Using social cues to estimate possible destinations when driving a robotic wheelchair — Source link ↗

Arturo Escobedo, Anne Spalanzani, Christian Laugier

Institutions: French Institute for Research in Computer Science and Automation, University of Grenoble

Related papers:

- Robotic wheelchair based on observations of both user and environment

- Human-Oriented Design of Autonomous Navigation Assisted Robotic Wheelchair for Indoor Environments

- A model-based, probabilistic framework for plan recognition in shared wheelchair control: experiments and evaluation

- Multi-layered map based navigation and interaction for an intelligent wheelchair

- Navigation of an autonomous wheelchair in a structured environment

Using social cues to estimate possible destinations when driving a robotic wheelchair

Jesús Arturo Escobedo Cabello, Anne Spalanzani, Christian Laugier

Using social cues to estimate possible destinations when driving a robotic wheelchair

Arturo Escobedo[1], Anne Spalanzani[2], Christian Laugier[1] [†]
[1]INRIA Rhône-Alpes, Grenoble, France
[2]Univ. Grenoble Alpes, Lab. LIG, Grenoble, France. Inria

Abstract— **Approaching a group of humans is an important navigation task. Although many methods have been proposed to avoid interrupting groups of people engaged in a conversation, just a few works have considered the proper way of joining those groups. Research in the field of social sciences have proposed geometric models to compute the best points to join a group. In this article we propose a method to use those points as possible destinations when driving a robotic wheelchair. Those points are considered together with other possible destinations in the environment such as points of interest or typical static destinations defined by the user's habits. The intended destination is inferred using a Dynamic Bayesian Network that takes into account the contextual information of the environment and user's orders to compute the probability for each destination.**

Index Terms— **Intention estimation, human aware navigation, semi-autonomous navigation, robotic wheelchair, face control, voice control.**

I. Introduction

Ensuring proper living conditions for an ever growing number of elderly people is a significant challenge in many countries. In this context, a very relevant application is to assist people with reduced mobility to preserve some of their independence.

Electric wheelchairs operated by a joystick are the most common mobility assistance device. However, elderly people often spend a large period of time learning to operate accurately those devices and even expert drivers loose the freedom of one of his hands in order to manipulate the joystick. This restricts the user's capacity to perform other task such as holding things, opening a door etc. Under the consideration that it is a natural human behavior to look where we are going [1], in this paper we propose the use of the direction of the face together with a voice recognition system as input methods to drive the robotic wheelchair.

A safe movement must be guaranteed by this type of robotic wheelchairs specially because they are intended to be used by fragile persons. Other important aspects that must be considered in the design are: The system must be compliant with the user's intentions, it has to behave in a socially acceptable way and it must interact in an intuitive way with the driver.

Our approach is based on and extends the work presented in [2]. The destination inference method was improved by considering non-static destinations to join groups of people in a socially acceptable way. The notion of personal space and interaction space is used by the navigation system to prevent uncomfortable situations when the wheelchair is moving among humans [3].

We propose to estimate the user's intention in order to reduce the number of necessary commands to drive the robotic wheelchair and deal with ambiguous or inaccurate input interfaces. In this way, the wheelchair can be in charge of some part of the navigation task and alleviate the user involvement.

The problem is approached from the hypothesis that it is possible to learn typical destinations where the user spends most of his time and uses this information to predict his desired goal when driving the robotic wheelchair. Those typical destinations may be learned and then used to predict motion on the basis of a user model.

A probabilistic framework is used to model the existent relationship between the intention of the user and the observed command. The main originality of the approach lies in modeling explicitly the intentions as typical destinations and then use it to check the reliability of the user's command to decide how much preeminence it should be assigned in the final robot control.

The structure of this paper is as follows: Section II offers an overview of related works. A general description of the system architecture is presented in III; section III-A focuses on the estimation of intended destination and section III-B explains the module to detect social interactions and compute adequate meeting points to be taken as possible destinations. In section IV our experimental setup is presented and an analysis of the results is provided. Section V presents conclusions about the work and perspectives for future improvements of the system.

