eschenbach, 2011).*

Definition 12 *The Affordance Space is related to the concept of affordances as potential activities that the environment provides to agents. Affordance Spaces are potential Activity Spaces.*

An example of activity space can be seen in figure 2.7 a) where a human is taking a picture. Normally people stop to do not interrupt the activity. Crossing an Affordance Space is generally unproblematic as compared to crossing an Activity Space but blocking an Affordance Space could be socially not accepted. An example of Affordance Space is shown in figure 2.7 b).

Related to the previous concepts, Borkowski et al. (2010) claims that perceived geometrical features of the environment must be linked with semantic information of objects in order to achieve a semantic robot navigation. However the task becomes complicated as the perception of the space done by sensors is objective (despite of uncertainty) while human abstraction of the space is very subjective. In any case, it is necessary to take that into account semantics of space when planning social acceptable navigation solutions.

2. SOCIAL CONVENTIONS: FROM COMFORT AND PROXEMICS TO SOCIAL ROBOTICS

Semantics of space and interaction with objects. Possible formalizations to the problem of understanding semantics of space can be found in a field called Qualitative Spatial Reasoning which is concerned with the acquisition, organization, utilization, and revision of knowledge about spatial environments, in particular, the framework to reason about dynamical spatial scenes presented in [Bhatt and Dylla \(2009\)](#) is of special interest. That approach combines existing qualitative theories about space with a formal logic-based calculus suited to modeling dynamic environments. According them, it is necessary to explicitly take into consideration the fact that perceivable changes in the surrounding space are typically the result of interaction within the environment. Humans, robots and systems that act, and interact, are embedded in space, and this change is often the result of actions and events. Drawback of this method is its explosive computational complexity.

Also in the framework of the space related to objects, [Lloyd \(2009\)](#) proposed the existence of a “practical space” created by the interaction of the space represented by human body and the space where the objects exist. It can be translated into a practical rule: look around the human body to infer possible interactions with the world.

Other groups have proposed similar ideas to that presented in this section, for example, [Greenberg et al. \(2011\)](#) explore the use of the concept of proxemics (see 2.3.1) between humans to be extended to proxemics with smart objects. They highlighted the importance that such concept will have to achieve ubiquitous computing, for example, devices can react to distance and orientation of user in a way that can be more intuitive for humans. Implementations of these ideas in form of Proxemics applied to objects are presented in ([Marquardt et al., 2011](#)) and in ([Wang et al., 2012](#)). In the same sense, [Gharpure and Kulyukin \(2008\)](#) presented an ontology to characterize space according to the ways we interact with it. Such approach was used to implement spatial cognition for robot-assisted shopping. Those works are complementary of our approach in the sense that they are focused on interaction with objects regarding proxemics and we are interested in not disturbing resulting interactions with objects.

Considerations taken into account in this thesis. The objective of this section was to present arguments to support the idea that humans can assign a meaning to the space inspired by the interaction of the human with an object in the environment. Recognizing the importance of that idea, in the present thesis the concept of Activity Space has been used to estimate the risk a robot could fall into by invading a specific empty space which could have a semantic importance depending on objects or actions surrounding it. The resulting model is presented in section 3.4.

2. SOCIAL CONVENTIONS: FROM COMFORT AND PROXEMICS TO SOCIAL ROBOTICS

Robots are a very special kind of intelligent objects, they have been for a while in the industry where its apparition has been signaled as the beginning of a new industrial revolution, now their arrival to our homes seems imminent. The next section discusses how humans will manage their space around robots.

2.3.4 Space related to robots

Assuming that it is not known still what will be the human spatial behavior with robots and that robots are more intelligent objects it is logic to wonder, how humans will manage space around robots? Some approaches in robotics literature, have been conducted to establish the rules that probably will govern the physical behavior of robots regarding interaction with humans. It can be assumed that people will engage in proxemic behavior with robots in much the same way that they interact with other people as supported by work done in (Takayama and Pantofaru, 2009). A classification of the reviewed works is listed below according to the next proxemics factors: speed, appearance, direction of approach and other factors.

Speed. In experiments presented by Butler and Agah (2001), indicate that the human subjects were comfortable with all speeds except the fast approach speed. The comfortable speeds were at 0.254m/s and 0.381m/s while the uncomfortable fast speed was 1m/s. Normal walking speed for young human is about 1m/s, then those results suggest that humans prefer slower speed for a robot.

Appearance. According to Butler and Agah (2001) the appearance and size of the robot will cause a reaction that must be considered when the behaviors of the robot are performed. They compared approaching and avoiding of a robot, first by using the base robot only and after by adding a humanoid body to the base. Their results showed that people who preferred the humanoid robot were more comfortable with closer distances than the other subjects. An empirical framework for Human-Robot proxemics was proposed in Walters et al. (2009) where, after multiple experiments it is suggested a method to calculate an approach distance estimate for a robot taking into account any combination of proxemic factors like robot appearance, human preferences or type of task. Such method consisted in to take a base distance (57 cm) and to calculate approach distance by adding the coefficient of proxemic factors which can be positive or negative. Values proposed for robot taking into account appearance factor are reproduced here:

2. SOCIAL CONVENTIONS: FROM COMFORT AND PROXEMICS TO SOCIAL ROBOTICS

- With a humanoid robot giving an object to human, comfortable distance is: (base distance =) 57cm + (Humanoid-RH Approach =) 3cm - (Giving Object HR Approach=) 7cm = 53cm.
- With a mechanoid robot giving an object to human, comfortable distance is: (base distance =) 57cm - (mechanoid-RH Approach =) 3cm - (Giving Object HR Approach=) 7cm = 47cm.

Investigations made by [Mumm and Mutlu \(2011\)](#) showed that participants who disliked the robot compensated for the increase in the robots gaze by maintaining a greater physical distance from the robot, while participants who liked the robot did not differ in their distancing from the robot across gaze conditions. Also, men maintained a greater distance from the robot than women did.

Direction of approach or gaze direction. According to [Butler and Agah \(2001\)](#), the indirect pattern of approach seemed to be the favorite of the subjects in their experiments. Indirect approach is a less threatening behavior because the threat of contact has been reduced. In experiments done by ([Dautenhahn et al., 2006](#)), participants evaluated the direct frontal approach as least comfortable for a bring object task by finding robots motion threatening and aggressive. Different conclusion was found in [Torta et al. \(2011\)](#) where an approaching scenario for a robot which takes into account the relationship between direction and distance of approach was evaluated, the user's evaluation showed that frontal (± 35 and 0 degrees with relation to the person orientation) directions of approach are perceived as comfortable while farthestmost (± 70) directions are perceived as uncomfortable. Models for close, optimal and far distance to have a comfortable communication were extracted.

In [Huettenrauch et al. \(2006\)](#) they focus in the spatial interaction between a robot and a user analyzing the interaction using variations in distance and spatial orientation. They implemented experiments based in the technique Wizard of Oz. Hall's interpersonal spatial zones and Kendon's formation were tested for applicability in the analysis of human-robot interaction episodes in a home tour where a user shows a robot where objects and places are located. Vis-a-Vis was found to be prevailing among the spatial configurations tested for. In [Kuzuoka et al. \(2010\)](#), they found that it is possible to reconfigure an arrangement between a robot and a human by changing the position of robot, when the robot is executing tasks of museum guide, which is more effective than only rotate its head.

Regarding gaze direction, when the robots head is oriented toward the person's face,

2. SOCIAL CONVENTIONS: FROM COMFORT AND PROXEMICS TO SOCIAL ROBOTICS

it decreases the minimum comfortable distance for men, but increases the minimum comfortable distance for women (Takayama and Pantofaru, 2009). Recently in Sciutti et al. (2012) proactive gaze and automatic imitation are proposed as tools to quantitatively describe if and how human actions adapt in presence of robotic agents, based on the concept of motor resonance.

Other factors. Takayama and Pantofaru (2009) found that personal experience with pets and robots decreases the personal space around robots. They also found that the personality trait of agreeableness decreases personal spaces when people approach robots, while the personality trait of neuroticism and having negative attitudes toward robots increase personal spaces when robots approach people.

Considering human-robot physical proximity, the taxonomy presented by Yanco and Drury (2004) defined six modes: none, avoiding, passing, following, approaching and touching. These values are ordered from less to more physical interaction. For cases when multiple types of physical interaction are applicable, the value chosen should be the type that involves the most physical proximity.

Guidelines for robot design Takayama and Pantofaru (2009) provided some guidelines for the design of robots capable of exhibit proxemic behavior:

- using recognition of familiar people (e.g., people the robot has identified several times in the past) in deciding how close to approach people,
- including the robots gaze behavior in deciding how close to approach people and
- considering pairing person identification with personal information (as personality traits) to better gauge the appropriate approach distances.

Other work presented in Lam et al. (2011) proposed rules that a single robot should obey in order to achieve not only a safe but also a least disturbance motion in a human-robot environment. It is considered the fact that both humans and robots have their sensitive zones, depending either on their security regions or on psychological feeling of humans. The first three of these rules are reproduced here because of its coincidence with the approach followed in this thesis:

- *Collision Free Rule:* The host robot has to maintain its safety and be able to reach the goal destination,

2. SOCIAL CONVENTIONS: FROM COMFORT AND PROXEMICS TO SOCIAL ROBOTICS

- *Interference Free Rule*: The host robot should not enter into the personal space of a human or the working space of any other robot unless its task is to approach any of them.
- *Waiting Rule*: Once the host robot enters into the personal space of a human carelessly or unwillingly, it has to stop and to wait for a threshold time.

There are still very few works focused in the human management of space around robots compared to the one around humans, however robotics close to humans is an increasing field and when successful cases of service robotics (like vacuum cleaners) become more numerous the new available scenarios will permit to corroborate or refuse the information presented in this section.

Until now we have discussed about the factors that must be taken into account in order to endow a robot with social capabilities, next section will present a review of the social robotics field where many of such factors has been already touched.

2.4 Social Robotics

In the literature we can observe the growing interest of the robotics community in research that includes behavior of humans and its impact in the development of tasks by the robot. In this section we discuss the aspect of sociality from the point of view of robotics literature with focus in mobile robots. There is an abundance of terms that includes the key word “social” in robotics and agents context consequently an unique and complete definition of social robot could not be possible to get. Instead, a list of desired abilities in a social robot are described with an special emphasis on the ability of socially-aware navigation at the end.

2.4.1 Social robot abilities

Sociality is concerned with all the aspects that make individuals interact with each other to satisfy needs that could not be achieved by individuals alone. As opposed to the mere aggregation of organisms around favorable environmental conditions, sociality implies interactions between individuals (Mehu and Scherer, 2012). A list of abilities for a social robot is listed below:

- A social robot participates pro-actively with the human players in its surroundings, and meet the humans’ expectations of how an able player within their so-

2. SOCIAL CONVENTIONS: FROM COMFORT AND PROXEMICS TO SOCIAL ROBOTICS

cial environment should act. The robot possesses an internal understanding and adaptable social model of human society (Ge, 2007).

- Social robots must engage in “natural” interaction with humans, i.e., interaction in the same way as humans do with other humans (Shi et al., 2011; Duffy, 2001) and develop relationships or a rapport with them (Kanda et al., 2009; Kahn et al., 2008). Robot should mimic human social norms and be able to provide a consistent set of behaviors (Bartneck and Forlizzi, 2004). Social robots must know how to initiate an interaction with a human Satake et al. (2009), for example by displaying availability to him/her (Yamazaki et al., 2007) or friendly attitude (Hayashi et al., 2011). For Michalowski et al. (2006) not only capability of the natural initiation, but also of the maintenance, and termination of social interactions with humans is important.
- A social agent or robot must exhibit common sense which is the collection of shared concepts and ideas that are accepted as correct by a community of people Barraquand and Crowley (2008).
- A social robot should be able to exhibit its status and its intentions and to deal with its human partner abilities and preferences (Alili et al., 2009).
- Social robots are able to coordinate its actions, plans and schedules with the patterns of human activities, giving them the ability to smoothly blend into the work-flows and daily routines of people. Robots exhibit its sociality by minimizing the interference with people in the same environment Tipaldi and Arras (2011); Sehestedt et al. (2010).
- A social robot is able to exhibit gestures synchronized with the dialog (Glas et al., 2011).

Special attention is put on the social abilities that are directly related with the motion and navigation of the robot. Together, such abilities compose the socially-aware robot navigation and are listed below.

- A social robot must have the ability to behave in such a way that a human partner does not feel aggravated or afraid by the robots movements and can easily infer the motion intentions of the robot (a.k.a. legibility), (Kruse et al., 2012).
- A social robot must take the social aspects of interaction with people into account when navigate Pacchierotti et al. (2007) .

2. SOCIAL CONVENTIONS: FROM COMFORT AND PROXEMICS TO SOCIAL ROBOTICS

- A social robot takes into account the comfort of the human as well as his/her preferences and needs [Sisbot et al. \(2010\)](#).
- A social robot in navigation have the knowledge about action permissions in social spaces and moves accordingly ([Lindner and Eschenbach, 2011](#)). Its navigation involves an awareness of other users who are currently present or have been there in the past ([Jeffrey and Mark, 2003](#)).
- A social robot is able to differentiate between obstacles and persons so that an appropriate behavior can be selected, for example, keeping comfortable distance from a person [Topp and Christensen \(2005\)](#)
- In [Kuderer et al. \(2012\)](#) it is stated that a robot which is able to predict the behavior of the pedestrian can navigate in a socially more compliant way.
- In [Gockley et al. \(2007\)](#) it is argued that moving in easily understood and predictable ways will both improve people’s trusting and comfort with the robot as well as will help to insure the safety of people moving near the robot.

2.4.2 Socially-aware robot navigation

Based on social robot notions and its abilities described above, it can be deduced that

Definition 13 *A socially-aware navigation is the strategy exhibited by a social robot which understands social conventions relatives management of space and conforms to them in order to preserve a comfortable interaction with humans. Resulting behavior is predictable, adaptable and easily understood by humans.*

Definition 13 implies, from robot’s point of view, that humans are no longer only dynamic obstacles but also social entities. This thesis supports the idea that robots will mimic human interaction in human environments even in the case of robot-robot interaction in order to keep their behavior easily understood.

2.4.3 Socially-aware robot navigation related work

Based on the key concepts proposed by [Goffman \(1963\)](#) (see definitions 6 and 7) the related work on socially-aware navigation has been divided into focused interaction and unfocused interaction according to the main characteristics of the study in question.

2. SOCIAL CONVENTIONS: FROM COMFORT AND PROXEMICS TO SOCIAL ROBOTICS

| Type of interaction | Task and References |
|-----------------------------------|--|
| Unfocused interaction | Minimizing probability of encounter. (Tipaldi and Arras, 2011; Sehestedt et al., 2010; Chung and Huang, 2010) Avoiding collisions. (Tamura et al., 2010; Ohki et al., 2010; Ratsamee et al., 2012; Lam et al., 2011; Kruse et al., 2012; Lamarche and Donikian, 2004) Passing people. (Pacchierotti et al., 2006; Kirby et al., 2009; Kuderer et al., 2012; Pandey and Alami, 2010) Staying in line. (Nakauchi and Simmons, 2000) |
| Focused interaction | Approaching humans. (Althaus et al., 2004; Yamaoka et al., 2010; Shi et al., 2011; Carton et al., 2012) Following people. (Gockley et al., 2007; Zender et al., 2007; Miller et al., 2008) Walking side-by-side. (Morales Saiki et al., 2012) |
| Focused and unfocused interaction | Combination of previous listed tasks. (Sisbot et al., 2007; Hansen et al., 2009; Svenstrup et al., 2010) |

Table 2.3: Related work on socially-aware robot navigation.

2.4.3.1 Unfocused interaction

In unfocused interactions people and robot share a common presence in a setting. Robots must negotiate their position with others by means of nonverbal behaviors or by the knowledge of rules in social spaces.

Minimizing probability of encounter. In Tipaldi and Arras (2011) a spatial affordance map is used to learn and predict spatio-temporal behavior of people in a house, such map serve as a cost model for planning robot paths which minimize the probability of encounter with people. Very similar approach is presented in Sehestedt et al. (2010) where motion patterns are learned in an office environment by means of Sampled Hidden Markov. Chung and Huang (2010) proposed Spatial Behavior Cognition Model (SBCM), a framework to describe the spatial effects existing between human-human and human-environment. SBCM was used to learn and predict (short-

2. SOCIAL CONVENTIONS: FROM COMFORT AND PROXEMICS TO SOCIAL ROBOTICS

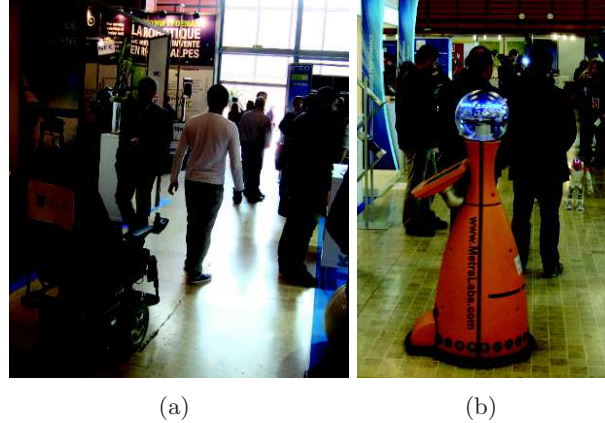


Figure 2.8: Examples of unfocused interaction. Robots must negotiate their position with others by means of nonverbal behaviors or by the knowledge of rules in social spaces

term and long-term) behaviors of pedestrians in an environment and to help a service robot to prevent potential collisions.

Avoiding collisions. [Tamura et al. \(2010\)](#) proposed a method for smooth collision avoidance of humans by using the social force model to determine whether a pedestrian intends to avoid collision with the robot or not. In [Ohki et al. \(2010\)](#) estimation of motion and personal space of humans are used by a rescue robot to avoid collisions with evacuees. A recent work [Ratsamee et al. \(2012\)](#) has extended the social force model by including a force due to face pose in order to perform human-like navigation with a robot which is able to avoid a human in a face-to-face confrontation. Based on their harmonious rules, [Lam et al. \(2011\)](#), developed a Human-Centered Sensitive Navigation which is capable of move a robot respecting human and other robots sensitive zones. Their experiments show a robot avoiding humans and other robots by respecting its sensitive zones in corridors and in surveillance tasks. In order to get legible strategies, experiments were realized by [Kruse et al. \(2012\)](#) to collect data from human avoiding collision with other human. Focused was put in velocity adaptation more than in path adaptation. They proposed a new cost model that takes into account the context in order to adjust velocity of the robot.

In [Lamarche and Donikian \(2004\)](#), visual optimization of the path along with personal space were used to achieve human-like collision avoidance for agents in virtual crowds. Agents speed was filtered by the personal space module in charge of respecting a given

2. SOCIAL CONVENTIONS: FROM COMFORT AND PROXEMICS TO SOCIAL ROBOTICS

minimal distance to humanoids and obstacles.

Passing people. (Pacchierotti et al., 2006) presented a robot motion control which included a module that achieves a people passing behavior in corridors (pass a person by the right) inspired by human spatial behavior. In Kirby et al. (2009) a generalized framework for representing social conventions as components of a constraint optimization problem was presented and it was used for path planning and navigation. Social conventions were modeled as costs to the A* planner with constraints like shortest distance, personal space and pass on the right. Simulation results showed the robot navigating in a “social” manner, for example by moving to its right when encountering an oncoming person, as it is socially expected.

In Kuderer et al. (2012) they propose a technique to reason about the joint trajectories that are likely to be followed by all the agents, including the robot itself. The approach learns a model of human navigation behavior that is based on the principle of maximum entropy from the observations of pedestrians. They implemented their technique on a mobile robot and carried out experiments in which a human and a robot pass each other while moving to their target positions.

Socially Aware Mobile Robot Motion is addressed in Pandey and Alami (2010). The framework is supported by adding, deleting or modifying milestones based on static and dynamic parts of the environment, the presence and the motion of an individual or group as well as various social conventions. Experiments show the robot dynamically adapting its navigation around humans according the factors already mentioned.

Staying in line. Authors in Nakauchi and Simmons (2000) developed a model for the personal space of people standing in line to build an strategy for a robot to do the same task. Their personal space model is used both to detect the end of a line and to determine how much space to leave between the robot and the person in front of it.

2.4.3.2 Focused interaction

In focused interactions people and robot share a common focus of attention when executing the main activity.

2. SOCIAL CONVENTIONS: FROM COMFORT AND PROXEMICS TO SOCIAL ROBOTICS

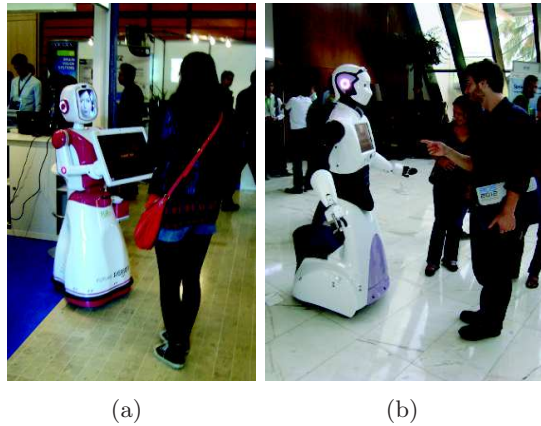


Figure 2.9: Examples of focused interaction. People and robot share a common focus of attention when realizing the main activity.

Approaching humans. [Carton et al. \(2012\)](#) investigated what abilities robots will need to successfully retrieve missing information from humans. Socially-aware navigation is employed to request help from human passers-by. An approach based on Bezier curves is implemented as a nonlinear optimization problem with the objective to find a velocity profile for the Bezier path under constraints enhancing social acceptance. Experiments were realized where the robot approaches a static human at different velocities and angles.

In [Yamaoka et al. \(2009, 2010\)](#) formations has been implemented to appropriately control robot position as it presents information to a human. They established a model for information-presenting robots to appropriately adjust their position. The model consisted of four constraints to establish O-space: proximity to a listener, proximity to an object, listeners field of view, and presenters field of view. Model was implemented for a humanoid robot with a motion-capturing system.

In [Althaus et al. \(2004\)](#) the authors proposed a method for a robot to join a group of people engaged in conversation. The results of the implementation and the experiments conducted with their platform show a human-like behavior as judged by humans. Robot just wants to preserve the formation of the group and doesn't know explicitly where the o-space is located.

[Shi et al. \(2011\)](#) studied natural human interaction at the moment of initiating conversation in a shopkeeper scenario where a salesperson meets a customer. Then they use the observed spatial formation and participation state to model the behavior of initiating a conversation between a robot and a human.

2. SOCIAL CONVENTIONS: FROM COMFORT AND PROXEMICS TO SOCIAL ROBOTICS

Following people and walking side-by-side. Gockley et al. (2007) proposed a model for person following behavior and evaluate two approaches: one following the exact path of the person and other following in the direction of the person, they conclude that the second one is the most human-like behavior. The factors taken into account for the design of the model proposed are human-likeness, personal space, reliability in tracking of person and safety.

The people following behavior presented in Zender et al. (2007) preserves socially acceptable distances from its human user, and gives readable social cues (gaze, speech) indicating how the robot tries to maintain engagement during following.

Miller et al. (2008) proposed an iterative planning technique that seeks for people moving to the same target location than the robot and follow them. An idea presented at that work is that robot can get space to pass by shooing someone away, to do that robot approaches the person, accelerates shortly in front of him/her before braking again. In most cases, this behavior causes people to intuitively free the path. It is interesting because even when this is not socially correct robot is aware of human reactions to space invasions and makes use of them to navigate.

In Morales Saiki et al. (2012) they developed a computational model for side-by-side walking in a social robot by using an utility model describing how people prefer to move. The model was built based on recorded trajectories of pairs of people walking side-by-side.

2.4.3.3 Focused and unfocused interaction

Some works have proposed techniques capable of fulfill the two kind of interactions. In Hansen et al. (2009) an adaptive system was introduced which detects whether a person seeks to interact with the robot based on the person's pose and position, this work was presented as a basis for human aware navigation. Navigation was implemented using human centered potential fields. This method was extended in Svenstrup et al. (2010) by including RRTs to minimize the invasion to social spaces of humans.

Closer to human aware navigation and management of physical space, we could mention Sisbot et al. (2007) where a motion planner is presented which takes explicitly into account its human partners. The authors introduced criteria based both on the control of the distance between the robot and the human, and on the control of the position of the robot within the human's field of view. The authors introduced the cri-

2. SOCIAL CONVENTIONS: FROM COMFORT AND PROXEMICS TO SOCIAL ROBOTICS

terion of visibility, which is simply based on the idea that the comfort increases when the robot is in the field of view of a person.

2.5 Conclusion and thesis positioning

This chapter presented an introduction to important concepts related to social conventions first, from the point of view of sociology and after, from the point of view of robotics. Human management of space was identified as the most important issue to be integrated in robot navigation. As reviewed in the section 2.3 human management of space is a very complex dynamic system involving special factors for each one of the studied cases: one person, a group of people interacting or humans interacting with objects. The factors reviewed in this work are expressed in terms of functions combining distance, direction and comfort.

The study of nonverbal behavior exhibited by humans can give cues for the robots to

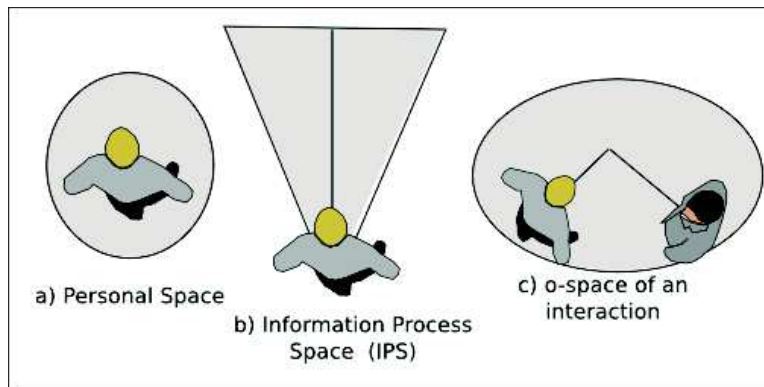


Figure 2.10: We consider as discomfort the invasion made to humans' space by the robot, specifically, a) Personal Space b) Information Process Space or c) o-space. Models for that spaces are presented in the next chapter under the concept of Social Filter.

mimic the unfocused interaction resulting of the navigation close to humans and also to perform focused interaction in a more human-like manner. In the present thesis we use the behavioral cues of distance and posture of people as a cue to estimate if two people are interacting and to determine the sensible space of a human.

Perception of non verbal behavior is, in general, a very challenging problem. Automatic techniques to collect social cues and methods to process them in order to get social signals are needed not only for robotics field but also for social sciences where it exists the requirement of both fairer judgment and more precise measurements. Tools

2. SOCIAL CONVENTIONS: FROM COMFORT AND PROXEMICS TO SOCIAL ROBOTICS

developed in this thesis are directed in that sense.

This thesis considers as discomfort the invasion made to humans' space, specifically Personal Space (Hall, 1966), O-Space (Kendon, 2010) and Information Process Space (Kitazawa and Fujiyama, 2010), by the robot. Models for those social concepts have been implemented and grouped under the concept of a Social Filter which is explained in chapter 3. A graphical representation of those spaces can be observed in Fig.2.10. The proposed models permit to estimate a risk of disturbance based on the magnitude of discomfort that the robot could cause, it is assumed that the discomfort will be higher in the spaces previously mentioned. The integration with navigation strategies is presented in chapter 4.

A robot also needs to understand how to identify an activity space. Based on the data collected from sensors, our model of the environment can decide that the probability of occupation in a particular area is zero (according to standard techniques of representation in robotics), i.e. empty or free space, becoming an option for navigate through it. In a socially-aware navigation approach that space still needs to pass some more tests. If we reflect in the way a human would take navigation decisions in the same empty space, we observe that he/she takes it decision not only based on collision issues but also by considering the meaning of that space associated to risk, discomfort or disturbance to others. The disturbance problem that our robot can create by entering the activity space between a human and an object will be detected by using the framework of formations, details are found in section 3.4.

One can wonder if it is necessary to have a more complex model of space around humans different from the standard technique which consists on inflating the obstacles detected from sensors by a predefined ratio, we think yes, it is, mainly because studies like Gifford (1983) support the idea that perceived interpersonal distance is different from physical one and also because, as explained in the sociological review, that inflation ratio responds to a complex system of regulation and communication use of the space. This work is a basic step to understand that complexity.

Regarding the related works listed in section 2.4.3 it is observed that the concept of personal space is present but the concept of o-space and F-formations have not been included explicitly. We think that the latter concepts can give us a clue to consider the interactions between the dynamic obstacles in the environment and to improve autonomous navigation by a better understanding of management of space realized by humans. In this work we are restricting the shape of o-space only to the number of

2. SOCIAL CONVENTIONS: FROM COMFORT AND PROXEMICS TO SOCIAL ROBOTICS

people involved and we assume only standing participants in conversation. We suppose that if a formation between individuals A and B is detected then an interaction exist between A and B instead to assign it a probability like (Groh et al., 2010). Inspired by (Ge et al., 2009) we propose to find groups of more than two people from detected formations, resulting method is presented in 3.3.4.

Our socially-aware risk-based navigation presented in section 4.4 is intended to deal with focused and unfocused interaction. Compared to the proposal of human aware navigation presented in (Sisbot et al., 2007), our assumption of discomfort is, in some way, the opposite of the visibility criterion: the field of view (in front of the human) shows the point of interest of a person then, if the robot enters to it, the activity of the person will be interrupted decreasing the comfort function.

Chapter 3

Modeling social behaviors : The Social Filter

3.1 Introduction

Our proposal to endow robots with the ability of socially-aware navigation is the Social Filter, which implements constraints inspired by social conventions in order to evaluate the risk of disturbance represented by a navigation decision.

The information flow between the Social Filter and the components of the navigation

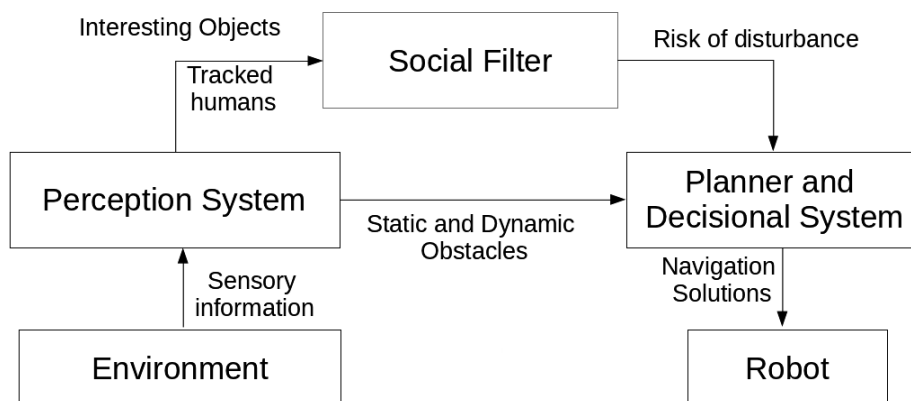


Figure 3.1: Flow of information related to the Social Filter. The original navigation solutions are “filtered” according to the social conventions taken into account.

system can be observed in the diagram shown in fig. 3.1. The Social Filter receives from the perception system a list of tracked humans and a list of interesting objects in

3. MODELING SOCIAL BEHAVIORS : THE SOCIAL FILTER

the environment. The interesting objects are designated manually according to their importance in a particular context, for example, an information screen in a bus station. After the process of such data, the Social Filter is able to output the risk of disturbance relative to people and interesting objects on request of the planner and the decisional system. Thus, the original navigation solutions are “filtered” according to the social conventions taken into account. Notice that the concept of social filter is built as a higher layer above the original safety strategy, the planner and the decisional system are responsible to include the new constraints.

The methods of the Social Filter were designed according to the theory of social conventions presented in the precedent chapter, including:

- estimation of disturbance due to invasion of Personal Space,
- estimation of disturbance due to invasion of O-Space in interactions,
- estimation of disturbance due to invasion of Activity Space and
- early detection of conversations to predict future O-space location.

Next sections explain each one of social conventions currently implemented in the Social Filter. It begins to present the models of Personal Space and Information Process Space relatives to the management of spaces of one person, then the model of O-Space space related to groups is addressed, followed by the explanation of the model of Activity Space used in the case of interactions of objects. Finally, the approach developed to early detection of O-Space is presented.

3.2 Space related to one person

3.2.1 Model of Personal Space

The model that we have implemented to represent Personal Space employed as reference the measures of personal zone and intimate zone defined by Hall (described in section 2.3.1.1), i.e about 1.2 meters from the position of body. The discomfort is higher at the front of the human than behind. Personal Space has shape of ellipse similar to the one proposed in the Social Force model. The size of Personal Space will not be modulated by pedestrian speed and, for simplicity reasons, will remain constant during all navigation situations. Personal Space is centered in human position and its front is bigger.

The model consists in blending two Gaussian functions Γ_f and Γ_b both of them centered

3. MODELING SOCIAL BEHAVIORS : THE SOCIAL FILTER

in the position of the person. The first one represents the personal space situated in front of a human and for this reason it's wider than the last one representing the back space. The Gaussian values represent the risk of disturbance associated with a point in the space around the pedestrian. The composition of Gaussians reflects well the fact that disturbance is more important in positions closer to the center than in the borders and stronger in the front than in the back.

Risk of disturbance computation is made by evaluating a two-dimensional Gaussian function Γ of covariance matrix Σ and centered in x , for each point p around the person:

$$\Gamma_{x,\Sigma}(p) = e^{-\frac{1}{2}(p-x)^t \Sigma^{-1} (p-x)} \quad (3.1)$$

where x, p are in \mathbb{R}^2 and Σ is a diagonal covariance matrix defined as:

$$\Sigma = \begin{pmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{pmatrix}. \quad (3.2)$$

The appropriate shape of the personal space is obtained by selecting the same values for σ_x in both Γ_f and Γ_b but different values for σ_y , being the one of Γ_f the double of the value for Γ_b . Fig. 3.2 shows an example of personal space for a human, the approximated shape is shown in the left where the Gaussian values are projected in the plane of the ground, observe in the right the high importance of disturbance at the center and front. The disturbance has a maximum value of one in human's position and a minimum value of zero in his/her public zone.

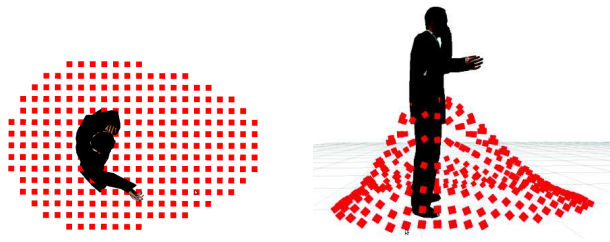


Figure 3.2: Personal space calculated by Social Filter Module. Left: top view, right: side view. Height of the Gaussian means Risk of disturbance then maximum disturbance is located at human position.

This model is static (similar to that presented by (Laga and Amaoka, 2009) which uses face orientation instead of body orientation) but could be easily extended with a dynamic behavior that corresponds to context information like crowdedness, relation-

3. MODELING SOCIAL BEHAVIORS : THE SOCIAL FILTER

ship or culture.

3.2.2 Model of Information Process Space

A second model was designed to estimate risk of disturbance around one single person. This time the representation was based on the Information Process Space (IPS) explored in (Kitazawa and Fujiyama, 2010) whose analysis proposes a size and shape of such space (IPS was presented in section 2.3.1.2). Remember that is in the IPS where psychological comfort is evaluated. The model presented in this section assumes that interference to the IPS will causes discomfort to person. A function is established to estimate such discomfort. To discourage the robot from invade IPS the discomfort estimated at each point is taken as the risk of disturbance in that point.

The idea is to estimate discomfort related to IPS by using the formulation of the Doppler effect phenomenon. The Doppler effect establishes that the perception in the frequency of a sound varies with the movement of source and observer according to the next equation:

$$f' = \frac{c}{c - v_s \cos \theta_s} f, \quad (3.3)$$

where f' is the frequency perceived by the observer, f is the frequency emitted by the source, c is the velocity of sound, v_s is the velocity of the source and θ_s is the angle between the direction of the source and the direction of the line that links observer and source.

