Intention Driven Human Aware Navigation for Assisted Mobility

Jorge Rios-Martinez, Arturo Escobedo, Anne Spalanzani, Christian Laugier

Intention Driven Human Aware Navigation for Assisted Mobility

Jorge Rios-Martinez[1], Arturo Escobedo[1], Anne Spalanzani[2], Christian Laugier[1] †
[1]INRIA Rhone-Alpes, Grenoble, France
[2]UPMF-Grenoble 2 - INRIA - LIG, France

***Abstract*— Ensuring proper living conditions for an ever growing number of elderly people is an important 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 particular, this paper studies the case of a robotic wheelchair, specifically its autonomous navigation and user adapted control. Integration of a technique to interpret user intention using head movements and a human aware motion planner is presented. Test results exhibit emerging behavior showing a robotic wheelchair interpreting gesture commands and taking the user to his desired goal, respecting social conventions during its navigation.**

***Index Terms*— Proxemics, Human aware navigation, User intention, Adapted control.**

I. Introduction

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 as it can help them to preserve their independence. For such systems, it is crucial to take into account the actual needs and characteristics of both its users and the people around them. This paper studies the case of a robotic wheelchair, specifically its autonomous navigation and user adapted control, whose operation has been designed around the following requirements:

- Usability: People with motor disabilities or aging people often have problems using joysticks and other standard control devices. The system should account for this, for example by favoring the most reasonable actions when presented with an ambiguous command.
- Safety: The system should avoid collisions with both static and dynamic entities.
- Sociability: When moving, a robot may considerably disturb people around it, especially when its behavior is perceived as unsocial. Even worse, the wheelchairs passenger may be held responsible for that behavior. It is thus important to produce socially acceptable motion.

Social capability of planner chosen is based on the simple idea that, in a human populated environment, when people interact, they often adopt spatial formations implicitly forming “interaction zones”. Thus, socially acceptable motion can be enforced not only by respecting personal space but also by detecting interaction zones and then computing the risk to invade them.

Usability can be improved by adding contextual information in order to ease interaction with the user, for example, the knowledge of interesting or frequently visited locations in a particular environment can be used by the system to infer the user's plan. Once the plan is identified the system can assist the user by executing the low level needed commands. The structure of this paper is as follows:

Section II offers an overview of related works. Section III presents our technique for achieving user adapted control. Section IV describes the RiskRRT method. In section V examples of execution on our real platform are exhibited. Section VI presents conclusions about the work and perspectives.

II. Related work

A. Semi-Autonomous Navigation

In latest years many efforts have been made to develop robotic wheelchairs that operate in a similar manner to an autonomous robot, where the user gives a final destination and supervises as the smart wheelchair moves (e.g.,NavChair [1], MIT Media Lab wheelchair [2]).

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.

Shared control is presented in situations in which the assisting device combines the control input coming from the robot and the user. This device may be a wheelchair, a teleoperated robot, a robotic travel aid for the visually impaired, or any other device where robot and human cooperate in a task [3].

The estimation of the user's plan is a key point in many shared control tasks because it allows to the automatic controller/robot to adequate its actions to the desire of its user. Inferring the user plan is necessary whenever the interface with the user doesn't allow him to explicitly dictate this to the robot as with many popular electric wheelchair interfaces.

Some methods aiming the implicit estimation of the users intention from simple joystick inputs have been proposed in [3], [4]. They model the users intent as possible trajectories

† {jorge.rios-martinez, jesus.escobedo-cabello, anne.spalanzani, christian.laugier}@inria.fr

Jorge Rios-Martinez has been partially funded by CONACYT 250140/308006.

to follow, then a probability distribution is maintained over the set of trajectories and finally the selection of the most probable one is done using the input from the user within a Bayesian framework.

In [5] a learned Partially Observable Markov Decision Process (POMDP) is used to estimate the intended destination into a predefined map of the environment in a high level topological manner. This means that the user focuses on driving the wheelchair from one spatial location to another without having to worry about all the low level control. Places of interest are selected as spatial locations in the environment where the user spends significantly most of his time.

The method presented in this article to infer the user's intended goal aims to build a model as simple as possible combining the Bayesian network approach of the first mentioned methods and the simple topological goal based representation of the environment used in the second one. A more natural human-machine interface based on a face tracking system is used to command the wheelchair (III-A) while the navigation is performed using a human aware planning algorithm (IV).