II. Related Work

A. Robotic Wheelchair Navigation Systems

Research in robotic wheelchairs can be classified in three different categories; robotic wheelchairs with fully autonomous navigation, semi-autonomous wheelchairs and manually controlled wheelchairs with collision avoidance.

Robotic wheelchairs with fully autonomous navigation are normally designed as mobile robots equipped with a chair. Those systems work so that the user gives a final destination

and supervises as the wheelchair is in charge of the complete navigation task [4]. In many cases, the user expresses explicitly his intended destination by pointing it into a map or choosing it from a menu. This type of interaction is not very natural and sometimes can be complicated for the user (specially elder users).

Semi-autonomous systems use the estimation of the user's plan as a key point because it allows the automatic controller to adjust its actions to the desire of its user. Some methods to perform an implicit estimation of the user's intention from joystick commands have been proposed in [5], [6]. They model the user's intention as a set of possible trajectories. A probability distribution is maintained over the set of trajectories and the most likely one is selected within a Bayesian framework.

In [7] a learned Partially Observable Markov Decision Process (POMDP) was used to estimate the user's intended destination in a topological map of the environment. Places of interest are selected as those locations in the environment where the user spends comparatively most of his time. The user drives the wheelchair from one spatial location to another while the robotic device avoids obstacles in the middle.

Other smart wheelchairs limit their assistance to collision avoidance and leave the majority of planning and navigation duties to the user. These systems do not normally require prior knowledge of an area or any specific alterations to the environment. They require instead more planning and continuous effort on the part of the user and are only appropriate for users who can effectively plan and execute a path to a destination [8].

B. Human Aware Navigation

Hall [9] classified human interactions based on a concept of distance, creating the concept of "public distance" to refer to situations in which people give a speech, and "social distance" that characterizes situations in which people talk to each other for the first time. Our approach uses the notion of social distances to find an appropriate location to place the wheelchair when the user wants to interact with the other people in a respectful and comfortable way.

[10] proposes a model of approaching behavior with which a robot can initiate a conversation with people who are walking. To prevent failures, their model includes a prediction of the walking behavior of people, choosing a target person, planning its approaching path, and nonverbally indicating its intention to initiate a conversation.

Other work, [11], introduced an adaptive system which detects whether a person seeks to interact with the robot based on the person's pose and position, that system was presented as a basis for human aware navigation. Their results showed that the system was capable of navigating according to past interaction experiences and to adapt to different behaviors.

A method for detecting and tracking groups of people was presented in [12]. They address the problem of detecting and learning socio-spatial relations between individuals and
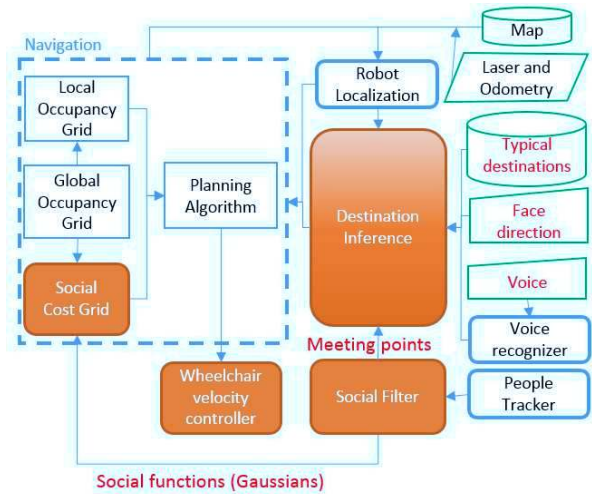

Fig. 1. System Architecture Overview

tracking their group formations using multiple social grouping hypotheses in a recursive way. This method also improves person-level tracking in two ways: the social grouping information is fed back to predict human motion over learned intra-group constraints and to support data association by adapting track-specific occlusion probabilities.

