Multimodal Control of a Robotic Wheelchair: Using Contextual Information for Usability Improvement

Arturo Escobedo[1], Anne Spalanzani[2], Christian Laugier[1] †
[1]INRIA Rhône-Alpes, Grenoble, France
[2]UPMF-Grenoble 2 - INRIA - LIG, France

***Abstract*— In this paper, a method to perform semi-autonomous navigation on a wheelchair is presented. The wheelchair could be controlled in semi-autonomous mode estimating the user's intention by using a face pose recognition system or in manual mode. The estimator was performed within a Bayesian network approach. To switch these two modes, a speech interface was used. The user's intention was modeled as a set of typical destinations visited by the user. The algorithm was implemented to one experimental wheelchair robot. The new application of the wheelchair system with more natural and easy-to-use human machine interfaces was one of the main contributions. as user's habits and points of interest are employed to infer the user's desired destination in a map. Erroneous steering signals coming from the user-machine interface input are filtered out, improving the overall performance of the system. Human aware navigation, path planning and obstacle avoidance are performed by the robotic wheelchair while the user is just concerned with "looking where he wants to go".**

***Index Terms*— Intention estimation, semi-autonomous navigation, Bayesian inference, wheelchair control, face control, voice control.**

I. Introduction

The aging of world's population is bringing the need to assist elder people to move when they lose the necessary motor skills, strength or visual acuteness to do it by themselves [1]. The development of smart robotic platforms capable of doing that can clearly improve the quality of life for elders. Patients and medical staff have a strong desire for the services that a smart wheelchair can offer. To be well accepted, the movement of the wheelchair must be reliable, safe and comfortable. The platform discussed in this paper has been designed in order to meet the following requirements:

- Usability: People with motor disabilities and elders often have problems using joysticks and other standard control devices. Under the consideration that it is a natural human behavior to look where we are going [2], we have equipped our experimental platform with a face pose interface that allows to control the wheelchair's motion by moving the face. A voice recognition system is used to give those commands that would be difficult to express using only the face (Stop, move, faster, etc). In semi-autonomous mode the robot navigates securely among obstacles while minimizing user's frustration by estimating his intentions.
- Safety: The system must avoid collisions with both static and dynamic entities.
- Respect of social conventions: When moving, the robot may considerably disturb people around it, especially when its behavior is perceived as unsocial. It is thus important to produce socially acceptable motion to reduce disturbances.

This article is structured as follows: Section II offers an overview of related works. A general description of the system architecture is presented in III while section IV focuses on the estimation of intended destination and control of the wheelchair, the multimodality of the system is explained in section V. In section VI our experimental setup is presented. Results and analysis is provided in VII. Section VIII presents conclusions about the work and perspectives for future improvements of the system.

II. Related Work

Different interfaces have been used to drive robotic wheelchairs. In [3] a wheelchair controlled by a 2D face pose recognition system is presented. The user is completely in charge of driving the wheelchair by moving his face without any assistance from the robotic controller.

Speech recognition is used in [4], [5]. In [4] the wheelchair has two possible driving modes: follow a wall and get into an elevator. The user switches between modes using the voice control. The method presented in [5] employs speech recognition to control the movement of the wheelchair in different directions (back, forward, left, right), they intend to use this system for people that cannot use any other user interface, however, the low bandwidth of this kind of interface makes it difficult for the user to drive if no further assistance is provided. In [6] the Google Speech Recognition Service and Microsoft SAPI are used in combination with a 2D face tracking to control a mobile robot. The speech recognition is used to switch between modes. In the first mode the user directs the robot using "execute backwards, forward, left and right" commands, while in the second mode the direction of the movement is controlled by moving the face up, down, left and right similar as a joystick, again no further assistance is provided by the robotic controller. Voice recognition is employed in [7] to select the desired navigation

† {jesus.escobedo-cabello, anne.spalanzani, christian.laugier}@inria.fr
This work has been funded by the INRIA large scale initiative project PAL, Personally Assisted Living

task or destination using commands as (follow me, go to the kitchen) and then the robot executes the task autonomously.