B. Human Aware Navigation

Human aware navigation is receiving an increasing attention in robotics community, this area of research appears once that robots navigate in human environments and safety solutions are not enough; now the main concern is related to produce solutions which also have to be understandable and acceptable by human beings. Next is a review of related works. A proposal of human aware navigation was presented in [6], where a motion planner takes explicitly into account its human partners. The authors introduced the criterion 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. Other work, [7], 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 navigate based in past interaction experiences and to adapt to different behaviors.
In [8] it was proposed a Spatial Behavior Cognition Model (SBCM) to describe the spatial effects existing between human-human and human-environment. SBCM was used to learn and predict behaviors of pedestrians in a particular environment and to help a service robot to take navigation decisions. Technique in [9] proposed an on-line method to learn generally occurring motion patterns in an office environment with a mobile robot. Navigation is realized by using these patterns, in form of sampled HMM, along with a Probabilistic Roadmap based path planning algorithm. Socially acceptable motion is achieved by minimizing social distractions, such as going through someone else's working space.

The work presented in [10] 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.

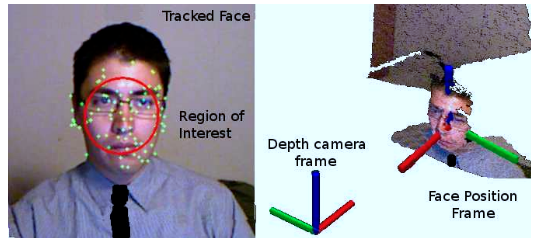


Fig. 1. In the presented approach the user drives the wheelchair by using his face. The face of the user is tracked by processing the RGB image received from a Kinect sensor (left image). The pose of the face is estimated from the depth data as shown in the right.

Rules define sensitive zones for both humans and robots, depending either on their security regions or on psychological feeling of humans.

Personal space, o-space and their relation to comfort were addressed in [11], where a risk based navigation was extended to include risk due to discomfort. Human's movement is supposed to be known by learning of typical trajectories in a particular environment. Optimization techniques that take into account social factors have been also proposed. In [12] a generalized framework for representing social conventions as components of a constrained optimization problem was presented and it was used for path planning, their results exhibited a more social navigation. In [13] an stochastic adaptive optimization method was used to minimize discomfort of humans in the environment, while robot navigate to the goal. Results show robot navigation respecting both information process space and personal space of people. Recently, legibility of robot navigation around humans was explored in [14]. A context depending cost model was developed to adjust robot behavior to human behavior in crossing scenarios.

III. Semi Autonomous Controller

A. Face Control Subsystem

The user controls the robotic wheelchair by using the movements of his head, this is accomplished by means of a face tracking system that estimates the direction of the sight of the user using data from a Microsoft's Kinect ©. The images taken by the 2D RGB camera are used to set up a region of interest over the depth data coming from the infrared sensor.

The face of the user is detected using a typical Haar detector then a set of SWIFT features are selected over the face and 2D tracking is performed using the Lucas-Kanade method as described in [15]. The identification of the face pose is done by a random forest classifier which takes as input the 3D data from the Kinect sensor and gives the estimated position of the face [16], the results of the face tracking are shown in figure 1.

B. User Intentions Estimation System

The user intentions are modeled as topological poses into a predefined map. Those locations are set by the user's

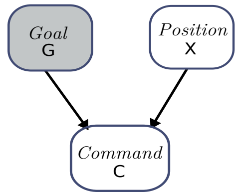


Fig. 2. The Bayesian network used to estimate the hidden user intention or goal G by knowing the current position X and the user command C.

habits (those places where the user spends most of his time are taken into account as probable goals) also interesting points taken from the map of the environment as doors, desks and other facilities are taken into account as probable destinations.

The reasoning method used is based on a Bayesian Network depicted in figure 2 that combines the information taken from the user interface input with the prior knowledge of the environment to infer a posterior probability distribution over the set of possible goals in the environment. This probabilistic model aims to take into account the uncertainty in the estimation of the desired goal and the inherent error over the user command read from the face tracking system.

To estimate the status of the *goal* variable, the command direction coming from the user-machine interface and the current user's position are applied as evidence. The posterior probability of the current goal given the position of the user and the direction of the command is expressed as $P(G^i_t|C_tX_t)$. The prior probability over each goal in the environment is denoted as $P(G^i_t|X_t)$.