Personal space, o-space and their relation to comfort were addressed in [3] where a risk based navigation was extended to include risk due to discomfort. Human movement is supposed to be known by learning of typical trajectories in a particular environment.

III. SYSTEM ARCHITECTURE

Fig. 1 presents an overview of the complete system. Several subsystems were developed by our team (shown in orange) while the other necessary modules were taken from the open source community (white blocks).

Destination Inference [2]: This subsystem estimates the desired destination within the map of the environment among a list of possible predefined goals and *meeting points* computed by the social filter. The probability for each destination is computed using a Dynamic Bayesian Network that considers the current position of the wheelchair, user's face direction and the initial voice command. The destination with the highest probability is selected and sent to the navigation module.

Social Filter [3]: Detects social interactions and creates virtual obstacles corresponding to those interaction zones in order to produce socially acceptable motion. The social filter integrates constraints inspired by social conventions to evaluate the risk of disturbance and takes it into account when performing the autonomous navigation planning.

Social Grid: Receives information from the social filter and projects it to an occupancy grid adequate to be used by standard motion planning algorithms.

Meeting Points: The social filter uses geometrical information from groups of people to compute good places for the wheelchair to be placed in order to join the group.
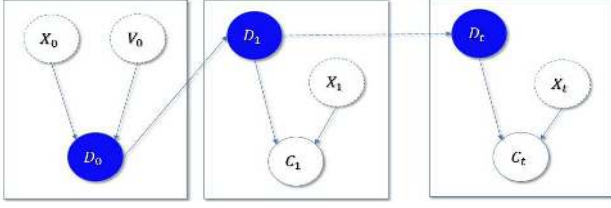
Fig. 2. Dynamic Bayesian Network used to estimate the user intended destination D_t.

Navigation: The navigation subsystem includes an A* motion-planner and Dynamic Window controller to compute safe trajectories to drive the wheelchair.

A. Destination Inference

The user's intention is modeled as a set of possible destinations. Those destinations may be defined by the user's habits (places where the user spends most of his time during the day), interesting points taken from the map of the environment as doors, desks and other facilities and dynamic destinations related to humans in the scene -For more details consider reading our previous work [2]. In this work we propose a method to detect and use destinations defined by social groups of people.

The presented reasoning method is based on a dynamic Bayesian network described in Fig. 2.

The joint probability distribution that corresponds to the graph (Fig. 2) is:

$$P(D_{0:t}, C_{0:t}, X_{0:t}) = P(X_0)\ P(V_0)\ P(D_0|X_0, V_0)*$$
$$\prod_{k=1:t} [P(D_k|D_{k-1})P(C_k|D_k, X_k)\ P(X_k)] \quad (1)$$

Where, X_t is the position of the wheelchair in a 2D plane with respect to the map reference frame. This position is discrete and its domain is defined by a grid whose size is according to the map of the environment. It is assumed that the position of the wheelchair X_t is a fully observable measurement given by the localization system.

V_0 is the starting vocal command. This variable can take two possible values $\{GO, JOIN\}$. C_t represents the direction pointed by the user at time t. This direction is centered on the user's local frame.

The variable D represents both the static destinations and the meeting points. The domain of the D variable (denoted as Ω_D) is defined as the union between the set of static destinations Ω_S and the set of meeting points Ω_M.

$$\Omega_D = \Omega_S \vee \Omega_M \quad (2)$$

Where, Ω_D is the domain of the D variable. $\Omega_s = \{d^{(1)}, d^{(2)}, ..., d^{(N)}\}$ is the set of N static destinations. $\Omega_m = \{d^{(N+1)}, d^{(N+2)}, ..., d^{(N+M)}\}$ is the set of M meeting points.