Some works go further adding more intelligent assistance in order to improve the driver's experience when using a robotic wheelchair. Wheelchairs that operate similarly to autonomous robots and navigate autonomously towards a destination given by the user were proposed in [7].

Other smart wheelchairs limit their assistance to collision avoidance while the user is in charge of most of the navigation task; planning and controlling the trajectory. These systems do not normally require prior knowledge of an area or any specific alterations to the environment, however, they require more planning and effort from the user [8], [9], [10] .

Shared control is presented in situations in which the assisting device combines the control input coming from the robot and the user in order to accomplish a given task [10].

The estimation of the user's plan is a key point in many shared control systems because it allows the automatic controller to adjust its actions to the desire of its user. Inferring the user plan is necessary whenever the interface with the user does not allow him to explicitly dictate it to the robot as with many popular electric wheelchair interfaces (brain control interface, face tracking, gaze tracking, sip and puff, joystick, etc). A robotic wheelchair can assist by taking over low-level control, requiring the user to use the input method only to give high-level directional commands.

Some methods to perform an implicit estimation of the user's intention from simple inputs have been proposed in [9], [10]. 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 probable destination is selected by he user within a Bayesian framework.

In [11] 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 allows the user to focus on driving the wheelchair from one spatial location to another without having to worry about the low level control.The spatial representation used is based on a topological graph representation of the environment, where vertexes are locations and edges represent a viable path connecting two locations as a result of performing an action. Places of interest are selected as spatial locations in the environment where the user spends comparatively most of his time.

Our approach is based on the Bayesian method presented in [9] but differs in two main aspects: the model of the user's intention and the user-machine interface. In [9] the intention is modelled as a set of possible trajectories while in this article it is modelled as a set of possible destinations. This change leads to a considerable reduction in the computation complexity of the Bayesian Network. The main contribution of the present work is the estimation of the user's intention by tracking the position of the face.

The integration of the method in the complete experimental platform granted to perform the practical tests to evaluate it. The software used for doing this research is completely open source which brings the possibility of a future implementation of a low cost add-on for normal wheelchairs. The navigation is performed using a new human aware planning algorithm that integrates a notion of social conventions and avoidance of dynamic obstacles to prevent uncomfortable situations when the wheelchair is navigating among humans [12], [13].

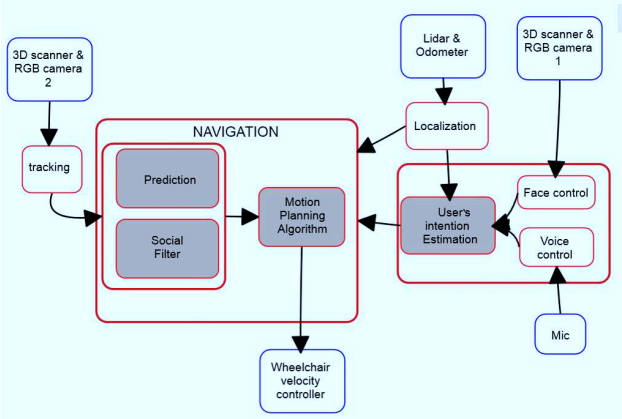


Fig. 1. System Architecture Overview

III. System Architecture

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

- **User's Intention Estimation**: The user's intention subsystem estimates the desired destination within the map of the environment among a list of possible predefined goals. Those locations can be previously selected by an expert caregiver, the user, or learned automatically by the system. The probability for each typical destination is computed using a Bayesian Network that considers the current position of the wheelchair and user's face direction. The destination with the highest probability is selected and sent to the navigation module.
- **Prediction**: Data from the trackers is used to make probabilistic predictions about the configuration of the free space in the future environment. The motion prediction subsystem receives tracking data (i.e. position, orientation and velocity) and outputs grids, representing the posterior probability of the space being occupied at a given time step in the future. Prediction itself is accomplished with a Growing Hidden Markov Model (GHMM) [14].
- **Social Filter**: 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 in order to evaluate the risk of disturbance and take it into account when making the autonomous navigation planning. We focus on detecting and predicting conversations in the environment surrounding the