Using Bayes' rule:

$$P(G^i_t|C_tX_t)) = \frac{P(C_t|X_t, G^i_t)P(G^i_t|X_t)}{P(C_t|X_t} \tag{1}$$

This can be simplified by using the normalizer η.

$$P(G^i_t|C_tX_t)) = \eta P(C_t|X_t, G^i_t)P(G^i_t|X_t) \tag{2}$$

We assume that the direction of the command given by the user C_t depends on the current position X_t and the intended goal G^i_t. Therefore this probability should be higher for goals that are in the direction of the current command input as shown in figure 3. $P(C_t|X_t, G^i_t)$ represents the probability of giving a command C_t when the user is located at position X_t and her goal is at position G^i_t at current time t. The notation G^i_t is used to express $G_t = g_i$ where g_i is one of the predefined goals in the environment as appear in Fig. 3.

$$P'(C_t|X_t, G^i_t) = \frac{1 - |a_i|}{\pi} \tag{3}$$

Where a_i term is the angle between both command and goal directions. The we normalize it as:

$$P(C_t|X_t, G^i_t) = \frac{P'(C_t|X_t, G^i_t)}{\sum_i P'(C_t|X_t, G^i_t)} \tag{4}$$

The prior probability table $P(G^i_t|X_t)$ was set manually for our current experimental set-up taking into account the user's

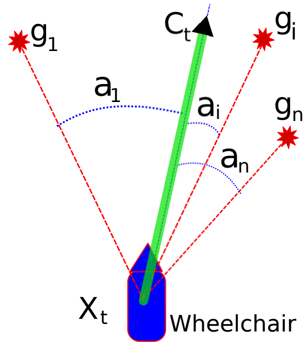


Fig. 3. The probability distribution for a given command C_t (big arrow) is proportional to the angle a_i formed with respect to each goal g_i in the environment.

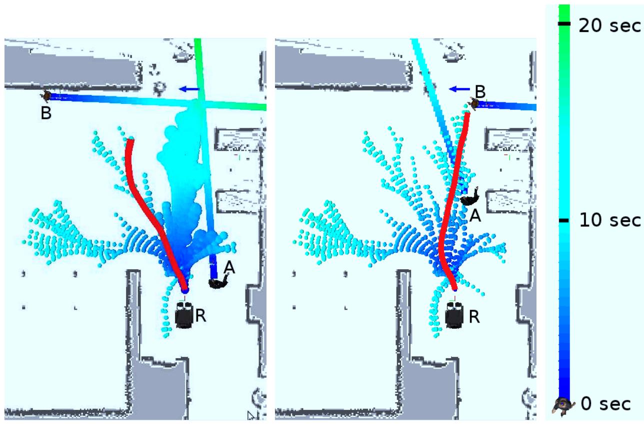
habits. However, for a real-environment it is mandatory to learn its values autonomously using some machine learning method.

IV. RiskRRT approach

A. Planner main loop

Our planner is based on RiskRRT [17], a partial motion planner which integrates motion predictions to provide safe trajectories. At each loop the algorithm checks if a new goal is available. When a new goal is received, RiskRRT collects information about the static obstacles and humans' position in the environment. Then a prediction of pedestrians trajectories takes place. At this moment RiskRRT has enough information to proceed with the exploration of the environment by means of a Rapidly-Exploring Random Tree which is constantly updated with the new perceived data. New nodes are created by selecting possible compliant controls that conduct the robot towards randomly selected points. The process of exploration has a depth threshold for the nodes and limited time in order to achieve real time performance. A probability of collision is assigned to each node taking into account static obstacles and human predicted trajectories.
Finally, a best path is selected by choosing the branch with the lowest probability of collision and with the closest distance to goal. It is important to mention that RiskRRT generated paths include information about space, time and robot dynamics which is very important advantage to navigate in dynamic environments. In the fig. 4 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. 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.

This method has been extended by including a mechanism to obtain socially acceptable behavior which is explained in next section.



(a) (b)

Fig. 4. Example of execution for the RiskRRT 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.

B. Social conventions in human navigation

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 [18], which characterizes the space around a human being in terms of comfort to social activity.