The inference question to be answer at time t is $P(D_t = d^{(i)}|C_{1:t}, X_{0:t}, V_0)$. It can be recursively computed Using Bayes rule as:

$$P(D_t^{(i)}|C_{01t}, X_{0:t}, V_0) = \alpha P(X_{0:t})\ P(V_0)*$$
$$P(D_0|X_0, V_0)\ P(C_t|X_t, D_t^{(i)})*$$
$$\sum_j [P(D_t^{(i)}|D_{t-1}^{(j)})P(D_{t-1}^{(j)}|C_{1:t-1}, X_{0:t-1}, V_0)] \quad (3)$$

The term $P(D_0|X_0, V_0)$ will be referred as the initial probability model, the term $P(C_t|X_t, D_t^{(i)})$ is the *command model*, $P(D_t^{(i)}|D_{t-1}^{(j)})$ is the transition model.

The initial probability model has two different modalities, when $V_0 = GO$ the value of this term is encoded in a probability table [2], whose values are obtained by learning typical destinations in the environment according to the user's habits. However , when $V_0 = JOIN$ the meeting points get an evenly distributed high probability while all the others get low values.

$$P(D_0^{(i)}|X_0, V_0 = GO) = P(D_0^{(i)}|X_0) \quad (4)$$

$$P(D_0^{(i)}|X_0, V_0 = JOIN) =$$
$$\begin{cases} \alpha/M & \text{if } d^{(i)} \in \Omega_m \\ (1-\alpha)/N & \text{otherwise} \end{cases} \quad (5)$$

In our tests we used $\alpha = 0.9$.

$P(C_t|X_t, D_t^{(i)})$ represents the probability that a command C_t will be sent by the user when he is located at position X_t and his destination is $d^{(i)}$ at current time t.

$$P(C_t|X_t, D_t^{(i)}) = \frac{P'(C_t|X_t, D_t^{(i)})}{\sum_i P'(C_t|X_t, D_t^{(i)})} \quad (6)$$

where,

$$P'(C_t|X_t, D_t^{(i)}) = \frac{1}{\sigma\sqrt{2\pi}} \exp^{-\frac{1}{2}(\frac{a_i}{\sigma})^2} \quad (7)$$

The a_i term is the angle between the command and the destination (the value of the standard deviation σ was set from experimental data to have a value of 0.2).

$P(D_t^{(i)}|D_{t-1}^{(j)})$ expresses the probability that the current estimated destination changes respect the last one.

$$P(D_t^{(i)}|D_{t-1}^{(j)}) =$$
$$\begin{cases} (gain-1)/(N+gain-1) & \text{if } D_t^{(i)} = D_{t-1}^{(j)} \\ 1/(N+gain-1) & \text{otherwise} \end{cases} \quad (8)$$

Where, N is the number of possible goals in the environment. $gain$ expresses how important the previous estimation will be with respect to the new data. Large values of this term will lead to slow responses whenever the user changes his intended destination. After practical experimentation it was defined a value of $P(D_t^{(i)}|D_{t-1}^{(j)})$ that is 10 times bigger if the last estimated goal was the same.

The selection of the goal is performed using exact inference. The selected destination is the one with highest probability.

$$d_{sel} = \arg\max_{d^{(i)}} P(D_t = d^{(i)} | C_t, X_t) \qquad (9)$$

This goal is then sent to the autonomous navigation system when the user gives one of the vocal commands "GO" or "JOIN".

B. Social conventions in human navigation

In normal situations, people try to maintain a minimum space between them to feel comfortable. Hall [9] modeled the level of discomfort experienced by the person related to the distance at which he is approached. This idea was formalized as the concept of *personal space*, which characterizes the space around a human being in terms of comfort during social activities. Similarly, interactions between two or more people have been modeled using the concept of *o-space* by [13].

1) Personal Space: The implemented model consists in blending two Gaussian functions Γ_f and Γ_b (Eq.10) both of them centered in the position of the person. The first one represents the personal space in front of a human and is 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.

$$\Gamma_{x,\Sigma}(p) = e^{-\frac{1}{2}(p-x)^t \Sigma^{-1}(p-x)} \qquad (10)$$

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

$$\Sigma = \begin{pmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{pmatrix}. \qquad (11)$$

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.