A moving pedestrian with constant velocity can be considered as a sound source and all the other neighboring points as observers. Think about the perceived frequency in one point as the discomfort in the same point, the higher frequency detected the more discomfort associated. The numerical values of the parameters in eq. (3.3) have been determined empirically to best adjust the results for IPS in (Kitazawa and Fujiyama, 2010). They are $c = 3.43$, $v_s = 3.0$ and f is determined in function of distance as stated in next equation:

$$f = \begin{cases} 1 & \text{if } d < d_e \\ 1 - \left(\frac{d-d_e}{d_l}\right) & \text{if } d_e \leq d \leq d_e + d_l \\ 0 & \text{if } d > d_e + d_l \end{cases} \quad (3.4)$$

where d is the distance from the human's position, d_e is the main radius of IPS effect and d_l is the range where the IPS loses its effect. In our current implementation

3. MODELING SOCIAL BEHAVIORS : THE SOCIAL FILTER

$d_e = 2.5m$ and $d_l = 1.0$. The fig. 3.3 shows the shape of the IPS around a human. The disturbance has a maximum value of one and a minimum value of 0. The model of IPS

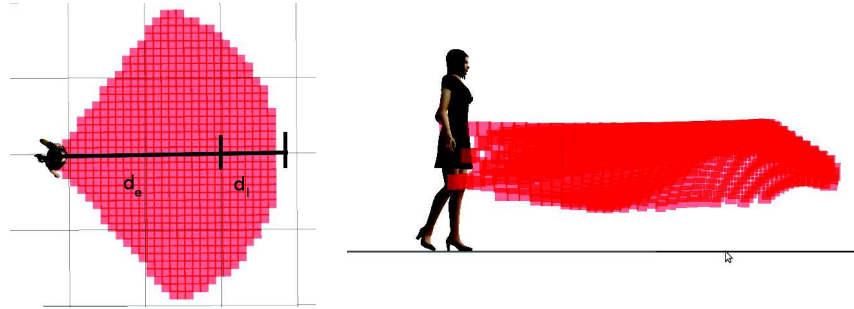


Figure 3.3: Information Process Space calculated by Social Filter Module. **Left:** top view, **right:** side view. The height of the function means Risk of disturbance, maximum disturbance is located in front of human.

presented in this section is not in conflict with the model of Personal Space, in fact both of them were used in the navigation approach described on section 4.3. However a deeper analysis is needed in order to minimize the redundancy that appears at the front of humans when both models are used.

Next section describes the models implemented for the case of groups of people.

3.3 Space related to groups of people

The interest of this section is centered on the O-space built by people in conversation and on the associated management of space done by humans around it. It was not pretended to build a model that fully explains the behavior of humans in interaction but a model that helps a robot to navigate around an interaction in a human-like manner. To that purpose the O-Space concept was helpful.

3.3.1 The shape of O-space: Preliminary psychophysical experiment

In this section it is described a preliminary experiment which was conducted to have an idea of the shape and measures of the O-Space in two-people formations, specifically, the Vis-a-vis was studied because it is the most common and basic arrangement for social interaction. Assuming that people incorporates wheelchairs as an extension to their own body (Higuchi et al., 2006), it is claimed that the management of space around a group made by a human driving a wheelchair must reflect his/her natural

3. MODELING SOCIAL BEHAVIORS : THE SOCIAL FILTER

behavior in absence of the wheelchair, i.e., the respect of O-Space is maintained.

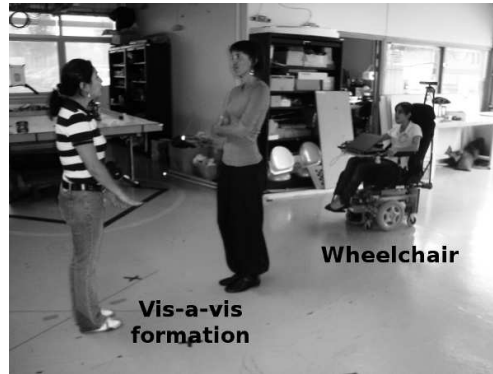


Figure 3.4: Psychophysics experiment. Drivers were asked to guide the wheelchair in a path that passed close to a Vis-a-vis formation.

Experiment. Our tests were realized at INRIA’s Robotics Lab. Eight subjects were asked to drive our robotic wheelchair which was entirely in manual mode. No one of them had driven a wheelchair before. Brief explanations about the wheelchair were followed by 5 minutes of free driving, after that time, they were asked to guide the wheelchair from a starting position to a goal position, both positions were clearly signaled in the floor. Two people were placed in a vis-a-vis formation at the middle of the room with a distance of 1 meter between them in such a way that the driver must avoid them to reach the goal. The interacting people were asked not to do visual contact with the drivers. A total of eight executions were registered for each participant. After the first four executions of the test the orientation of the vis-a-vis formation was changed and the test continued. Figures 3.4 and 3.5 show the arrangement of the scenario for our tests.

Data collection and process. A laser-based localization module was executed on the powered wheelchair, it permitted to reconstruct the trajectories followed by drivers during the test and to estimate the shape of the O-Space by adjusting the avoided space around the interaction to a known geometric shape. Resulting trajectories are shown in fig.3.6 a). We observed that the O-Space can be approximated with an ellipse which is presented in fig. 3.6 b) for the first configuration of Vis-a-vis and in fig. 3.6 d) for the second configuration, such ellipses are the result of a least-squares orthogonal distance fitting¹ realized over the trajectories collected. In the first case the major axis

¹It was used a modified version of the code in:
<http://computation.cs.cinvestav.mx/~anzures/vision/ellipse.php>

3. MODELING SOCIAL BEHAVIORS : THE SOCIAL FILTER

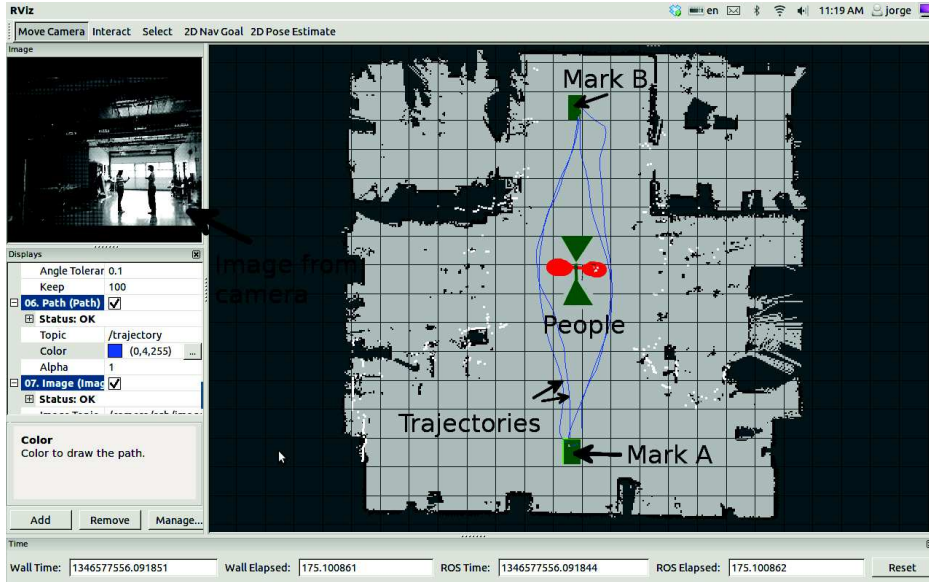


Figure 3.5: Scenario of the experiment at INRIA’s Robotics Lab. The map was made using laser information from wheelchair. Red ellipses represent the position of people in the first part of the test and green triangles the positions for the second part. Examples of registered trajectories joining marks A and B are shown in blue.

had a length of 5.99m and the minor axis had a length of 2.37m. In the second case the major axis had a length of 6.49m and the minor axis had a length of 1.91m. In both cases the estimated center was located very close to the center of the conversation. This information was useful to complete the parameters of our model presented in the next section. The speed profile registered for each driver was useful to adjust the speed configuration of the algorithm RiskRRT presented in next chapter.

Discussion It is clear that a more controlled and extensive experiment is needed to understand the shape of O-Spaces. It was noticed that when drivers had an acquaintance in the Vis-a-vis formation, in some executions they tried to establish visual contact or talk with him, then in future tests relationship between drivers and people in formation must be controlled. The resulting measures could have a significant variation from the real ones because it seems that people without experience driving wheelchair underestimates spatial requirements, for example, in passability tests (Higuchi et al., 2006). Also, the effect of the distance to the walls would have a significative effect on the shape we obtained, we think that in a more spacious area the shape of O-Space would be wider. Then our results only would be valuable for conversations in corridors.

3. MODELING SOCIAL BEHAVIORS : THE SOCIAL FILTER

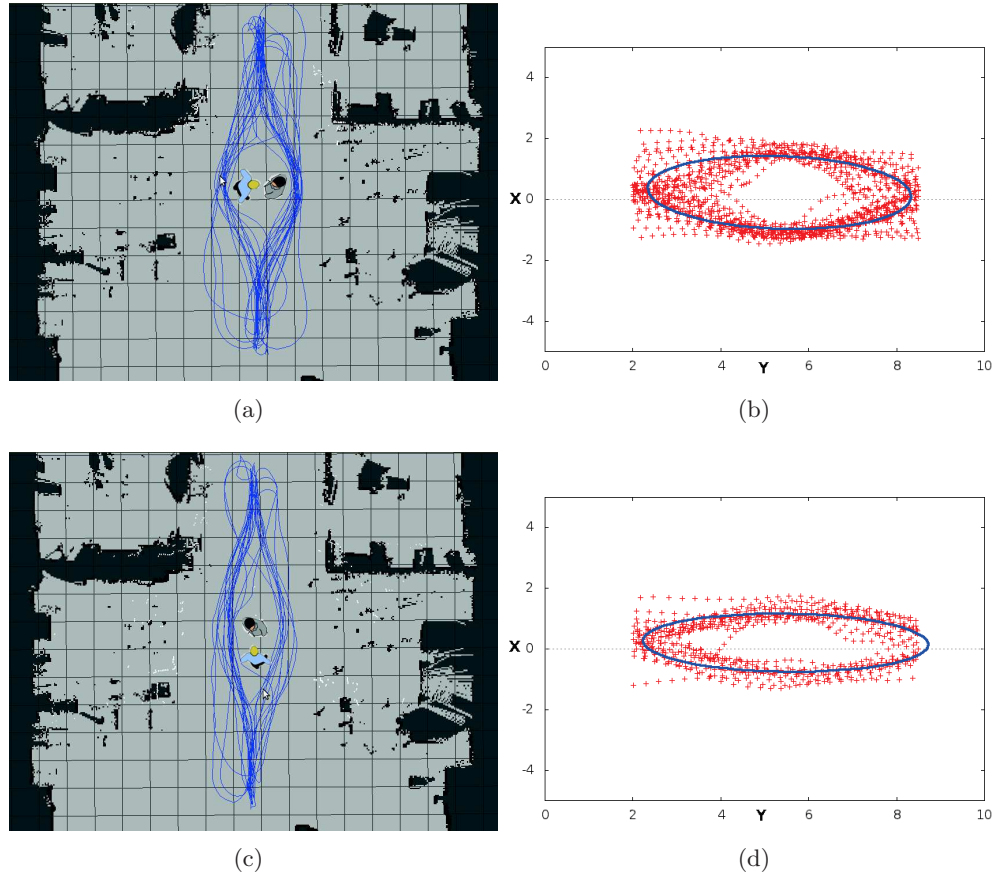


Figure 3.6: Analysis of trajectories. In a), c) observed trajectories of the wheelchair in the test and in b), d) by adjusting ellipses to the observed trajectories, parameters were obtained for O-Spaces shape. Only positions in a region of 3 meters around the interaction were included.

In any case this preliminary study can serve as a basis for future experiments.

3.3.2 Model of O-Space in two-people formations

In this section it will be described how the O-space location has been estimated by taking into account the information of the previous section and the concepts related to interaction spaces exposed in section 2.3.2.1. First, the model for two-people in interaction is addressed, after, the model for bigger groups is presented in the next section.

The assumption is that a formation between individuals A and B is enough evidence

3. MODELING SOCIAL BEHAVIORS : THE SOCIAL FILTER

to detect an interaction between them.

Considering two people some formations have been identified as the most frequent, see section 2.3.2 for a review, four of them have been implemented in the current model: Vis-a-vis, L-Shape, C-Shape and V-Shape. The shape of O-space will be dependent on the particular formation identified.

A geometric representation has been designed for each formation, the model is based on the position and orientation of the body of participants.

Given the positions of pedestrians $H_1 = (x_1, y_1)$ and $H_2 = (x_2, y_2)$ in the plane of the ground and their respective orientations ϕ_1 and ϕ_2 around the normal to that plane, D_H is calculated as the euclidean distance between H_1 and H_2 . Calculate a point V_i as the intersection of the vectors beginning in H_1 and H_2 in the direction of ϕ_1 and ϕ_2 , respectively. Let H_{12} be the mean point between H_1 and H_2 . Let C be the mean point between V_i and H_{12} . Calculate D_i as the distance between C and H_{12} .

The O-space is represented by a two-dimensional Gaussian function Γ_c of covariance matrix S and centered in C , then for each point Q around the center we have:

$$\Gamma_{C,S}(Q) = e^{-\frac{1}{2}(Q-C)^t S^{-1}(Q-C)} \quad (3.5)$$

where S is a diagonal covariance matrix defined as:

$$S = \begin{pmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{pmatrix}. \quad (3.6)$$

To get the shape of the O-space depending on the formations, some values has been chosen for the parameters which are shown in the next table:

| Formation | σ_x | σ_y |
|-----------|------------|------------|
| Vis-a-vis | $D_H/3$ | $2 * DH/3$ |
| L-Shape | $D_H/2$ | D_i |
| C-Shape | $D_H/4$ | $D_i/3$ |
| V-Shape | $D_H/3$ | $2D_i$ |

The Gaussian is oriented in the direction of the segment $\overrightarrow{H_{12}C}$ in order to be consistent with the location of the point of interest exhibited by humans by using the orientation of their bodies. The O-space was implemented by mean of a grid, the value for the risk of disturbance of each cell corresponds to the value of its center evaluated in the Gaussian.

All the elements defined can be seen in fig. 3.7 for the four cases of formations im-

3. MODELING SOCIAL BEHAVIORS : THE SOCIAL FILTER

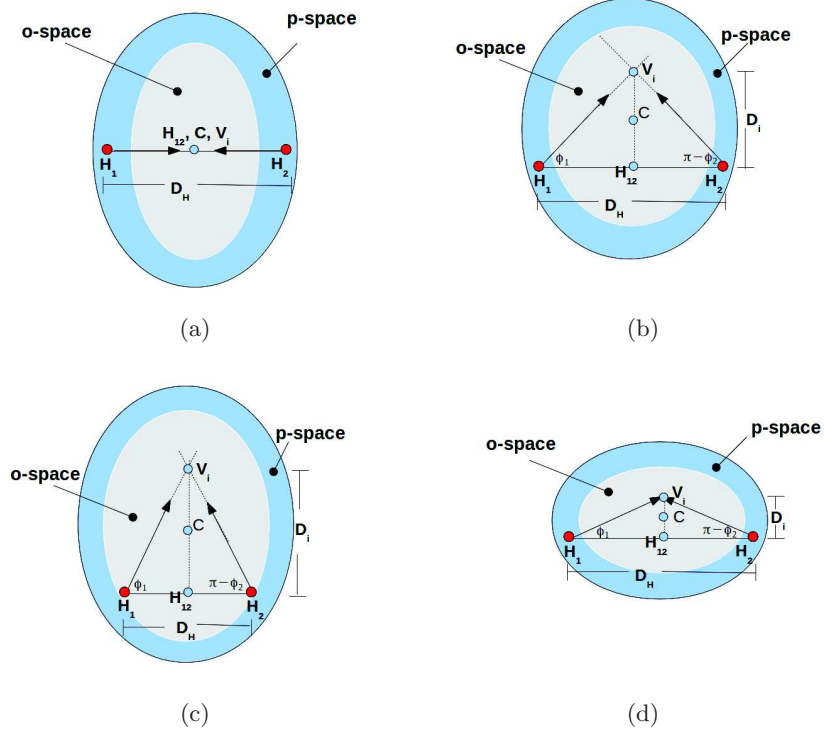


Figure 3.7: Scheme showing the elements of O-space model depending on formations: a) Vis-a-vis b) L-Shape c) C-Shape and d) V-Shape

plemented in the Social Filter. In the Fig. 3.8, the calculated O-space for a Vis-a-Vis interaction is shown, observe that the shape is adjusted according the results of the experiment described in section 3.3.1, i.e., the ellipse is longer and wider when the orientation of the Gaussian coincides with the orientation of the wheelchair.

Until this point we have stated that the formations of people can help us to develop strategies to avoid interference between robot and people interacting, however it is natural to ask if we can use our models to give a cue to the robot to approach a formation in order to interact. Next section discusses this subject.

3. MODELING SOCIAL BEHAVIORS : THE SOCIAL FILTER

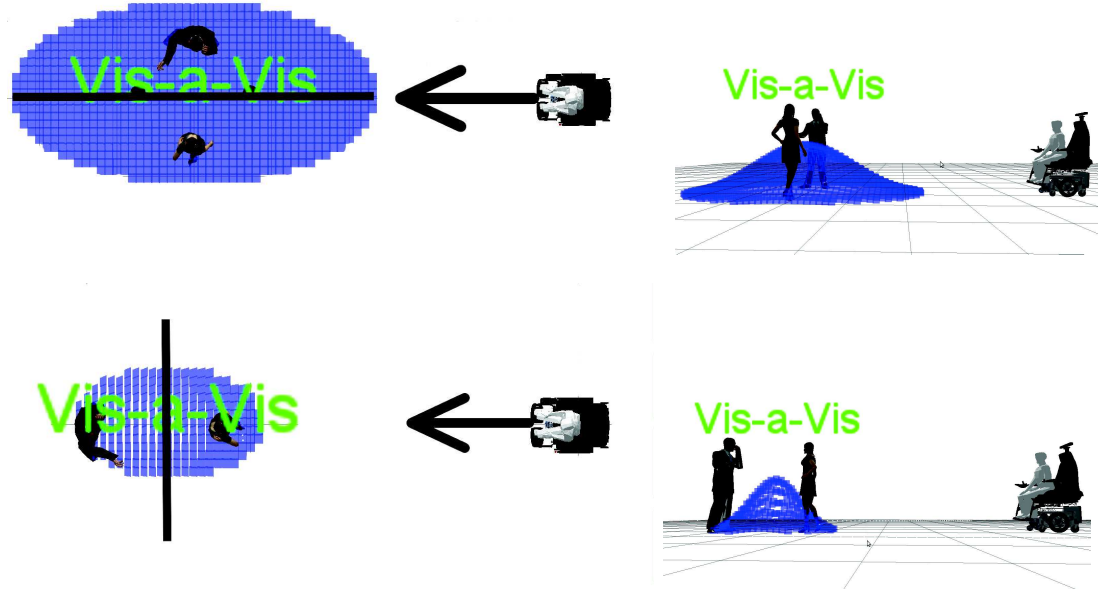


Figure 3.8: O-space calculated by Social Filter Module for a Vis-a-Vis F-formation. Maximum risk of disturbance is located at O-space center, in the picture the disturbance is represented by height of Gaussian. O-space shape is adapted according the orientation of wheelchair and human interaction.

3.3.3 Approaching a two-people formation

In this section it is presented how our models of O-space can be used to give a cue to the robot to approach a two-people formation in order to interact. The knowledge of the risk of disturbance distribution can be used to determine where a robot can be more effective to attract the attention of interactants and become part of a group. According to the theory presented in section 2.3 it is possible to determine an area, roughly coincident with p-Space of an interaction, where the robot can be placed to show its interest to the group in the same way as humans do. Such area is formed by the border of the O-space and the same border enlarged by the average size of a human body.

Some points, called meeting points, have been calculated as positions in the p-Space where a robot can share the space in an equitable way with the other two humans already present. A Meeting point must be placed on the line formed by V_i and H_{12} (defined in precedent section).

Fig. 3.9 shows the location of the meeting points, calculated by the Social Filter. We can pass these points to our robot as goals, when the robot reaches the Meeting point it

3. MODELING SOCIAL BEHAVIORS : THE SOCIAL FILTER

will be part of the group according to our models. Observe that the location of Meeting points is related to the focus of attention showed by the formation.

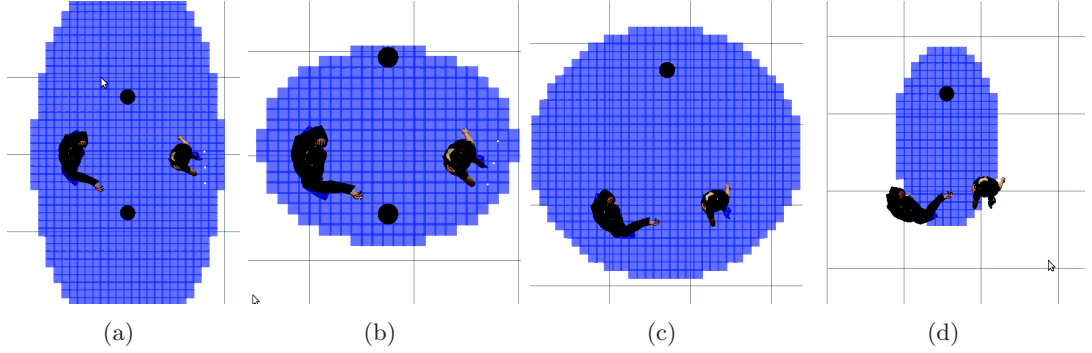


Figure 3.9: Meeting points (black dots) calculated by the Social Filter for each formation, viewed from upside. If a robot reaches that position, a group of three will be formed according to our models. The location of Meeting points depends on the focus of attention.

3.3.4 Groups of more than two people

It was presented in section 2.3.2.2 that when more than two people are in conversation they exhibit a formation with circular shape. It is more obvious when the number of participants grows. In this thesis the O-space in groups of more than two humans has the shape of a circle whose center coincides with that of the inner space. The approach followed consists in detect a group by observing formations between the participants according to the following rule: a person p_i can be added to an existing group of size k if and only if p_i has a formation with at least the half of the existing group members (Ge et al., 2009).

The detection algorithm takes as input the two-people formations previously detected and applies the *intra-group tightness criterion* (equation 3.7) to decide which of them must be combined. The process operates on a list of groups which is initialized with the formations received as input. At the beginning the first element is designated as candidate to group p or CG_p . The CG_p is compared with the next on the list designated as CG_q , if the two groups satisfy the criterion then they are merged into a new group. If a merge occurs the original groups are deleted from the list, the merged group is added to the list in the position of CG_p becoming the new CG_p .

When the criterion is not satisfied, the next in the list to CG_q is chosen as CG_q and the process continues. Once that CG_q reaches the end of the list a new iteration starts.

3. MODELING SOCIAL BEHAVIORS : THE SOCIAL FILTER

The algorithm finishes when CG_p reaches the end of the list and the output are the groups with more than two elements.

Definition 14 Let G_p and G_q be two groups of p and q elements, intra-group tightness criterion is formulated as:

$$e(G_p \cup G_q) \geq \hat{e}(G_p \cup G_q) + e(G_p) - \hat{e}(G_p) + e(G_q) - \hat{e}(G_q) \quad (3.7)$$

where $e(G_x)$ is a function that returns the number of formations in the group G_x and $\hat{e}(G_x)$ is the minimum number of formations that a set of size x must have to be considered as a group of size x , (Ge et al., 2009).

Once that a group of more than two people is detected, the location of the O-Space, c is calculated based on the positions of the people using the formula:

$$c = \frac{1}{n} \sum_1^n p^i \quad (3.8)$$

where n is the number of people in the group and p^i is the position of person i . The value sd_g is defined as the minimum distance from p^i to c .

As in the other cases a two-dimensional Gaussian is used to estimate the risk of disturbance around the O-space. The Gaussian is centered in c and the standard deviation for both dimensions is sd_g .

Example of output from the Social Filter is shown at fig. 3.10 where the system is able to recognize a group of four people instead of three single formations between them, also a C-shape and a Vis-a-vis formations were detected in the scene.

A similar approach in (Maisonasse et al., 2006), took advantage of contextual elements like saliency of objects in the scene to estimate shared attention in order to detect interaction groups. They also agree in that arrangement of interaction group is the best detector of activity presence in a scene. Unlike them the present thesis is concerned with the use of space made by humans during interaction more than to classify the type of interaction.

Next section discusses our implementation done in order to take into account interaction space between a human and one object.

3. MODELING SOCIAL BEHAVIORS : THE SOCIAL FILTER

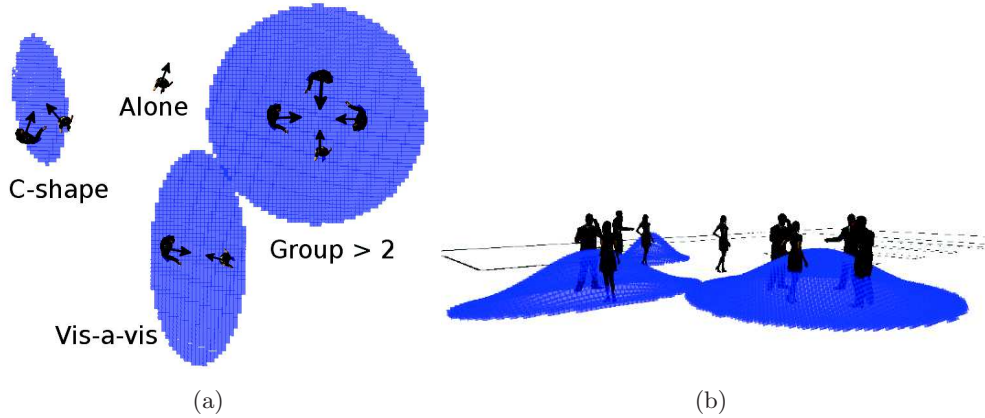


Figure 3.10: Output of Social Filter, C-shape and Vis-a-vis formations together with a group of four people have been detected. In a) up-view, in b) side-view. O-Space is colored in blue, the height of the Gaussian represents the Risk of disturbance to each interaction.

3.4 Space related to objects

Spaces of interaction resulting of the activity made by a human with a fixed object in the environment can be classified under the name of Activity Spaces, concept defined and discussed in 2.3.3.

The detection of interaction with objects was done by perceiving Vis-a-vis formations between interest objects and humans, when a Vis-a-vis occurs the Activity Space will be represented by the same model to the used in O-spaces in Vis-a-vis formations, the difference is that the object is thought as a fixed human participant with orientation and position defined by hand. This last assumption poses a strong constraint to the system but we argue that in service applications the needed information about interesting objects locations is easily available. For example, the location and orientation of an automatic ticket machines in a train station can be easily known. Still the model is dependent on human position and orientation and also on robot direction of approach.

Risk of disturbance estimation. In a similar way that with Vis-a-vis formation between humans, the center of the Activity Space is placed at the middle of the line joining the center of the human body and the object. It is at the center where the discomfort is maximal as expressed by the bi-dimensional Gaussian. However due to the fact that the object does not have a Personal Space, the combined effect of Personal Space of the human and the Activity Space model made that in the case that space is big

3. MODELING SOCIAL BEHAVIORS : THE SOCIAL FILTER

enough the tendency of the robot will be to pass closer to the object than to the human.

In the fig. 3.11 we can see the output of the Social Filter when a human is looking the information of train departures in a screen, at this moment the activity space is created. Robot is discouraged to pass in the activity space because there the risk of disturbance is higher.

In some cases even when a person is facing an information panel he/she is not in-

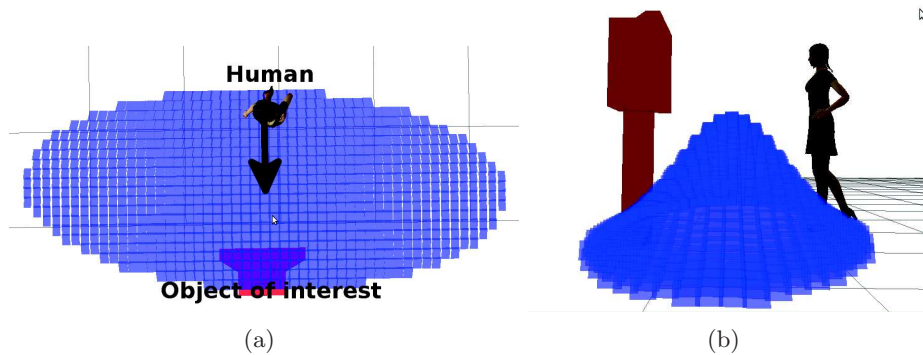


Figure 3.11: Output of Social Filter when the interaction of a human with an object of interest is detected: a) up view , b) side view. In the activity space the risk of disturbance (in blue) is higher.

teracting with it then an activity space must not be created. In order to take into account these cases the present model can be extended by including multiple features, like in (Prest et al., 2012) where a weakly supervised approach was proposed to learn interactions between humans and objects from a set of images.

3.5 Prediction of O-Space location in future conversations

In previous sections it has been presented how to detect interactions between humans, the proposed models are geometric-based and can be considered static because they did not take into account their evolution in time. When humans navigate in dynamic environments they anticipate social situations and include this information in their decisions of motion. The non-holonomic constraints of many robots, like our robotic wheelchair, make more evident the need of anticipation in order to keep a comfortable behavior. A socially-aware robot that respects O-space will suddenly break if it detects an interaction too late on its way and cannot find a feasible alternative. In that case,

3. MODELING SOCIAL BEHAVIORS : THE SOCIAL FILTER

stopping is a safe behavior but not a socially intelligent one. Also, to avoid all regions where people are approaching it is not efficient. Figure 3.12 shows a pair of common situations appearing in human contexts, in both of them two people are getting closer. In the “meeting situation”, they are intentionally approaching in order to initiate a conversation. In the “passing” situation they come near because their trajectories cross the same space at same time. We think that if a robot is able to distinguish between these situations its navigation will become more comfortable and efficient.

Present section exposes an approach to decide if a standing conversation will take place



Figure 3.12: Two common situations appearing in human environments. In a) a “meeting” situation where two people are close to begin an interaction, in b) a “passing” situation where two people will pass each other.

in the short future between two people. Section begins to describe the experiment done to collect data in order to observe the two situations introduced above. After related variables are described and analyzed. Finally, the Hidden Markov Models (HMM) used to classify the behavior of two agents will be presented.

3.5.1 Analysis of interaction scenarios

In order to understand the dynamics of the “meeting” and “passing” situations, an interaction scenario and a no-interaction scenario, respectively, were designed. These two scenarios at the beginning are ambiguous, it is necessary to observe for a while to distinguish between them. Definitely, humans use gesture, visual behavior and verbal behavior to succeed in the same task. This section shows that even without the full incorporation of these variables it is possible to realize classification.

Data acquisition in a psychophysical experiment. The robotics lab of FMAT¹ was used to film people in interaction and no-interaction scenarios in order to collect

¹www.matematicas.uady.mx

3. MODELING SOCIAL BEHAVIORS : THE SOCIAL FILTER

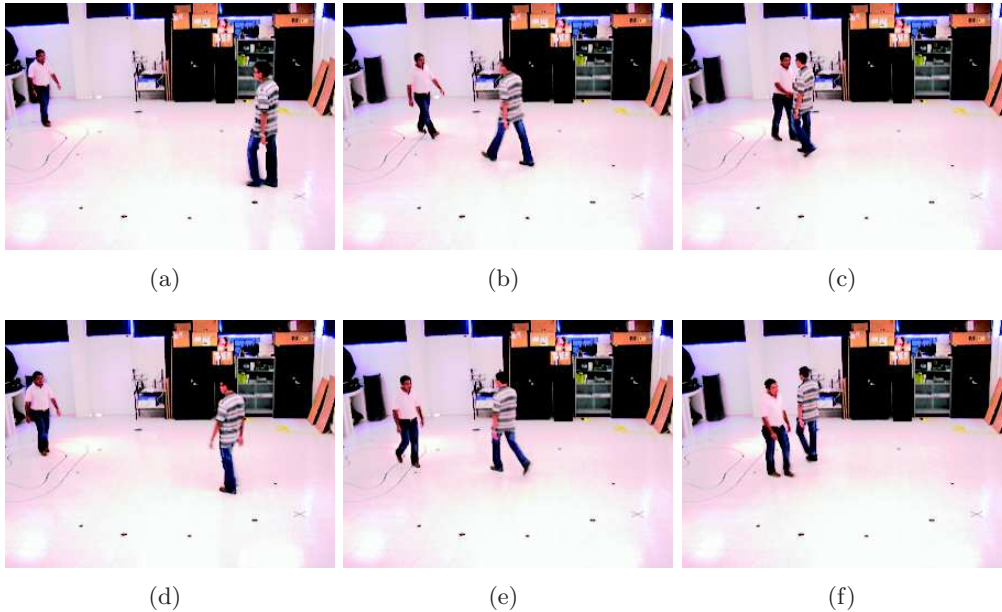


Figure 3.13: Designed scenarios to study “meeting” and “passing” situations, **up**: interaction scenario shown in three consecutive moments: a) wander b) approach c) conversation, **down**: no-interaction scenario at three consecutive moments: d) wander e) come near f) pass.

data about them. Twenty four different pairs of people were recorded in these two scenarios shown in fig. 3.13. For the no-interaction scenario participants were asked to walk from a mark in the floor to the other located 6 meters away, as the demanded trajectories were aligned, participants were constrained to avoid them mutually. For the second scenario they were asked to greet each other and, depending on the role assigned, ask for directions to the library, finish the conversation and continue to their corresponding marks. The sequences collected were manually labeled using the ViPER Ground Truth Authoring Tool¹, at each frame we marked the position of the center of people’s body projected in the ground plane together with the orientation, projected in the same plane, as shown in the figure 3.14.

Such data were recorded in XML format and later processed to get the next variables in function of time: relative distance (rd), relative speed (rs), alignment (al), formation (fo), change in position (cp) and change in angle (ca). These variable were selected because of its relation to the situations we are studying and because they could give better results to distinguish the two scenarios.

¹<http://vipер-toolkit.sourceforge.net>

3. MODELING SOCIAL BEHAVIORS : THE SOCIAL FILTER

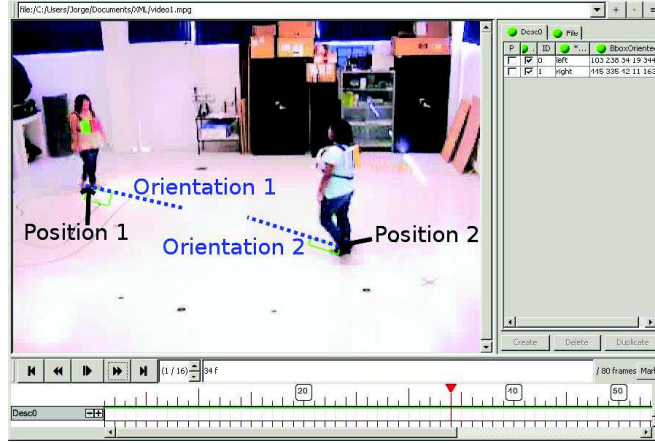


Figure 3.14: The sequences recorded were labeled, each frame has position and orientation of people.

To begin the definition of the framework, suppose that x_i^t is the 2D vector which contains the ground plane coordinates for person i at frame t . Next we will describe each variable and the equation to calculate it. Graphics are proportioned to facilitate the analysis. The analysis was done in a window of time of five seconds between 2.0 and 7.0 seconds. It is signaled in the graphics with a shadow. It was defined like this because sequences started when people were still not moving and because interaction situations were longer than no-interaction situations. The beginning of a interaction was defined as the moment when people shake hands, it happened in average at second 4.5.