Another interesting social situation arises when two or more of the persons in the environment are interacting. We model interactions using the concept of *o-space* which has been developed by sociologists [19]. 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 a module called **"Social Filter"** 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 done by considering the probability of not encountering a collision along the path and not entering in a personal space or an o-space. Detailed explanation can be found in [11].

1) Modeling Personal Space: We have modeled personal space as a composition of two human centered Gaussians, one for the front and one for the back of the space, the front is larger as people is more sensitive to this space. Fig. 5 shows an example of personal space as provided by the Social Filter.

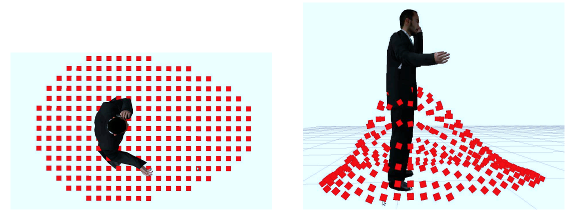


Fig. 5. Personal space calculated by Social Filter Module. Height of the Gaussian means Risk of disturbance then maximum disturbance is located at human position.

2) Modeling o-Space: When more than two people are in conversation, they tend to make a formation with circular shape. 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 [19]. The social filter identifies individual F-formations (Vis-a-vis, L-Shape, C-Shape or V-Shape) and builds the corresponding o-space. in Fig. 6, the calculated o-space for a Vis-a-Vis interaction is shown.

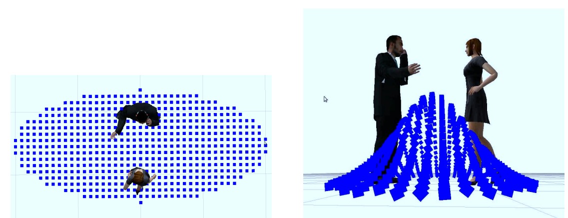


Fig. 6. O-space calculated by Social Filter Module for a Vis-a-Vis F-formation. Maximum risk of disturbance is located at o-space center, in the picture the disturbance is represented by height of Gaussian.

C. Planning under social constraints

This section explains how the social constraints are included in the RiskRRT framework, see [11], [17] for specific details. First we define PZ_i as the probability of disturbing by passing inside the o-space (sec. IV-B.2) of interaction i, and we calculate it as the maximum value of o-space model for that interaction evaluated in the intersection with the robot's path. PZ_i can be thought as a collision with a dynamic obstacle:

$$P_{cd} = 1 - \prod_{m=1}^{M} [1 - P_{cd}(o_m)] \prod_{i=1}^{r} [1 - PZ_i] \tag{5}$$

where P_{cd} is the probability of dynamic collision considering the M humans in the environment and $P_{cd}(o_m)$ is the probability of collision with the human o_m taking into account the personal space. $P_{ps}(o_m)$ is the risk of disturbing by passing in such personal space and can be approximated as the probability that A, the area swept by the robot's path, intercepts the one represented by the personal space:

$$P_{ps}(o_m) = \int_A PS(o_m(t)) \tag{6}$$

where $PS(o_m(t))$ is the model of personal space centered in $o_m(t)$ at time t as described in IV-B.1. To take into account this last constraint we use:

$$P_{cd}(o_m) = P_{dyn}(o_m) + P_{ps}(o_m)(1 - P_{dyn}(o_m)) \quad (7)$$

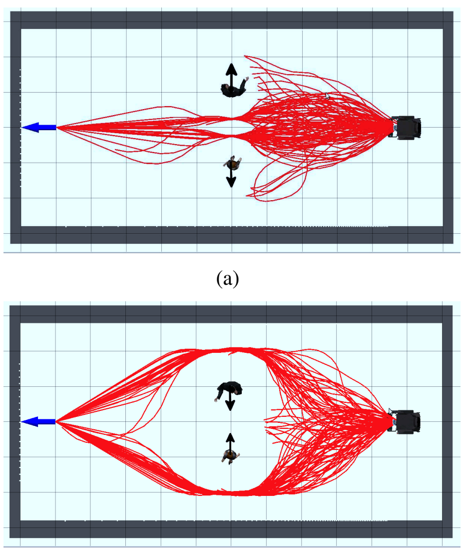
where $P_{dyn}(o_m)$ is the probability of dynamic collision between the robot and o_m considering only their trajectories. Last, total probability of collision, $collP$ is calculated for each node using:

$$collP = P_{cs} + (1. - P_{cs}) * P_{cd}; \quad (8)$$

Finally, also for each node we calculate a weight based on probability of collision and distance to goal. In order to compare paths worst node weight is chosen to represent a particular path.