2) O-Space: When two or more people are in a conversation, they tend to adopt formations with circular shapes. The o-space could be taken as a circle whose center coincides with that of the inner space. For the specific case of two people, some formations, called F-formations, have been identified as being particularly frequent [13]. The social filter identifies individual F-formations (Vis-a-vis or V-Shape) and builds the corresponding o-space. In Fig. 3, the calculated o-space for a Vis-a-Vis interaction is shown.
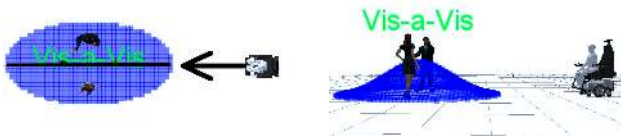

Fig. 3. O-space calculated by the Social Filter Module for a Vis-a-Vis formation. The maximum risk of disturbance is located at o-space center, in the picture the disturbance is represented by the Gaussian function height.

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:

$$\Gamma_{C,S}(Q) = e^{-\frac{1}{2}(Q-C)^t S^{-1}(Q-C)} \qquad (12)$$

where S is a diagonal covariance matrix as defined in Eq.11. 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	$D_H/3$	$2 * DH/3$
V-Shape	$D_H/3$	$2D_i$

C. Social Cost Grid

Many planning algorithms use occupancy grids as its basic data structure to represent the state of the world. Practical implementations of such algorithms typically used in robot navigation systems (in this work A* and dynamic window) consider occupancy grids.

In order to adequate the data computed by the social filter we project the values of the personal space and interaction space into a 2D occupancy grid. Two different thresholds are used as shown in Fig. 4. Cells in the occupancy grid with personal and/or interaction space values higher than the threshold are marked as occupied zones so the planner will avoid passing through those regions, while regions with lower values are marked as free space.

Occupied regions computed in the social grid are then combined with the occupancy grid created with the robot's laser range finder and the map of the environment. In order to consider the size of the wheelchair the regions marked as occupied are inflated by an 'inflation radius' according to the size of the robotic wheelchair (Fig. 4(b)).
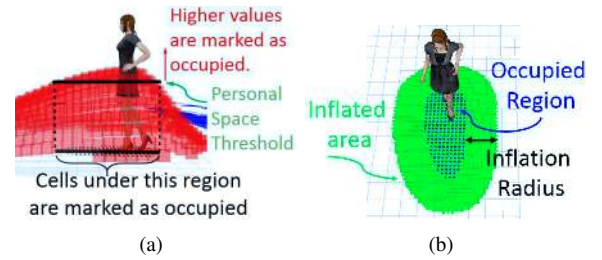

Fig. 4. Threshold applied to the Gaussian functions delivered by the social filter to created the occupancy-cost grid. (b) The inflated area (green) considers the dimensions of the wheelchair. In order to guarantee a path free of collisions the center of the wheelchair should never be inside this area.
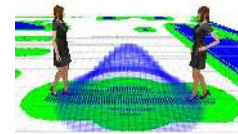

Fig. 5. Projection of both p-space and o-space on the occupancy grid used by the planning algorithm. The Gaussian in the middle represents the value of the interaction space, which is projected on the occupancy grid as occupied space (blue area) and then inflated to consider the size of the robot (green area).

D. Computing Meeting Points

A geometrical model is used to determine those places where the robotic wheelchair should be placed to attract the attention and become part of an interacting group. Those points are located in an area, roughly coincident with the o-space of an interaction where the robot can share the space in an equitable way with the other humans already present.

Fig. 6 shows different types of social formations and where the meeting points should be located. Fig. 6(a) represents a vis-vis or frontal formation where two humans H_1 and H_2 are talking face to face. In this case two meeting points are placed on the line perpendicular to the H_{12} line that joins H_1 and H_2. the points are located at a distance $D_H/2$ from the center of the o-space. D_H is the distance between the two persons.