Relative distance (rd) is the euclidean distance between people positions at time t calculated as in eq. 3.9. The fig. 3.15 shows the graphics of relative distance against time for twelve sequences from our collected data. It can be observed that the difference between the two scenarios is evident and it can be traduced as: in interaction scenarios the relative distance decreases until participants begin to interact and after it remains constant. In no-interaction scenarios the distance decreases until it reaches its minimum and immediately it increases again.

$$rd = \|x_i^t - x_j^t\| \quad (3.9)$$

Relative speed (rs) Is the difference between the speed of person i and person j at time t as shown in eq.3.11. Speed of each person is calculated as in eq. 3.10, where δ is the time between two consecutive observations. The fig. 3.16 shows the graphics of relative speed against time for twelve sequences from our collected data. It

3. MODELING SOCIAL BEHAVIORS : THE SOCIAL FILTER

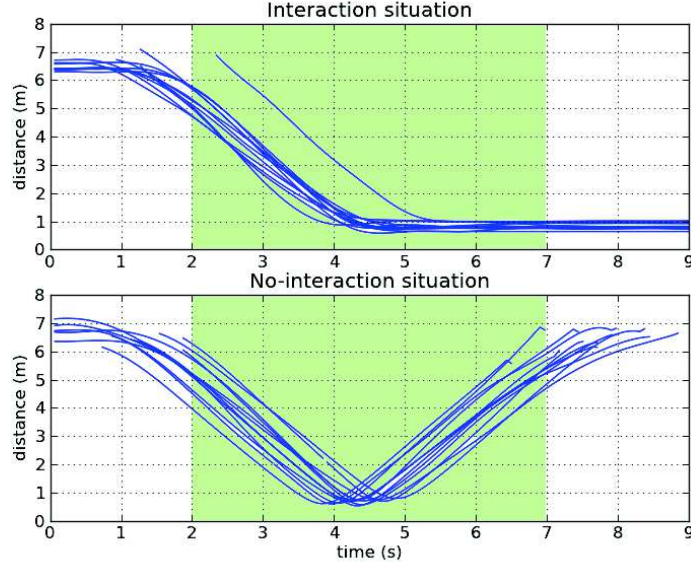


Figure 3.15: Behavior of the variable relative distance in twelve sequences.

can be observed that in an interaction situation, at the beginning, the relative speed kept around 2.0, then it begins to decrease until stabilization near 0.0, signaling that a conversation has started. In the other case the relative speed kept around 2.25 during all the window of time analyzed.

$$sx_i = (x_i^t - x_i^{t-1})/\delta \quad (3.10)$$

$$rs = \|sx_i - sx_j\| \quad (3.11)$$

Alignment (al) is the angle formed between the orientations of the two people at time t , it is calculated as shown in 3.12 where α_i is the orientation of person i , its behavior is plotted in the fig. 3.17. The graphics for each case are very unstable which is explained because the orientation was labeled by hand with precision of degrees. The important fact to observe is that in an interaction situation the alignment stabilize around 180 degrees once that the interaction has begun. One graphic stabilize in a different value because the participants formed a formation distinct to Vis-a-vis. In the case of no-interaction the graphics never stabilize.

$$al = V_1 \cdot V_2, \text{ where } V_i = [\cos(\alpha_i), \sin(\alpha_i)] \quad (3.12)$$

3. MODELING SOCIAL BEHAVIORS : THE SOCIAL FILTER

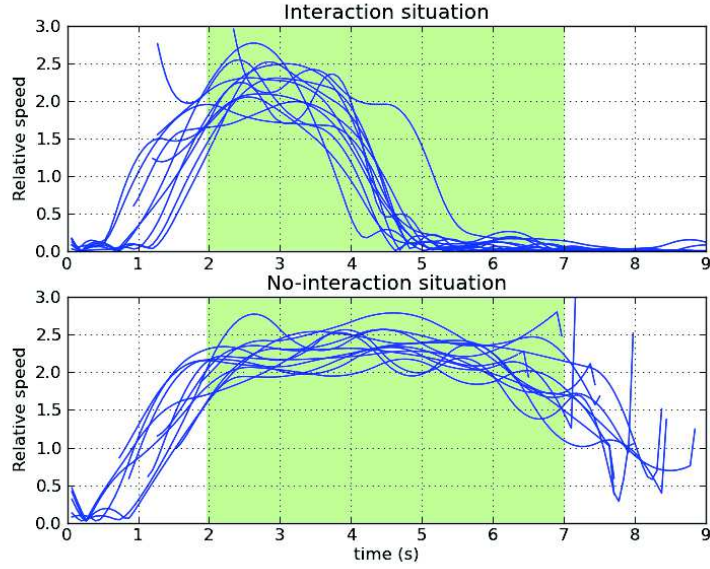


Figure 3.16: Behavior of the variable relative speed in twelve sequences.

The alignment variable was first chosen because it was expected that the avoiding

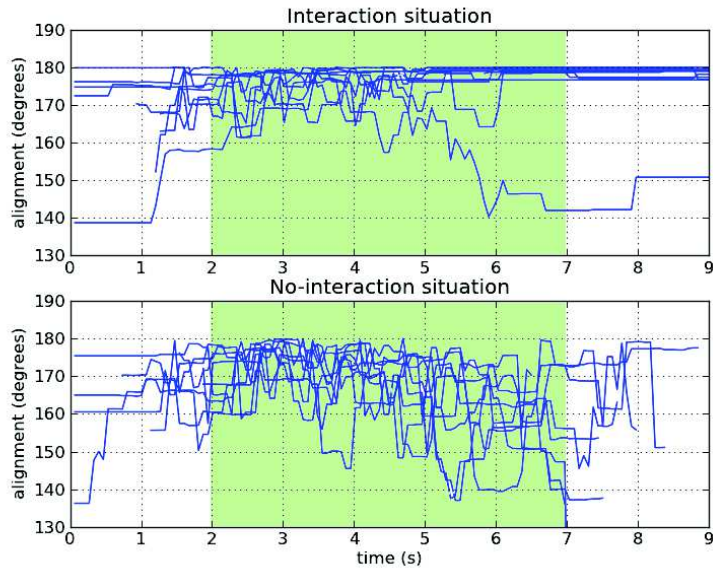


Figure 3.17: Behavior of the variable alignment in twelve sequences, the difference between the two cases is that no-interaction never stabilizes.

3. MODELING SOCIAL BEHAVIORS : THE SOCIAL FILTER

behavior would exhibit a constant pattern between sequences, but it was not the case.

Change on distance (cd) is a discrete variable that shows if the distance is increasing, decreasing or remains constant between consecutive frames, it is calculated as follows:

$$cd = \begin{cases} -1 & \text{if } rd^t - rd^{t-1} < 0 \\ 1 & \text{if } rd^t - rd^{t-1} > 0 \\ 0 & \text{otherwise} \end{cases} \quad (3.13)$$

The behavior of this variable is shown in the fig. 3.18 which confirmed our analysis done on the variable relative distance alone. The additional information was that this relative distance can not be used as predictor because the observable difference occurs at almost the same time in both situations.

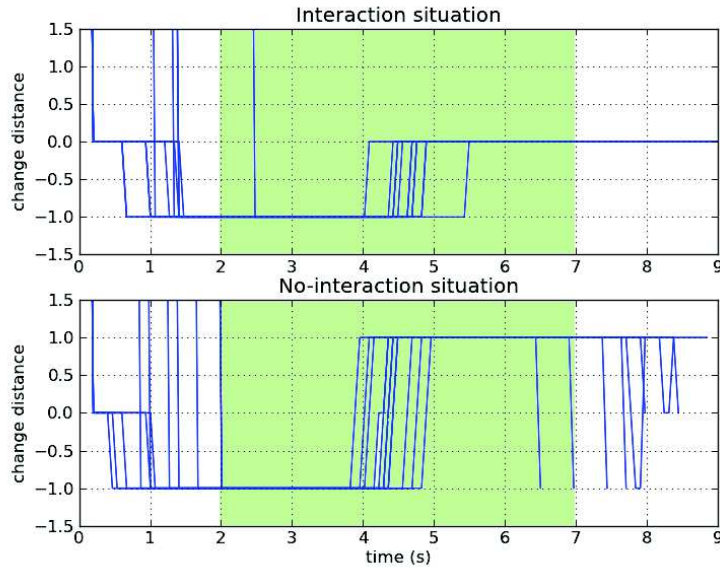


Figure 3.18: Behavior of the change in relative distance in twelve sequences.

Formation (fo) was inspired by the formations between people, it has a value of 1.0 if people exhibit a formation, 0.5 if we are not sure and 0.0 if they are totally out of formation. Following paragraphs will explain how to calculate fo .

The space around a human can be discretized in slices of 22.5 degrees about the common line, L , from one person to other, as shown in the figure 3.19. The orientation of the body corresponds to a unique slice and the label of the slice can be used to designate

3. MODELING SOCIAL BEHAVIORS : THE SOCIAL FILTER

such orientation. By using that discretization we can identify a formation according the pair of labels, for example a Vis-avis formation is a (A, I) pair. It was found that this

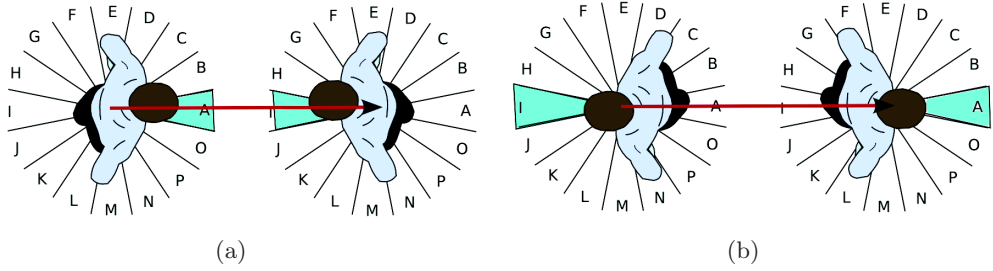


Figure 3.19: Discretization of the space around people. In a) is shown a pair (A, I) which coincides with a Vis-a-vis formation and have a value of $fo = 1$, in b) there is a pair (I, A) which have a value of $fo = 0$. Line L is in red.

discretization is very similar to the Oriented Point Relation Algebra, $OPRA_4$ presented in (Moratz et al., 2005) but its properties as Spatial Calculus tool were not analyzed in this work.

It is important to remember that the discretization is oriented in the direction of L , the line joining the two people's position, in other words (A, I) is not equal to (I, A) .

A function is defined to categorize the current discretization. Let θ be the angle, in degrees, formed by a person with L , θ is used to get a label. For example, if $\theta \in (-11.5, 11.5]$ its label will be A , if $\theta \in (168.75, 191.25]$ its label will be I .

To complete the definition, use the next equation to get fo :

$$fo(r_1, r_2) = \begin{cases} 1.0 & \text{if } (r_1, r_2) \in FFORMSET \\ 0.5 & \text{if } (r_1, r_2) \in CFFORMSET \\ 0.0 & \text{otherwise} \end{cases}$$

where r_i is a region from the discretization of space shown in fig. 3.19 and the sets FFORMSET and CFFORMSET are listed in the following table:

| FFORMSET | CFFORMSET |
|--|--|
| $(A, I), (B, H), (C, G), (D, F), (E, E), (M, M), (N, L), (O, J), (P, K)$ | $(A, H), (A, J), (B, G), (B, I), (C, F), (C, H), (D, E), (D, G), (E, F), (M, L), (N, K), (N, M), (O, K), (O, L), (P, I), (P, J)$ |

The behavior of this variable is shown in the fig. 3.20. It can be observed that in the case of interaction situation before the beginning of the conversation, formations

3. MODELING SOCIAL BEHAVIORS : THE SOCIAL FILTER

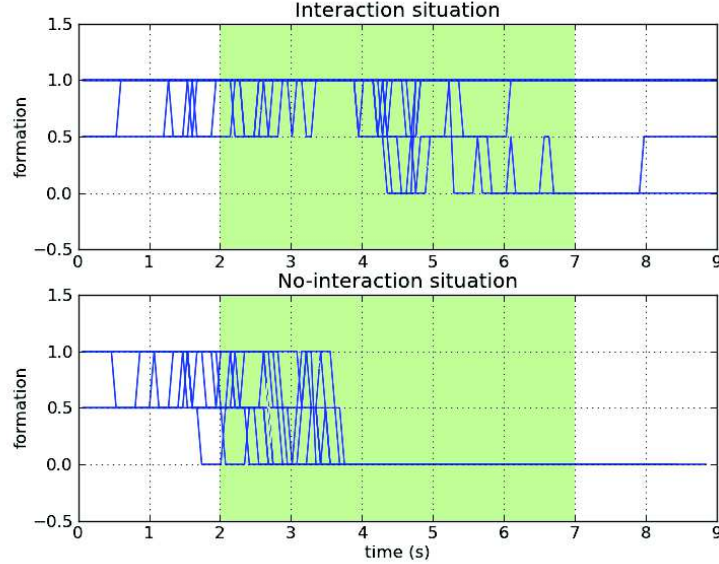


Figure 3.20: Behavior of the formation variable in twelve sequences.

were detected in the most of cases. Between the second 3.0 and 4.0, in all the sequences formations were detected. At the end for almost all the sequences formations were detected. In two sequences during the first seconds of interaction one participant adopted an orientation that could not be classified as formation then fo got a value of zero. The second case resulted more interesting because it was observed that people break the formation in order to avoid the other before they get too closer. It can be seen that few time before the fourth second all sequences were in no formation state.

Discussion. A set of variables in function of time were selected to do the analysis: relative distance (rd), relative speed (rs), alignment (al), formation (fo), change in position (cp) and change in angle (ca). After the analysis of the chosen variables it was decided that *relative speed* and *formation* were the most viable to learn in order to have an early detection of interaction. They were not only different between the two situations but also the difference can be observed before the interaction occurs. fo variable includes the effect of the variable alignment and is less noisy. Distance variable is of course very important, but for the prediction task did not result very informative. In many cases the simple statement of thresholds on the variables rs and fo could work to detect interaction, for example, with the fulfillment of the next three conditions:

$$rd < 2.0 \quad rs < 1.5 \quad fo = 1.0$$

3. MODELING SOCIAL BEHAVIORS : THE SOCIAL FILTER

However we think that a specific human behavior is composed by a sequence of events and a method that takes into account the evolution of those variables in time will be more robust. Moreover the inclusion of variables related to visual, gestural and verbal behavior would conduct to establish more thresholds. Instead of thresholds the present work will be based on the technique called Hidden Markov Models which is well suited to work with classification of sequences as will be described in next sections.

3.5.2 Model of interaction using Hidden Markov Models (HMM)

This section will present Hidden Markov Models (HMM) used to classify the behavior of two agents. The HMM models proposed permit to decide if a conversation will take place based on its higher probability of occurrence done a sequence of observations. If a conversation is predicted, the observations are used to estimate O-Space location and calculate associated risk of disturbance. Next we will present the framework of HMM and the models we have developed based on them.

3.5.2.1 HMM definition and notation

HMM techniques can deal with time-sequential data and can provide time-scale invariability as well as learning capability for recognition. HMM have been used with success in many applications, for example in speech recognition and motion prediction. Recently, they have been also proposed to realize automatic analysis of social interactions. Many books and tutorials are easily available to learn HMMs, here we have followed the presentation done in (Vazquez Govea, 2010).

Hidden Markov Models are a specialization of the Bayes Filter for discrete variables, they can be defined in terms of three variables:

- S_t, S_{t-1} the current and previous states, which are discrete variables with value $S_t, S_{t-1} \in 1, \dots, N$ for a fixed N .
- O_t , the observation variable, which is a multidimensional vector in \mathbb{R}^M .

HMM is a class of stochastic state transition model, a probability of transition is assigned from one state to another, as time goes, state transitions occur stochastically. An additional conditional independence assumption is made with respect to the standard Bayes filter: both the observation and the transition probabilities are considered to be stationary, that is, independent of time. The next three probabilities complete the definition of an HMM model:

3. MODELING SOCIAL BEHAVIORS : THE SOCIAL FILTER

- $P([S_1 = i]) = \pi_i$. The state prior is stored as a vector $\pi = \{\pi_1, \dots, \pi_N\}$ where each element contains the prior probability for the corresponding discrete state.
- $P([S_t = j] | [S_{t-1} = i]) = a_{i,j}$. Transition probabilities are represented with a $N \times N$ transition matrix A where each element $a_{i,j}$ represents the probability of reaching state j in the next time step, t , given that the system is in state i at time $t - 1$.
- $P(O_t | [S_t = i]) = G(O_t; \mu_i, \sigma_i)$. The observation probability density function is represented by a Gaussian distribution for every state. The set of all the Gaussians' parameters are denoted as $b = \{(\mu_1, \sigma_1^2), \dots, (\mu_N, \sigma_N^2)\}$.

We denote the whole set of parameters for an HMM by $\lambda = \{\pi, A, b\}$. Next section presents a simple example to illustrate the HMM model.

3.5.2.2 Example of HMM

Consider the graph in fig. 3.21, a directed graph which have its edges labeled with transition probabilities. The sum of probabilities over the outgoing edges from each node is one, therefore we can say now that this graph represents a first-order Markov Model. A single realization of a Markov model is a random path that moves from state to state according to the model's probabilities. The transitions probabilities are organized in the transition matrix:

$$A = \begin{bmatrix} 0.2 & 0.8 \\ 0.9 & 0.1 \end{bmatrix} \quad (3.14)$$

whose element A_{12} is the probability of moving to state S_2 , given state S_1 as starting point. The hidden property of HMM's comes from the fact that states are not directly observable (they are "hidden") so the observations give us an indirect information in terms of the probability of observing O_t when current state is S_i . In the case of discrete observations usually we work with symbols that correspond to each observation, in a continuous case we suppose we have a Gaussian distribution for observations in each state. For the example, suppose our parameters for observation probability are $b = (20, 25), (15, 16)$, and that for the definition of the state prior, we can assume that both states are equiprobable for the first observation, then $\pi = [0.5, 0.5]$. Finally, our HMM definition, $\lambda = \{\pi, A, b\}$, is complete.

In practice, the parameters of A , B , and π are determined by a learning procedure, the standard approach is to use the Forward-Backward algorithm.

3. MODELING SOCIAL BEHAVIORS : THE SOCIAL FILTER

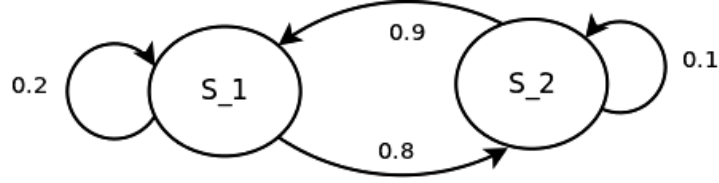


Figure 3.21: Graph used to represent an HMM model

Once that we have our model defined and the parameters estimated, we are interested in answer the question, what is the probability of a complete observation sequence given the model parameters?. Such probability, $P(O_{1:T} | \lambda)$, is called joint observation probability or simply likelihood. It can be calculated from the forward probabilities defined in equation 3.15 as it is shown in equation 3.16.

$$\alpha_t(i) = \left[\sum_{j=1}^N \alpha_{t-1}(j) P([S_t = i] | [S_{t-1} = j]) \right] P(O_t | [S_t = i]) \quad (3.15)$$

$$P(O_{1:T} | \lambda) = \sum_{i=1}^N \alpha_T(i) \quad (3.16)$$

In this thesis the interest is to classify a sequence of observations coming from two humans as interaction or not interaction. An HMM model has been defined for each situation. Relative speed and formation has been selected as variables. The parameters of the model have been obtained by training on sequences from the dataset presented in section 3.5.1 using the Baum-Welch algorithm. This is an approximative iterative optimization technique for maximizing the likelihood of the data. The algorithm takes an initial estimate of the parameters and greedily improves it by following the likelihood gradient.

Finally the learned models have been used to classify test sequences from the mentioned dataset based on the likelihood. In next sections, details and results of our method are presented.

3.5.2.3 Model of interaction

As first step to design an HMM model for our problem, it was necessary to decide what the states were, after an analysis of the two cases we wanted to test and based on experience about the situation, it was determined that our model for interaction was composed of three states “wander”, “approach” and “conversation” and the model for no-interaction had three states “wander”, “come near” and “pass”. In fig. 3.13

3. MODELING SOCIAL BEHAVIORS : THE SOCIAL FILTER

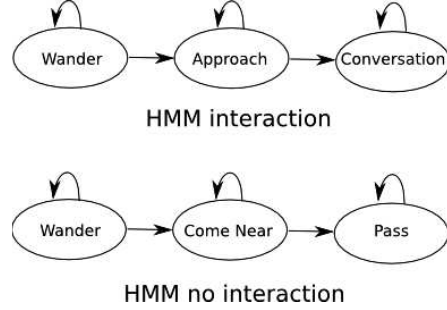


Figure 3.22: HMM models for interaction and no-interaction

examples of these states can be observed. Observe that this model is of type “left-right” because from an state done only two options are possible: stay or continue to the unique state at the right. The resulting graphs representing our models can be observed in the fig. 3.22.

The relative speed was converted to discrete values as follows:

| | |
|-----------|------------------------|
| H:High | $rs > 1.5$ |
| A:Average | $0.5 \geq rs \leq 1.5$ |
| L:Low | $0.0 \geq rs \leq 0.5$ |

Formation variable can take three values:

| | |
|---------------|------------|
| W:well-formed | $fo = 1.0$ |
| U:uncertain | $fo = .5$ |
| B:bad-formed | $fo = 0.0$ |

Each symbol of the observation was formed for a pair $rs \times fo$, i.e., nine possible symbols were considered. After learning our model for interaction, λ^{INT} , got the following parameters:

$$A = \begin{bmatrix} 0.8766 & 0.1233 & 0 \\ 0 & 0.9522 & 0.0477 \\ 0 & 0 & 1 \end{bmatrix} \quad (3.17)$$

$$b = \begin{bmatrix} 0.7 & 0.01 & 0.2 & 0.01 & 0.01 & 0.01 & .01 & 0.01 & .01 \\ 0.9 & 0.01 & 0.02 & 0.01 & 0.01 & 0.01 & 0.01 & 0.01 & 0.02 \\ 0.02 & 0.01 & .02 & 0.01 & 0.01 & 0.01 & 0.7 & 0.01 & 0.2 \end{bmatrix} \quad (3.18)$$

3. MODELING SOCIAL BEHAVIORS : THE SOCIAL FILTER

and the model for no interaction, λ^{NI} got the following parameters:

$$A = \begin{bmatrix} 0.7997 & 0.2002 & 0 \\ 0 & 0.8399 & 0.16 \\ 0 & 0 & 1 \end{bmatrix} \quad (3.19)$$

$$b = \begin{bmatrix} .02 & 0.01 & 0.9 & 0.01 & 0.01 & 0.01 & 0.01 & 0.01 & 0.01 \\ 0.9 & 0.01 & 0.02 & 0.01 & 0.01 & 0.01 & 0.01 & 0.01 & 0.01 \\ 0.02 & 0.01 & 0.9 & 0.01 & 0.01 & 0.01 & 0.01 & 0.01 & 0.01 \end{bmatrix} \quad (3.20)$$

The length of observations for each sequence had a fixed value of 60 which is equivalent to four seconds. As sequences had different length the rule used both in training and test to get the observations was: the first observation is taken once that the relative distance between pedestrians is less than five meters.

3.5.2.4 Results of interaction vs no interaction classification

Assuming that the sequences of the two competing situation can be differentiated and that the learned models accurately characterize them, $P(O_{1:T} | \lambda^k)$, can be used to select the one that provides the better explanation for a sequence. Sixteen sequences of interaction and seven of no-interaction from the dataset were selected to test our model. They were distinct from those used for training. For each sequence 45 consecutive observations were collected according to the rule mentioned above. The likelihood was calculated using λ^{INT} and λ^{NI} . The bigger likelihood was selected and the respective model was taken as winner. The classification was correct for all the sequences.

The graphics presented in fig. 3.23 show the average of the likelihood (loglikelihood) calculated for the proposed models. In a) sixteen interaction sequences were tested with different length of observations, in b) seven no interaction sequences were tested. It can be observed that with few observations the system is not able to classify accurately a sequence. However, after two seconds (30 observations), the classification is correct in both cases. Interestingly, a no interaction scenario can be correctly detected with 20 observations (approximately 1.3 seconds). In this work it was defined a shaking-hands action between people as the beginning of a conversation such action was observed around 40 observations from the beginning of interactions.

According to the results, our method is capable of doing an early detection of interactions.

3. MODELING SOCIAL BEHAVIORS : THE SOCIAL FILTER

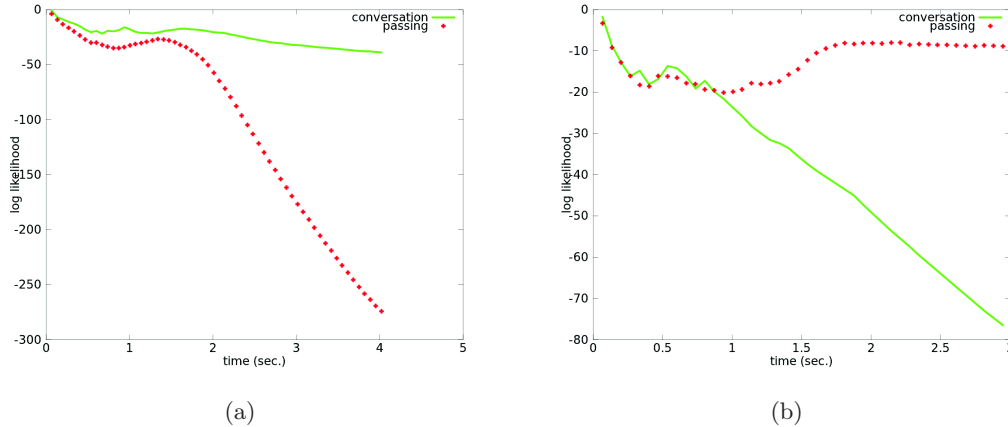


Figure 3.23: Average of the likelihood calculated for two models: conversation (green continue line) and passing (red dotted line). In a) sixteen interaction sequences were tested. In b) seven no interaction sequences were tested.

3.6 Conclusions

In this section we have presented the implementation of social conventions in a module called Social Filter. The methods of the Social Filter were designed according to the theory of social conventions presented in the previous chapter, the models developed in this thesis are listed below:

- One person. A model of Personal Space was presented which enables the estimation of disturbance made for a robot entering such a space. In the same way a model of Information Process Space was developed in order to estimate human psychological discomfort.
- Groups of people. A model of detection of the O-Space exhibited by interacting people based on the observed formation among participants was proposed. Position and orientation information from humans in the environment can be used now to estimate the risk of disturbance that robot could cause by interfering interactions. The same model can be used to promote socially correct approach from robot to humans by using defined Meeting Points. Psychophysics experiments were realized in order to find adequate parameters for our O-Space model.
- An strategy to estimate disturbance due to invasion of the Activity Space was also proposed.

3. MODELING SOCIAL BEHAVIORS : THE SOCIAL FILTER

- Finally, a strategy for early detection of conversations was proposed based on Hidden Markov Models and trained with data coming from real humans. The objective is to predict future O-space location and provide this information to navigation strategies.

Chapter 4

Including social conventions in robot navigation

4.1 Introduction

Nowadays the applications in robotics are moving more and more to human environments. As areas of mobile service robotics and robotic assistance of humans are becoming more common in everyday life, humans need to share the physical space with robots and robots need to take into account the presence of humans. To be accepted, robots must behave in a socially acceptable way. Their trajectories must be safe but also predictable. Their behavior should follow social conventions like respecting proximity constraints, avoiding people interacting or joining a group engaged in conversation without disturbing. As discussed in the previous chapter the abilities to understand and to express non verbal cues are associated with social adjustment. Similarly, the perception of territorial boundaries established by a group of humans and the respect to these bounds are evidence of social behavior. In this chapter, two proposals of socially-aware navigation are described, both of them are able to move in dynamic environments. Section 4.2 presents a background on navigation in dynamic environments. Section 4.3 is about an optimization-based navigation strategy aimed to minimize discomfort of humans in the environment while the robot navigates to its goal, previously presented in (Rios-Martinez et al., 2012). Section 4.4 reviews our proposal of combining a Risk-Based navigation approach with the models of social conventions developed in this thesis, such work was published in (Rios-Martinez et al., 2011).

4.2 Navigation in dynamic environments: state of the art

Navigation in dynamic environments is still a challenging problem despite of the many available approaches that have addressed the problem.

The standard approach consists in employing a planning algorithm which takes as input start state and goal states for a robot, together with an abstraction of the environment, and outputs a sequence of valid intermediary states linking start and goal. After, robot must execute the planned sequence until it arrives to the goal by using a control strategy adapted to robot kinematic and dynamic features.

Planning algorithms can be classified in local planners and global planners. Local planners perform short-term reasoning based in the knowledge of local surrounding of robots (normally equivalent to the reach of on-board sensors) while global planners produce the complete plan for the robot from its current position to the desired goal, including states falling out of the reach of the on-board sensors.

Frequently, discrete versions of the environment, like occupancy grids, are used in global planners which are in charge of static obstacles avoidance. In occupancy grid maps, each cell $\langle x, y \rangle$ of the grid stores the probability that the corresponding area in the environment is occupied. Many current navigation systems use A* algorithm (Hart et al., 1968) or its variants to calculate an optimal solution (at grid discretization size) which is communicated to a local planner. Following the present framework, very often, the local planners are responsible of the avoidance of dynamical obstacles. Examples of popular local planners are potential field techniques (Khatib, 1985) or dynamic window approaches (Fox et al., 1997).

Sampling-based path planning algorithms can be used both as local planners or global planners, depending on the design and the available time. Examples of those algorithms are Probabilistic Roadmaps (PRM) (Kavraki et al., 1996) and Rapidly-exploring Random Trees (RRT) (LaValle and Kuffner, 1999) which have been very successful in recent years for solving problems from robotics, manufacturing, and biological applications.

Next, recent navigation strategies which have been proposed to move in dynamic environments are listed:

- **Navigation using patterns of motion**

Bennewitz et al. (2005) applied A* algorithm to determine the minimum-cost path in the time-space configuration of a robot. The environment is represented using an occupancy grid map. The cost for traversing a cell is proportional to its occupancy probability. The robot's belief about future movements of detected dynamic obstacles is also reflected in the cost by taking into account the probability

4. INCLUDING SOCIAL CONVENTIONS IN ROBOT NAVIGATION

that one of the persons covers cell $\langle x, y \rangle$ at time t .

- **Navigation supported by smart buildings**

Fusing mobile robots and sensor networks, to solve the problem of efficient navigation in dynamic human-centered indoor environments has been proposed in (Steinhaus et al., 2007). They proposed a navigation architecture that includes:

- use of a heterogeneous distributed sensor network consisting of stationary and mobile components (scalability, completeness);
- stationary sensors build up a homogeneous color camera network;
- use of 3D laser range sensors as mobile sensors;
- distributed sensing, environment modeling, data fusion and, path planning;
- data fusion to 3D environment models;
- hierarchical real-time path planning and path adaptation approach.

An initial path is determined with help of a wave propagation algorithm. The generated path is permanently modified and adapted due to moving obstacles by using elastic band method.

- **Navigation with lattices.**

The state lattice is a construct that reformulates the nonholonomic motion planning problem into a graph search and may reduce the search space in an intelligent way so as to retain compliance with vehicle kinematic and dynamic constraints. In (Rufli and Siegwart, 2009) a navigation scheme is proposed based on multi-resolution state lattices in four dimensions: 2D position, heading, and velocity. The resolution of lattices is adapted depending on environmental and task characteristics. After the construction of a graph of possible edges, D* search is used to find the plan. As the robot can only reach a finite set of states via a limited number of transitions, a time-viable heuristic is created off-line and used on-line by the search algorithm. The total cost of motion between two states is composed by a combination of the traversal time and the collision risk with dynamic obstacles. Main drawback is that with complex scenarios the re-planning time does not allow real time execution.

In the similar approach presented by (Kushleyev and Likhachev, 2009) the authors proposed a time-bounded lattice which merges together dynamically-feasible six dimensional planning (2D position, heading, angular velocity, linear velocity, time) and fast kinematic planning in 2D. The key idea is that when time passes

4. INCLUDING SOCIAL CONVENTIONS IN ROBOT NAVIGATION

the uncertainty in the prediction of dynamic obstacles motion grows until become useless, then a bound in time is fixed. Weighted A* is used to search the solution. Representation of any particular moving obstacle is a small set of predicted time-parameterized trajectories, each associated with a confidence and a continuous uncertainty distribution.

Authors reported a fast re-planning time, tens of milliseconds with thirty obstacles. More recent work has been proposed to address the loss of optimality when time is added, in (Phillips and Likhachev, 2011) a method is proposed that exploits the idea of safe intervals, such method offers theoretical guarantees on optimality. Still the scalability to larger environments needs to be improved.

- **Navigation using POMDPs.**

A Markov Decision Process (MDP) permits to model problems where an action must be chosen among a number of different alternatives. Main idea is that the choice must reflect not only the immediate reward but also the future consequences. The rewards for each action are expressed in terms of probabilities. A Partially Observable MDP (POMDP) is a MDP, the difference is that the current state of the process cannot be observed directly, instead the current observation gives a hint about what state it is in. The observations can be probabilistic as well.

In (Foka and Trahanias, 2010) A POMDP model is utilized as a unified model for robot navigation which incorporates modules for localization, planning and obstacle avoidance. The approach uses future short-term and long-term motion prediction of humans to decide what is the best approach the robot should employ and if the robot should increase or decrease its speed to avoid an obstacle more effectively. Reward function is implemented as two reward grid maps: one static and one dynamic. Static one keeps information about distance to the goal and about static obstacles. Dynamic one is based on prediction of a human's final destination position. Their approach of long-term prediction is based on the so called *hot-points*, i.e., points where people would have interest in visiting them. It has been implemented in a robot which is capable of avoid humans by executing a detour, changing completely the planned path or increasing/decreasing its speed.

- **Navigation with RRTs.**

Sampling-based algorithms have the advantage that they are able to find a feasible motion plan relatively quickly (when a feasible plan exists), even in high-dimensional state spaces. Furthermore, the RRT, in particular, effectively han-

4. INCLUDING SOCIAL CONVENTIONS IN ROBOT NAVIGATION

dles systems with differential constraints. These characteristics make the RRT a practical algorithm for motion planning. CC-RRT approach presented in (Luders et al., 2011) uses chance constraints to evaluate the risk of constraint violation at each time-step, embedding uncertainty directly within the planner. CC-RRT grows a tree of state distributions which are known to satisfy an upper bound on probability of collision. To include the ability of safety avoidance of dynamic obstacles, a trajectory prediction based on Gaussian Processes and RRTs is combined with the CC-RRT.

Other extensions to RRT framework are suitable to implement strategies of navigation in dynamic environments, for example, one recent approach (Jaillet et al., 2010) combines RRT algorithm with transition tests based on the mechanical work criterion to measure path quality in a space that is mapped by a given cost function, this method has been already applied to Human Robot Interaction problem in Mainprice et al. (2011). Similar algorithm, called RRT*, was presented in (Karaman et al., 2011), which implements committed trajectories and branch-and-bound tree adaptation with the aim of making it more efficient in the use of computation time on-line. Resulting approach is presented as an anytime algorithm suitable for real-time implementation.

Hybrid architectures (global + local) seems to be more promising dealing with autonomous navigation in dynamic environments problem. The addition of time in the motion planning is also critical as pointed out in the reviewed works. Very often researchers face the dilemma of responsiveness/completeness that is posed in the presence of runtime variance and unpredictable environments. Regarding this subject the work reported in (Hauser, 2012) presented an adaptive time-stepping technique based on the observation that running time in sample-based planners is variable across runs on a single query, and can vary by orders of magnitude with the width of narrow passages in the feasible space.

4.3 Optimization-based navigation with Information Process Space and Personal Space conventions

The objective of this section is to present a socially-aware strategy to safely move a robot in an unknown and complex environment where people are moving and inter-

4. INCLUDING SOCIAL CONVENTIONS IN ROBOT NAVIGATION

acting. The robot, by using only its sensor data, must navigate respecting humans' comfort. To obtain good results in such a dynamic environment, a prediction on humans' movement is also crucial, in this approach an implicit strategy is used. Taking into account the aforementioned problems a suitable cost function has been introduced. In this approach we formulate the problem of socially-aware robot navigation as an optimization problem where the objective function includes, in addition to the distance to goal, information about comfort of present humans. A new stochastic and adaptive optimization method was used, the CAO methodology. CAO is able to efficiently handle optimization problems for which an analytical form of the function to be optimized is unknown, but the function is available for measurements at each iteration. As a result, it perfectly suits the problem of respecting psychological comfort whose equation would be very difficult to get analytically (see section 2.2). Additionally, the proposed method can easily incorporate any dynamical and environmental constraints.

This work has been developed in collaboration with other members of our team and was presented in (Rios-Martinez et al., 2012). The theoretic adaptation of the CAO method to the problem of robot navigation in human environments was done by Renzaglia (2012) and was presented as an application on his PhD thesis. The main ideas, discomfort model formulation, software integration and simulations results were developed as part of this thesis.

Next sections explain the details of the approach. To validate the performance of the proposed solution, several simulation results are provided at section 4.3.3.

4.3.1 Environment and robot model

The method does not depend on the model of the environment but it is assumed that it is possible to know the position and direction of all the pedestrians in the scene and that a collision with the objects in environment can be correctly determined in execution time. As the algorithm was first thought for home environments the size of the scenario is not big and not cluttered.

The kinematic and dynamic properties of the robot are not taken into account, it is assumed that the robot is capable to reach all the positions calculated by the planner. To the algorithm the robot is a circular one whose center is passed as the position of the robot. Size of the robot is taken into account in the collision step by inflating obstacles and social spaces.

4. INCLUDING SOCIAL CONVENTIONS IN ROBOT NAVIGATION

4.3.2 Navigation strategy

The navigation strategy formulated is based on the Cognitive-based Adaptive Optimization (CAO) approach chosen by its particular characteristics explained below. CAO, recently proposed (Kosmatopoulos, 2009; Kosmatopoulos and Kouvelas, 2009), is a new stochastic optimization algorithm very useful if the analytical expression of the function to optimize is unknown but numerical values are available for any state configuration. That kind of situations are very common in robotics and the power of the CAO algorithm to handle such a problem is already shown in (Renzaglia et al., 2010), (Renzaglia et al., 2011) for multi-robot cooperative coverage. The CAO approach extends the popular Simultaneous Perturbation Stochastic Approximation (SPSA) algorithm (Spall, 1992).