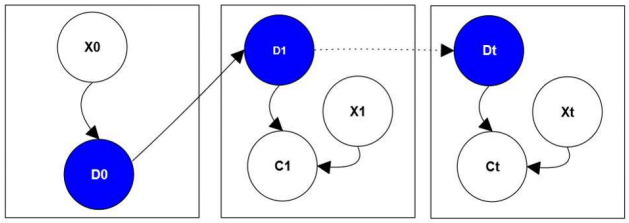


Fig. 2. The Bayesian Network used to estimate the current user's intended *destination* D_t. At each *time step* t the posterior probability is updated by using the current *position* X_t and current *command* C_t. D_t also depends on the value of the last estimation D_{t-1} to take into account the history of given commands. Prior knowledge is expressed as the probability of going from each *starting position* X_0 to any of the possible destinations D_0

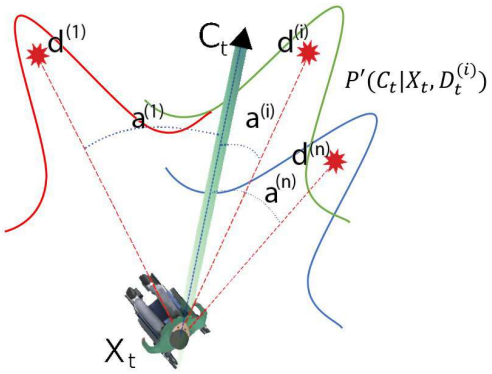


Fig. 3. The probability value for a given command C_t is proportional to the angle $a^{(i)}$ between the sight-line and each typical destination $d^{(i)}$ in the environment.

wheelchair [13].

- **Motion Planning**: The navigation subsystem includes a laser-based landocalization module and a motion-planner which integrates predictions to compute safe trajectories that are fed to the execution module. The motion planner is based on a risk based motion planner [15], a partial motion planner which integrates motion predictions to provide safe trajectories. This algorithm was thought to operate in dynamic, uncertain environments, it supposes that the moving pedestrians detected in the environment follow typical motion patterns that are represented by Growing Hidden Markov Model (GHMM). This motion planner generates human friendly paths respecting people's personal and interaction spaces, as provided by the social filter.

IV. User's Intention Estimation

The user's intent is modelled as a set of topological poses into a predefined map. Those poses may be defined by the user's habits (places where the user spends most of his time during the day) and interesting points taken from the map of the environment as doors, desks and other facilities.

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

The initial probability $P(D_0|X_0)$ provides the prior information about the user's habits. If there is no previous knowledge, this term will be expressed by a uniform distribution, however, if the user is in a known environment; the most frequented places can be extracted by observation so that this probability distribution will be learned from the environment.

The approach presented here, consisted in learning those typical destinations by previous observation of the experimental scenario (explained later in section VI). The probability $P(D_0|X_0)$ was learned by tracking the number of times that a destination was visited departing from any of the others. The computed values are shown in table II.

The *command* variable C_t is dependent on the *destination* D_t and *position* X_t variables (this is a consequence of the natural human behaviour of *"looking where we are going"* [2]).

The probability for each *destination* variable $P(D_t^{(i)}|C_t, X_t)$ is updated every time step as follows:

$$P(D_t^{(i)}|C_t, X_t) = P(C_t|X_t, D_t^{(i)}) \sum_j [P(D_t^{(i)}|D_{t-1}^{(j)}) * P(D_{t-1}^{(j)}|C_{t-1}, X_{t-1})] \quad (1)$$

The notation $D_t^{(i)}$ is used to express $D_t = d^{(i)}$ where $d^{(i)}$ is one of the typical destinations in the current scenario as appear in Fig. 3.

$P(C_t|X_t, D_t^{(i)})$ 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.

Under the assumption that the user will try to give commands directed straight forward to the destination position as shown in Fig. 3, the non normalized probability $P'(C_t|X_t, D_t^{(i)})$ is computed as a normal distribution.