In the case of a v-shape formation Fig. 6(b) a meeting point is placed on the line formed by V_i and H_{12}. V_i is the intersection of the two lines of sight of interacting humans. The computed meeting points are used by the inference algorithm as possible destinations as it will be described later.
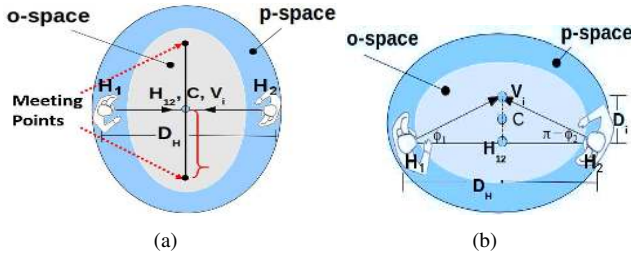

Fig. 6. Geometrical placement of meeting points in different social formations. (a) For a vis-vis formation. (b) For a v-shape formation. The location of Meeting points depends on the focus of attention defined by the orientation of the bodies ϕ.
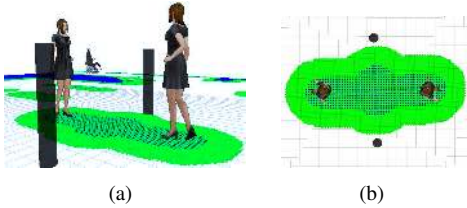

Fig. 7. The resulting meeting points for a vis-vis formation (black markers). If the robot reaches that position, a group of three will be formed. (a) Lateral view. (b) Top view.

IV. SIMULATION EXPERIMENTS

The proposed approach was designed considering the experimental scenario shown in Fig. 8. People in the scene were tracked to learn the typical destinations that are then placed in the map of the environment (red circles in Fig. 8(a)). Each destination has a related probability value as described in section III-A. The values for the presented experimental scenario where presented in a previous publication [2].

The user can start the movement at any location of the experimental scenario, he drives the wheelchair by seeing towards his desired destination and saying a vocal command as JOIN or GO.
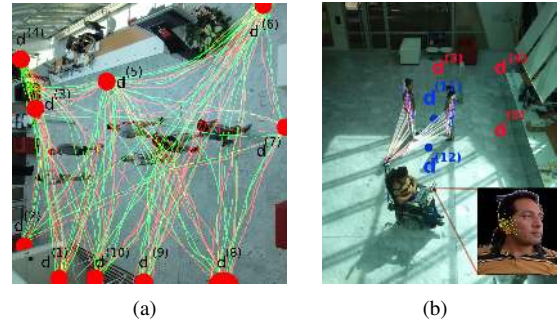

Fig. 8. Possible experimental scenario is located at INRIA main entry hall where typical destinations have been detected in previous works [2]. (a) The lines in the figure represent some of the trajectories normally followed by people in the scene. Typical destinations (end of a trajectory) are marked with circles. (b) In the intended application, the position of the user's face and the vocal command are used to infer his desired destination. Social formations are detected and the corresponding meeting points (d(10), d(11)) are considered as possible destinations by the destination inference algorithm.

A. Face Control Subsystem

The user can control the robotic wheelchair by using the movements of his face. The direction of the face is estimated by a random forest classifier which takes as input a 3D point cloud (for further information on this subject please refer to [15]).

B. Voice Control Subsystem

As it is impossible to deduce the full range of user intentions only analyzing the user's face direction; it is necessary the system to be multi-modal. For this reason, we used the Google voice recognition service with a reduced set of instructions -related to our application- to give those commands that it would be difficult to express using only the face as: stop, move, faster, etc.