The difference between the SPSA and the CAO approach is that SPSA employs an approximation of the gradient of an appropriate cost function using only the most recent experiments, while the CAO approach employs linear-in-the-parameters approximators that incorporate information of a user specified time window of the past experiments together with the concept of stochastic candidate perturbations for efficiently optimizing the unknown function.

Using this method it was possible also to obtain an indirect prediction on the people movement, which is a very crucial point to get good results for a similar task. Next, the formulation of CAO is presented.

4.3.2.1 Cognitive-based Adaptive Optimization formulation, (Renzaglia, 2012)

Let us suppose to have an optimization function depending on a set of variables $x_k^{(1)}, \dots, x_k^{(M)}$ (e.g., the robot's positions):

$$J_k = \mathcal{J} \left(x_k^{(1)}, \dots, x_k^{(M)} \right) \quad (4.1)$$

where $k = 0, 1, 2, \dots$ denotes the time-index, M the state's dimension, J_k the numerical value of the optimization function at the k -th time-step and \mathcal{J} is a nonlinear function which depends, apart from the explicit variables, on the particular environment where the robot lives. Due to a lack of information, like for example particular environment characteristics or a very complex nature of phenomenon, the explicit form of the function \mathcal{J} is not known in most practical situations; as a result, standard optimization algorithms (e.g. steepest descent) are not applicable to the problem in hand. However,

4. INCLUDING SOCIAL CONVENTIONS IN ROBOT NAVIGATION

in most practical cases the current value can be estimated, e.g. from the robot's sensor measurements. In other words, at each time-step k , an estimate of J_k is available through sensor measurements,

$$J_k^n = \mathcal{J} \left(x_k^{(1)}, \dots, x_k^{(M)} \right) + \xi_k \quad (4.2)$$

where J_k^n denotes the estimate of J_k and ξ_k denotes the noise introduced in the estimation of J_k due to the presence of noise in the robot's sensors.

Apart from the problem of dealing with a criterion for which only a sensor-based information is available, an efficient algorithm for real applications has additionally to deal with the problem of restricting the state variables so that obstacle avoidance as well as dynamical constraints are met. In other words, at each time-instant k , the vectors $x_k^{(i)}, i = 1, \dots, M$ should satisfy a set of constraints which, in general, can be represented as follows:

$$\mathcal{C} \left(x_k^{(1)}, \dots, x_k^{(M)} \right) \leq 0 \quad (4.3)$$

where \mathcal{C} is a set of nonlinear functions of the state variables. As in the case of \mathcal{J} , the function \mathcal{C} depends on the particular environment characteristics (e.g. location of obstacles, terrain morphology) and an explicit form may be not known in many practical situations; however, it is natural to assume that during the task is possible to get information whether a particular selection of state variables satisfies or violates the set of constraints (4.3).

Hence, the optimization problem can be described as the problem of changing $x_k^{(1)}, \dots, x_k^{(M)}$ to a set that solves the following constrained optimization problem: maximize (4.1) subject to (4.3). As already noticed, the difficulty in solving in real-time and in real-life situations this constrained optimization problem lies in the fact that explicit forms for the functions \mathcal{J} and \mathcal{C} are not available.

Algorithm operation. As first step, the CAO approach makes use of function approximators for the estimation of the unknown objective function \mathcal{J} at each time-instant k according to

$$\hat{J}_k \left(x_k^{(1)}, \dots, x_k^{(M)} \right) = \vartheta_k^T \phi \left(x_k^{(1)}, \dots, x_k^{(M)} \right). \quad (4.4)$$

Here \hat{J}_k denotes the approximation/estimation of \mathcal{J} generated at the k -th time-step, ϕ denotes the nonlinear vector of L regressor terms, ϑ_k denotes the vector of parameter estimates calculated at the k -th time-instant and L is a positive user-defined integer denoting the size of the function approximator (4.4). The parameter estimation vector

4. INCLUDING SOCIAL CONVENTIONS IN ROBOT NAVIGATION

ϑ_k is calculated according to

$$\vartheta_k = \underset{\vartheta}{\operatorname{argmin}} \frac{1}{2} \sum_{\ell=\ell_k}^{k-1} \left(J_\ell^n - \vartheta^\tau \phi \left(x_\ell^{(1)}, \dots, x_\ell^{(M)} \right) \right)^2 \quad (4.5)$$

where $\ell_k = \max\{0, k - L - T_h\}$ with T_h being a user-defined nonnegative integer. Standard least-squares optimization algorithms can be used for the solution of (4.5).

In order for the proposed methodology to guarantee with efficient performance, special attention has to be paid in the selection of the regressor vector ϕ . The particular choice adopted in this paper is described in section 4.3.2.2.

As soon as the estimator \hat{J}_k is constructed according to (4.4), (4.5), the set of new state variables is selected as follows: firstly, a set of N candidate state variables is constructed according to

$$x_k^{i,j} = x_k^{(i)} + \alpha_k \zeta_k^{i,j}, \quad i \in \{1, \dots, M\}, j \in \{1, \dots, N\}, \quad (4.6)$$

where $\zeta_k^{i,j}$ is a zero-mean, unity-variance random vector with dimension equal to the dimension of $x_k^{(i)}$ and α_k is a positive real sequence which satisfies the conditions:

$$\lim_{k \rightarrow \infty} \alpha_k = 0, \quad \sum_{k=1}^{\infty} \alpha_k = \infty, \quad \sum_{k=1}^{\infty} \alpha_k^2 < \infty. \quad (4.7)$$

Among all N candidate new variables $x_k^{1,j}, \dots, x_k^{M,j}$, the ones that correspond to non-feasible variables, i.e. the ones that violate the constraints (4.3), are neglected and then the new state is calculated as follows:

$$\begin{aligned} \left[x_{k+1}^{(1)}, \dots, x_{k+1}^{(M)} \right] = & \underset{j \in \{1, \dots, N\}}{\operatorname{argmin}} \quad \hat{J}_k \left(x_k^{1,j}, \dots, x_k^{M,j} \right) \\ & x_k^{i,j} \text{ not neglected} \end{aligned}$$

The idea behind the above logic is simple: at each time-instant a set of many candidate new state variables is stochastically generated and the candidate, among the ones that provide with a feasible solution, that provides the “best” estimated value \hat{J}_k of the optimization function is selected as the new set of state variables. The random choice for the candidates is essential and crucial for the efficiency of the algorithm, as such a choice guarantees that \hat{J}_k is a reliable and accurate estimate for the unknown function \mathcal{J} ; see (Kosmatopoulos, 2009; Kosmatopoulos and Kouvelas, 2009) for more details. On the other hand, the choice of a slowly decaying sequence α_k , a typical choice of adaptive

4. INCLUDING SOCIAL CONVENTIONS IN ROBOT NAVIGATION

gains in stochastic optimization algorithms, is essential for filtering out the effects of the noise term ξ_k . The next theorem summarizes the properties of the CAO algorithm described above:

Theorem 1 *Let $x^{(1^*)}, \dots, x^{(M^*)}$ denote any local minimum of the constrained optimization problem. Assume also that the functions \mathcal{J}, \mathcal{C} are either continuous or discontinuous with a finite number of discontinuities. Then, the CAO algorithm as described above guarantees that the state $x_k^{(1)}, \dots, x_k^{(M)}$ will converge to one of the local minima $x^{(1^*)}, \dots, x^{(M^*)}$ with probability 1, provided that the size L of the regressor vector ϕ is larger than a lower bound \bar{L} .*

The proof of this theorem, not presented here for brevity purposes, is among the same lines as the main results of (Kosmatopoulos, 2009; Kosmatopoulos and Kouvelas, 2009); the main difference is that while in that case it is established that the CAO algorithm is approximately a *gradient-descent* algorithm, the CAO algorithm used in this paper is proven to be approximately a *projected gradient-descent* algorithm.

As already noticed in precedent section, the CAO algorithm requires only a local approximation of the unknown function \mathcal{J} and as a result the lower bound \bar{L} has not to be large (as opposed to methods that construct a global approximation of the function \mathcal{J}). Although, there exist no theoretical results for providing the lower bound \bar{L} for the size of the regressor vector ϕ , practical investigations on many different problems (even in cases where the dimension of the variables to be optimized is as high as 500; see (Kosmatopoulos, 2009)-(Kosmatopoulos and Kouvelas, 2009) for more details) indicate that for the choice of the regressor vectors such a bound is around 20.

4.3.2.2 Social conventions

In this section it is formulated the problem of social robot navigation and it is shown how the proposed optimization algorithm can be applied in practice. Fig. 4.1 presents the general idea of the proposal:

First, the robot position, people position and goal position are passed to CAO, the algorithm creates local stochastic perturbations and choose the best value for the optimization function at each iteration. Minimization is done in both distance to goal and discomfort until reaching the goal.

Our intent is to safely move a robot in a complex and unknown environment respecting the comfort of the people moving in. Let $x_0^{(R)}$ be the robot start position and let $x^{(G)}$ be the goal position. Our intent is to move the robot from $x_0^{(R)}$ to $x^{(G)}$ minimizing

4. INCLUDING SOCIAL CONVENTIONS IN ROBOT NAVIGATION

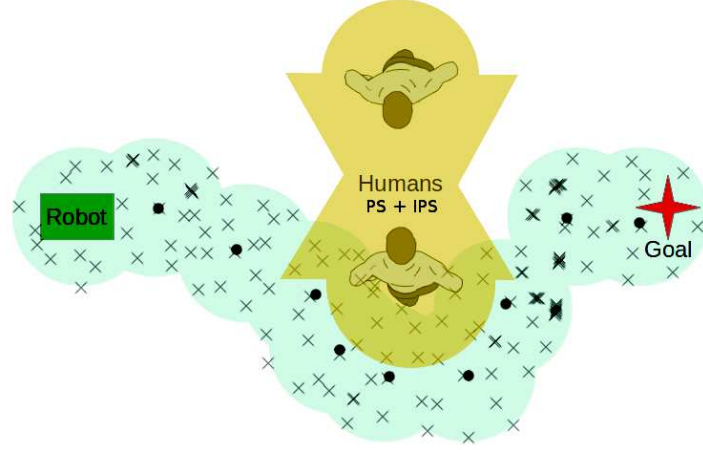


Figure 4.1: CAO generates random perturbations (crosses) at each iteration, then the “best” estimated value of the optimization function is selected (black dots) as the new robot position and the process is repeated until reaching the goal (red star). The size of the exploration area and the number of iterations can be adjusted.

the discomfort of humans located at positions $\{p^{(i)}\}$. The discomfort function has two components, one for the invasion of Personal Space ($dis(PS)$) and the other for invasion of Information Process Space ($dis(IPS)$), both of them have been explained in the chapter 3. To fulfill both the tasks of reaching the goal and respecting the people, we define the optimization function in the following way:

$$J = \lambda * (dis(PS) + dis(IPS)) + D(x^{(G)}) \quad (4.8)$$

where λ is a constant parameter and $D(x^{(G)})$ is a function depending on the distance to the goal. In our case it is the Euclidean distance.

The difference with respect to the general presentation of the algorithm, provided in section 4.3.2.1, is that now the cost function depends on both active variables (the robot’s position $x^{(R)}$) and passive variables (humans’ positions $\{p^{(i)}\}$). This means that now the cost function can be expressed in the form:

$$J = J(x^{(R)}; \{p^{(i)}\}) \quad (4.9)$$

and only the controllable components $x^{(R)}$ are perturbed to generate the candidate new positions.

4. INCLUDING SOCIAL CONVENTIONS IN ROBOT NAVIGATION

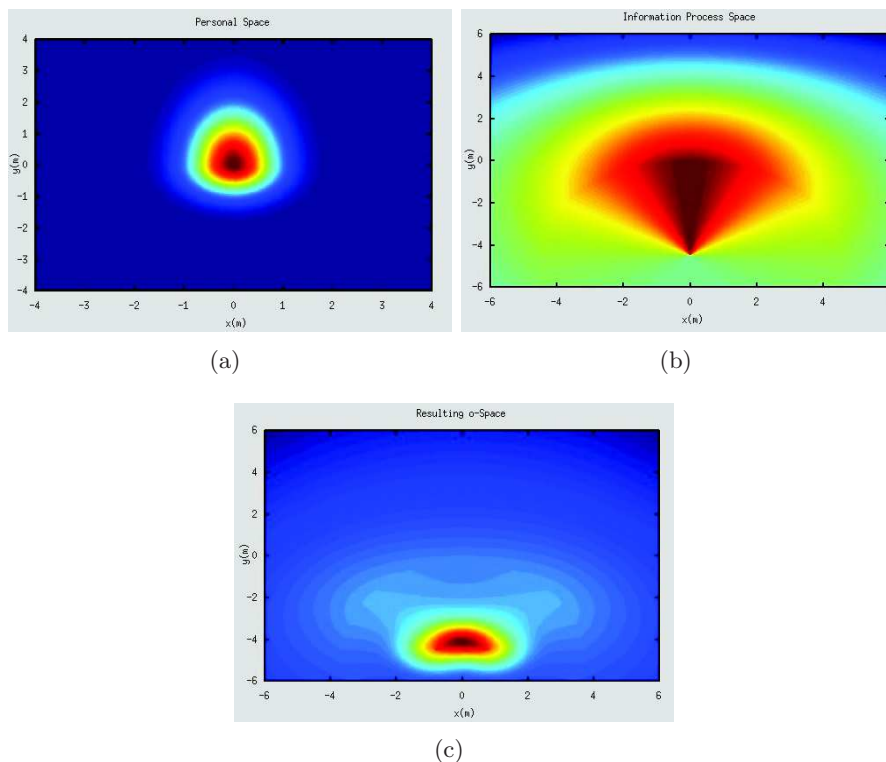


Figure 4.2: Models implemented to represent discomfort in humans' spaces: (a) Personal Space of a human in $(0,0)$ and orientation of 90 degrees. (b) IPS for a human in $(0,-4.5)$ and orientation of 90 degrees. (c) O-space for two humans in positions $(-0.85,-4.5)$ and $(0.85,-4.5)$ and orientations of 30 and 150 degrees, respectively. Higher discomfort in darker red, lower in lighter blue.

In this first approach, when two people are interacting the O-Space is automatically created by the intersection of the two IPS, as we can see in the case presented in Fig. 4.2(c). Graphics for the discomfort in our models are shown in Fig. 4.2 : the first one is the Personal Space for a pedestrian walking in the direction of y-axis, the second one the IPS for the same case and the third one shows the resulting o-space for two pedestrians in conversation. The robot must avoid the higher discomfort regions (in red) while navigating.

4.3.2.3 Movement Prediction, (Renzaglia, 2012)

As already stated, our intent is to consider a dynamic environment where the people $\{p^{(i)}\}$ are moving. The objective function is then time-dependent and in general it will

4. INCLUDING SOCIAL CONVENTIONS IN ROBOT NAVIGATION

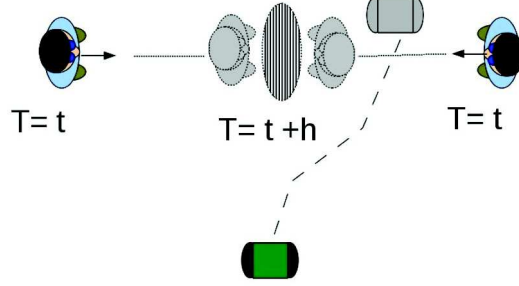


Figure 4.3: An example of prediction: the robot anticipates humans' movements and avoids them.

be different for each time step:

$$J_t = J(x^{(R)}; \{p_t^{(i)}\}). \quad (4.10)$$

In this case, in order to solve the optimization problem, i.e. finding the optimal next robot position, the result can be considerably improved if we consider the function J_{t+1} instead of J_t , where:

$$J_{t+1} = J(x^{(R)}; \{p_{t+1}^{(i)}\}). \quad (4.11)$$

This function is obviously unknown at time t but it could be approximated if a prediction model is available. Indeed, we can express the positions $\{p_{t+1}^{(i)}\}$ by means of a limited set of q past configurations

$$\{p_{t+1}^{(i)}\} = g(\{p_t^{(i)}\}, \dots, \{p_{t-q}^{(i)}\}) \quad (4.12)$$

where the new function g represents the prediction model. In our case we do not assume any particular model and the function g is to consider completely unknown. Hence also the function

$$J_{t+1} = J(x^{(R)}; g(\{p_t^{(i)}\}, \dots, \{p_{t-q}^{(i)}\})) \quad (4.13)$$

is now unknown. The strategy to approach the problem is not to explicitly predict the humans' movement but try to directly approximate the cost function (4.13) using its available past values. To do this in practice, we construct at each time step an approximator \hat{J}_t , like in (4.4), of the unknown function J_{t+1} using the last $m > q$ numerical values of J_t such that:

$$\hat{J}_t(x_t^{(R)}; \{p_{t-1}^{(i)}\}, \dots, \{p_{t-q-1}^{(i)}\}) \approx J(x_t^{(R)}; \{p_t^{(i)}\}). \quad (4.14)$$

4. INCLUDING SOCIAL CONVENTIONS IN ROBOT NAVIGATION

In this way, using the last available set of humans' positions, we have an indirect approximation of the humans' movement prediction and we obtain

$$\hat{J}_t(x^{(R)}; \{p_t^{(i)}\}, \dots, \{p_{t-q}^{(i)}\}) \approx J_{t+1} \quad (4.15)$$

i.e., the function we want to optimize.

Once the optimization problem is defined, a fundamental point for a good behavior of the algorithm is an appropriate choice of the form of the regressor vector ϕ , introduced in equation (4.4). Several different choices for its explicit expression are admissible and, for the particular application treated in this paper, it was found that it suffices to choose the regressor vector as follows:

1. choose the size of the function approximator L to be an odd number;
2. select the first term of the regressor vector ϕ to be the constant term;
3. select randomly the next $(L - 1)/2$ terms of ϕ to be any 2nd-order terms of the form $x_a^{(i)} \cdot x_b^{(j)}$ [with $a, b \in \{1, \dots, \dim(x^{(i)})\}, i, j \in \{1, \dots, M\}$ randomly-selected positive integers];
4. select the last $(L - 1)/2$ terms of ϕ to be any 3rd-order terms of the form $x_a^{(i)} \cdot x_b^{(k)} \cdot x_c^{(j)}$ [with $a, b, c \in \{1, \dots, \dim(x^{(i)})\}, i, k, j \in \{1, \dots, M\}$ randomly-selected positive integers].

After the setting of the regressor vector ϕ and once the values of the cost function are available for measurement, it is possible to find at each time step the vector of parameter estimates θ_k and thus the approximation of the cost function \hat{J}_k . Then, another important choice in order to assure the convergence of the algorithm is the expression of α_k , defined in equation (4.6). A typical choice for such a sequence is given by

$$\alpha_k = \frac{\gamma}{(k + 1)^\eta}, \quad (4.16)$$

where γ is a positive user-defined constant and $\eta \in (0, 0.5)$.

Please note that the CAO algorithm's computational requirements are dominated by the requirement for solving the least-squares problem (4.5). As the number of free parameters in this optimization problem is L , most popular algorithms for solving least-squares problems have, in the worst case, $\mathcal{O}(L^3)$ complexity.

4. INCLUDING SOCIAL CONVENTIONS IN ROBOT NAVIGATION

4.3.3 Simulation results

In this section several scenarios are presented to show the execution of our algorithms in simulation. The first scenario is shown in Fig. 4.4: in this case five humans are present, three of them are moving and two interacting. The robot starts at (1,1) and reaches its goal while avoiding people and o-space of interaction.

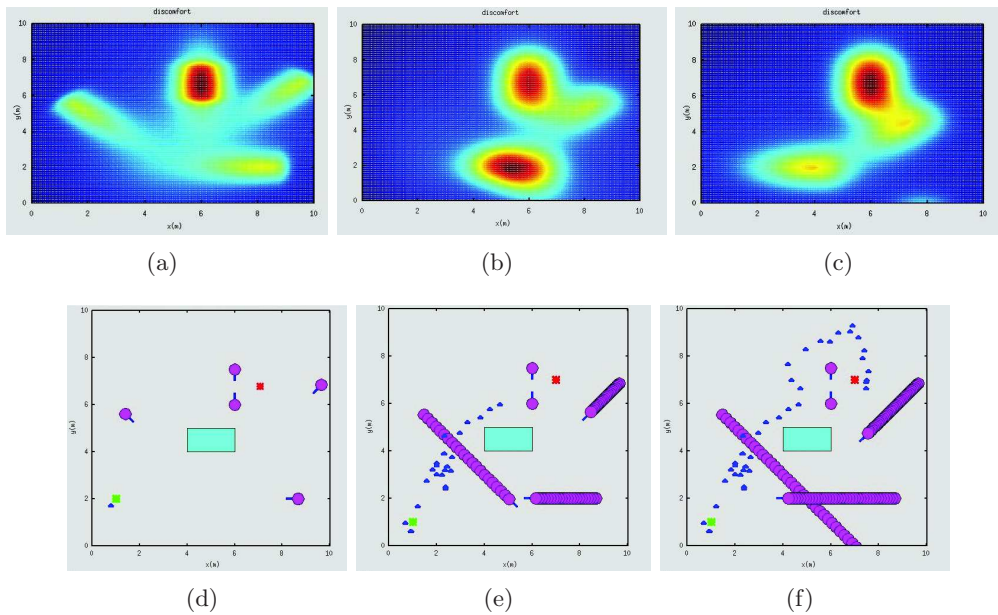


Figure 4.4: Simulation of the robot navigating in an environment populated by people at three different times. Three humans walking and two in conversation. The discomfort function is shown on the top. People are represented by circles, robot's positions by small triangles, in green and red initial and goal position respectively.

In Fig. 4.5 four different and more complex scenarios are presented. In (a) a robot has to pass through a corridor while two humans are chatting in the middle. It is possible to see how the robot is able to understand the interaction and to avoid them without disturbing. We can notice how the method evaluates many points that fall in the shortest path but finally can find a more comfortable way. In Fig. 4.5(b), the robot start position is aligned with the goal position but as one people is looking to the walls the chosen path guides the robot toward the middle of the corridor and then to the goal. We can remark that in this case, since the two people are not interacting, the robot can pass between them without trouble.

A representation of a room with people inside is exhibited in Fig. 4.5(c). Here the chosen path does not interrupt any human. Last example is shown in Fig. 4.5(d),

4. INCLUDING SOCIAL CONVENTIONS IN ROBOT NAVIGATION

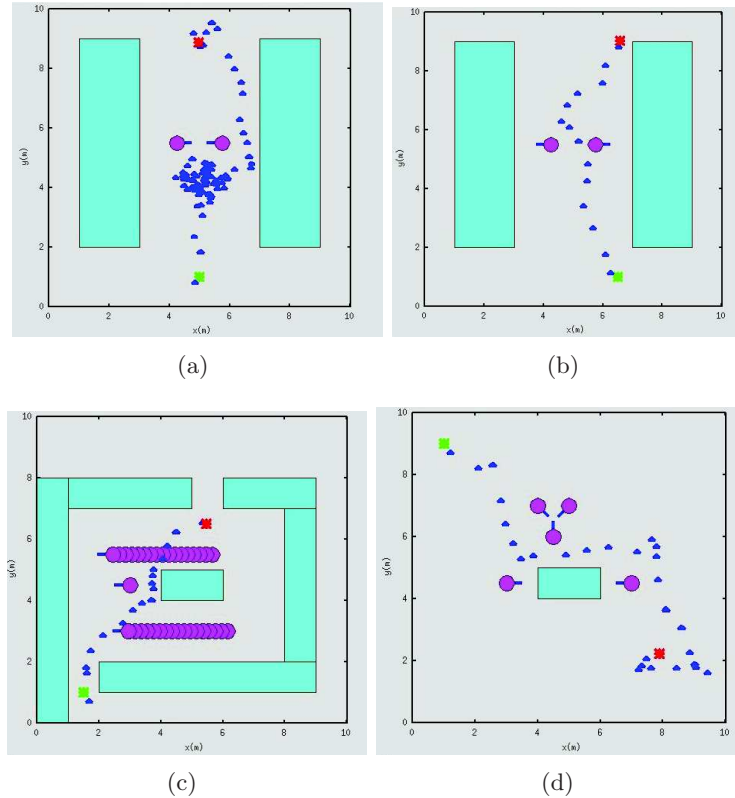


Figure 4.5: More simulations with different scenarios. Start positions are in green, goal positions in red. In (a) the robot decides to take a path that minimizes discomfort of interacting humans. In (b) a similar configuration but humans are not interacting. In (c) and (d), two different complex scenarios where the robot’s trajectories respect people comfort.

where the robot respects o-space of the group and p-space of humans. Note that in every simulation the presence of obstacles does not create any problem to the robot navigation. Additionally, the proposed algorithm, due to the random generation of next state configuration, is able to overcome many local minima generated by obstacle avoidance problems though this ability is dependent by the parameter α_k and the size of obstacles.

ROS implementation. The algorithm presented here has been ported to Robot Operating System ¹ framework with the main consequence being that now the method can be easily tested, extended and distributed for the robotics community. Fig. 4.6 show a simulation in the ROS visualizer for two scenarios. First scenario shows the

¹ROS (Robot Operating System) provides libraries and tools to help software developers create robot applications. Web: www.ros.org

4. INCLUDING SOCIAL CONVENTIONS IN ROBOT NAVIGATION

plan returned by CAO-based planner which avoids the O-Space of people interacting. Second scenario shows a plan that passes in the middle as the social interaction is not present. The test of collision with the environment for disturbances was done by using the occupancy grid map generated by ROS using a simulated laser.

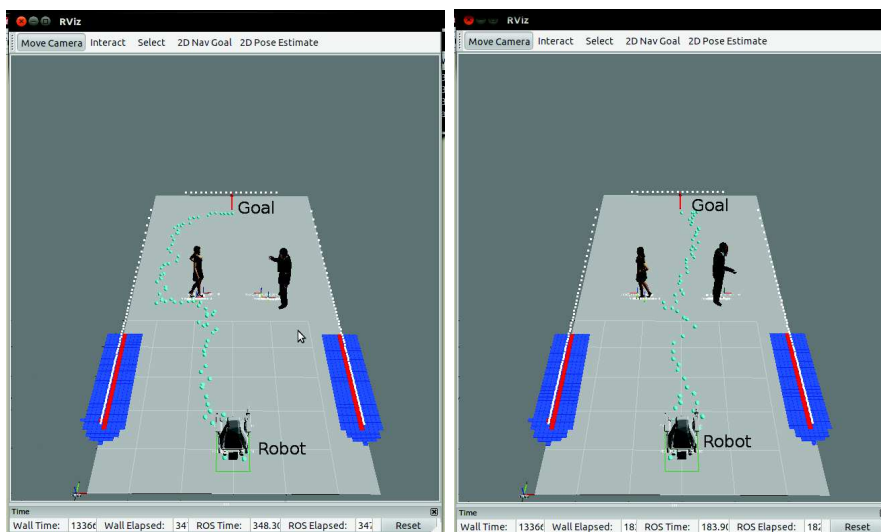


Figure 4.6: CAO-based Navigation implemented in ROS framework: a) interaction scenario, solution avoids to disturb social space. b) no interaction scenario, solution prefers to pass in the middle because it is more optimal.

Obstacle avoidance. The ability of the method to avoid obstacles can be observed in the sequence of images displayed in Fig.4.7. In a) the robot has received a goal and uses the CAO-based planner algorithm to generate a first solution which is passed to execution. A time limit is defined to re-plan, then in the consecutive images we can observe how new plans has been generated as the robot approaches its goal. The observation of obstacles was done employing a ROS module which also permits to inflate the obstacles according to the size of the robot in order to keep a safe navigation. Obstacles information is represented in the method as restrictions to the optimization method.

4. INCLUDING SOCIAL CONVENTIONS IN ROBOT NAVIGATION

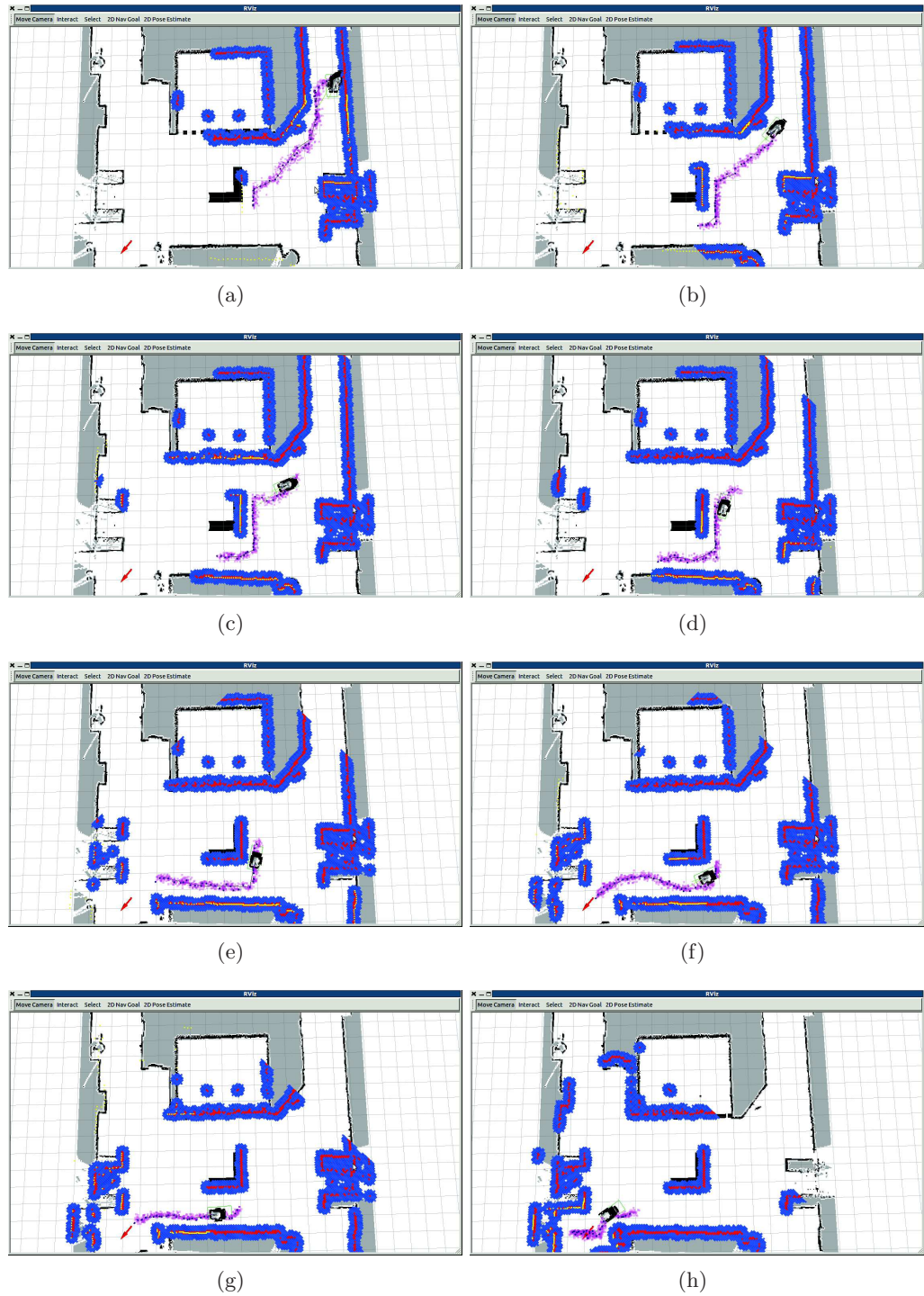


Figure 4.7: Trajectory generated by the algorithm in the case of avoiding of obstacles, in the image it can be observed the inflated obstacles. Goal is signaled with an arrow.

4. INCLUDING SOCIAL CONVENTIONS IN ROBOT NAVIGATION

Human avoidance. The model of IPS enables the robot to avoid moving humans in advance as observed in Fig. 4.8. In a) a direct path to the goal has been planned by the algorithm. In b), c) and d) as the human moves re-estimations of the plan are realized with different proposed solutions. In e) it can be seen that the only solution found was to turn. In f) a solution avoiding the human was found and the robot continues with its route to the goal.

Conclusion: We have presented a new stochastic optimization algorithm to move a robot in a complex, dynamic and unknown environment taking into account the respect of humans' comfort. In particular, the proposed approach presents the following advantages:

- It does not require any a priori map of the environment
- It can include a prediction of the humans' movement
- It can easily incorporate any kind of dynamical and environmental constraints
- The random next-state searching allows us to overcome many local minima
- Low computational complexity, allowing real time implementations

By its nature CAO-Based navigation with social conventions is best adapted to local decisions because the approximation done to the future state of the environment is limited. In this approach it was used a model based on distance to have a measure of discomfort in a sort of virtual sensor. However, the ability of the CAO to work with direct sensor information will be fully exploited when sensors able to measure comfort by means of social signals (like face expressions) be available.

Discussion: The case of conversation with people was not addressed at this point but we think it could be done by fixing a permitted distance to the center of O-space according to the distance exhibit by the other participants (to give equal access to robot). At same time the IPS would be inverted to maintain the robot in the field of view of humans as in this case discomfort has changed of meaning due to the task. Note that CAO method do not create an explicit approximation or estimation of obstacles location, humans' movement prediction and other unknown information; instead, it on-line produces a local approximation only of the unknown cost function to optimize. For this reason, they require simple, and thus scalable, approximation schemes to be employed. Taking advantage in the fact that our robotic platform (presented in the

4. INCLUDING SOCIAL CONVENTIONS IN ROBOT NAVIGATION

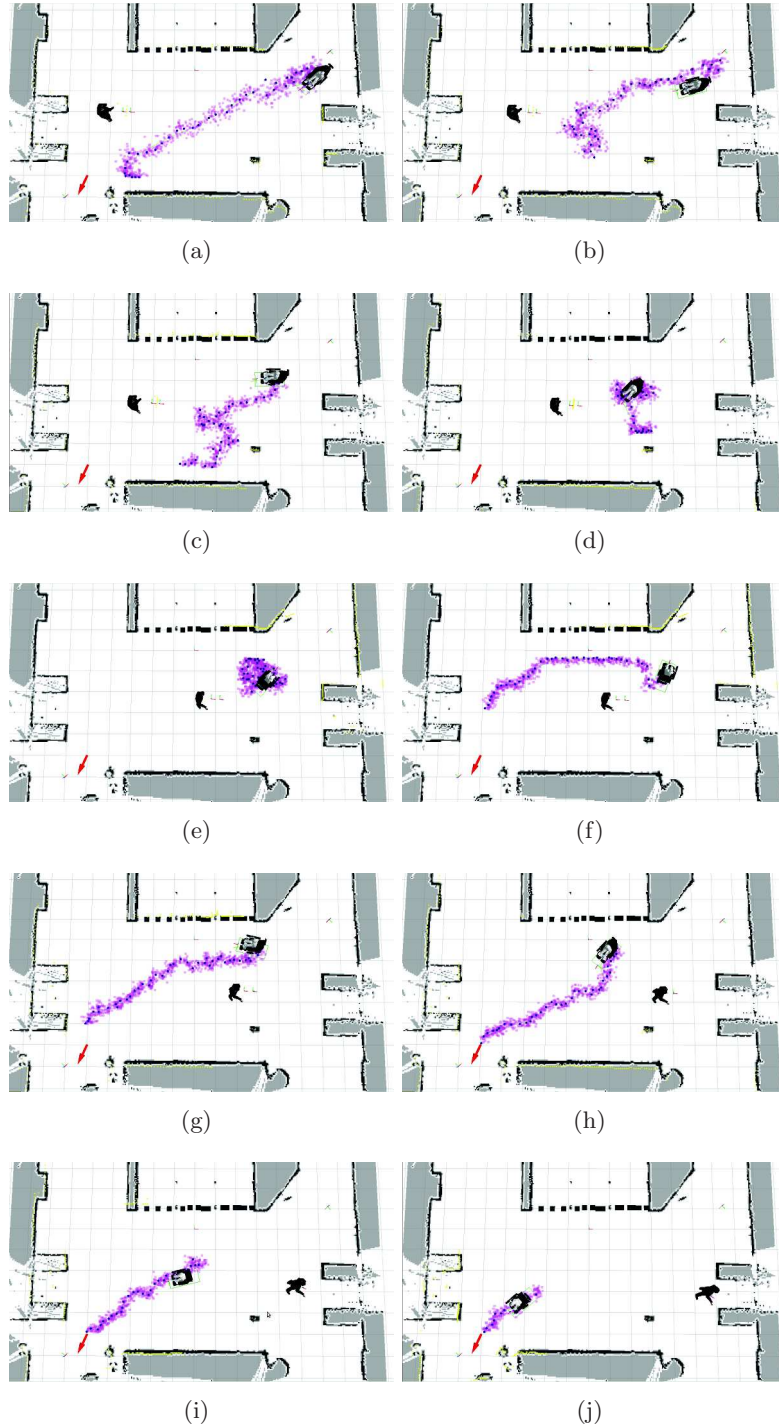


Figure 4.8: Trajectory generated by the algorithm in the case of avoiding a moving human that causes a re-estimation of robot original plan. Goal is signaled with an arrow.