D. Simulation results

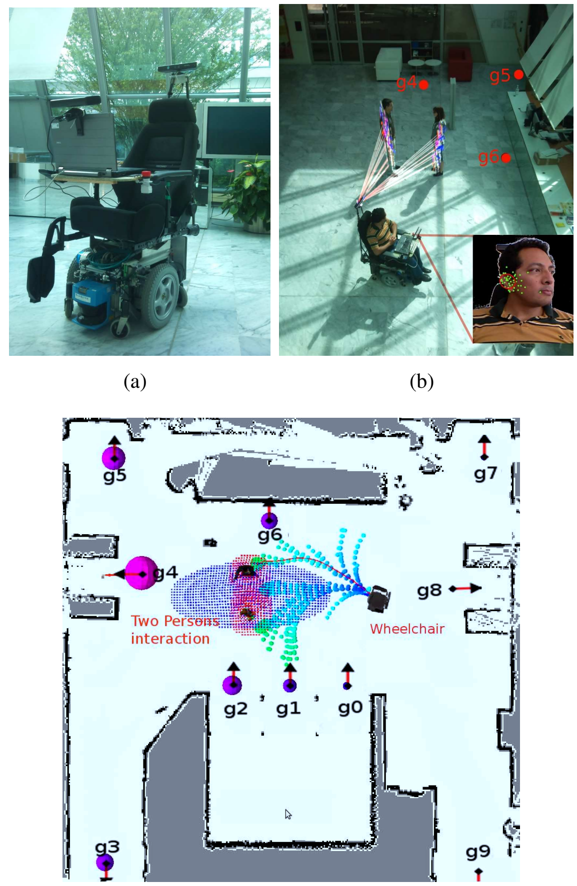
In order to test socially acceptable behavior, we conducted several simulation tests.



(b)

Fig. 7. Socially acceptable navigation in RiskRRT. Each figure shows a sample of generated plans (in red) for one hundred executions of RiskRRT, each execution run for twenty iterations of the algorithm. Goal is the blue arrow. In a) static persons are looking towards walls, therefore there is no social interacting zone, then navigation respects only their personal spaces. In b) both social spaces are respected.

Our first test scenario consisted of two people in the same corridor together with the wheelchair whose task is to navigate towards the goal. We realized one hundred executions of the planner in very similar conditions, as it can be seen in Fig. 7, when the social filter is turned on, all the plans managed to respect both personal space and interaction space without disturbing the involved people. We can observe that because of the random exploration of the environment some executions (almost ten percent) select as best path one that stops before to invade social space. Tests with the social filter off showed that thirty percent of plans passed in the middle of an interaction.



(c)

Fig. 8. Experimental scenario, in a) The wheelchair used for testing in b) the INRIA's hall. Possible goals are marked with a circle, the persons present in the scene are tracked by one of the Kinect sensors while the other Kinect is used to track the face of the user. c)Results of the system in a real scenario. The size of the spheres in the environment represent the value of the computed posterior probability for each destination. The computed trajectory and interaction regions are also shown.

V. Integration and Evaluation

Integration was done by using the semi-autonomous controller presented in sec. III as the source of the goals for the RiskRRT planner. Solution plans are executed in our experimental platform, details follow.

A. Experimental Platform

Our platform, is an automated wheelchair (Fig.8(a)) equipped with one Sick laser and two 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 Fig. 8(c). Data coming from the upper Kinect allow us to have position and orientation of pedestrians in the scene while data from frontal Kinect collect face features to feed our intention recognizer algorithm.

B. Evaluation in a real scenario

The system was evaluated at scenario shown in Fig.8(b). User can start the movement at any location of the experimental scenario, he is asked to drive the wheelchair by seeing towards his desired destination. In the example, the user is seeing to the left so that it is more probable that he is aiming to go to the coaches located in that direction. The direction of his face is computed as previously explained in section III-A. Typical destinations were defined into the map, they are marked with small arrows in Fig.8(b).