Some words have an strong context related to the type of destination that they can be applied to, for example the verb GO is more general than the verb JOIN which means that it does not give so much information to infer the user's desired destination. However, when the user uses the keyword JOIN, it is more likely that he intends to approach a group of people if there is any present on the scene. This behavior is encoded in the probability distribution of Eq. 5.

C. Multimodal Interface

In the example (Fig. 9), the user is looking in the direction of the arrow (blue) and there are two people in the middle of the simulated INRIA-hall. The task of the wheelchair is to navigate towards the goal that gets the highest posterior probability.

When the user gives the vocal command to start the movement (GO or JOIN) the prior probabilities are loaded and the posterior probability is computed considering the direction of his face, the probability for each goal is depicted in Fig. 9 as the size of the sphere. The goal with the highest posterior probability is sent to the navigation module, the wheelchair plans the path and compute the necessary
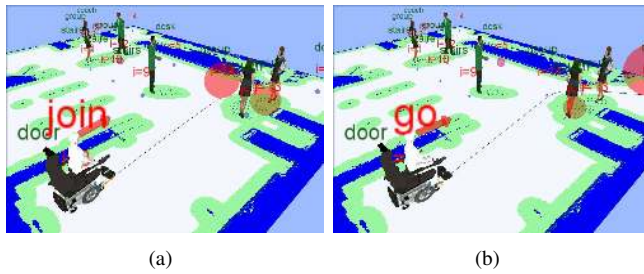
(a)   (b)

Fig. 9.   In this example we show the different modalities that can be achieved by our method. In the first example (a); given the strong context related to the use of the verb JOIN the probabilities are computed in such a way that those destinations defined by a group of people get a higher value. However the use of the verb GO, (b); does not give such contextual information; therefore, the destination that gets the maximum probability is mainly defined by the direction in which the user is pointing and the prior probability model.

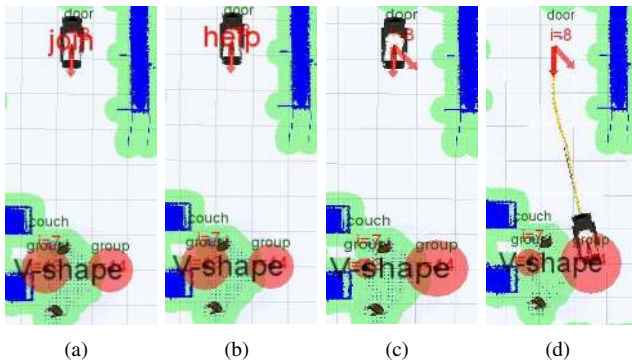(a)   (b)   (c)   (d)

Fig. 10.   Error handling. When the system detects an error in the inferred destination (it is blocked or there is not enough evidence to decide) (a) it asks for help to the user (b). The user gives a new command (c) and if the new destination is valid it will move towards it.

velocity commands to arrive to the destination while avoiding obstacles and respecting the personal and interaction spaces.

The destination inference method considers both meeting points defined by the two persons in vis-vis formation. In the first case Fig. 9(a): the user gives a JOIN order. Given the strong context related to this verb the probabilities are computed in such a way that the meeting points get a higher probability value than all the other possible destinations.

In the second case Fig 9(b) the user gives a GO order which does not give as much contextual information as the JOIN command, so the probability distribution is more evenly distributed, therefore, the destination that gets the maximum probability value is mainly defined by the direction in which the user is looking.

In case of an error, for example when the inferred destination can not be reached because it is occupied or there is not enough evidence to decide which destination is better. The system will ask for help and wait for the user to give a new command. In the example Fig. 10 the system selects a meeting point that is not reachable by the wheelchair, then it asks for help and waits for the user to give a new command. When the new command is considered a new destination could be selected and if the new destination is valid the wheelchair will start moving.

V. Conclusions and Future Work

The approach presented in this paper integrates a human aware cost grid with an standard autonomous navigation controller. The system was designed to improve both usability taking advantage of the user intention inference method and sociability by including the concepts of personal space and o-space in the navigation system.