4. INCLUDING SOCIAL CONVENTIONS IN ROBOT NAVIGATION

next chapter) was also working with ROS some tests have been conducted on the real platform however the control algorithm was not able to smoothly follow the navigation plans generated by this approach. Many times the algorithm produces solution faster than the robot could follow and the constant changes produced a shaky movement. This was mainly because the algorithm has no information about the dynamic constraints of the robot and also because in real-time applications it needs a very precise schedule to produce good results. The probabilistic approach was good to avoid many local minima but it was evident that a more structured strategy of exploration of the space was needed. Next section presents a different technique which is also probabilistic but with a more complete planning algorithm, including a model of the environment and a model of the robot. Moreover it has the ability to deal with events in a bigger window of time.

4.4 Risk-based navigation with Personal Space, O-Space and Activity Space conventions

Risk is a danger of loss or harm, that could be more or less predictable. When humans walk they will choose the optimal route, unless a risk in the environment push them to adapt their original trajectory (a compliant model was presented in (Moussad et al., 2011)). In this thesis, risk of collision, risk of disturbance to an individual, risk of disturbance to an interaction between human-human and human-object were considered. The order used to mention them reflects the priority assigned to those risks.

As starting point for socially-aware navigation, the strategy proposed in (Fulgenzi et al., 2010; Fulgenzi, 2009) was chosen because it explicitly uses estimation of risk as heuristic for taking decisions. That algorithm, called RiskRRT, was thought to operate in dynamic, uncertain environment. It supposes that the moving pedestrians, detected in the environment, follow typical motion patterns which can be learned by an off-board platform before navigation and be used in execution time by the robot.

RiskRRT is an extension of the Rapidly-exploring Random Tree algorithm (LaValle and Kuffner, 1999), where the likelihood of the obstacles future trajectory and the probability of collision is explicitly taken into account. The tree is grown in a random fashion but a bias is included to direct the search to the goal. Best trajectory (path in the tree) is chosen using as heuristic the “probability of success” and the distance to the goal from its nodes. Fig. 4.9 shows an example of the tree’s growth in an environment populated by humans.

4. INCLUDING SOCIAL CONVENTIONS IN ROBOT NAVIGATION

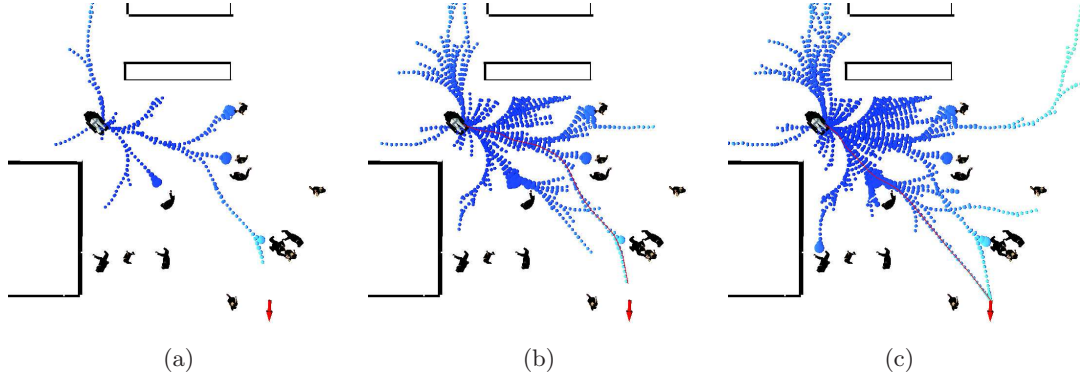


Figure 4.9: Growth of the tree in RiskRRT at three different times, the exploration is biased to the goal (red arrow), branches with big risk are not grown, best branch is chosen as solution.

In this section it is explained how our models of social conventions were combined with RiskRRT by including the knowledge of Personal Space of pedestrians, the O-Space for possible interactions between them and the Activity Space for interactions between humans and objects. The particular considered interaction was the conversation between pedestrians which was missed in the most part of related works.

4.4.1 Environment and robot model

At a given instant, the robot knowledge about the state of the world, as proposed by (Fulgenzi et al., 2010), is represented by:

1. An occupancy grid representing the structure of the static environment around the robot, according to the previous observations;
2. A list of moving objects, their estimated position, velocity and previous observations;
3. An estimation of the state of the robot (position, velocity);
4. A set of typical patterns which represent the motion models for the obstacles;
5. A set of Gaussian Processes representing the typical patterns of the dynamic obstacles;
6. A goal state.

To take into account the new social constraints it was included to the list:

4. INCLUDING SOCIAL CONVENTIONS IN ROBOT NAVIGATION

7. A model of personal space $PS(o_m)$ attached to each dynamic obstacle o_m , according to section 3.2.1
8. A list $LI = \{Z_i\}_{i=1..r}$ of interactions detected in the environment, each interaction Z_i has a model of O-Space attached to it, according to section 3.3.2. Interactions can be of type Human-Human or Human-Object.

The robot model is that of the two-wheeled differential drive whose parameters for maximum and minimum linear velocity as well as maximum and minimum angular velocity are used to generate nodes for the tree.

4.4.2 Navigation strategy

The complete strategy of navigation includes three tasks: one dedicated to perception (of static and moving obstacles), a task for planning partial but safe trajectories and a task for navigating safely along planned trajectories. The prediction done for forecasting the position of moving obstacles is based on a linear prediction for very close future and a set of learned Gaussian Processes for near future (Fulgenzi et al., 2010). In this section, Risk-RRT algorithm and the collision risk assessment including the new social constraints are presented.

4.4.2.1 The Algorithm RiskRRT, (Fulgenzi, 2009)

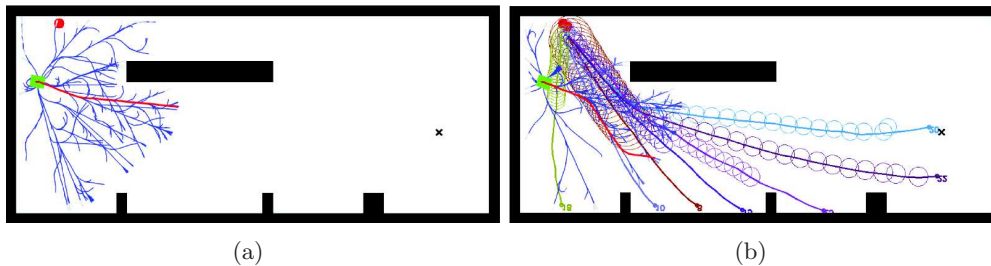


Figure 4.10: Example of execution of Risk-RRT algorithm. In a) the robot navigation system has created a tree of possible paths to follow, robot is the green rectangle, the chosen path is in red. In (b) we can observe how the robot has adapted its trajectory trying to avoid a possible collision with pedestrian (in red) by considering the predictions of typical pedestrian trajectories.

The goal oriented operation of Risk-RRT is described in Algorithm 1. The main loop is controlled by the distance to the goal, once that robot achieves its goal algorithm finishes and robot brakes.

4. INCLUDING SOCIAL CONVENTIONS IN ROBOT NAVIGATION

Algorithm 1 Risk-RRT

```
1: procedure RISK-RRT
2:    $Plan = \text{empty}$ 
3:    $Tree = \text{empty}$ 
4:    $Goal = \text{read}()$ 
5:    $t = \text{clock}()$ 
6:   while Goal not reached do
7:      $t = \text{clock}()$ 
8:      $X = \text{robot localization}$ ;
9:     delete unreachable trajectories in  $Tree$  for  $(X, t)$ 
10:     $(Map, humans, interactions) = \text{observe environment}$ 
11:    predict  $humans$  for time  $t, t + 1, \dots, t + N\tau$ 
12:    if environment different then
13:      update risk in nodes using  $(Tree, Map, humans, interactions, t)$ 
14:    end if
15:    while  $\text{clock}() < t + \text{loopRate}$  do
16:      grow branches in  $T$  with less risk and depth  $\leq N$ 
17:    end while
18:     $Plan = \text{Choose best trajectory in } T$ 
19:     $t = \text{clock}()$ 
20:    if  $Plan$  is empty then
21:      brake
22:    else
23:      put  $Plan$  available to execute
24:    end if
25:  end while
26:  brake
27: end procedure
```

The first step in the main loop is to update robot position by using a localization strategy in order to prune the tree. A new root for the tree is chosen according to robot position then nodes that have become unreachable are eliminated. After, an observation of the environment is done in order to collect information of humans, interactions and map. Such information is used to predict the future until time $t + N\tau$, where τ corresponds to the time step and N is the maximum length of branches (parameters `timeStep` and `maxDepth` in table 4.1).

If a significant change in the environment is perceived then risk is updated in the current nodes of the tree (Algorithm 1, line: 13). Once this step has finished the remaining time is used to grow the tree (Algorithm 1, line: 16) according to the standard RRT process but it is biased (bias parameter in table 4.1) to the current goal. Branches

4. INCLUDING SOCIAL CONVENTIONS IN ROBOT NAVIGATION

with less risk of collision are privileged.

Finally, the best trajectory is chosen (Algorithm 1, line: 18) to be the solution plan. The criteria to choose the best path and the candidate branches to grow are described in section 4.4.2.2.

After each planning cycle, the planned trajectory is generally just a partial trajectory which is made available to the execution algorithm. Execution and planning are done in parallel: while the robot moves along the previous planned partial path a new loop updates the environment information with the information coming from the perception algorithm, the tree is updated and grown and the new partial path is passed for execution when the time is over.

Risk-RRT takes explicitly into account the real-time constraint and limits the time available for planning to a fixed interval as expressed by the parameter `loopRate` (described in table 4.1).

In fig. 4.10 we can observe an example of navigation employing Risk-RRT in the case of one pedestrian entering in the environment and robot going to its goal. At the beginning the robot has explored the environment and then decides to follow one trajectory, some steps ahead when it detects the presence of pedestrian, a prediction is realized based in the Gaussian processes and it must adjust its previous choice to avoid a collision with the human.

RiskRRT parameters. The algorithm has many parameters that can be set in order to have a better performance in a particular environment. A list of them with their descriptions are shown in table 4.1.

4.4.2.2 Probabilistic Risk of Collision

When searching for a safe path, the algorithm must determine the amount of risk involved by taking an action $u \in U$ in configuration $q(t_1)$ at time t_1 . U is the set of feasible actions for the robot. That risk can be written as $P(\text{coll}(q(t_1), u) = 1)$, i.e, the probability of collision if u is chosen, such probability will be referred as P_c in the rest of the thesis. The risk is computed on the basis of the probability of occupation on the surface A which is swept by the robot moving from $q(t_1)$ under control u in the interval

4. INCLUDING SOCIAL CONVENTIONS IN ROBOT NAVIGATION

| Parameter | Description | Default value |
|------------|---|---------------|
| timeStep | Time step between nodes in the Tree, big values favor exploration but has an impact in the smoothness of paths. | 0.5 seconds |
| maxDepth | It is the maximum depth a branch can get when planning, big values favor static obstacles avoiding. Smaller values favor reactivity | 25 nodes |
| nv | Resolution of the discretization of the admissible velocities interval, big values favor smoothness but increments process time. | 12 |
| nphi | Resolution of the discretization of the admissible steering angles interval, big values favor smoothness but increments process time. | 20 |
| pThreshold | Threshold used to filter nodes with very risky values, a value of zero means that all nodes are accepted, a value of one means that only nodes without any risk are accepted. | 0.9 |
| loopRate | This parameter tell the system how fast a plan must be available for solution | 4hz |
| bias | It is the probability to choose the goal as point in the process of RRT exploration, a value of one | 0.01 |

Table 4.1: Important parameters for the RiskRRT algorithm

of time $[t_1, t_2]$:

$$q(t_2) = f(q(t_1), u, \tau) \quad (4.17)$$

$$A = \int_{t_1}^{t_2} q(t) dt \quad (4.18)$$

where $f(\cdot)$ is the motion model of the robot and $\tau = t_2 - t_1$ is the time step. The risk of collision must incorporate both the static and the moving obstacles. Even when two humans in conversation don't exhibit a significant motion they must be treated as dynamic ones because they represent more risk than static obstacles. The space occupied by personal space and o-space can't be detected by sensors, these spaces will be linked to the dynamic obstacles and their costs will be reflected on the corresponding probability of collision. We keep the hypothesis that moving obstacles and static obstacles cannot overlap, and consequently that collision with static obstacles and collision with

4. INCLUDING SOCIAL CONVENTIONS IN ROBOT NAVIGATION

moving obstacles are mutually exclusive events, which yields:

$$P_c = P_{cs} + (1 - P_{cs}) \cdot P_{cd} \quad (4.19)$$

$$P_{cd} = 1 - \prod_{m=1}^M [1 - P_{cd}(o_m)] \quad (4.20)$$

where P_{cs} is the probability of collision due to the static obstacles, $P_{cd}(o_m)$ is the probability of collision due to the dynamic obstacle o_m and P_{cd} is the probability of collision due to all the dynamic obstacles.

The static obstacles are represented in the occupancy grid which is assumed to be stationary. Given $\mathcal{M}(t_0)$ with $t_0 \leq t_1$ the most recent estimation of the static map and $\zeta \subset \mathcal{M}(t_0)$ the subset of cells which is the minimal approximation of surface A , the risk of collision with a static obstacle is given by the max probability over the subset ζ :

$$P_{cs} = \max_{\zeta} (P(\text{Occ}(\text{Cell}_{x,y}) = 1)) \quad (4.21)$$

where $\text{Cell}_{x,y}$ is the cell of the occupancy grid in (x, y) position. The risk of collision with a moving obstacle o_m is approximated by the probability of the area swept by the robot intercepts the one swept by the obstacle in the considered interval:

$$P_{cd}(o_m) = P(o_m(t) \cap A \neq \emptyset, \forall t \in [t_1, t_2]) \quad (4.22)$$

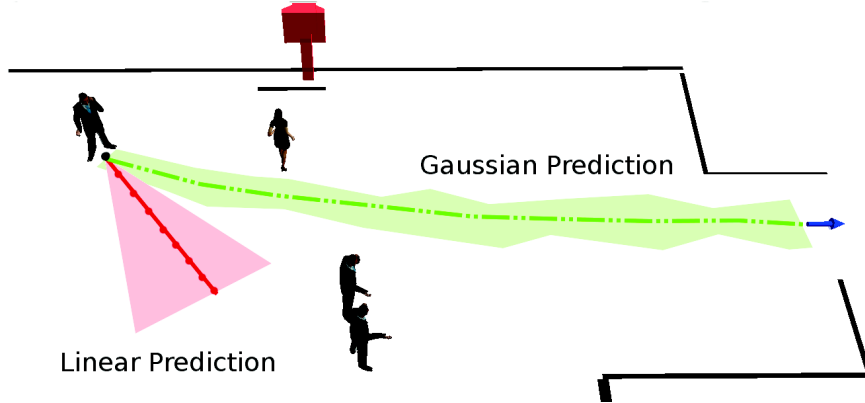


Figure 4.11: Example of prediction for a pedestrian. A linear prediction is used to estimate short-term future states and a Gaussian Processes prediction for long-term future states. Around each prediction it can be observed the area representing the variance in each case. In the case of linear prediction uncertainty grows in the future. In the other case the uncertainty has been adjusted according the learned models.

4. INCLUDING SOCIAL CONVENTIONS IN ROBOT NAVIGATION

Prediction Models. In the case of the short-term future states a simple linear prediction based on past observations of the pedestrian is used. When a set of learned typical trajectories is available it is used to get a long-term prediction $o_m(t)$ by means of a weighted sum (mixture) of Gaussian Processes. Each individual path is represented by a Gaussian Process. An example of both models is represented in the figure 4.11. A Gaussian Process is a generalization of the Gaussian probability distribution in function space, see (Meng Keat Christopher, 2009) for a more detailed explanation and equations for Gaussian Processes. From a Gaussian Process it is possible to obtain probability distribution for unobserved portions of the path as in the case of path prediction. Each Gaussian Process component k is considered separately, then all the components are summed:

$$P_{cd}(o_m, k) = \int_A G(o_m(t), \mu_k, \Sigma_k) \quad (4.23)$$

$$P_{cd}(o_m) = \sum_{k=1}^K w_{m,k} P_{cd}(o_m, k) \quad (4.24)$$

where $P_{cd}(o_m, k)$ is the probability of collision with the obstacle m moving along pattern k ; $G(o_m(t), \mu_k, \Sigma_k)$ is the Gaussian Process representing pattern k , given the observation history of object o_m . The probability is marginalized over the set of possible patterns to yield $P_{cd}(o_m)$, where $w_{m,k}$ is the weight of the k component for object m . When a set of typical paths is not available, the same equations are used with $K = 1$, $w_{m,1} = 1$.

Nodes weight. Each node has an assigned weight which is a combined measure of its risk and its distance to the goal, d_g . The original equation for the calculus of weight was having problems with big distances, because in that cases the gain on distance was always being more important than risk of disturbance. The equation then was changed by a simpler weighted-sum aggregate objective function:

$$W = 1 - WP_c + 1/d_g \quad (4.25)$$

Where WP_c stands for the worst P_c found among the ancestors of the node. New equation performs better but still it remains a problem to find the aggregate objective function that can capture desired optimal points. A deeper analysis could be done, for example, by following suggestions found in (Messac et al., 2000).

Finally, best path returned is the partial branch that finishes in the node with maximum

4. INCLUDING SOCIAL CONVENTIONS IN ROBOT NAVIGATION

W.

Next section explains in detail how the Social conventions were included in RiskRRT framework.

4.4.2.3 Social conventions

In order to include social conventions in RiskRRT, the concepts of Personal Space, O-Space and Activity Space were considered.

First, it was defined PZ_i as the probability of disturbing by passing inside the O-Space (defined in sec. 2.3.2) of interaction i , and we calculate it as:

$$PZ_i = \max_{\zeta}(\Gamma_{C_i, S_i}(Cell_{x,y})) \quad (4.26)$$

To reflect the fact of disturbing an interaction we think of it as a collision with a dynamic obstacle and modify the equation 4.20 to get:

$$P_{cd} = 1 - \prod_{m=1}^M [1 - P_{cd}(o_m)] \prod_{i=1}^r [1 - PZ_i] \quad (4.27)$$

In the case of the personal space we define P_{ps} as the probability of disturbing by passing in the personal space of the human o_m . We can approximate P_{ps} as the probability that A , the area swept by the robot, intercepts the one represented by the personal space:

$$P_{ps}(o_m, k) = \int_A PS(o_m(t)) \quad (4.28)$$

Where $PS(o_m(t))$ is the model of personal space centered in $o_m(t)$ as described in section 2.3.1. Again, to take into account this last constraint the original equation 4.24 was modified to get:

$$P_{cd}(o_m) = \sum_{k=1}^K w_{mk} [P_{cd}(o_m, k) + P_{ps}(o_m, k)(1 - P_{cd}(o_m, k))] \quad (4.29)$$

After those extensions the risk calculated for every partial path is given by the risk of collision along the path and the risk of disturbance to a Personal Space, an o-Space or an Activity Space. It is important to mention that the risk of disturbance to an Activity Space is calculated exactly in the same way that in the case of O-Space, the difference is in the way the Social Filter computes the parameters of the models.

4. INCLUDING SOCIAL CONVENTIONS IN ROBOT NAVIGATION

4.4.3 Simulation results

A number of tests have been performed to assess the effect of including prediction in our motion planning algorithm.

Case 1. Fig. 4.12 shows one iteration of the navigation main loop. As it can be seen, the resulting trajectory differs from optimal trajectories obtained by traditional planning algorithms, the robot actually opts for a larger trajectory that avoids obstructing the moving pedestrians.

In all our simulations the speed of pedestrians has been fixed to one m/s¹ and maximum speed of our wheelchair is also one m/s.

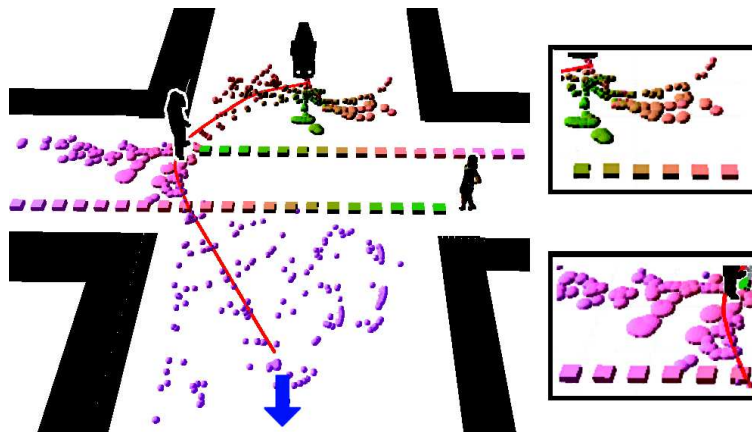


Figure 4.12: Predictive navigation example. RiskRRT selected a plan (red line) to the goal (blue arrow). The chosen path leads the robot to pass by the back of the first person, and then reduces the speed to let the second person to pass as well. With this strategy, the robot minimizes the risk of collision and the discomfort caused for the two pedestrians. Once second person has passed, the algorithm choses a straighter path to the goal. Frames at the right of the figure show that estimated risk is bigger at future positions of the wheelchair (circles) which are close to predicted positions of pedestrians (squares).

Case 2. In the fig. 4.14 planner execution at two distinct iterations can be observed. Tree shows the portion of the environment already explored. Nodes are represented by colored spheres, their color represents the time at which they would be reached by the robot, same interpretation is done for color in predicted trajectories of humans.

¹field study ((Knoblauch et al., 1996)) suggested that values of 1.25m/s for younger pedestrians and 0.97m/s for older pedestrians are appropriate

4. INCLUDING SOCIAL CONVENTIONS IN ROBOT NAVIGATION

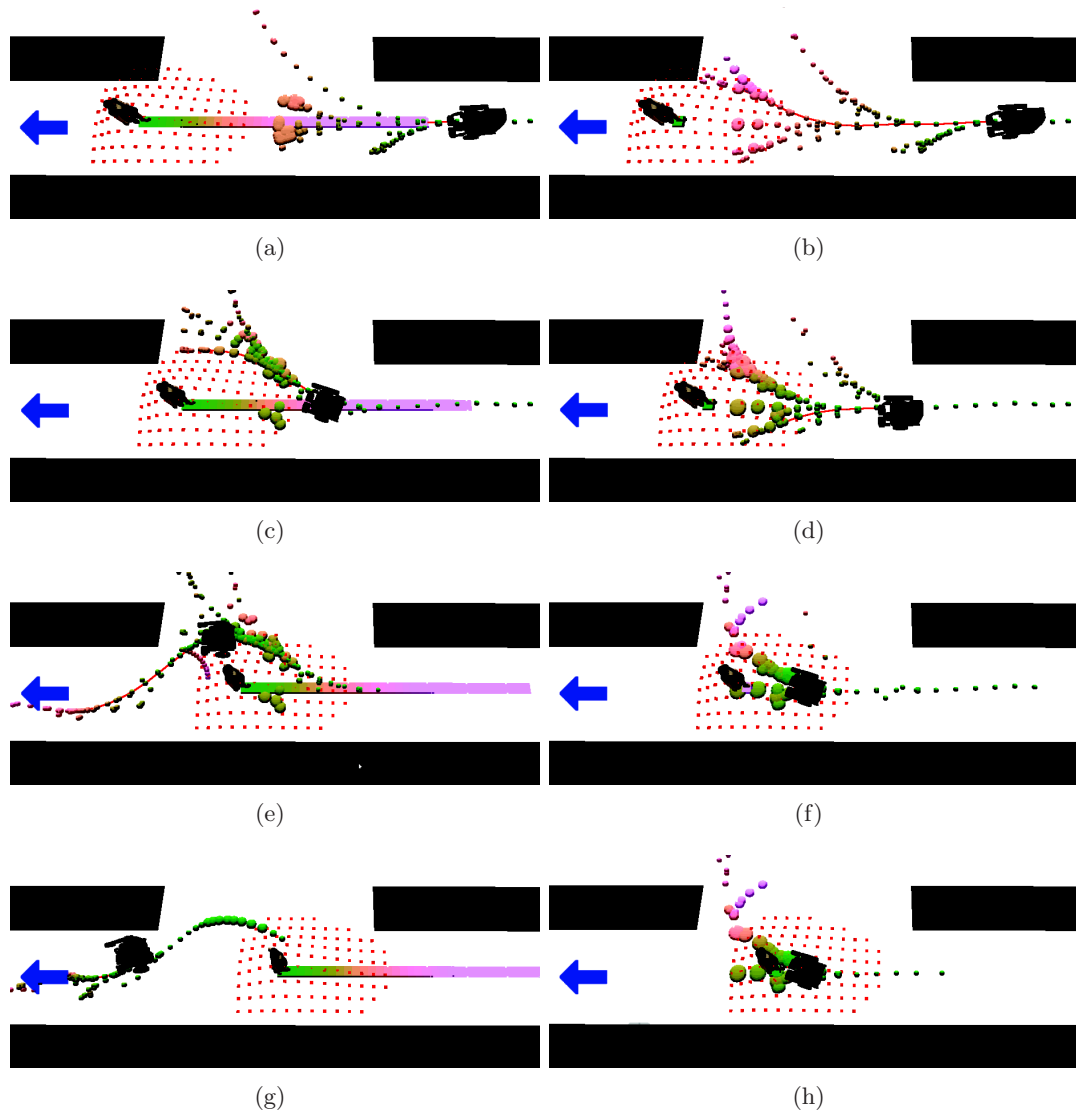


Figure 4.13: Qualitative comparison of predictive navigation (first column) vs non-predictive navigation (second column). Prediction helps to discover future high-risk states (a) and anticipate avoidance paths to finally reach the goal (g). Without prediction avoidance begins too late (f) and the disturbance is unavoidable (h).

Size of nodes represents the estimated risk, therefore a node very close to a predicted pedestrian position of the same color will have a big size. Robotic wheelchair (R) has been asked to go to goal marked with a blue arrow. In a) partial path solution was found (red line) avoiding high risk zones identified in the image by the bigger size of nodes. Wheelchair is deliberately not moving then, after some instants, new observations and

4. INCLUDING SOCIAL CONVENTIONS IN ROBOT NAVIGATION

predictions produce different risk estimation and permits to select a better path.

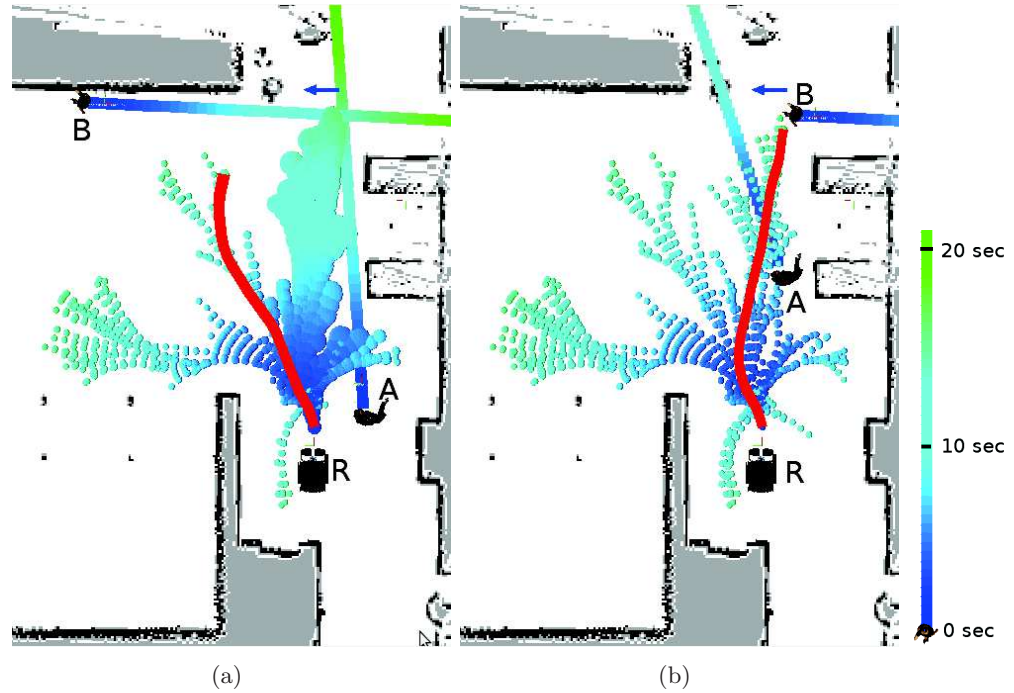


Figure 4.14: Example of execution for the algorithm at two distinct iterations. One goal (blue arrow) has been passed to the wheelchair (R), while two people (A,B) are walking around. In a) partial path solution was found (red line) avoiding high risk zones identified in the image by the bigger size of nodes. Wheelchair is deliberately not moving then, after some instants, new observations and predictions produce different risk estimation and permits to select a better path.

Case 3. Fig. 4.13 compares the paths that were obtained using predictions of pedestrian movements (left column) with those obtained without predictions (right column). The robot's initial position is on the left end of the corridor while the goal is at right end. Since the corridor is narrow, the only way to avoid colliding or disturbing the pedestrian is by moving aside in the corridor opening before continuing to the goal. In the figure, it is possible to see how, by using predictions, the wheelchair is capable to detect a possible collision in the middle of the corridor (4.13 a)) and to choose an alternative path to reach the goal. Without prediction it takes a straight path to reach the goal which, at first does not seem to pose any risk (4.13 b)) later, when the wheelchair detects the future collision (4.13 d)) it stops but it has already invaded the Personal Space in front of the human.

4. INCLUDING SOCIAL CONVENTIONS IN ROBOT NAVIGATION

Case 4. To test our models of interaction we have chosen a scenario that shows one conversation between two humans standing. The pedestrians are placed in a Vis-a-Vis F-formation, that is, facing each other in their social zone (sec. 2.3.1). The space between them is big enough to let the robot passing. Detection of conversations is proportioned by the Social Filter module.

Thirty executions of the planner were conducted in very similar conditions, as it can be seen in Fig. 4.15, when the Social Filter is off, the plans do avoid people but do not respect social space. When the Social Filter is turned on again, all the plans managed to respect interaction space without disturbing the involved people.

Case 5. In this scenario, shown in fig. 4.16, it was represented a typical scene of an environment populated by humans, it can be observed conversation between two or more participants, humans standing in a line and also humans walking. The system is able to recognize interactions and social spaces. The robotic wheelchair must reach a goal located many meters away. The socially-aware navigation based on RiskRRT output plans to guide the wheelchair until its goal. Close to the goal a conflicting situation appears where the robot must decide how to pass close to two humans. In the first case the prediction helps the robot to estimate that it is able to pass in front of a human without entering to her Personal Space. In the second case the decision was to wait until the human has passed to continue to the goal.

Discussion. In a dynamic environment it is not enough detecting interactions because it could be too late to take a decision, we need to predict when and where an interaction will take place. Some studies (Kelley et al., 2008) include the analysis of interaction between humans to get information that could be useful for robots to mimic that behavior, future work must be focused on adding a technique for predicting better the apparition of an O-Space in the path of the robot.

4.5 Conclusions

The present chapter has presented our proposals to socially-aware robot navigation in dynamic environments by integrating social conventions in the planning loop. A review on the state of art techniques on robot navigation in dynamic environments was exposed. Also, two techniques were proposed: one considering optimization-based navigation and the other Risk-based navigation.

The optimization-based navigation strategy is based on the Cognitive-based Adaptive Optimization (CAO) approach applied to robots (Renzaglia et al., 2010; Renzaglia,

4. INCLUDING SOCIAL CONVENTIONS IN ROBOT NAVIGATION

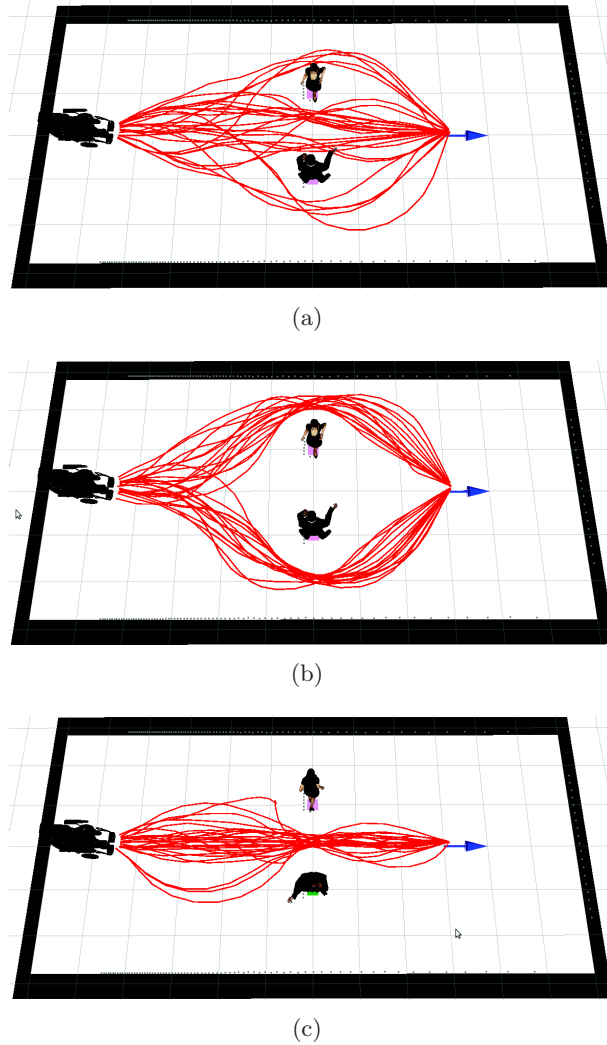


Figure 4.15: Socially-Aware navigation in RiskRRT test. Each figure shows a sample of generated plans (in red) for thirty executions of RiskRRT: a) without social filter social spaces are not respected, b) and c) with social filter, navigation is socially acceptable. In c) people are looking towards walls, therefore there is no social interacting zone, then navigation respects only their personal spaces.

2012). We formulate the problem of socially-aware robot navigation as an optimization problem where the objective function includes, in addition to the distance to goal, information about comfort of present humans. CAO is able to efficiently handle optimization problems for which an analytical form of the function to be optimized is unknown, but the function is available for measurements at each iteration. As a result, it perfectly suits the problem of respecting psychological comfort whose equation would

4. INCLUDING SOCIAL CONVENTIONS IN ROBOT NAVIGATION

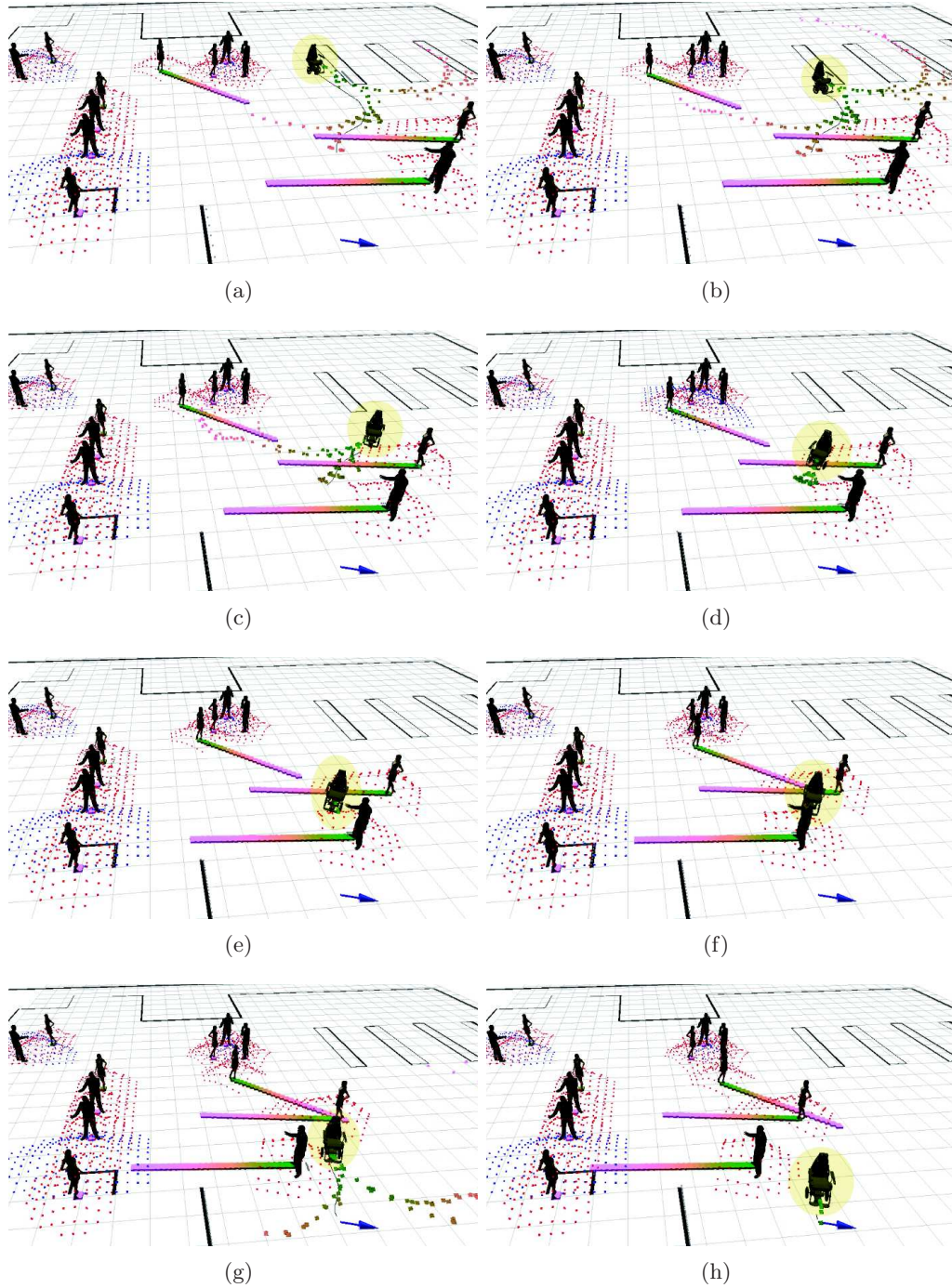


Figure 4.16: Socially-Aware navigation in dynamic environments populated by humans. A goal (blue arrow) has been set, the robotic wheelchair successfully reaches it by following plans generated by RiskRRT, while it avoids obstacles, respecting humans and social spaces.