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

This probability is then normalized:

$$P(C_t|X_t, D_t^{(i)}) = \frac{P'(C_t|X_t, D_t^{(i)})}{\sum_i P'(C_t|X_t, D_t^{(i)})} \quad (3)$$

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. This works as a smoothness term which avoids abrupt changes in the estimated destination due to involuntary movements.

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.

$$P(D_t^{(i)}|D_{t-1}^{(j)}) = \begin{cases} 10/(n+9) & \text{if } D_t^{(i)} = D_{t-1}^{(j)} \\ 1/(n+9) & \text{else} \end{cases} \quad (4)$$

where n is the number of possible goals in the environment.

The selection of the goal is performed by looking for the goal with the maximum posterior probability.

$$d_{sel} = \arg\max_{d^{(i)}} P(D_t = d^{(i)} | C_t, X_t) \quad (5)$$

The selected destination is just sent to the autonomous navigation system if its probability value is bigger than a given threshold ϕ, which was fixed by experimentation.

$$d_{wheelchair} = \begin{cases} d_{sel} & \text{if } P(D_t = d_{sel} | C_t, X_t) >= \phi \\ null & \text{else} \end{cases} \quad (6)$$

V. Multimodal Control

The wheelchair can be controlled in *semi-autonomous* mode employing the user's intention prediction module or in *manual* mode. In manual mode the user controls the wheelchair's angular speed by moving his face. The linear speed is controlled with vocal commands as explained in section VI-C.

In semi-autonomous mode; the user shows the direction to his desired destination by looking towards it. The user's intention module computes the destination with the highest posterior probability, depicted in Fig.7 (a) as the biggest circle. The navigation module receives the map of the environment, the list of humans present in the scene and the currently estimated goal to compute the necessary trajectory to the goal as it is shown there.

When moving the user does not have to worry about the necessary planning to avoid obstacles because the autonomous navigation system is in charge of that, however, he can stop or start the wheelchair by using the speech recognition system.

The main function of the speech interface is to switch between manual and semi-autonomous modes by saying the *"manual"* and *"autonomous"* vocal commands.

- Autonomous: The wheelchair computes the most likely intended destination of the user and navigates autonomously towards it.
- Manual: The linear speed of the wheelchair is regulated using vocal commands while the angular speed is controlled by moving the face. No obstacle avoidance or path planning is provided by the wheelchair.

The speech recognition system is used in combination with the face pose estimator to provide different operation modes as detailed in Table I

VI. Experimental Setup

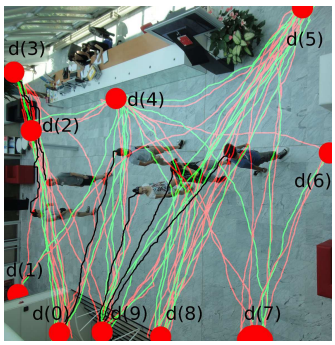
The experimental scenario is shown in Fig. 4(a). People in the scene was tracked to learn the typical destinations that are then located on the map of the environment. Each destination has a related probability value as described in section IV. The values for the presented experimental scenario are shown in Table II.

The robotic wheelchair has some on-board sensors and a computing unit as presented in Fig. 4(b).

TABLE I

Vocal commands and associated driving behaviours in manual and autonomous modes.

Autonomous Mode	
go + face direction	Computes the most probable destination in that direction and navigates autonomously against it.
one brake + face direction	Stops the wheelchair and keeps turning in the same position.
brake (2 times)	Stops completely the wheelchair.
Manual Mode	
go + face direction	Moves the wheelchair with a constant linear speed; the angular speed is proportional to the face direction.
faster	Increases the linear speed
slower	Decreases the linear speed
one brake + face direction	Stops the wheelchair and keeps turning in the same position.
brake (2 times)	Stops completely the wheelchair.



(a)



(b)

Fig. 4. Experimental setup. (a) Scenario: The lines in the figure represent some of the trajectories normally followed by people. Typical destinations (end of a trajectory) are marked with circles. (b) The robotic wheelchair used for the described experiments.The mobile base includes all the electronic components and the computer in charge of the low level control.