The main contributions presented in this article were; first of all, the method to compute socially accepted points to join a group of people and then use them as possible destinations for the user. Second, the use of information from several user-machine interfaces - voice and face position-, to reduce ambiguities in the inference model.

References

[1] D. Bernardin, H. Kadone, D. Bennequin, T. Sugar, M. Zaoui, and A. Berthoz, "Gaze anticipation during human locomotion." *Experimental brain research*, Sept. 2012.

[2] J. A. Escobedo Cabello, A. Spalanzani, and C. Laugier, "Multimodal Control of a Robotic Wheelchair: Using Contextual Information for Usability Improvement," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, Nov. 2013.

[3] J. Rios-Martinez, A. Spalanzani, and C. Laugier, "Understanding human interaction for probabilistic autonomous navigation using Risk-RRT approach," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, Sept. 2011, pp. 2014–2019.

[4] Y. Wang and W. Chen, "Hybrid map-based navigation for intelligent wheelchair." in *ICRA*.   IEEE, 2011, pp. 637–642.

[5] E. Demeester, A. Huntemann, D. Vanhooydonck, G. Vanacker, A. Degeest, H. Brussel, and M. Nuttin, "Bayesian Estimation of Wheelchair Driver Intents: Modeling Intents as Geometric Paths Tracked by the Driver," *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*, no. Iros, pp. 5775–5780, Oct. 2006.

[6] E. Demeester, A. Huntemann, J. del R. Millan, and H. Van Brussel, "Bayesian plan recognition for Brain-Computer Interfaces," *2009 IEEE International Conference on Robotics and Automation*, pp. 653–658, May 2009.

[7] T. Taha, J. V. Miro, and G. Dissanayake, "POMDP-based long-term user intention prediction for wheelchair navigation," *2008 IEEE International Conference on Robotics and Automation*, pp. 3920–3925, May 2008.

[8] G. Vanacker, J. del R Millán, E. Lew, P. W. Ferrez, F. G. Moles, J. Philips, H. Van Brussel, and M. Nuttin, "Context-based filtering for assisted brain-actuated wheelchair driving." *Computational intelligence and neuroscience*, vol. 2007, p. 25130, Jan. 2007.

[9] E. T. Hall, *The hidden Dimension: Man's Use of Space in Public and Private*.   The Bodley Head Ltd, London, UK, 1966.

[10] S. Satake, T. Kanda, D. F. Glas, M. Imai, H. Ishiguro, and N. Hagita, "How to Approach Humans?-Strategies for Social Robots to Initiate Interaction-," *J. Robot. Soc. Japan*, vol. 28, no. 3, pp. 327–337, 2010. [Online]. Available: http://joi.jlc.jst.go.jp/JST.JSTAGE/jrsj/28.327?from=CrossRef

[11] S. T. Hansen, M. Svenstrup, H. J. Andersen, and T. Bak, "Adaptive human aware navigation based on motion pattern analysis," *The 18th IEEE International Symposium on Robot and Human Interactive Communication*, 2009.

[12] M. Luber and K. O. Arras, "Multi-Hypothesis Social Grouping and Tracking for Mobile Robots," in *Robot. Sci. Syst.*, 2013.

[13] A. Kendon, "Spacing and orientation in co-present interaction," in *Development of Multimodal Interfaces: Active Listening and Synchrony*, ser. Lecture Notes in Computer Science.   Springer Berlin / Heidelberg, 2010, vol. 5967, pp. 1–15.

[14] J. Rios-Martinez, "Navigation de robots avec conscience social: entre l'evaluation des risques et celle des conventions sociales," THESE, Université de Grenoble, Jan. 2013. [Online]. Available: http://hal.inria.fr/tel-00837525

[15] G. Fanelli, J. Gall, and L. Van Gool, "Real time head pose estimation with random regression forests," *Computer Vision and Pattern Recognition*, 2011.