4. INCLUDING SOCIAL CONVENTIONS IN ROBOT NAVIGATION

be very difficult to get analytically (see section 2.2). The model of Information Process Space developed in this thesis was integrated in order to work as a “virtual” sensor providing comfort measures. Simulation results in different scenarios with moving humans were presented.

Our models of social conventions were combined with RiskRRT (Fulgenzi, 2009; Fulgenzi et al., 2010) by including the knowledge of Personal Space of pedestrians, the O-Space for possible interactions between them and the Activity Space for interactions between humans and objects. The particular considered interaction was the conversation between pedestrians which was missed in the most part of related works. The approach presented shows a way to take into account social conventions in navigation strategies providing the robot with the ability to respect the social spaces in its environment when moving safely towards a given goal. Due to the inclusion of our social models, the risk calculated for every partial path produced by RiskRRT algorithm is given by the risk of collision along the path and the risk of disturbance to a Personal Space, an o-Space or an Activity Space.

Chapter 5

Socially-aware navigation for assisted mobility

5.1 Introduction

In contrast with static or controlled environments where global path planning approaches are suitable, highly dynamic environments present many difficult issues: the detection and tracking of the moving obstacles, the prediction of the future state of the world and the on-line motion planning and navigation. The decision about motion must be related with the on-line perception of the world, and take into account all the sources of uncertainty involved. In the last few years, the problem of incomplete, uncertain and changing information in the navigation problem domain has gained more and more interest in the robotic community, and probabilistic frameworks aiming to integrate and elaborate such information have been developed.

In the work presented in (Spalanzani et al., 2012) we proposed to adopt solutions where navigation decisions are based on a risk evaluation. The risk function can rely on safety but also on comfort and socially-aware navigation.

The aim is to give a robot the possibility to exploit the fact that pedestrians and vehicles usually do not move at random in the given environment but often engage in typical behaviors or motion patterns. The robot may use this information to better predict the future position of humans and adapt its behavior accordingly. The robot may also follow social conventions to be well integrated in the human-populated environment.

To develop methods for safe autonomous navigation among static and moving obstacles, it is important to consider the following points:

- The fact that the environment is **dynamic** cannot be ignored: the robot per-

5. SOCIALLY-AWARE NAVIGATION FOR ASSISTED MOBILITY

formance is influenced by obstacles moving in the environment and the robot must be able to take safe and good decisions at anytime and act promptly in the dynamic environment.

- The **uncertainty** and **incompleteness** of the information perceived by the robot is not negligible and some mean to take it into account into the decision process should be introduced;

The main differences between methods presented here and classical planning methods are:

- Finding the shortest path is not the main objective,
- Navigation decisions are based on a risk evaluation. The risk function can rely on safety but also on comfort and socially-aware navigation.

Motivation for developing assisted mobility services

Many developed countries are facing the phenomena of ageing: older population is growing faster than the total population. The percentage of older persons is projected to double worldwide by the middle of this century. The challenge for the future is “to ensure that persons everywhere are able to age with security and dignity and to continue to participate in their societies as citizens with full rights”¹. Ensuring proper living conditions for an ever growing number of elderly people is a significant challenge for many countries. The difficulty and cost of hiring and training specialized personnel has fostered research in assistive robotics as a viable alternative. In this context, an ideally suited and very relevant application is to transport people with reduced mobility. Some works, like [Harmo et al. \(2005\)](#), have shown that aids for mobility and lifting are among the most important general assistive solutions in the framework of home automation and service robots for the elderly and disabled.

A service robot is understood as a robotic system with a certain level of autonomy in performing service operations for given tasks within a specified environment and interaction with human users ([Ceccarelli, 2011](#)). An example of service where robots could have a positive impact is the transport and assistance for people with reduced mobility inside airports (See fig. 5.1).

This chapter is intended to describe how the capabilities of the socially-aware navigation approach presented in section 4.4 were integrated into a complete system architecture aimed to provide assisted mobility.

¹World Population Ageing 2009 Report in <http://www.un.org/esa/population/publications>



Figure 5.1: Transport and assistance for people with motor difficulties inside airports is a good opportunity for service robots endowed with socially-aware navigation properties.

Next section discusses some important and challenging problems which have to be faced in order to succeed in human populated environments.

5.2 Perception of human motion and activity

The necessary change in the point of view from moving obstacle to moving human is not as simple as it could seem. This thesis has supposed that a system able to detect and track reliably humans in dynamic environment is available. Also that the posture of the body can be known at execution time. However the capabilities of such a system include are still very active research areas, moreover, those activities are only the basis to reach an understanding of human behavior which is really the key point to take into account in a socially-aware navigation.

It can be deduced that perception of the space and objects done by sensors is objective (despite of uncertainty) while human abstraction of the space is very subjective, therefore it is important to keep that factor in mind in the design of social acceptable solutions for navigation.

Done the importance that the Perception component has for the complete framework we have depicted, this section briefly reviews some techniques that could be used to achieve detection, tracking, motion prediction and activity recognition which are pre-

5. SOCIALLY-AWARE NAVIGATION FOR ASSISTED MOBILITY

vious required stages in order to reach automatic understanding of human behavior.

5.2.1 Detection techniques

An extended and very well studied approach to detect humans is by using vision. With fixed monocular cameras, it can be supposed that only humans are moving in the scene and it can be used as hint to separate foreground from background. In (Bellardi et al., 2010) an statistical by pixel approach is capable to learn what pixels belong to the background in an adaptive way, by calculating the difference between the observed image and the background the pixels of foreground can be detected and after clustered by means of a Self Organizing Network. Each cluster must correspond to a human, and the center of mass of the cluster can be used to estimate his/her position. Same clustering method was used for doing object extraction in (Laugier et al., 2008) but this time using an occupancy grid build with laser scans as input.

Since the point of view of a mobile robot, background and humans are both moving then it is necessary to use features that can be robust in that setting. One approach using video that showed good results is the presented in (Dalal et al., 2006) where appearance descriptors extracted from single frames are combined with motion descriptors extracted from either optical flow or spatio-temporal derivatives against the subsequent frame. Appearance descriptors used are Histogram of Oriented Gradient (HoG) while differential flow descriptors helped to cancel effects of camera motion. A faster implementation of HoG was presented in (Zhu et al., 2006), where AdaBoost was used as a feature selection.

Also popular are the approaches that fuse data coming from two or more sensors, for example in (Luo et al., 2007) they used data coming from a Laser Range Finder to detect body candidates and data coming from a camera to realize face detection, finally they fuse both data to obtain the final detection. Data from three sensors: a color camera, a time-of-flight camera and a laser range finder were fused in the system proposed in (Knoop et al., 2006) which relies in a human body model to estimate 3d position and have the interesting characteristic that it runs in real-time (20Hz). Detection step normally produces a bounding box containing the estimate space where the human is however more complex tasks of Human Robot interaction need to have a further understanding of human body configuration, i.e. more precise body parts poses. A method to do upper body estimation and its temporal association in video sequences was presented in (Ferrari et al., 2008), results shown how their method detects and estimate torso, lower arms, upper arms and head of humans in the scene. Pose estimation

5. SOCIALLY-AWARE NAVIGATION FOR ASSISTED MOBILITY

of one human with laser data is done in (Svenstrup et al., 2009). Using a monocular camera and a 3d model of a human it is possible to estimate human position as presented in (Lin et al., 2009). A review on 3D human pose discrimination from images using silhouette matching is found in (Chen et al., 2010). The sensory system proposed in (McColl et al., 2011) for body pose identification consists of a thermal camera and a Time of Flight range camera, they used information to assess the accessibility of a human to interact with a robot.

5.2.2 Motion: models, tracking and prediction

Tracking. There is an abundance of literature on people tracking, in particular using laser range finders and color based vision. Tracking aims at associating an identity to detected humans in consecutive time steps. Most simple approaches are based in the well-known Kalman Filter applied on the detected features along time. Useful information that we can get from a tracker is current position and speed.

A very robust and multi-person tracking approach was presented in (Ess et al., 2009) which used only visual information from a stereo camera, its main drawback was the processing time which was reported as about 30 seconds per frame. Real-time trackers from a moving vehicle approaches using stereo vision were presented in (Bajracharya et al., 2009) and (Abd-Almageed et al., 2007). Head tracking, facial gesture interpretation and upper body tracking were realized using vision in (Siddiqui et al., 2009) as strategy to get short range interaction between robot service and humans.

Motion models. Be aware of humans in the environments also implies that we are capable to assign them a navigation behavior, to say that a human follow a straight line when walking is enough in many cases but not for realistic ones. There are many models of pedestrian navigation already proposed, next some of them are listed. A simple model which introduced social distances to the Cellular Automata pedestrian dynamics model was presented in (Was et al., 2006), their approach is reactive and discretization of environment is needed. Regarding the models for motion of pedestrians, (Shao and Terzopoulos, 2007) presented a very complete in urban environments which considered behaviors like standing in place, moving forward, turning in different directions, speeding up and slowing down. (Arechavaleta et al., 2008) proved that a nonholonomic model could be used to model human locomotion, they suggested an analogy between the steering wheels and the torso for the control of locomotion differently an approach based on utility optimization was presented in (Hoogendoorn and

5. SOCIALLY-AWARE NAVIGATION FOR ASSISTED MOBILITY

Bovy, 2004). Given dense enough crowds, the robot could believe the environment to be unsafe, i.e., every path is expected to collide with an agent in the crowd due to massive uncertainty. Thus, the robot either makes no forward progress, or takes extreme evasive action to avoid collisions within the crowd. This problem is referred as the “freezing robot problem” (FRP). A recent work (Trautman and Krause, 2010) demonstrated that even in the case of perfect prediction of human motion, the FRP can still appear. They proposed a model based in Gaussian Processes that can estimate crowd interaction from data arguing that humans cooperatively make room to create feasible trajectories. (Kulic and Nakamura, 2010) proposed an HMM based approach for extracting a model of movement primitives and their sequential relationships during on-line observation of human motion. In (Althoff et al., 2012) an overall collision probability is used for safety assessment. Two kind of objects are distinguished related to a trajectory τ_i : passive objects which ignore the trajectory τ_i and active objects which react to τ_i in order to reduce risk collision. The probabilistic safe assessment takes into account the perception and collision avoidance capabilities of both passive and active objects.

Prediction. Some studies have shown that better prediction of human motion is obtained by taking into account interaction among pedestrians (Pellegrini et al., 2009), others used context information like in (Luber et al., 2011) where it is proposed a method that encodes spatial priors on human behavior and showed a place-dependent motion model whose predictions follow the space-usage patterns that people take and which are described by the learned spatial Poisson process. Probabilistic grids of human motion were built in (Thompson et al., 2009) based on the idea that knowledge about functional places and navigational way points in an environment can provide information about the intent of humans. As obstacles can hide moving humans techniques like the ones presented in (Krishna et al., 2006) and (Chung et al., 2009) can be used to increase safety.

A method to learn and predict pedestrian movement was presented in (Scovanner and Tappen, 2009) which is formulated as a series of continuous optimizations. Typical behavior can be learned by observing the motion of humans in determined environment, models based on GHMM (Vasquez et al., 2009), POMDPs (Karnad and Isler, 2012), Gaussian Processes (Meng Keat Christopher, 2009), principle of maximum entropy (Kuderer et al., 2012) or the “sub-goal” concept (Ikeda et al., 2012) have been proposed. In (Wang et al., 2008) represented human activities using Gaussian Process dynamical models, learning was made by collecting data of both motion and human

5. SOCIALLY-AWARE NAVIGATION FOR ASSISTED MOBILITY

pose from a motion capture system. By knowing the motion patterns of people we can improve the collision avoidance with them. Also such patterns can be followed by a robot (Yuan et al., 2010) in order to profit of the human spatial understanding of the environment.

5.2.3 Activity recognition

Methods based on vision are also popular to do activity recognition, for example, recognition of abnormal behaviors such as violence or vandalism in video scenes of the metro was done in (Cupillard et al., 2002). In (Park and Aggarwal, 2000) Human interactions like shaking hands and pointing to the opposite person were detected from gray-scale images. Regions of spatio-temporal salient points on images were detected by measuring the variations in the information content of pixel neighborhoods in order to classify and detect human actions (Oikonomopoulos et al., 2005). In (Park and Oh, 2007) the semantic segmentation method was suggested in order to predict the next action and recognize the intention by recording and analyzing daily life behaviors in a real indoor environment. An ontology-based approach was proposed in (Chen et al., 2004) to analyze Social Interaction Patterns in a nursing home from video sequences. Ontology was implemented as a dynamic Bayesian network. Another similar approach using hierarchical durational-state dynamic Bayesian network was proposed in (Du et al., 2006) to do Interacting Activity Recognition. A method to detect changes in small group configuration and activity by analyzing audio and video of human meetings was presented in (Brdiczka et al., 2009).

Social interaction patterns can be studied by using other devices like active Radio Frequency Identification tags which sends information when face-to-face contacts are detected as in (Barrat et al., 2008). People in need of guidance or assistance was recognized in (Kemotsu et al., 2008) by combining: the current position of the person, the path traversed by the person, the persons pose, the position and motion of the persons hands, and the position and direction of the persons face. These features were collected by a network of fixed video cameras, range sensors, radio frequency identification (RFID) tags, and motion sensors in the environment. A combination of the output from many sensors was used in Groh et al. (2011) in a method to verify the existence of social situations in the environment, specifically they combined audio and geometry of interactions in a Subjective Logic framework. Sound and structured light-based depth sensing were used in (Loper et al., 2009) to achieve mobile peer-to-peer interaction between humans and robots.

5. SOCIALLY-AWARE NAVIGATION FOR ASSISTED MOBILITY

It was observed that HMMs have been proposed to deal with the task of behavior learning in many works. In (Mead et al., 2011b) where an HMM was trained to recognize initiation, acceptance and termination cues in social encounters. Data was collected using a Microsoft Kinect. Total distance, relative orientation and visibility were the variables selected to train the model. *Following*, *Meeting* and *Passing by* activities were modeled using HMMs in (Kelley et al., 2008), by selecting angle and distance between two humans as observations. They were capable to infer, for each observed agent, the intent of the action it is most likely to perform, from the previously trained HMMs. Interactions between humans were modeled and after detected by mean of CHMM's in (Oliver et al., 2000). In (Sukthankar and Sycara, 2006) it is enunciated that team plans involving physical movement possess a distinctive spatial structure, characterized by the relative positions of teammates and external landmarks, that can be exploited to classify team behaviors, as a consequence the authors proposed to use team templates and spatially-invariant HMMs to classify team formations. A review on interpretation schemes in high level understanding of human actions and interactions can be found in (Aggarwal and Park, 2004).

5.2.4 Measuring human activity

Many sensors has been already mentioned as helpful to deal with the task of understanding the scene. As new sensors are becoming available, it is important to take advantage of them to measure social signals.

Wearable sensors are a very interesting type, an example of them is shown at fig. 5.2 (c). In (Bajcsy et al., 2009) wearable sensors (accelerometers and gyroscopes) connected to a wireless network were used to study the type of interaction between two subjects. Same type of sensor were used in (Zhu and Sheng, 2009) for learning and recognizing regular daily activities like standing, sitting, walking upstairs and downstairs, running and sleeping. In particular wearable sensors like Sociometric Badges (Olguin et al., 2009) could be directly used by our method to get high level descriptions of human behavior, mainly in the case of face-to-face interaction and physical proximity, an image of Sociometric badge is shown at fig. 5.2 (a).

Non invasive sensors must be encouraged to use, for example, the intelligent communicating tiles presented in (Pepin et al., 2009) offer an opportunity to collect useful information in order to model human activities. Unfortunately, their cost and difficulty of deployment do not permit, still, an easy access to them.

In the course of the thesis work it has been observed how one sensor is attracting the

5. SOCIALLY-AWARE NAVIGATION FOR ASSISTED MOBILITY

attention of the robotics community by its potential in Human Robot Interaction, this is the Kinect (fig. 5.2 (b)), a motion sensing input device commercialized by Microsoft. Some researchers (Mead et al., 2011a) have proposed the use of this sensor to get Proxemics features and investigate individual, attentional, interpersonal and physiological factors that contribute to social spacing. Also the Kinect was used in (Sung et al., 2012) to learn and recognize kitchen, living room and office human activities.

Authors in (Bainbridge et al., 2012) recommended the use of sensors to detect bio-signals in order to understand human subjective factors. Their results showed that higher hand temperature, higher temperature growth, higher tactile measurements, and farther face distances indicate more positive feelings towards a robot.

In the social sciences side, many studies about interpersonal distances were realized by observations and by answering questionnaires. Now there is a need of more accurate techniques to get that information, in that sense Hempel and Westfeld (2009) proposed photogrammetry techniques from intensity and range images to measure distance between interacting persons.

In the fig. 5.2 some examples of sensors offering big potential to be used in Social Signal Processing are shown.

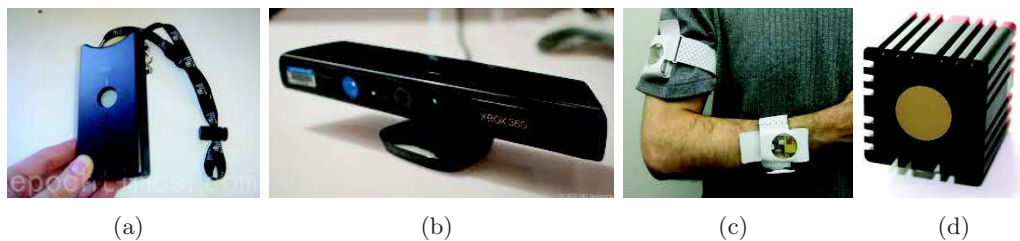


Figure 5.2: Sensors offering big potential to be used in Social Signal Processing. a) Sociometric badge by Sociometric Solutions, b) Kinect by Microsoft, c) Shimmer wireless sensor by Shimmer Research and d) SwissRanger 4000 by MESA Imaging

5.3 The assisted mobility system

5.3.1 Requirements

In an assisted mobility system, it is crucial to take into account the actual needs and characteristics of both its users and the people around them. In particular, this thesis studied the case of a system which includes a robotic wheelchair, such system has been designed around the following requirements:

5. SOCIALLY-AWARE NAVIGATION FOR ASSISTED MOBILITY

- *Safety*: The system should avoid collisions with both static and dynamic entities.
- *Usability*: People with motor disabilities often have problems using joysticks and other standard control devices in navigation tasks requiring fine input. The system should account for this, for example by favoring the most “reasonable” actions when presented with an ambiguous command. Usability is also linked to the acceptance of robotic systems by users/operators.
- *Comfort*: Strong accelerations can be intolerable and even dangerous for a wheelchair user, this imposes an additional constraint on how the robot may move.
- *Respect of social conventions*: When moving, a robot may considerably disturb people around it, especially when its behavior is perceived as unsocial. Even worse, the wheelchair’s passenger may be held responsible for that behavior. It is thus important to produce socially acceptable motion.

From the technical standpoint those requirements imply that, in addition to the conventional robot tasks (e.g. localization, path execution) the following points should be specifically addressed:

- *Integrated motion-planning and long-term motion prediction*: Most human-populated environments are highly dynamic, requiring considerable look-ahead about how other objects will move in order to ensure collision-free robot motion under “comfortable” accelerations. This motivates the proposed integration of a long-term motion prediction algorithm based on the idea of typical behavior with a risk-based motion planning algorithm.
- *Interaction detection for socially acceptable robot-motion*: Our approach is based on the simple idea that, when people interact, they often adopt spatial formations implicitly forming “interaction zones”. Thus, socially acceptable motion can be enforced by first detecting interaction zones and then computing the risk to invade them.

Discomfort and motion sickness caused by strong accelerations in human-transport robots is a very important problem to solve however it is not addressed in the present work, we think that techniques like the one presented in (Solea and Nunes, 2009) or the graceful motion proposed in (Gulati and Kuipers, 2008) could be implemented. Usability is an ongoing effort, a current PhD student in our team is working with new gestural interfaces to control the wheelchair. The points concerned to the present thesis

5. SOCIALLY-AWARE NAVIGATION FOR ASSISTED MOBILITY

are the safety and the respect of social conventions.

One of the main ambitions with this platform is to provide an open benchmark that could be used to compare and evaluate different approaches. This is an important task given the diversity of existing wheelchairs (Simpson, 2005), including autonomous (Kollar et al., 2010), semi-autonomous (Parikh et al., 2005) and social aware systems (Kirby et al., 2009; Rios-Martinez et al., 2011). Work in (Grasse et al., 2010), is placed in the same framework than ours, the difference with them is that their focus is on assisting the user by recognizing the desired path among a set of frequent paths and by following it autonomously. Their system does not have a motion planning module nor a social filter module. A recent review of wheelchair projects is also found in that article. Our approach could be classified in the emergent field of Socially Assistive Robotics (Feilseifer and Matari, 2005).

5.3.2 Structure of the proposed system

Fig. 5.3 presents an overview of our system’s architecture. It is divided into several

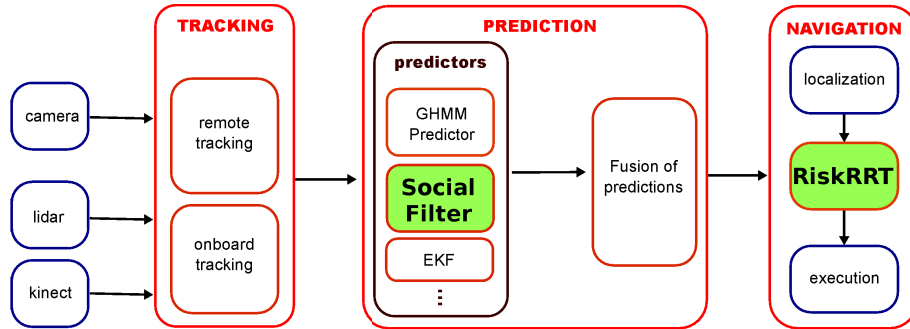


Figure 5.3: Achitecture overview.

subsystems:

1. *Tracking subsystem*: mobile objects are tracked both off-board and on-board the robotic wheelchair.
2. *Prediction subsystem*: the prediction subsystem processes data from the trackers and transforms it into probabilistic predictions about the configuration of the free space in the environment. It also features a “Social Filter”, which detects present and future interactions and creates virtual obstacles corresponding to interaction zones.

5. SOCIALLY-AWARE NAVIGATION FOR ASSISTED MOBILITY

3. *Navigation subsystem*: the navigation subsystem includes a laser-based localization module and a motion-planner which integrate predictions to compute safe trajectories that are fed to the execution module.

5.3.2.1 Risk-Based Socially-Aware Navigation

The navigation subsystem is based on the approach presented in section 4.4 which combines Risk-RRT (Fulgenzi et al., 2009), a motion planner which integrates motion predictions to provide safe trajectories, and the “Social Filter”, the mechanism to obtain socially acceptable behavior by following social conventions of human management of space.

When the wheelchair is transporting a human, it will have to move in a populated environment where an “optimal” behavior may be perceived as unsocial. People will become uncomfortable if they are approached at a distance that is deemed to be too close, where the level of discomfort experienced by the person is related to the importance of his or her space. This simple idea was formalized introducing the concept of *Personal Space*, first proposed by Hall (Hall, 1966), which characterizes the space around a human being in terms of comfort to social activity, see section 3.2.1 for details. Another interesting social situation arises when two or more of the persons in the environment are interacting. Interactions are modeled by using the concept of *O-Space* which has been studied and developed in human sciences (Kendon, 2010), see section 3.3.2 for details. This space can be observed in casual conversations among people where participants’ position and orientation are used to establish boundaries of the space. This space is respected by other people and only participants are allowed to access to it, therefore the intrusion of a stranger causes discomfort. In our path planner, human friendly paths are generated by including the “Social Filter” described in chapter 3 which transforms those spaces into corresponding cost functions which lead the robot to avoid them. As a result, the choice of a best path done by RiskRRT is based on the “minimum risk” calculated for every partial path considering the probability of not encountering a collision along the path and not invading a Personal Space or an O-Space. The focus is put on detecting and predicting conversations in the environment surrounding the wheelchair (Rios-Martinez et al., 2011).

5.3.2.2 Pedestrians tracking and motion prediction

The off-board tracker provides global information about moving obstacles and learning input for our motion prediction module. At this point, our tracking systems are under

5. SOCIALLY-AWARE NAVIGATION FOR ASSISTED MOBILITY

developing and testing. Meanwhile, several tests have been realized by using augmented reality markers (see fig. 5.10 (a)) that people wear as hats. This has allowed us to validate the overall architecture, even if it is not a viable solution in the long run.

The on-board system will provide detailed information about the objects that appear in the robot’s perceptual field. Its main use is to identify interactions between people (e.g. two persons shaking hands). The on-board tracking performs leg detection using a LIDAR sensor and people detection and tracking using the Kinect sensor, according to the technique proposed in *openni_tracker* ROS package¹, results are shown in section 5.5.2.

The motion prediction subsystem takes tracking data (i.e. position, orientation and velocity) and outputs K grids, representing the posterior probability of the space being occupied at times $\{t_1, \dots, t_K\}$ in the future. Prediction itself is accomplished with a Growing Hidden Markov Model (Vasquez et al., 2009) and an Extended Kalman Filter but the grid representation makes it easy to experiment with other prediction algorithms. The prediction grids are then processed by a fusion module, which uses an entropy-based measure of prediction quality to merge predictions coming from different algorithms.

5.4 Experimental platform

The current approach is being implemented in our experimental platform, an automated wheelchair (Fig. 5.4(a)) equipped with two Sick lasers and a Microsoft Kinect, running ROS (Robotic Operating System) for achieving semi-autonomously mobility actions commanded by the wheelchair’s user. Laser permits us to build a map of the environment, like shown on the bottom of Fig. 5.4(b) and at time of execution it provides information to localize the wheelchair in the environment according to the map. Data coming from the Kinect will allow us to have position and orientation of pedestrians in the scene at short distance (no more than four meters). The wheelchair is also equipped with an on-board computer to take care of the low-level hardware control tasks, on top of that it also carries a notebook computer with the navigation, prediction and tracking software.

In addition to the mobile platform, there is also an external camera (Fig. 5.5) overlooking the whole test environment. It is connected to an external computer that communicates with the wheelchair via wireless network. This was motivated due to the short range of Kinect sensor to detect people then the two systems are complementary,

¹http://www.ros.org/wiki/openni_tracker

5. SOCIALLY-AWARE NAVIGATION FOR ASSISTED MOBILITY

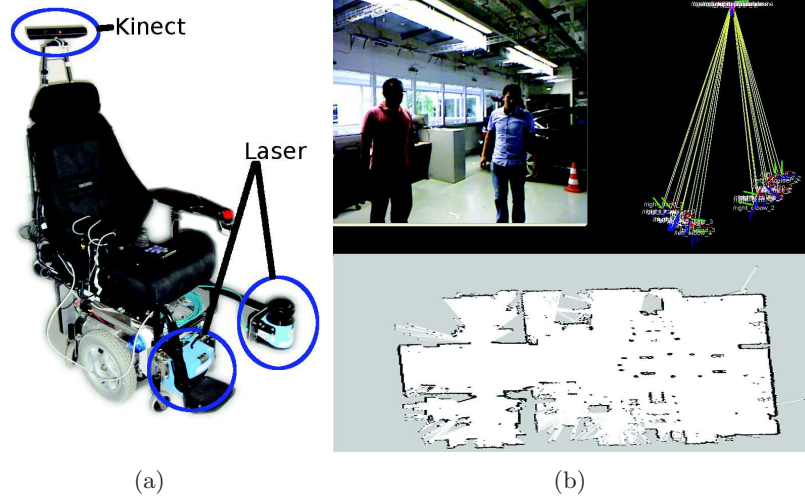


Figure 5.4: Experimental platform: in (a) the robotic wheelchair, on the top of (b) the data provided by the Kinect tracking human body, on the bottom the final map built by using laser data.

one is more precise at close range, the other is better at long range. At present the two tracking systems on-board and off-board do not share information but the fusion of them is expected to give an increase on the performance of the complete system.

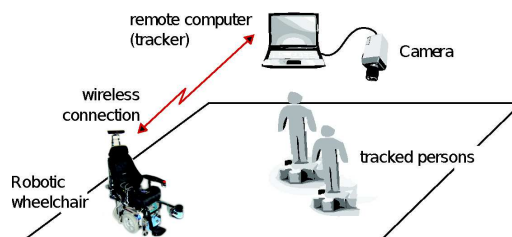


Figure 5.5: An external camera was fixed at INRIA's hall in order to overlooking the whole test environment. Humans were tracked by looking at markers, at any moment, position and orientation from each person was available.

From the software perspective, the system has been implemented as a number of independent modules using the Robot Operating System (ROS) presented in [Quigley et al. \(2009\)](#) which currently is receiving a lot of attention from robotics community.

5.5 Experimental Results

5.5.1 Perception of interactions with external camera

Using the off-board camera described in Fig. 5.5 position and orientation from each human was collected. That information was passed as input to Social Filter module which was able to identify interactions and also to estimate O-Space location. In Figs. 5.6 a) and 5.7 a) it can be observed how detection and tracking of humans is done by using markers (hats). A vis-a-vis formation is detected in fig. 5.6 b) while a group of three people is detected in fig. 5.7 b).

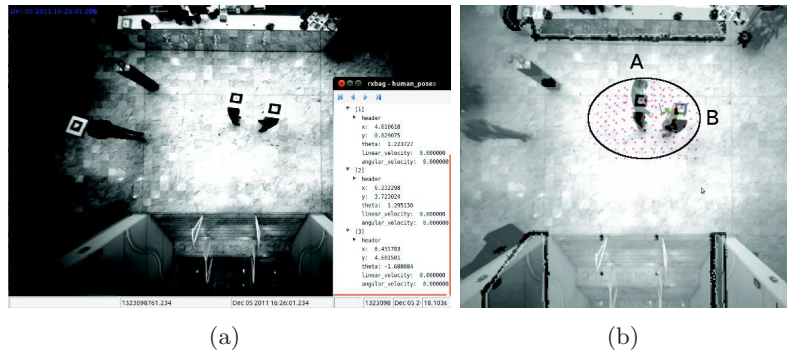


Figure 5.6: Interaction between two people detected by using markers with a fixed video camera. In a) real image and information of orientation and position from humans, in b) map and interaction area detected superimposed on real image. Humans are labeled with A,B and C letters.

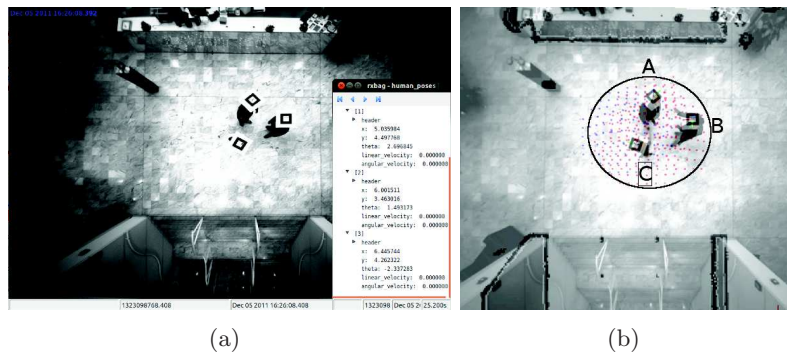


Figure 5.7: Interaction between three people detected by using markers with a fixed video camera. In a) real image and information of orientation and position from humans, in b) map and interaction area detected superimposed on real image.

5. SOCIALLY-AWARE NAVIGATION FOR ASSISTED MOBILITY

5.5.2 Perception of interactions with kinect sensor

The on-board Kinect attached to the wheelchair was used to track people and to detect interactions. The Kinect sensor permits to get the position and orientation of the torso for each identified human. That information is passed to the Social Filter described in chapter 3. Images in Fig. 5.8 show the importance of formation between humans in order to estimate focused interaction. In a) and b) humans are very close but our model does not identify any interaction because their formation do not favor a conversation. Some resulting human models and interaction models are shown in Fig. 5.9. There the

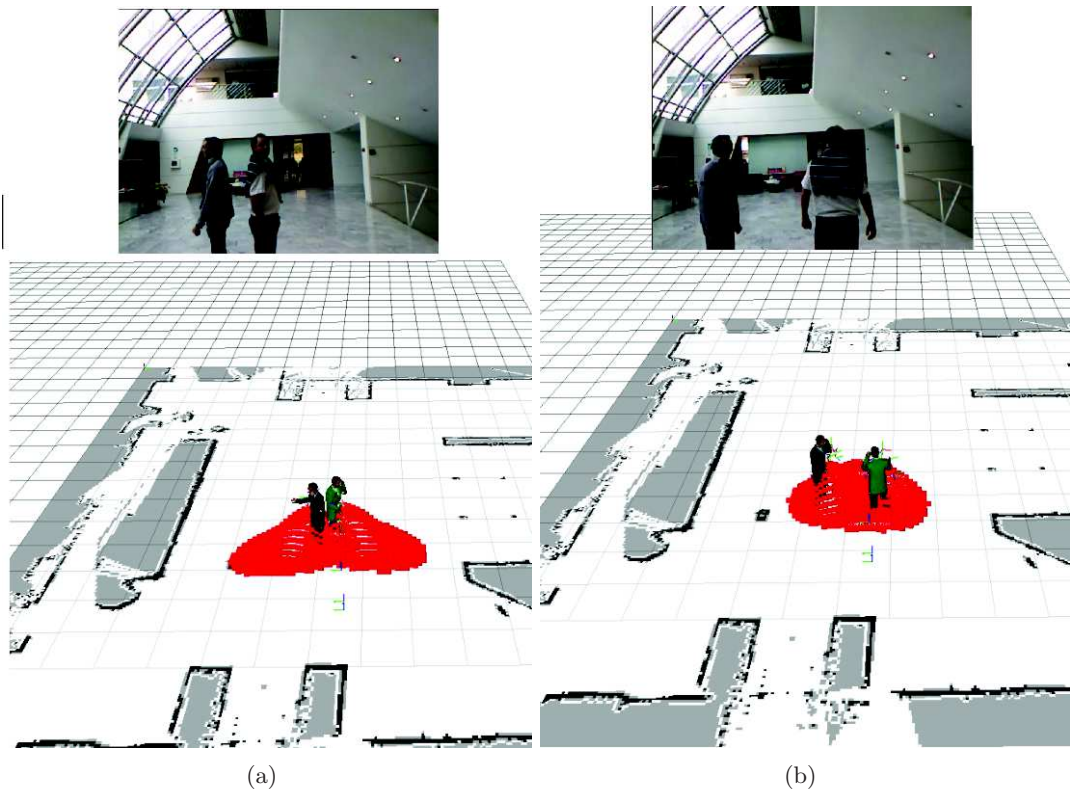


Figure 5.8: Social Filter model detects formations that favor the focused interaction. In a) and b) even when relative distance is adequate to conversation the orientation is not.

orientation of torso is used to estimate the main focus of interest and together with distance and relative speed information, Vis-a-vis formations ((a), (b)) and L-Shape formation ((c)) were detected. In the Fig. 5.9 d) a group in conversation is detected based on the formations observed between participants.