Whenever a new command is read from the face control, the user estimation module computes the goal with the highest posterior probability, depicted in Fig.8(c) as the size of the sphere marking each goal, then it is sent to the navigation module to start the movement.

The navigation module receives the map of the environment, the currently computed goal and the list of people present in the field of view of the frontal viewing Kinect to compute the necessary trajectory to the goal as shown in Fig.8(c). In the example there are two persons in conversation, standing in the middle of the path between the wheelchair and the current estimated goal. Even if the user is pointing to the goal located in the other side of the two persons he does not have to worry about the necessary planning and commands to avoid interrupting the conversation because the autonomous navigation system is in charge of that.

VI. Conclusions and future work

The approach presented in this paper has integrated a human aware motion planner and a semi-autonomous controller adapted to the user with the aim to preserve independence of people with reduced mobility. The system was designed to improve both usability by reasoning about user intentions and sociability by including concepts like personal space and o-space in the planning algorithm. Experiments with a real wheelchair show that integration has been successfully achieved. The user intention algorithm has proved to be useful to translate simple input commands into high level orders or goals for our autonomous navigation system.

Using the pose of the face as input can be advantageous to assist the elderly because it provides a more natural way of interaction (we usually see where we are going) so they can be more confident when using the wheelchair. In this work we explored the use of head direction to infer the user desired destination in a navigation task. The system is capable of estimating the desired goal among a set of interesting points and to transport the user to it while avoiding humans in a social way.

Current work will be extended in two directions. First, by including as interesting points positions where the user should be located when he wants to join a group of people. Second, by adapting autonomously the user intention algorithm to user disability. Moreover other user-machine interfaces, like voice based, will be included to minimize ambiguities like that of the movements of the head that are intended to give a navigation command from those resulting from the natural observation of the environment.

References

[1] S. P. Levine, D. a Bell, L. a Jaros, R. C. Simpson, Y. Koren, and J. Borenstein, "The NavChair Assistive Wheelchair Navigation System." *IEEE transactions on rehabilitation engineering*, vol. 7, no. 4, pp. 443–51, Dec. 1999.

[2] T. Kollar, S. Tellex, D. Roy, and N. Roy, "Grounding Verbs of Motion in Natural Language Commands to Robots," in *International Symposium on Experimental Robotics*, 2010.

[3] E. Demeester, M. Nuttin, D. Vanhooydonck, and H. Van Brussel, "Assessing the user's intent using bayes' rule: application to wheelchair control," in *Proceedings of the 1st International Workshop on Advanced in Service Robotics*, 2003, pp. 117–124.

[4] 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*, pp. 5775–5780, 2006.

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

[6] E. A. Sisbot, L. F. Marin-Urias, R. Alami, and T. Simeon, "A human aware mobile robot motion planner," *IEEE Transactions on Robotics*, vol. 23, 2007.

[7] 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.

[8] S.-Y. Chung and H.-P. Huang, "A mobile robot that understands pedestrian spatial behaviors," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, oct. 2010, pp. 5861 –5866.

[9] S. Sehestedt, S. Kodagoda, and G. Dissanayake, "Robot path planning in a social context," in *Robotics Automation and Mechatronics (RAM), 2010 IEEE Conference on*, june 2010, pp. 206 –211.

[10] C. P. Lam, C. T. Chou, K. H. Chiang, and L. C. Fu, "Human-centered robot navigation, towards a harmoniously human-robot coexisting environment," *IEEE Transactions on Robotics*, vol. 27, no. 1, 2011.

[11] 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*, Sep. 2011, pp. 2014–2019.

[12] R. Kirby, R. Simmons, and J. Forlizzi, "Companion: A constraint-optimizing method for person acceptable navigation," *The 18th IEEE International Symposium on Robot and Human Interactive Communication*, 2009.

[13] 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.

[14] T. Kruse, P. Basili, S. Glasauer, and A. Kirsch, "Legible robot navigation in the proximity of moving humans," in *Advanced Robotics and its Social Impacts (ARSO), 2012 IEEE Workshop on*, may 2012, pp. 83 –88.

[15] "Pi face tracker webpage," 2012. [Online]. Available: www.pirobot.org

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

[17] C. Fulgenzi, A. Spalanzani, and C. Laugier, "Probabilistic motion planning among moving obstacles following typical motion patterns," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, oct. 2009, pp. 4027 –4033.

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

[19] 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.