TABLE II

Values for $P(D_0|X_0)$ obtained by observing the human activity in the experimental scenario.

$d_0 =$	$d^{(0)}$	$d^{(1)}$	$d^{(2)}$	$d^{(3)}$	$d^{(4)}$	$d^{(5)}$	$d^{(6)}$	$d^{(7)}$	$d^{(8)}$	$d^{(9)}$
$x_0 =$										
$d^{(0)}$	.01	.03	.06	.45	.1	.26	.03	.02	.02	.02
$d^{(1)}$	.04	.01	.02	.45	.25	.1	.03	.02	.03	.02
$d^{(2)}$	.1	.05	.01	.45	.25	.25	.02	.03	.02	.03
$d^{(3)}$	.45	.02	.02	.01	.1	.02	.06	.02	.03	.25
$d^{(4)}$	.03	.02	.03	.35	.01	.35	.02	.02	.03	.1
$d^{(5)}$	.15	.01	.01	.01	.1	.01	.02	.03	.45	.2
$d^{(6)}$	.03	.02	.02	.03	.1	.25	.01	.04	.45	.02
$d^{(7)}$	.03	.02	.03	.1	.25	.45	.03	.01	.02	.03
$d^{(8)}$	.03	.02	.03	.03	.25	.45	.1	.02	.01	.02
$d^{(9)}$	.03	.02	.03	.35	.1	.35	.02	.02	.04	.01

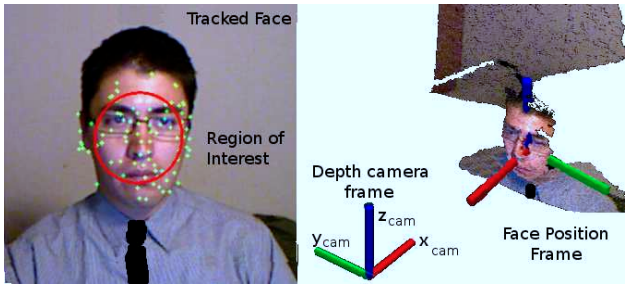


Fig. 5. 3D information of the user's face is used to measure its direction. The 2D RGB image (left) is used to detect the face and fix a region of interest in the 3D point cloud. [16]

A. Wheelchair

The equipment used is the robotic wheelchair shown in Fig. 4(b) that consists of a mobile base equipped with a seat, all the on-board electronics and different sensors (Light Detection and Ranging Unit, quadrature encoders for odometry measurements, emergency bumpers and 2 Kinect sensors). One Kinect camera is used for the user-machine interface and the other is intended to be used for perception in future works.

B. Face Control Subsystem

The user can control the robotic wheelchair by using the movements of his face . A face tracking system that estimates the direction of the face from the 3D data gathered by the Kinect camera Fig. 5. 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. For further information on this subject please refer to [16] and [17].

C. Voice Control Subsystem

The voice interface is used to fulfill some lack in functionality that the face pose interface can not supply. This voice recognition system was developed at Carnegie Mellon University and described in *Pocketsphinx*, [18]. The system allows to specify a set of syntactic rules (or grammar) which constrains the ordering of words within a sentence. This grammar enhances the speech recognition quality by reducing the hypothesis space. In general a small vocabulary makes speech recognition more accurate, therefore a dictionary that focuses in a very small fixed set of tasks was considered (go, brake, faster, slower, autonomous, manual)

VII. Experimental Evaluation

To evaluate the performance of the method, four different persons were asked to drive the wheelchair in *"manual"* and *"autonomous"* modes described in section V.

The user is asked to start the movement from one of the labeled destinations shown in Fig. 4(a) and approach a list of targets e.g (go to the reception, then go to the door in the right and then come back).