5. SOCIALLY-AWARE NAVIGATION FOR ASSISTED MOBILITY

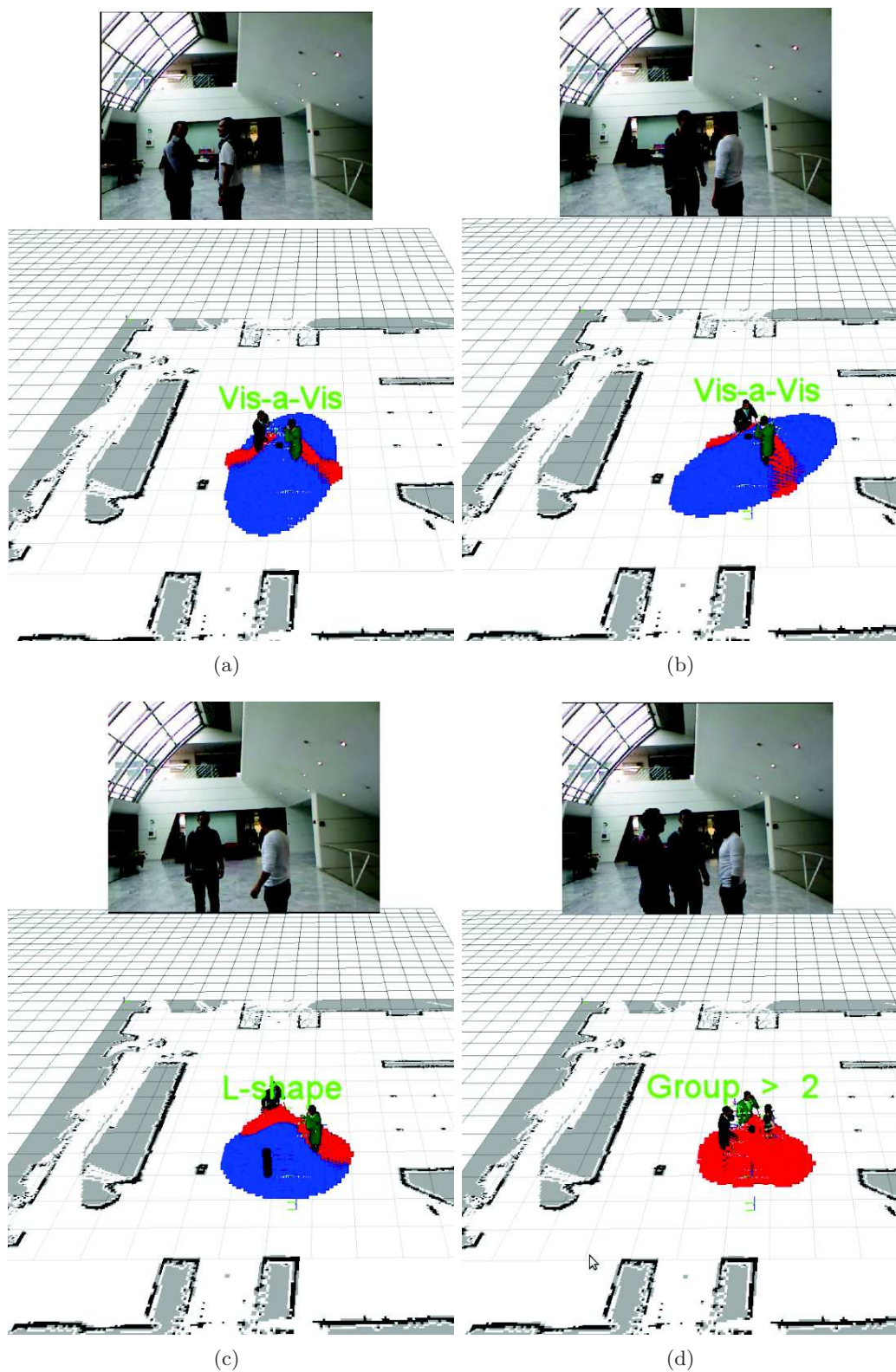


Figure 5.9: Interactions detected with Social Filter from Kinect input for two different pairs of humans in a), b) and c). A group is detected in d). Torso direction is used to estimate the main focus of interest.

5.5.3 Navigation results

Tests were conducted in the INRIA hall, linking together the tracking, Social Filters and navigation modules, previously presented. The tracking module fed information to the social filter module which computed social interaction zones, according to the orientation and position of humans in the scene. Fig. 5.10(a) shows one image of two persons interacting while the robot passes by, with a researcher closely following. Fig. 5.10(b) shows the same situation but taken from the overhanging camera linked to the tracker computer, where the robot position, its plan and intended destination can be seen. Several tests were conducted to evaluate the capability of the robot to avoid

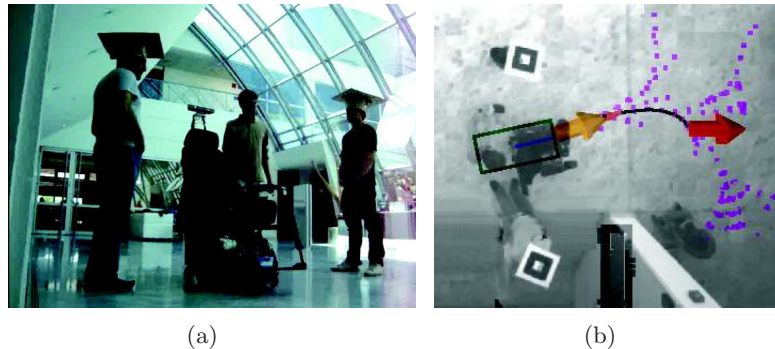


Figure 5.10: (a) Experimental test with two interacting humans and a robotic wheelchair navigating among them. (b) Overview camera image of the test scenario with the robot plan overlaid.

zones that would cause discomfort to the people interacting with each other. Tests were realized with and without the Social Filter module, to demonstrate that not taking into account the zones of social interaction would result in paths that are shorter but “rude” or even frightening. Fig. 5.11 and Fig. 5.12 are examples of the two experiments that were performed. Sequences presented in those Figures show roughly the same initial configuration for the robot and the interacting persons, as well as the same goal. The only difference is that, in the Fig. 5.12 the social filter has been disabled while in Fig. 5.11 it is active, which is illustrated by the point cloud between persons.

Due to the absence of a social space, in Fig. 5.12, the planning algorithm treats the humans as simple obstacles, and the chosen path is the one that moves straight to the goal. However, in the other Figure, when the social filter is active, nodes that are generated inside the interaction zone are penalized with a high risk, and then are excluded during the path search.

This example clearly shows that although a straight path to the goal can be con-

5. SOCIALLY-AWARE NAVIGATION FOR ASSISTED MOBILITY

sidered to be more efficient in terms of total distance that was traveled and energy, it moves in such a way that it causes discomfort to interacting groups of people in the environment. On the other hand, the example shown in the right column, manages to avoid the zone of interaction, at the cost of traveling a longer distance.

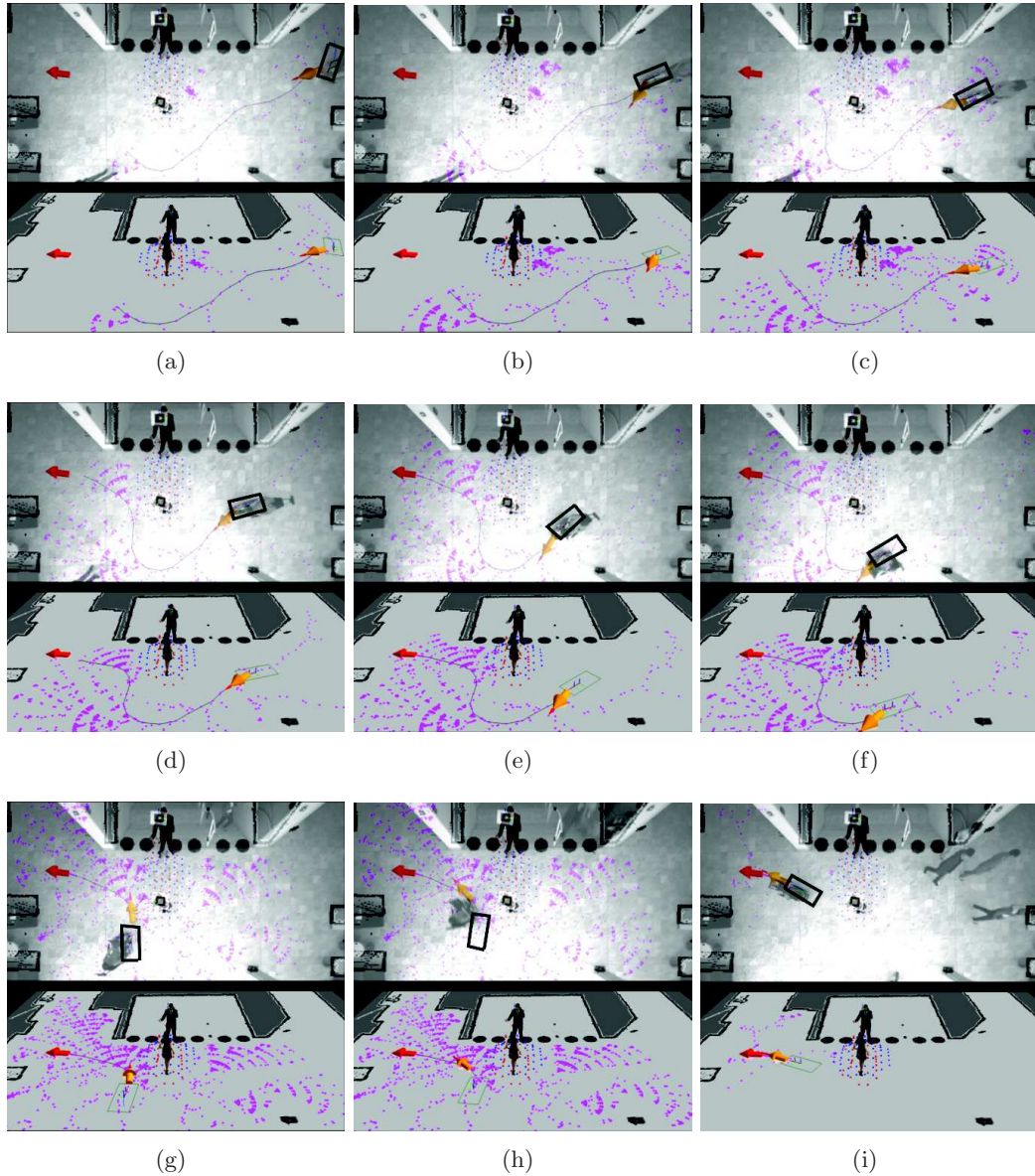


Figure 5.11: Socially-Aware navigation in dynamic environments populated by humans with Social Filter activated. The robotic wheelchair avoids to pass in the middle of interaction.

5. SOCIALLY-AWARE NAVIGATION FOR ASSISTED MOBILITY

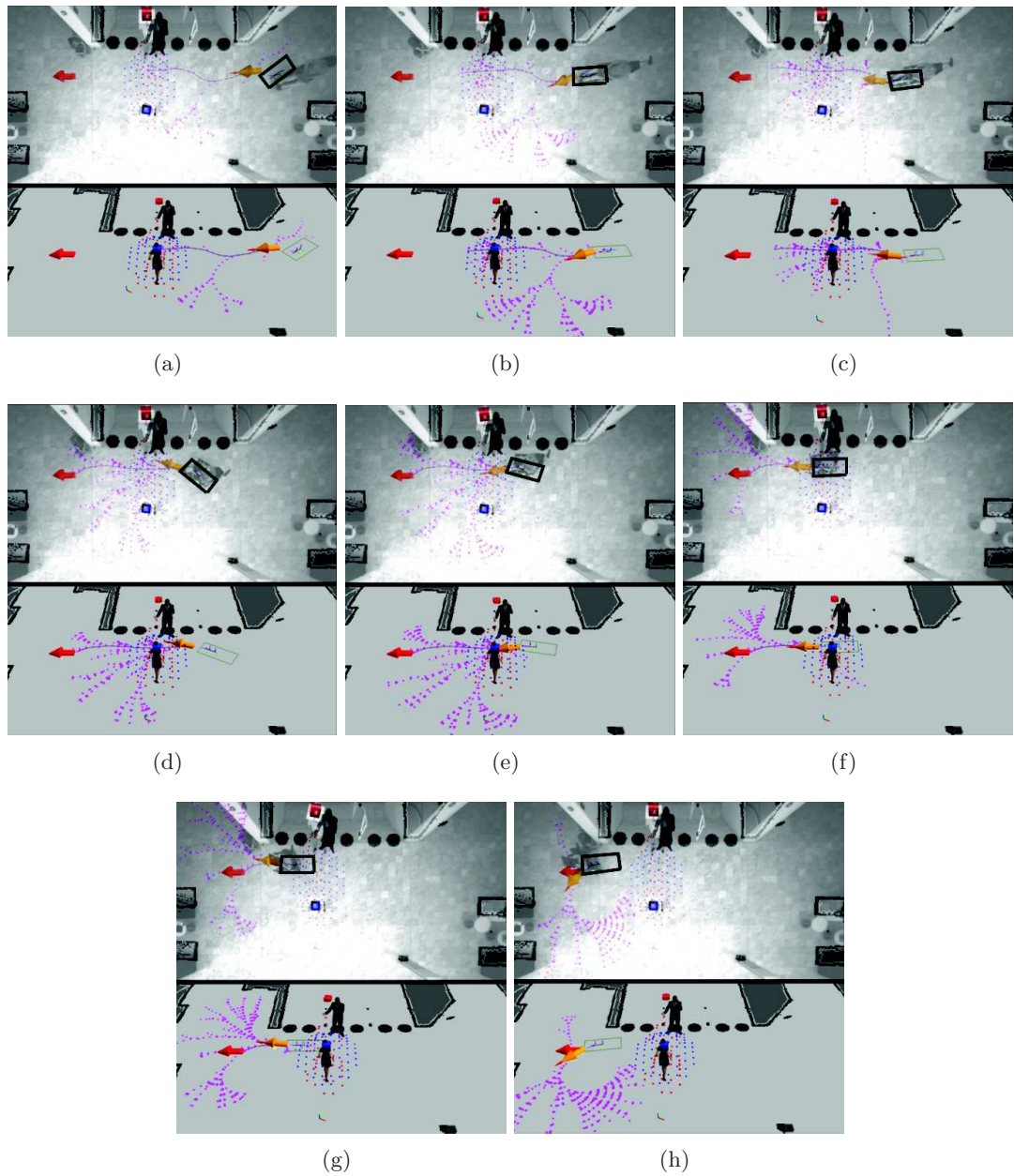


Figure 5.12: Navigation in dynamic environments populated by humans with Social Filter deactivated. Navigation is not socially-aware, main concern is only minimizing distance to goal.

5.6 Conclusions

The preliminary experiments and the experience of putting together the experimental platform have been really instructive in relation to several aspects of the problem at the application and the technical level:

- *Socially acceptable behavior is very important.* Even in our scripted tests, both interacting people and the wheelchair’s user reported that they felt very uncomfortable when the robot passed right through the middle of a talking group.
- *Predictive behavior and socially acceptable behavior are often similar.* For example, when pedestrians were passing through the robot’s path, it often happened that it stopped (knowing that the path was going to be free) to let the person pass. This seems to indicate that in many cases, knowing how people will move, the most reasonable thing to do is to be polite. It also suggests game theory as a possible way to analyze these interactions.
- *Evaluation criteria are crucial but difficult to define* How do you evaluate comfort, safety or social compliance? in our experiments we have tried both objective (e.g. acceleration, heartbeat rate, sweat) as well as subjective variables (e.g. questionnaires, interviews). Both have significant advantages and drawbacks, but it is hard to come with a representative subset that can lead to significative, replicable and unbiased experiments and conclusions, especially when comparing dissimilar approaches (e.g. deterministic vs. probabilistic).

Considerable effort is being devoted to design quantitative tools to evaluate our system’s performance in terms of the stated goals. On the other hand, we have already conducted experiments on two scenarios to qualitatively assess our system.

The tests focused on two main functions: predictive navigation and socially acceptable navigation. In the first case, we asked people to actively interfere with the robot’s plans by either following the same path than the robot in the opposite direction or intersecting it at some point. In both cases the robot had to anticipate the human trajectories and generate an alternative collision-free plan.

In the second case, we aimed to assess the capability of the robot to avoid disturbing or causing discomfort to persons that were not moving but were interacting with each other. Groups of people were arranged in a manner that the direct path to the robot’s goal would be inside a social interaction zone, so a straight movement to the goal would violate the interaction zone and therefore, the robot had to find alternative paths.

5. SOCIALLY-AWARE NAVIGATION FOR ASSISTED MOBILITY

The RiskRRT algorithm has been extensively tested in simulation. It is now implemented in our real platform, where it will be tested against similar approaches to assess the actual impact that integrating risk estimation and trajectory prediction has in terms of safety.

Chapter 6

Conclusions and perspectives

Autonomous navigation techniques for robots moving among humans must consider the social aspect involved, i.e., they must be socially-aware. Common expectations about robot behavior and capabilities are high, maybe explained by the information generated from science fiction or because in a broad sense robots are the materialization of artificial intelligence. One main expectation to consider is that because of the sharing of the same physical spaces, robots and people must follow the same social conventions in the management of space. For example, robots must respect proximity constraints but also respect people interacting as humans do. Addressing this expectation, the work in this thesis proposed two methods which take into account social conventions. One was based on the minimization of the estimated discomfort generated by robot solutions of navigation. The other follow the same objectives but this time the focus was on minimizing the risk of disturbance on the paths produced as solutions.

The problem of designing socially-aware robot navigation systems requires effort from robotics and social psychology at least, the new emergent field of social robotics is an example of this need. In our work concepts of “comfort” and “proxemics” were studied in references of social sciences field with the objective to identify potentially useful theories to integrate in robot navigation. Inspired on that review, models of human spaces were developed. Particularly, the interaction spaces related concepts of O-Space and f-formation had not been exploited in navigation for dynamic environments literature.

Main Conclusions

1. The social conventions are useful for two main purposes: safety and communication. Avoiding the invasion of social spaces contributes to the safety of humans because it helps to guarantee that sensitive zones will not be touched and, at the

6. CONCLUSIONS AND PERSPECTIVES

same time, robots communicate the intention of the robot to be social.

2. Robotics and social psychology must collaborate. Work done in social sciences is a rich source of ideas and theories for expected robot behavior. At the same time social psychology could benefit of models and tools developed in robotics to validate and extend its own methods in a sort of cycle where the two parts win. When realizing the psychophysics experiments it was observed that working with humans is a very complex task, the multiple factors affecting human behavior are mixed and the only way to continue is by taking advice of experts in other fields.
3. Unfocused robot interaction with humans is important and feasible. The non verbal behavior in humans communicates a big amount of information and its main social signals can be detected by actual sensors. In unfocused interaction scenarios the robot must negotiate its position by exhibiting understandable behavior which is possible by following social conventions. Social signal processing is the technique that will need a stronger effort to be integrated.
4. Decisional systems must manage dynamic uncertain factors and be adaptive. Both methods of socially-aware robot navigation presented are capable to include multiple factors in order to decide the solution and to adapt to changes observed in the environment. However, the uncertainty of the social cues was not taken into account. Moreover social conventions are very dynamic and can be applied or not depending on many factors, for example the Personal Space can grow if the situation is threatening, decrease if an acquaintance is recognized or almost disappear if there is a crowd. Therefore, decisional system must be highly adaptive.

Perspectives

- A very important extension to social filter is the inclusion of dynamic adjustment on the risk estimation for the variables according to the situation recognized. The experiments realized permitted to have a preliminary idea of the O-Space extension in one specific case but more experiments must be realized to understand how it varies in other cases depending on the environment configuration.
- An important study that must be addressed is the effect of the shape and appearance in the risk perceived by humans, where risk can be seen as function of social expectation and observed behavior. It is important to take it into account for

6. CONCLUSIONS AND PERSPECTIVES

at least two reasons. The first one is that the expectation about social behavior seems to be related with the shape of the robot, for example, a humanoid robot is expected to be more intelligent than a robotic wheelchair even when the two follow the same models of social conventions. The second one is more practical and is related to the acceptability of a robot as a product, a very powerful and dexterous robot able to carry a person could not be accepted among humans because it looks scary.

- Socially-aware robot navigation is dependent on the model of the environment, then the Perception system plays a fundamental role in the whole approach. In the thesis markers have been employed but this is not a solution viable in realistic scenarios. For the definitive version of the platform, a basic detect-then-track system is being developed, where moving objects are first detected using a Self-organizing network (Bellardi et al., 2010), after this, objects are encoded as a color histogram, and then detected in later frames using the mean-shift algorithm (Comaniciu and Meer, 2002). Finally, the different detections are used as input for a tracker based on the Joint Probabilistic Data Association Filter.
- Assistance to mobility in airports is a right for people with reduced mobility. Regulations in Europe and United States (see Appendix A.1) establish as an obligation for the airports to provide services and facilities to offer disabled persons and persons with reduced mobility, access to air travel comparable to that of any other passengers.

Our platform presented in chapter 5 is a direct response to the problem of assistance to mobility. The ideas presented in this thesis will be implemented in a robotic wheelchair to transport people with reduced mobility in airports, (see more details at Appendix A.1). In this case the environment is known but the obstacles are very dynamic, there exists people walking with bags of many sizes and forms, interacting in many ways, following social conventions, displaying and breaking formations, etc., characteristics that offer many possible applications for navigation in the presence of humans.

6. CONCLUSIONS AND PERSPECTIVES

Appendix A

Assisted Mobility application

A.1 Accessible Air Travel regulation

European regulation ^{1 2}

The aim of European Regulation 1107/2006 is to offer disabled persons and persons with reduced mobility access to air travel comparable to that of any other passengers flying from airports in the European Union (EU) or on a EU based airline.

Article 8 of the Regulation says that the managing body of an airport shall be responsible for ensuring the provision of assistance without additional charge to persons with reduced mobility. The Regulation uses the term persons with reduced mobility to include “disabled persons” and “people with reduced mobility”: namely any person whose mobility is reduced due to physical disability (sensory or locomotory, permanent or temporary), intellectual disability or impairment, or any other cause of disability, or age. Mobility assistance may consist of:

- A guide to assist vision impaired passengers through security and to the correct gate
- Transport and assistance through the airport
- A lift to the aircraft doors

US regulation ³

Air carriers are required to provide assistance when requested by a passenger with

¹<http://www.caa.co.uk>

²<http://eur-lex.europa.eu/>

³<http://airconsumer.ost.dot.gov/publications/disabled.htm>

<http://www.unitedspinal.org/pdf/2012-accessible-air-travel-brochure.pdf>

A. ASSISTED MOBILITY APPLICATION

a disability to transport the passenger between gates to a connecting flight, as well as from the terminal entrance, or vehicle drop-off point, through the airport to the gate for a departing flight, and from the gate to the terminal exit or a vehicle pick-up point. This includes providing assistance in accessing key functional areas of the terminal, like ticket counters and baggage claim. It also includes a brief stop, at the passengers request, at the entrance to a rest room on the route. Air carriers must provide personnel, ground wheelchairs, boarding wheelchairs, and ramps or mechanical lifts. Level-entry boarding platforms or accessible mobile lounges must be used where they are available. When level-entry boarding is not available, airlines must use ramps or lifts. This applies to all US airports with 10,000 or more annual enplanements.

A passenger who has requested assistance must not be left unattended by the personnel responsible for enplaning, deplaning, or connecting assistance in a ground wheelchair or other device in which the passenger is not independently mobile for more than 30 minutes even if another person (i.e., family member) is accompanying the passenger, unless the passenger explicitly waives the obligation.

A.2 Problem description and proposed application

Assistance to mobility in airports is a right for people with reduced mobility. Regulations in Europe and United States establish as an obligation of the airports providing services and facilities to offer disabled persons and persons with reduced mobility, access to air travel comparable to that of any other passengers. In (Authority, 2010) many scenarios of application could be found mainly to restore to the users autonomy and privacy, a list of some problems related to People with Reduced Mobility are:

- Assistance desks or correct queue to join could be difficult to locate due to their disability.
- People with Reduced Mobility could feel abandoned when they are placed in some areas in the departure lounge to wait. Their lack of autonomy causes they worry about missing their flights.
- Allocating staff when People with Reduced Mobility are spread over a wide area is a difficult task.
- Assistants can face the problem to transport people in the wheelchair at the same time as transporting their luggage.

A. ASSISTED MOBILITY APPLICATION

Possible solutions described in that document point out to encourage innovation in developing new procedures and equipment, allowing passengers to remain in their own wheelchairs for as long as possible. Moreover it was evident the need for a transport system that allows several People with Reduced Mobility to be collected and transported collectively rather than individually. We think that techniques as vehicle platooning could be explored in this last case.

Our platform presented in chapter 5 is a direct response to the problem of assistance to mobility depicted in this section.

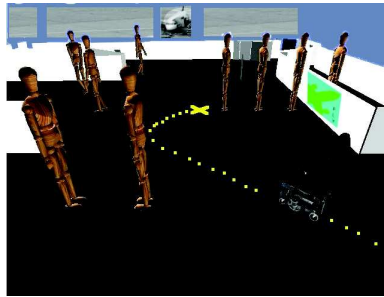


Figure A.1: Autonomous transport of people with reduced mobility in airports is a clear application of our strategy which takes in account human-human interaction. In this figure we can see a representation of the depart lounge of an airport, some social conventions could be observed, face to face interaction and staying in a queue. The wheelchair must navigate in the environment while respecting the cited conventions.

The ideas presented in this thesis will be implemented in a robotic wheelchair to transport people with reduced mobility in airports (fig. A.1).



Figure A.2: Topological map of the airport

A. ASSISTED MOBILITY APPLICATION

One particularity of an application developed for a known environment is that a high level topological map can be included to our system. Such representation of high level for the airport could be created based on a topological graph where vertex are interesting locations in the environment and edges represent a physic and viable path connecting two locations. The resulting map could be something like the one observed in fig. A.2. This map was designed by hand but an automatic generation could be done, for example, by following the spatial subdivision presented in (Lamarche and Donikian, 2004).

Based on the needs observed some services could be implemented, for example:

1. Toilet service. When the user wants to go to the bathroom, he orders this in an intuitive way (voice, for example) and the wheelchair go to the closest (accessible) bathroom, the goal is to place the person in the entrance of the bathroom. Collision with objects and with people must be avoided. An only safe approach could result in the wheelchair not moving at all, the strategy must show to pedestrians its intentions and they must grant the wheelchair with the space needed with the goal of maximizing at any moment the comfort of people and user.
2. Search staff service. The goal of this button is locate staff people in charge of the passenger. The idea is to approach the staff in a safety and understandable way. If the staff is in conversation with somebody else then the strategy of navigation must encourage the acceptance in this group.
3. Emergency strategy service. In case of an event of emergency, this service can help the person to find a route to the closest emergency exit.
4. Shopping Tour button. The user wants to take a look or to buy some in the shops of the airport, then he can press this button to initiate the tour, he can stop at any moment and return to the gate. The capabilities of the wheelchair permits only to arrive at shop and ask, it could be complicated to enter because of the space limits.

We stated that this services can help the people to recover some privacy and to offer them a more independent experience in airports.

References

- Abd-Almageed, W., Hussein, M., and Abdelkader, M. (2007). Real-time human detection and tracking from mobile vehicles. In *Intelligent Transportation Systems Conference, 2007. ITSC 2007. IEEE*, pages 149–154. 111
- Aggarwal, J. and Park, S. (2004). Human motion: modeling and recognition of actions and interactions. In *3D Data Processing, Visualization and Transmission, 2004. 3DPVT 2004. Proceedings. 2nd International Symposium on*, pages 640–647. 114
- Aiello, J. (1987). *Handbook of environmental psychology.*, chapter Human spatial behavior., pages 359–504. John Wiley & Sons, New York. 13
- Aiello, J. R. (1977). A further look at equilibrium theory: Visual interaction as a function of interpersonal distance. *Environmental Psychology and Nonverbal Behavior*, 1:122–140. 13
- Alili, S., Alami, R., and Montreuil, V. (2009). A task planner for an autonomous social robot. In Asama, H., Kurokawa, H., Ota, J., and Sekiyama, K., editors, *Distributed Autonomous Robotic Systems 8*, pages 335–344. Springer Berlin Heidelberg. 31
- Althaus, P., Ishiguro, H., Kanda, T., Miyashita, T., and Christensen, H. (2004). Navigation for human-robot interaction tasks. *Proc. of the IEEE International Conference on Robotics and Automation (ICRA)*, 2:1894–1900. 33, 36
- Althoff, D., Kuffner, J. J., Wollherr, D., and Buss, M. (2012). Safety assessment of robot trajectories for navigation in uncertain and dynamic environments. *Auton. Robots*, 32(3):285–302. 112
- Arechavaleta, G., Laumond, J.-P., Hicheur, H., and Berthoz, A. (2008). On the nonholonomic nature of human locomotion. *Autonomous Robots*, 25:25–35. 10.1007/s10514-007-9075-2. 111

REFERENCES

- Argyle, M. and Dean, J. (1965). Eye-contact, distance and affiliation. *Sociometry*, 28(3):289–304. [13](#), [18](#)
- Authority, C. A. (2010). Accessible air travel workshop. Electronic Report. [134](#)
- Bailenson, J. N., Blascovich, J., Beall, A. C., and Loomis, J. M. (2001). Equilibrium theory revisited: Mutual gaze and personal space in virtual environments. *Presence: Teleoper. Virtual Environ.*, 10(6):583–598. [13](#)
- Bailenson, J. N., Blascovich, J., Beall, A. C., and Loomis, J. M. (2003). Interpersonal Distance in Immersive Virtual Environments. *Personality and Social Psychology Bulletin*, 29:819–833. [13](#)
- Bainbridge, W., Nozawa, S., Ueda, R., Okada, K., and Inaba, M. (2012). A methodological outline and utility assessment of sensor-based biosignal measurement in human-robot interaction. *International Journal of Social Robotics*, 4:303–316. [115](#)
- Bajcsy, R., Giani, A., Tomlin, C., Borri, A., and Di Benedetto, M. (2009). Classification of physical interactions between two subjects. In *Wearable and Implantable Body Sensor Networks, 2009. BSN 2009. Sixth International Workshop on*, pages 187–192. [114](#)
- Bajracharya, M., Moghaddam, B., Howard, A., Brennan, S., and Matthies, L. H. (2009). A fast stereo-based system for detecting and tracking pedestrians from a moving vehicle. *Int. J. Rob. Res.*, 28(11-12):1466–1485. [111](#)
- Bar-Haim, Y., Aviezer, O., Berson, Y., and Sagi, A. (2002). Attachment in infancy and personal space regulation in early adolescence. *Attachment human development*, 4(1):68–83. [16](#)
- Barraquand, R. and Crowley, J. L. (2008). Learning polite behavior with situation models. In *Proceedings of the 3rd ACM/IEEE international conference on Human robot interaction, HRI '08*, pages 209–216, New York, NY, USA. ACM. [31](#)
- Barrat, A., Cattuto, C., Colizza, V., Pinton, J.-F., den Broeck, W. V., and Vespignani, A. (2008). High resolution dynamical mapping of social interactions with active rfid. *CoRR*, abs/0811.4170. [113](#)
- Bartneck, C. and Forlizzi, J. (2004). A design-centred framework for social human-robot interaction. In *Robot and Human Interactive Communication, 2004. ROMAN 2004. 13th IEEE International Workshop on*, pages 591 – 594. [31](#)

- Bellardi, T., Rios-Martinez, J., Vasquez, D., and Laugier, C. (2010). A novel approach for object extraction from video sequences based on continuous background/foreground classification. In *Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on*, pages 2456–2461. [110](#), [131](#)
- Bennewitz, M., Burgard, W., Cielniak, G., and Thrun, S. (2005). Learning motion patterns of people for compliant robot motion. *International Journal of Robotics Research*, 24:31–48. [72](#)
- Bhatt, M. and Dylla, F. (2009). A qualitative model of dynamic scene analysis and interpretation in ambient intelligence systems. *International Journal of Robotics and Automation*, 24(3). [26](#)
- Borkowski, A., Siemiatkowska, B., and Szklarski, J. (2010). Graph transformations and model-driven engineering. chapter Towards semantic navigation in mobile robotics, pages 719–748. Springer-Verlag, Berlin, Heidelberg. [25](#)
- Brdiczka, O., Maisonnasse, J., Reignier, P., and Crowley, J. L. (2009). Detecting small group activities from multimodal observations. *Applied Intelligence*, 30(1):47–57. [113](#)
- Butler, J. T. and Agah, A. (2001). Psychological effects of behavior patterns of a mobile personal robot. *Auton. Robots*, 10(2):185–202. [27](#), [28](#)
- Carton, D., Turnwald, A., Wollherr, D., and Buss, M. (2012). Proactively approaching pedestrians with an autonomous mobile robot in urban environments. In *13th International Symposium on Experimental Robotics*. [33](#), [36](#)
- Ceccarelli, M. (2011). Problems and issues for service robots in new applications. *I. J. Social Robotics*, pages 299–312. [108](#)
- Chen, C., Zhuang, Y., and Xiao, J. (2010). Silhouette representation and matching for 3d pose discrimination - a comparative study. *Image Vision Comput.*, 28(4):654–667. [111](#)
- Chen, D., Yang, J., and Wactlar, H. D. (2004). Towards automatic analysis of social interaction patterns in a nursing home environment from video. In *Proceedings of the 6th ACM SIGMM international workshop on Multimedia information retrieval, MIR '04*, pages 283–290, New York, NY, USA. ACM. [113](#)

REFERENCES

- Chung, S.-Y. and Huang, H.-P. (2010). A mobile robot that understands pedestrian spatial behaviors. In *Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on*, pages 5861–5866. 33
- Chung, W., Kim, S., Choi, M., Choi, J., Kim, H., bae Moon, C., and Song, J.-B. (2009). Safe navigation of a mobile robot considering visibility of environment. *IEEE TRANSACTIONS ON INDUSTRIAL ELECTRONICS*, 56:3941–3941–3501. 112
- Ciolek, M. and Kendon, A. (1980). Environment and the spatial arrangement of conversational encounters. *Sociological Inquiry*, 50:237–271. 21, 22, 23
- Comaniciu, D. and Meer, P. (2002). Mean shift: A robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(5):603–619. 131
- Costa, M. (2010). Interpersonal distances in group walking. *Journal of Nonverbal Behavior*, 34:15–26. 24
- Cristani, M., Paggetti, G., Vinciarelli, A., Bazzani, L., Menegaz, G., and Murino, V. (2011). Towards computational proxemics: inferring social relations from interpersonal distances. In *3rd IEEE International Conference on Social Computing*, pages 290–297. 21
- Cupillard, F., Bremond, F., and Thonnat, M. (2002). Tracking groups of people for video surveillance. In Jones, G. A., Paragios, N., and Regazzoni, C. S., editors, *Video-Based Surveillance Systems*. Springer. 113
- Dalal, N., Triggs, B., and Schmid, C. (2006). Human detection using oriented histograms of flow and appearance. In *In ECCV*, pages 428–441. 110
- Dautenhahn, K., Walters, M., Woods, S., Koay, K. L., Nehaniv, C. L., Sisbot, A., Alami, R., and Siméon, T. (2006). How may i serve you?: a robot companion approaching a seated person in a helping context. In *Human-Robot interaction*, pages 172–179. 28
- Du, Y., Chen, F., Xu, W., and Zhang, W. (2006). Interacting activity recognition using hierarchical durational-state dynamic bayesian network. In *Proceedings of the 7th Pacific Rim conference on Advances in Multimedia Information Processing, PCM’06*, pages 185–192, Berlin, Heidelberg. Springer-Verlag. 113