Every command and computed intent was stored at each time step. This allows to compare the output of the estimation

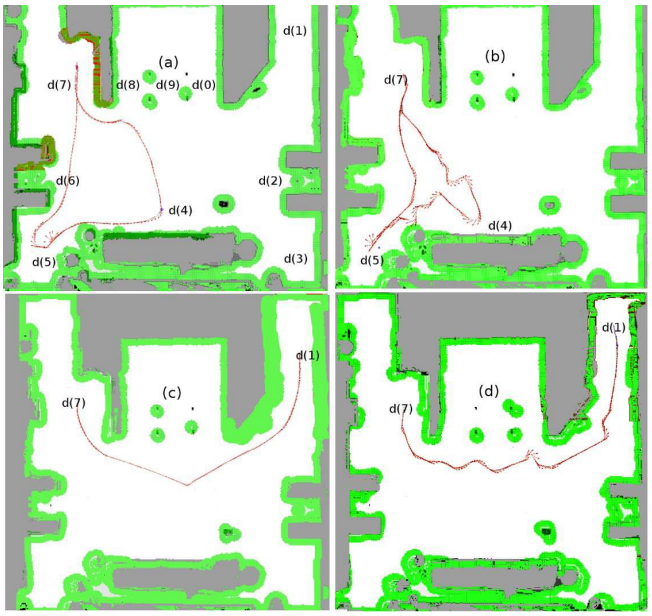


Fig. 6. Some samples of the logged trajectories are presented. (a) and (c) show the results when using the assistance of the user's intention estimation system. (b) and (d) were achieved by driving the wheelchair using the face without any assistance. Here we can observe the oscillations in the trajectory due to involuntary changes in the face direction produced when observing the surroundings. *A video showing the performance of the wheelchair can be found at http://www.youtube.com/watch?v=IObaXHXmiBQ.*

TABLE III
Performance metrics when driving the wheelchair with and without assistance

Semi-autonomous mode	Time [s]	Distance [m]	Num. of collisions
User 1	32.1	24.4	0
User 2	34.5	25.8	0
User 3	35.3	26.2	0
User 4	32.5	25.1	0
Manual Mode	Time [s]	Distance [m]	Num. of collisions
User 1	54.0	28.2	0
User 2	68.3	29.3	0
User 3	75.4	29.4	1
User 4	63.2	27.0	1

with the actual intention of the human. Second, when driving towards the destination, global measures such as the total time needed, the total distance travelled, and the followed trajectory were logged to quantify the performance.

In Fig. 6 some of the resulting trajectories are presented, the first thing that can be noticed is the improved performance accomplished when using the user's intention estimation algorithm. This avoids abrupt modifications in the trajectory due to involuntary changes in the direction of the face produced when the user is exploring the surroundings. The time when executing the task was also improved as shown in the following table. The most important fact to be noticed is that collisions were completely avoided when using the user's intention algorithm, on the other hand, in manual mode some collisions were produced.

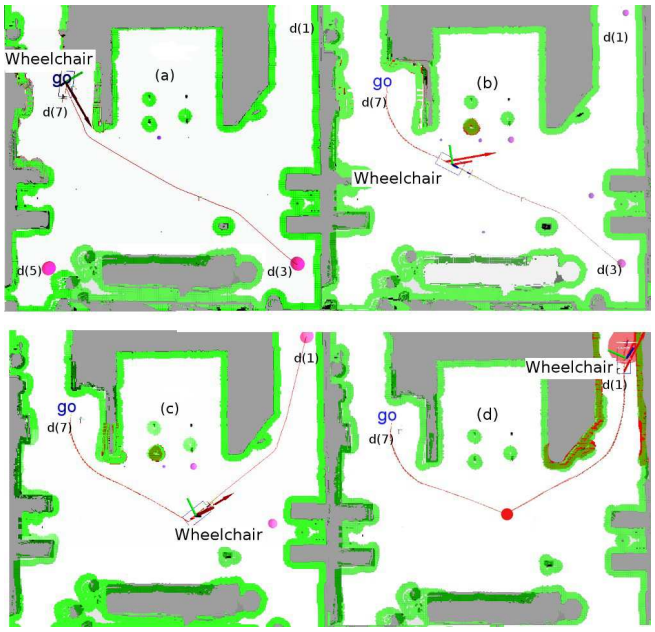


Fig. 7. Experimental Evaluation, the user is asked to go from $d^{(7)}$ to $d^{(1)}$ using the user's intentions system. (a) As the destination with the highest prior probability from $d^{(7)}$ is $d^{(5)}$ in the beginning of the test the results are biased to that side that is why $d(3)$ is initially chosen as the most likely goal. (b) As the user keeps looking to the left all the destinations in that direction become more likely. (c) When the wheelchair has enough evidence, it changes the desired destination to $d^{(1)}$. (d) The user arrives to his desired destination, the spot in the middle of the trajectory marks the place were the change of destination was produced.

VIII. Conclusions and Future Work

During experiments the user's intention estimation algorithm was used to translate simple input commands (face direction) into high level orders (the desired destination). This destination is used to feed the autonomous navigation system so that the wheelchair can move autonomously towards it. Adding contextual information of the environment improves the usability of the presented face tracking interface.

The system avoids abrupt changes in the trajectory due to involuntary face movements. It also avoids the problem of driving the wheelchair against obstacles due to the natural human behavior of looking towards risky zones.

One of the main issues of the method is the decrease of freedom in the number of possible movements that the user can perform, this is a challenge that remains as an open question for future works.

For future improvement it is necessary to take into account cases where there exist ambiguity in the possible desired goals. As it can be seen in Fig. 7(a) even if at the beginning the left door (left circle in image) was selected as the most probable destination over the right door, both of them had similar probability values (size of the circle). In those cases it should be requested some extra information from the user to make a better choice.

To work in a non-supervised environment, the user's intention algorithm must be extended and perform an online learning to add new important places and user's habits updating the prior probabilities described in section III.

Users consulted in this study claimed that the use of the face direction interface is more comfortable and practical than other conventional methods as the joystick.

References

[1] R. C. Simpson, "How many people would benefit from a smart wheelchair? the journal of rehabilitation research and development," *Rehabilitation Research and Development*, vol. 45, pp. 53–72, 2008.

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

[3] C. Banerjee, H. Gupta, and K. Sushobhan, "Low cost speech and vision based wheel chair for physically challenged," *2010 The 2nd International Conference on Computer and Automation Engineering (ICCAE)*, vol. 1, pp. 706–709, Feb. 2010.

[4] A. Murai, M. Mizuguchi, T. Saitoh, T. Osaki, and R. Konishi, "Elevator available voice activated wheelchair," *RO-MAN 2009 - The 18th IEEE International Symposium on Robot and Human Interactive Communication*, pp. 730–735, Sept. 2009.

[5] S. Poonguzhali, "Design and implementation of low cost intelligent wheelchair," *2012 International Conference on Recent Trends in Information Technology*, pp. 468–471, Apr. 2012.

[6] R. Berjon, M. Mateos, a. L. Barriuso, I. Muriel, and G. Villarrubia, "Alternative Human-Machine Interface system for powered wheelchairs," *2011 IEEE 1st International Conference on Serious Games and Applications for Health (SeGAH)*, pp. 1–5, Nov. 2011.

[7] 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.

[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. 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.

[10] 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.

[11] 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.

[12] V. Dizan, P. Silveira, A. Escobedo, J. Rios-Martinez, A. Spalanzani, and C. Laugier, "Human aware navigation for assistive robotics," in *13th International Symposium on Experimental Robotics*, 2012.

[13] 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.

[14] D. Vasquez, T. Fraichard, and C. Laugier, "Growing hidden markov models: a tool for incremental learning and prediction of motion," *International Journal of Robotics Research*, vol. 28, no. 11–12, pp. 1486–1506, 2009.

[15] C. Fulgenzi, A. Spalanzani, C. Laugier, and C. Tay, "Risk based motion planning and navigation in uncertain dynamic environment," INRIA," Research Report, 2010.

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

[18] "Cmu sphinx webpage," 2012. [Online]. Available: http://cmusphinx.sourceforge.net