- Duffy, B. R. (2001). Towards social intelligence in autonomous robotics: A review. In *Robotics, Distance Learning and Intelligent Communication Systems 2001 (RODLICS 2001)*, pages 1–6. 31
- Efran, M. G. and Cheyne, J. A. (1973). Shared space: the cooperative control of spatial areas by two interacting individuals. *Canadian Journal of Behavioral Sciences*, 5:201–210. 19
- Ess, A., Leibe, B., Schindler, K., and van Gool, L. (2009). Moving obstacle detection in highly dynamic scenes. In *Robotics and Automation, 2009. ICRA '09. IEEE International Conference on*, pages 56–63. 111
- Feilseifer, D. and Matari, M. J. (2005). Defining socially assistive robotics. In *in Proc. IEEE International Conference on Rehabilitation Robotics (ICORR05)*, pages 465–468. 117
- Ferrari, V., Marin-Jimenez, M., and Zisserman, A. (2008). Progressive search space reduction for human pose estimation. In *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pages 1–8. 110
- Foka, A. F. and Trahanias, P. E. (2010). Probabilistic autonomous robot navigation in dynamic environments with human motion prediction. *I. J. Social Robotics*, pages 79–94. 74
- Fox, D., Burgard, W., and Thrun, S. (1997). The dynamic window approach to collision avoidance. *Robotics Automation Magazine, IEEE*, 4(1):23–33. 72
- Fulgenzi, C. (2009). *Navigation autonome en environnement dynamique utilisant des modèles probabilistes de perception et de prédiction du risque de collision*. PhD thesis, Institut National Polytechnique de Grenoble - INPG. 91, 93, 106
- Fulgenzi, C., Spalanzani, A., and Laugier, C. (2009). Probabilistic motion planning among moving obstacles following typical motion patterns. In *Intelligent Robots and Systems, 2009. IROS 2009. IEEE/RSJ International Conference on*, pages 4027–4033. 118
- Fulgenzi, C., Spalanzani, A., Laugier, C., and Tay, C. (2010). Risk based motion planning and navigation in uncertain dynamic environment. Research report, INRIA. 91, 92, 93, 106

REFERENCES

- Ge, S. S. (2007). Social robotics: Integrating advances in engineering and computer science. In *4th Annual International Conference organized by Electrical Engineering/Electronics, Computer, Telecommunication and Information Technology (ECTI) Association*. 31
- Ge, W., Collins, R., and Ruback, B. (2009). Automatically detecting the small group structure of a crowd. In *Applications of Computer Vision (WACV), 2009 Workshop on*, pages 1–8. 24, 40, 52, 53
- Gharpure, C. and Kulyukin, V. (2008). Robot-assisted shopping for the blind: issues in spatial cognition and product selection. *Intelligent Service Robotics*, 1:237–251. 26
- Gifford, R. (1983). The experience of personal space: Perception of interpersonal distance. *Journal of Nonverbal Behavior*, 7(3):170–178. 39
- Glas, D. F., Satake, S., Kanda, T., and Hagita, N. (2011). An interaction design framework for social robots. In *Robotics: Science and Systems'11*, pages –1–1. 31
- Gockley, R., Forlizzi, J., and Simmons, R. (2007). Natural person following behavior for social robots. *HRI07*. 32, 33, 37
- Goffman, E. (1963). *Behavior in public places*. Free Press. 18, 19, 32
- Grasse, R., Morère, Y., and Pruski, A. (2010). Assisted navigation for persons with reduced mobility: Path recognition through particle filtering (condensation algorithm). *J. Intell. Robotics Syst.*, 60(1):19–57. 117
- Greenberg, C. I., Strube, M. J., and Myers, R. A. (1980). A multitrait-multimethod investigation of interpersonal distance. *Journal of Nonverbal Behavior*, 5:104–114. 10.1007/BF00986513. 13
- Greenberg, S., Marquardt, N., Ballendat, T., Diaz-Marino, R., and Wang, M. (2011). Proxemic interactions: the new ubicomp? *interactions*, 18(1):42–50. 26
- Groh, G., Fuchs, C., and Lehmann, A. (2011). Combining evidence for social situation detection. In *International Confernece on Social Computing (SocialCom)*. 113
- Groh, G., Lehmann, A., Reimers, J., Friess, M., and Schwarz, L. (2010). Detecting social situations from interaction geometry. In *Social Computing (SocialCom), 2010 IEEE Second International Conference on*, pages 1–8. 22, 40

- Gulati, S. and Kuipers, B. (2008). High performance control for graceful motion of an intelligent wheelchair. In *ICRA '08*, pages 3932–3938. [116](#)
- Grin-Lajoie, M., Richards, C. L., Fung, J., and McFadyen, B. J. (2008). Characteristics of personal space during obstacle circumvention in physical and virtual environments. *Gait & Posture*, 27(2):239 – 247. [16](#), [17](#)
- Hall, E. T. (1966). *The hidden Dimension: Man's Use of Space in Public and Private*. The Bodley Head Ltd, London, UK. [14](#), [15](#), [16](#), [39](#), [118](#)
- Hansen, S. T., Svenstrup, M., Andersen, H. J., and Bak, T. (2009). Adaptive human aware navigation based on motion pattern analysis. *The 18th IEEE International Symposium on Robot and Human Interactive Communication*. [33](#), [37](#)
- Harmo, P., Taipalus, T., Knuuttila, J., Vallet, J., and Halme, A. (2005). Needs and solutions - home automation and service robots for the elderly and disabled. In *Intelligent Robots and Systems, 2005. (IROS 2005). 2005 IEEE/RSJ International Conference on*, pages 3201 – 3206. [108](#)
- Hart, P., Nilsson, N., and Raphael, B. (1968). A formal basis for the heuristic determination of minimum cost paths. *Systems Science and Cybernetics, IEEE Transactions on*, 4(2):100 –107. [72](#)
- Hauser, K. (2012). On responsiveness, safety, and completeness in real-time motion planning. *Autonomous Robots*, 32:35–48. [10.1007/s10514-011-9254-z](#). [75](#)
- Hayashi, K., Shiomi, M., Kanda, T., and Hagita, N. (2011). Friendly patrolling: A model of natural encounters. In *Robotics: Science and Systems'11*, pages –1–1. [31](#)
- Hayduk (1981a). The permeability of personal space. *Canadian Journal of Behavioural Science*, 13:274–287. [13](#)
- Hayduk, L. (1981b). The shape of personal space: An experimental investigation. *Canadian Journal of Behavioral Science*. [16](#)
- Hayduk, L. (1994). Personal space: Understanding the simplex model. *Journal of Nonverbal Behavior*, 18:245–260. [10.1007/BF02170028](#). [17](#)
- Hayduk, L. A. (1978). Personal space: An evaluative and orienting overview. *Psychological Bulletin*, 85:117–34. [14](#)

REFERENCES

- Helbing, D. and Molnar, P. (1995). Social force model for pedestrian dynamics. *PHYSICAL REVIEW E*, 51:4282. [16](#)
- Hempel, R. and Westfeld, P. (2009). Multimodal signals: Cognitive and algorithmic issues. chapter Statistical Modeling of Interpersonal Distance with Range Imaging Data, pages 137–144. Springer-Verlag, Berlin, Heidelberg. [115](#)
- Higuchi, T., Imanaka, K., and Patla, A. E. (2006). Action-oriented representation of peripersonal and extrapersonal space: Insights from manual and locomotor actions1. *Japanese Psychological Research*, 48(3):126–140. [17](#), [19](#), [45](#), [47](#)
- Hogan, K. and Stubbs, R. (2003). *Can't Get Through: Eight Barriers to Communication*. Pelican Publishing. [11](#)
- Hoogendoorn, S. P. and Bovy, P. H. L. (2004). Pedestrian route-choice and activity scheduling theory and models. *Transportation Research Part B: Methodological*, 38(2):169–190+. [111](#)
- Huettenrauch, H., Eklundh, K., Green, A., and Topp, E. (2006). Investigating spatial relationships in human-robot interaction. In *Intelligent Robots and Systems, 2006 IEEE/RSJ International Conference on*, pages 5052 –5059. [28](#)
- Ikedo, T., Chigodo, Y., Rea, D., Zanlungo, F., Shiomi, M., and Kanda, T. (2012). Modeling and prediction of pedestrian behavior based on the sub-goal concept. In *Robotics: Science and Systems*. [112](#)
- Jaillet, L., Cortes and, J., and Simeon and, T. (2010). Sampling-based path planning on configuration-space costmaps. *Robotics, IEEE Transactions on*, 26(4):635 –646. [75](#)
- Jeffrey, P. and Mark, G. (2003). Navigating the virtual landscape: Coordinating the shared use of space. In Hk, K., Benyon, D., and Munro, A. J., editors, *Designing Information Spaces: The Social Navigation Approach*, Computer Supported Cooperative Work, pages 105–124. Springer London. [32](#)
- Kahn, P. H., Freier, N. G., Kanda, T., Ishiguro, H., Ruckert, J. H., Severson, R. L., and Kane., S. K. (2008). Design patterns for sociality in human-robot interaction. *HRI08*. [31](#)
- Kanda, T., Shiomi, M., Miyashita, Z., Ishiguro, H., and Hagita, N. (2009). An affective guide robot in a shopping mall. In *Proceedings of the 4th ACM/IEEE international*

- conference on Human robot interaction*, HRI '09, pages 173–180, New York, NY, USA. ACM. [31](#)
- Karaman, S., Walter, M. R., Perez, A., Frazzoli, E., and Teller, S. J. (2011). Anytime motion planning using the rrt. In *ICRA '11*, pages 1478–1483. [75](#)
- Karnad, N. and Isler, V. (2012). Modeling human motion patterns for multi-robot planning. In *Robotics and Automation (ICRA), 2012 IEEE International Conference on*, pages 3161–3166. [112](#)
- Kavraki, L., Svestka, P., Latombe, J.-C., and Overmars, M. (1996). Probabilistic roadmaps for path planning in high-dimensional configuration spaces. *Robotics and Automation, IEEE Transactions on*, 12(4):566–580. [72](#)
- Kelley, R., Nicolescu, M., Tavakkoli, A., Nicolescu, M., King, C., and Bebis, G. (2008). Understanding human intentions via hidden markov models in autonomous mobile robots. *HRI*. [103](#), [114](#)
- Kemmotsu, K., Koketsu, Y., and Iehara, M. (2008). Human behavior recognition using unconscious cameras and a visible robot in a network robot system. *Robot. Auton. Syst.*, 56(10):857–864. [113](#)
- Kendon, A. (2010). Spacing and orientation in co-present interaction. In *Development of Multimodal Interfaces: Active Listening and Synchrony*, volume 5967 of *Lecture Notes in Computer Science*, pages 1–15. Springer Berlin / Heidelberg. [13](#), [20](#), [22](#), [39](#), [118](#)
- Kennedy, D. P., Glascher, J., Tyszka, J. M., and Adolphs, R. (2009). Personal space regulation by the human amygdala. *Nat Neurosci*, 12(10):1226–1227. [17](#)
- Khatib, O. (1985). Real-time obstacle avoidance for manipulators and mobile robots. In *Robotics and Automation. Proceedings. 1985 IEEE International Conference on*, volume 2, pages 500–505. [72](#)
- Kirby, R., Simmons, R., and Forlizzi, J. (2009). Companion: A constraint-optimizing method for person acceptable navigation. *The 18th IEEE International Symposium on Robot and Human Interactive Communication*. [33](#), [35](#), [117](#)
- Kitazawa, K. and Fujiyama, T. (2010). Pedestrian vision and collision avoidance behavior: Investigation of the information process space of pedestrians using an eye

REFERENCES

- tracker. In *Pedestrian and Evacuation Dynamics 2008*, chapter 7, pages 95–108. Springer, Berlin, Heidelberg. [18](#), [39](#), [44](#)
- Knoblauch, R., Pietrucha, M., and Nitzburg, M. (1996). Field studies of pedestrian walking speed and Start-Up time. *Transportation Research Record*, 1538(-1):27–38. [100](#)
- Knoop, S., Vacek, S., Steinbach, K., and Dillmann, R. (2006). Sensor fusion for model based 3d tracking. In *Multisensor Fusion and Integration for Intelligent Systems, 2006 IEEE International Conference on*, pages 524–529. [110](#)
- Knowles, E. S., Kreuser, B., Haas, S., Hyde, M., and Schuchart, G. E. (1976). Group size and the extension of social space boundaries. *Journal of Personality and Social Psychology*, 33:647–654. [19](#)
- Kollar, T., Tellex, S., Roy, D., and Roy, N. (2010). Grounding verbs of motion in natural language commands to robots. In *International Symposium on Experimental Robotics*. [117](#)
- Kosmatopoulos, E. (2009). An adaptive optimization scheme with satisfactory transient performance. *Automatica*, 45(3):716 – 723. [77](#), [79](#), [80](#)
- Kosmatopoulos, E. and Kouvelas, A. (2009). Large scale nonlinear control system fine-tuning through learning. *IEEE Transactions on Neural Networks*, 20(6):1009–1023. [77](#), [79](#), [80](#)
- Krishna, K. M., Alami, R., and Simeon, T. (2006). Safe proactive plans and their execution. *Robotics and Autonomous Systems*, 54(3):244 – 255. [112](#)
- Krueger, J. (2011). Extended cognition and the space of social interaction. *Consciousness and Cognition*, 20(3):643 – 657. [14](#), [19](#), [22](#)
- Kruse, T., Basili, P., Glasauer, S., and Kirsch, A. (2012). Legible robot navigation in the proximity of moving humans. In *Advanced Robotics and its Social Impacts (ARSO), 2012 IEEE Workshop on*, pages 83–88. [31](#), [33](#), [34](#)
- Kuderer, M., Kretzschmar, H., Sprunk, C., and Burgard, W. (2012). Feature-based prediction of trajectories for socially compliant navigation. In *Proceedings of Robotics: Science and Systems*, Sydney, Australia. [32](#), [33](#), [35](#), [112](#)

- Kulic and, D. and Nakamura, Y. (2010). Incremental learning of human behaviors using hierarchical hidden markov models. In *Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on*, pages 4649–4655. 112
- Kushleyev, A. and Likhachev, M. (2009). Time-bounded lattice for efficient planning in dynamic environments. In *Robotics and Automation, 2009. ICRA '09. IEEE International Conference on*, pages 1662–1668. 73
- Kuzuoka, H., Suzuki, Y., Yamashita, J., and Yamazaki, K. (2010). Reconfiguring spatial formation arrangement by robot body orientation. In *Proceedings of the 5th ACM/IEEE international conference on Human-robot interaction, HRI '10*, pages 285–292, Piscataway, NJ, USA. IEEE Press. 28
- Laga, H. and Amaoka, T. (2009). Modeling the spatial behavior of virtual agents in groups for non-verbal communication in virtual worlds. *IUCS '09*. 43
- Lam, C.-P., Chou, C.-T., Chiang, K.-H., and Fu, L.-C. (2011). Human-centered robot navigation,towards a harmoniously human-robot coexisting environment. *Robotics, IEEE Transactions on*, 27(1). 29, 33, 34
- Lamarche, F. and Donikian, S. (2004). Crowd of virtual humans: a new approach for real time navigation in complex and structured environments. *Computer Graphics Forum*, 23:509–518. 33, 34, 136
- Laugier, C., Vasquez, D., Yguel, M., Fraichard, T., and Aycard, O. (2008). Geometric and bayesian models for safe navigation in dynamic environments. *Intelligent Service Robotics*, 1:51–72. 110
- LaValle, S. and Kuffner, J.J., J. (1999). Randomized kinodynamic planning. *International Conference on Robotics and Automation*, 1:473–479 vol.1. 72, 91
- Lin, H.-Y., Chen, T.-W., Chen, C.-C., Hsieh, C.-H., and Lie, W.-N. (2009). Human pose estimation from monocular image captures. In *Multimedia and Expo, 2009. ICME 2009. IEEE International Conference on*, pages 994–997. 111
- Lindner, F. and Eschenbach, C. (2011). Towards a formalization of social spaces for socially aware robots. In *Proceedings of the 10th international conference on Spatial information theory, COSIT'11*, pages 283–303, Berlin, Heidelberg. Springer-Verlag. 25, 32

REFERENCES

- Lloyd, D. M. (2009). The space between us: A neurophilosophical framework for the investigation of human interpersonal space. *Neuroscience & Biobehavioral Reviews*, 33(3):297 – 304. [22](#), [26](#)
- Loper, M. M., Koenig, N. P., Chernova, S. H., Jones, C. V., and Jenkins, O. C. (2009). Mobile human-robot teaming with environmental tolerance. In *Proceedings of the 4th ACM/IEEE international conference on Human robot interaction, HRI '09*, pages 157–164, New York, NY, USA. ACM. [113](#)
- Luber, M., Diego Tipaldi, G., and Arras, K. O. (2011). Place-dependent people tracking. *Int. J. Rob. Res.*, 30(3):280–293. [112](#)
- Luders, B. D., Aoude, G. S., Joseph, J. M., Roy, N., and How, J. P. (2011). Probabilistically safe avoidance of dynamic obstacles with uncertain motion patterns. Technical report, MIT, ACL11-02. [75](#)
- Luo, R., Chen, Y., Liao, C., and Tsai, A. (2007). Mobile robot based human detection and tracking using range and intensity data fusion. In *Advanced Robotics and Its Social Impacts, 2007. ARSO 2007. IEEE Workshop on*, pages 1 –6. [110](#)
- Mainprice, J., Akin Sisbot, E., Jaillet, L., Cortes, J., Alami, R., and Simeon, T. (2011). Planning human-aware motions using a sampling-based costmap planner. In *Robotics and Automation (ICRA), 2011 IEEE International Conference on*, pages 5012 –5017. [75](#)
- Maisonnasse, J., Gourier, N., Brdiczka, O., and Reignier, P. (2006). Attentional model for perceiving social context in intelligent environments. In Maglogiannis, I., Karpouzis, K., and Bramer, M., editors, *Artificial Intelligence Applications and Innovations*, volume 204 of *IFIP International Federation for Information Processing*, pages 171–178. Springer Boston. [53](#)
- Marquardt, N., Diaz-Marino, R., Boring, S., and Greenberg, S. (2011). The proximity toolkit: prototyping proxemic interactions in ubiquitous computing ecologies. In *Proceedings of the 24th annual ACM symposium on User interface software and technology, UIST '11*, pages 315–326, New York, NY, USA. ACM. [26](#)
- Marshall, P., Rogers, Y., and Pantidi, N. (2011). Using f-formations to analyse spatial patterns of interaction in physical environments. In *Proceedings of the ACM 2011 conference on Computer supported cooperative work, CSCW '11*, pages 445–454, New York, NY, USA. ACM. [21](#)

- McColl, D., Zhang, Z., and Nejat, G. (2011). Human body pose interpretation and classification for social human-robot interaction. *International Journal of Social Robotics*, 3(3):313–332. [111](#)
- Mead, R., Atrash, A., and Mataric, M. J. (2011a). Proxemic feature recognition for interactive robots: automating metrics from the social sciences. In *Proceedings of the Third international conference on Social Robotics*, ICSR'11, pages 52–61, Berlin, Heidelberg. Springer-Verlag. [21](#), [115](#)
- Mead, R., Atrash, A., and Mataric, M. J. (2011b). Recognition of spatial dynamics for predicting social interaction. In *Proceedings of the 6th international conference on Human-robot interaction*, HRI '11, pages 201–202, New York, NY, USA. ACM. [114](#)
- Mehu, M. and Scherer, K. R. (2012). A psycho-ethological approach to social signal processing. *Cognitive Processing*, 13(2). [11](#), [30](#)
- Meng Keat Christopher, T. (2009). *Analysis of Dynamic Scenes: Application to Driving Assistance*. PhD in computer science, Institut National Polytechnique de Grenoble (INPG). [98](#), [112](#)
- Messac, A., Puemi-Sukam, C., and Melachrinoudis, E. (2000). Aggregate objective functions and pareto frontiers: Required relationships and practical implications. *Optimization and Engineering*, 1:171–188. [10.1023/A:1010035730904](#). [98](#)
- Michalowski, M., Sabanovic, S., and Simmons, R. (2006). A spatial model of engagement for a social robot. In *Advanced Motion Control, 2006. 9th IEEE International Workshop on*, pages 762–767. [31](#)
- Morales Saiki, L. Y., Satake, S., Huq, R., Glas, D., Kanda, T., and Hagita, N. (2012). How do people walk side-by-side?: using a computational model of human behavior for a social robot. In *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction*, HRI '12, pages 301–308, New York, NY, USA. ACM. [33](#), [37](#)
- Moratz, R., Dylla, F., and Frommberger, L. (2005). A relative orientation algebra with adjustable granularity. In *IN: PROCEEDINGS OF THE WORKSHOP ON AGENTS IN REAL-TIME AND DYNAMIC ENVIRONMENTS (IJCAI)*. [62](#)

REFERENCES

- Morgado, N., Muller, D., Gentaz, E., and Palluel-Germain, R. (2011). Close to me? the influence of affective closeness on space perception. *Perception*, 40:877 – 879. [22](#)
- Moussad, M., Helbing, D., and Theraulaz, G. (2011). How simple rules determine pedestrian behavior and crowd disasters. *Proceedings of the National Academy of Sciences*. [19](#), [91](#)
- Mumm, J. and Mutlu, B. (2011). Human-robot proxemics: physical and psychological distancing in human-robot interaction. In *Proceedings of the 6th international conference on Human-robot interaction, HRI '11*, pages 331–338, New York, NY, USA. ACM. [28](#)
- Mller, J., Stachniss, C., Arras, K., and Burgard, W. (2008). Socially inspired motion planning for mobile robots in populated environments. In *International Conference on Cognitive Systems (CogSys)*. [33](#), [37](#)
- Nakauchi, Y. and Simmons, R. (2000). A social robot that stands in line. In *Intelligent Robots and Systems, 2000. (IROS 2000). Proceedings. 2000 IEEE/RSJ International Conference on*, volume 1, pages 357 –364 vol.1. [33](#), [35](#)
- Ohki, T., Nagatani, K., and Yoshida, K. (2010). Collision avoidance method for mobile robot considering motion and personal spaces of evacuees. In *Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on*, pages 1819 –1824. [33](#), [34](#)
- Oikonomopoulos, A., Patras, I., and Pantic, M. (2005). Spatiotemporal salient points for visual recognition of human actions. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, 36(3):710 –719. [113](#)
- Olguin, D. O., Gloor, P. A., and Pentland, A. S. (2009). Capturing individual and group behavior with wearable sensors. In *AAAI Spring Symposium on Human Behavior Modeling*. [114](#)
- Oliver, N. M., Rosario, B., and Pentland, A. P. (2000). A bayesian computer vision system for modeling human interactions. *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*, 22:831–843. [114](#)
- Pacchierotti, E., Christensen, H. I., and Jensfelt, P. (2006). Design of an office-guide robot for social interaction studies. *Proceedings of IEEE/RSJ IROS*. [33](#), [35](#)

- Pacchierotti, E., Jensfelt, P., and Christensen, H. (2007). Tasking everyday interaction. In Laugier, C. and Chatila, R., editors, *Autonomous Navigation in Dynamic Environments*, volume 35 of *Springer Tracts in Advanced Robotics*, pages 151–168. Springer Berlin / Heidelberg. 31
- Pandey, A. and Alami, R. (2010). A framework towards a socially aware mobile robot motion in human-centered dynamic environment. In *Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on*, pages 5855–5860. 33, 35
- Parikh, S. P., Jr, V. G., Kumar, V., and Jr, J. O. (2005). Usability study of a control framework for an intelligent wheelchair. In *IEEE International Conference on Robotics and Automation*, page 47454750. 117
- Park, J.-E. and Oh, K.-W. (2007). Human behavior analysis for human-robot interaction in indoor environments. In Orgun, M. and Thornton, J., editors, *AI 2007: Advances in Artificial Intelligence*, volume 4830 of *Lecture Notes in Computer Science*, pages 762–768. Springer Berlin / Heidelberg. 113
- Park, S. and Aggarwal, J. (2000). Recognition of human interaction using multiple features in gray scale images. In *Pattern Recognition, 2000. Proceedings. 15th International Conference on*, volume 1, pages 51–54 vol.1. 113
- Park, S. and Trivedi, M. M. (2007). Multi-person interaction and activity analysis: a synergistic track- and body-level analysis framework. *Mach. Vision Appl.*, 18(3):151–166. 17
- Patterson, M., Iizuka, Y., Tubbs, M., Ansel, J., Tsutsumi, M., and Anson, J. (2007). Passing encounters east and west: Comparing japanese and american pedestrian interactions. *Journal of Nonverbal Behavior*, 31:155–166. 10.1007/s10919-007-0028-4. 18
- Pellegrini, S., Ess, A., Schindler, K., and Gool, L. J. V. (2009). You’ll never walk alone: Modeling social behavior for multi-target tracking. In *ICCV’09*, pages 261–268. 112
- Pepin, N., Simonin, O., and Charpillet, F. (2009). Intelligent Tiles: Putting Situated Multi-Agents Models in Real World. In AAI, A., editor, *International Conference on Agents and Artificial Intelligence - ICAART’09*, Porto, Portugal. 114
- Phillips, M. and Likhachev, M. (2011). Sipp: Safe interval path planning for dynamic environments. In *ICRA ’11*, pages 5628–5635. 74

REFERENCES

- Prest, A., Schmid, C., and Ferrari, V. (2012). Weakly supervised learning of interactions between humans and objects. *IEEE Trans. Pattern Anal. Mach. Intell.*, 34(3):601–614. [55](#)
- Quigley, M., Gerkey, B., Conley, K., Fausty, J., Foote, T., Leibs, J., Berg, E., Wheeler, R., and Ng, A. (2009). Ros: an open-source robot operating system. In *ICRA Workshop on Open Source Software*. [120](#)
- Ratsamee, P., Mae, Y., Ohara, K., Takubo, T., and Arai, T. (2012). Modified social force model with face pose for human collision avoidance. In *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction, HRI '12*, pages 215–216, New York, NY, USA. ACM. [33](#), [34](#)
- Renzaglia, A. (2012). *Optimisation stochastique et adaptative pour surveillance cooperative par une equipe de Micro-Vehicules Aeriens*. PhD thesis, Universite de Grenoble. [76](#), [77](#), [82](#), [103](#)
- Renzaglia, A., Doitsidis, L., Martinelli, A., and Kosmatopoulos, E. (2010). Cognitive-based adaptive control for cooperative multi-robot coverage. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 3314–3320. [77](#), [103](#)
- Renzaglia, A., Doitsidis, L., Martinelli, A., and Kosmatopoulos, E. (2011). Adaptive-based distributed cooperative multi-robot coverage. In *American Control Conference*, pages 468–473. [77](#)
- Rios-Martinez, J., Renzaglia, A., Spalanzani, A., Martinelli, A., and Laugier, C. (2012). Navigating between people: A stochastic optimization approach. In *Robotics and Automation (ICRA), 2012 IEEE International Conference on*, pages 2880–2885. [71](#), [76](#)
- Rios-Martinez, J., Spalanzani, A., and Laugier, C. (2011). Understanding human interaction for probabilistic autonomous navigation using Risk-RRT approach. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 2014–2019. [71](#), [117](#), [118](#)
- Ruffi, M. and Siegwart, R. (2009). On the application of the d* search algorithm to time-based planning on lattice graphs. In *ECMR'09*, pages 105–110. [73](#)
- Satake, S., Kanda, T., Glas, D. F., Imai, M., Ishiguro, H., and Hagita, N. (2009). How to approach humans. strategies for social robots to initiate interaction. *HRI*. [31](#)

- Sciutti, A., Bisio, A., Nori, F., Metta, G., Fadiga, L., Pozzo, T., and Sandini, G. (2012). Measuring human-robot interaction through motor resonance. *International Journal of Social Robotics*, 4:223–234. 29
- Scovanner, P. and Tappen, M. F. (2009). Learning pedestrian dynamics from the real world. In *ICCV'09*, pages 381–388. 112
- Sehestedt, S., Kodagoda, S., and Dissanayake, G. (2010). Robot path planning in a social context. In *Robotics Automation and Mechatronics (RAM), 2010 IEEE Conference on*, pages 206–211. 31, 33
- Shao, W. and Terzopoulos, D. (2007). Autonomous pedestrians. *Graph. Models*, 69(5-6):246–274. 111
- Sheflen, A. E. (1976). *Human territories: how we behave in space and time*. Prentice Hall. 24
- Shi, C., Shimada, M., Kanda, T., Ishiguro, H., and Hagita, N. (2011). Spatial formation model for initiating conversation. In *Robotics: Science and Systems'11*, pages –1–1. 31, 33, 36
- Siddiqui, M., Liao, W.-K., and Medioni, G. (2009). Vision-based short range interaction between a personal service robot and a user. *Intelligent Service Robotics*, 2:113–130. 10.1007/s11370-009-0041-z. 111
- Simpson, R. C. (2005). Smart wheelchairs: A literature review. *The Journal of Rehabilitation Research and Development*, 42(4):423. 117
- Sisbot, E. A., Marin-Urias, L. F., Alami, R., and Simeon, T. (2007). A human aware mobile robot motion planner. *IEEE Transactions on Robotics*, 23. 33, 37, 40
- Sisbot, E. A., Marin-Urias, L. F., Broqure, X., Sidobre, D., and Alami, R. (2010). Synthesizing robot motions adapted to human presence - a planning and control framework for safe and socially acceptable robot motions. *I. J. Social Robotics*, pages 329–343. 32
- Solea, R. and Nunes, U. (2009). Robotic wheelchair trajectory control considering user comfort. In Cetto, J., Ferrier, J.-L., and Filipe, J., editors, *Informatics in Control, Automation and Robotics*, volume 37 of *Lecture Notes in Electrical Engineering*, pages 113–125. Springer Berlin Heidelberg. 116

REFERENCES

- Sommer, R. (2002). From personal space to cyberspace, serie: Textos de psicologia ambiental, no. 1. brasilia. [15](#)
- Spalanzani, A., Rios-Martinez, J., Laugier, C., and Lee, S. (2012). *Handbook of Intelligent Vehicles*, chapter Risk Based Navigation Decisions, pages 1459–1477. Springer. [107](#)
- Spall, J. C. (1992). Multivariate stochastic approximation using a simultaneous perturbation gradient approximation. *IEEE Transactions on Automatic Control*, *37*:332–341. [77](#)
- Steinhaus, P., Strand, M., and Dillmann, R. (2007). Autonomous robot navigation in human-centered environments based on 3d data fusion. *EURASIP J. Appl. Signal Process.*, 2007(1):224–224. [73](#)
- Sukthankar, G. and Sycara, K. (2006). Robust recognition of physical team behaviors using spatio-temporal models. In *Proceedings of the fifth international joint conference on Autonomous agents and multiagent systems, AAMAS '06*, pages 638–645, New York, NY, USA. ACM. [114](#)
- Sung, J., Ponce, C., Selman, B., and SaxenaJaeyong, A. (2012). Unstructured human activity detection from rgb-d images. In *International Conference on Robotics and Automation (ICRA)*. [115](#)
- Svenstrup, M., Bak, T., and Andersen, H. (2010). Trajectory planning for robots in dynamic human environments. In *Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on*, pages 4293–4298. [33](#), [37](#)
- Svenstrup, M., Tranberg, S., Andersen, H., and Bak, T. (2009). Pose estimation and adaptive robot behaviour for human-robot interaction. In *Robotics and Automation, 2009. ICRA '09. IEEE International Conference on*, pages 3571–3576. [111](#)
- Takayama, L. and Pantofaru, C. (2009). Influences on proxemic behaviors in human-robot interaction. *Proceedings of IEEE/RSJ IROS*. [27](#), [29](#)
- Tamura, Y., Fukuzawa, T., and Asama, H. (2010). Smooth collision avoidance in human-robot coexisting environment. In *IROS'10*, pages 3887–3892. [33](#), [34](#)
- Thompson, D. E., Aiello, J. R., and Epstein, Y. M. (1979). Interpersonal distance preferences. *Journal of Nonverbal Behavior*, *4*:113–118. 10.1007/BF01006355. [13](#)

- Thompson, S., Horiuchi, T., and Kagami, S. (2009). A probabilistic model of human motion and navigation intent for mobile robot path planning. In *Autonomous Robots and Agents, 2009. ICARA 2009. 4th International Conference on*, pages 663–668. [112](#)
- Tipaldi, G. D. and Arras, K. O. (2011). Please do not disturb! minimum interference coverage for social robots. In *IROS'11*, pages 1968–1973. [31](#), [33](#)
- Topp, E. and Christensen, H. (2005). Tracking for following and passing persons. In *Intelligent Robots and Systems, 2005. (IROS 2005). 2005 IEEE/RSJ International Conference on*, pages 2321 – 2327. [32](#)
- Torta, E., Cuijpers, R., Juola, J., and van der Pol, D. (2011). Design of robust robotic proxemic behaviour. In Mutlu, B., Bartneck, C., Ham, J., Evers, V., and Kanda, T., editors, *Social Robotics*, volume 7072 of *Lecture Notes in Computer Science*, pages 21–30. Springer Berlin / Heidelberg. [28](#)
- Trautman, P. and Krause, A. (2010). Unfreezing the robot: Navigation in dense, interacting crowds. In *Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on*, pages 797–803. [112](#)
- Turner, A. and Penn, A. (2002). Encoding natural movement as an agent-based system: an investigation into human pedestrian behaviour in the built environment. *Environment and Planning B: Planning and Design*, 29(4):473–490. [19](#)
- Vasquez, D., Fraichard, T., and Laugier, C. (2009). Growing hidden markov models: a tool for incremental learning and prediction of motion. *International Journal of Robotics Research*, 28(11–12):1486–1506. [112](#), [119](#)
- Vazquez Govea, A. (2010). Growing hidden markov models. In *Incremental Learning for Motion Prediction of Pedestrians and Vehicles*, volume 64 of *Springer Tracts in Advanced Robotics*, pages 71–82. Springer Berlin / Heidelberg. [64](#)
- Vinciarelli, A., Pantic, M., Bourlard, H., and Pentland, A. (2008). Social signal processing: state-of-the-art and future perspectives of an emerging domain. In *Proceedings of the 16th ACM international conference on Multimedia*, MM '08, pages 1061–1070, New York, NY, USA. ACM. [11](#), [12](#)
- Walters, M. L., Dautenhahn, K., te Boekhorst, R., Koay, K. L., Syrdal, D. S., and Nehaniv, C. L. (2009). An empirical framework for human-robot proxemics. *Proceedings New Frontiers in Human-Robot Interaction*. [27](#)

REFERENCES

- Wang, J., Fleet, D., and Hertzmann, A. (2008). Gaussian process dynamical models for human motion. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 30(2):283–298. [112](#)
- Wang, M., Boring, S., and Greenberg, S. (2012). Proxemic peddler: a public advertising display that captures and preserves the attention of a passerby. In *Proceedings of the 2012 International Symposium on Pervasive Displays, PerDis '12*, pages 3:1–3:6, New York, NY, USA. ACM. [26](#)
- Was, J., Gudowski, B., and Matuszyk, P. J. (2006). Social distances model of pedestrian dynamics. In *Proceedings of the 7th international conference on Cellular Automata for Research and Industry, ACRI'06*, pages 492–501, Berlin, Heidelberg. Springer-Verlag. [111](#)
- Yamaoka, F., Kanda, T., Ishiguro, H., and Hagita, N. (2009). Developing a model of robot behavior to identify and appropriately respond to implicit attention-shifting. In *Proceedings of the 4th ACM/IEEE international conference on Human robot interaction, HRI '09*, pages 133–140, New York, NY, USA. ACM. [36](#)
- Yamaoka, F., Kanda, T., Ishiguro, H., and Hagita, N. (2010). A model of proximity control for information-presenting robots. *Robotics, IEEE Transactions on*, 26(1):187–195. [33](#), [36](#)
- Yamazaki, K., Kawashima, M., Kuno, Y., Akiya, N., Burdelski, M., Yamazaki, A., and Kuzuoka, H. (2007). Prior-to-request and request behaviors within elderly day care: Implications for developing service robots for use in multiparty settings. In *ECSCW'07*, pages 61–78. [31](#)
- Yanco, H. and Drury, J. (2004). Classifying human-robot interaction: an updated taxonomy. In *Systems, Man and Cybernetics, 2004 IEEE International Conference on*, volume 3, pages 2841–2846 vol.3. [29](#)
- Yuan, F., Twardon, L., and Hanheide, M. (2010). Dynamic path planning adopting human navigation strategies for a domestic mobile robot. In *Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on*, pages 3275–3281. [113](#)
- Zender, H., Jensfelt, P., and Kruijff, G.-J. (2007). Human- and situation-aware people following. In *Robot and Human interactive Communication, 2007. RO-MAN 2007. The 16th IEEE International Symposium on*, pages 1131–1136. [33](#), [37](#)

- Zhu, C. and Sheng, W. (2009). Human daily activity recognition in robot-assisted living using multi-sensor fusion. In *Robotics and Automation, 2009. ICRA '09. IEEE International Conference on*, pages 2154 –2159. 114
- Zhu, Q., Yeh, M.-C., Cheng, K.-T., and Avidan, S. (2006). Fast human detection using a cascade of histograms of oriented gradients. In *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, volume 2, pages 1491 – 1498. 110