



**HAL**  
open science

## Geometric Tools for Perspective Taking for Human-Robot Interaction

Luis Félipé Marin-Urias, Emrah Akin Sisbot, Rachid Alami

► **To cite this version:**

Luis Félipé Marin-Urias, Emrah Akin Sisbot, Rachid Alami. Geometric Tools for Perspective Taking for Human-Robot Interaction. Seventh Mexican International Conference on Artificial Intelligence, 2008, Atizapan de Zaragoza, Mexico. hal-01976921

**HAL Id: hal-01976921**

**<https://laas.hal.science/hal-01976921>**

Submitted on 10 Jan 2019

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

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

# Geometric Tools for Perspective Taking for Human-Robot Interaction

Luis Felipe Marin-Urias, E. Akin Sisbot, Rachid Alami  
LAAS-CNRS ; Université de Toulouse ;  
7 avenue du Colonel Roche, F-31077 Toulouse, France

## Abstract

*Psychological studies on human interactions made surface the notion Perspective Taking. This notion, referring to reason on other persons' points of view, eases the communication between interacting individuals. For a robot interacting with people, we believe that Perspective Taking can play an important role that will allow the robot to understand, reason and act according to human's point of view thus allowing an efficient communication and cooperation.*

*In this paper, we propose a group of algorithms to evaluate and generate robot configurations that are not only collision free but also obeying HRI constraints by reasoning explicitly on human's perspective.*

## 1. Introduction

One of the challenges of human-robot interaction (HRI), is the environment sharing and spatial placement between the robot and the human. This, added to the problem of understanding human intentions to improve the robot's cognitive capabilities in order to result a better interaction. For these purposes, robot has to adopt different human behaviors, like perspective taking.

The notion of *perspective taking* comes from psychological studies on human to human interactions. It refers essentially to the fact of reasoning from other persons point of view. It is also interpreted as taking its own perspective from a different place on the space by applying what is called a *mental rotation* to perceive the environment from different place. These sets of actions are used by humans in their everyday lives, and are intended to ease communication between individuals and to help to have shorter and faster interactions.

Perspective taking can be used by the robot to generate configurations to approach human or to compute a geometric way to place itself where it can perceive an object referred by a person. The introduction of perspective taking concepts on robot's spatial reasoning for placement is called here *perspective placement*.

In this paper, we introduce a group of algorithms that introduces perspective taking to the geometrical reasoning of the robot. In section 2, we present the foundations of perspective taking as well state of the art approaches using this notion. Section 3 introduces PerSpective Placement (PSP), a system that generates and evaluates robot configurations according to human's perspective. Section 4 and 5 illustrates simulation and real world results of this system applied in various scenarios. Finally, the section 6 concludes this paper and gives perspectives on perspective taking in robotics.

## 2. Related work

Although human-robot interaction is a very active research field, there is no extensive amount of research on perspective taking on robotics.

Teversky et al. [1] [2] show the effectiveness of changing perspectives between persons in face-to-face communication scenarios for spatial descriptions. They conclude that the perspective switching is more effective than trying to avoid it.

Akerman [3], on the other hand, explains how a person can estimate his relative position in the world by 3D mental transformation of an object or of an entire place. As mental rotation and perspective taking are important for interacting with humans, they are often taken into account in 1) computer graphics for simulating human-like view [4][5], 2) in virtual reality for home or car design [6] [7], 3) in human training [8] and 4) in interface design for human-computer interaction [9].

Mental rotation has also found its place in robotics, precisely in mobile robot simulators [10], where a virtual environment simulates sensor data for the robot. With this, researchers can test different robot programs without moving real robots. Next best view problems are also applications where mental rotation is calculated in order to obtain a position to construct 3D models of objects [11] [12] or to obtain an automatic surface acquisition [13].

Perspective taking has begun to be used in HRI research field in the last years. Richarz et al.[14] use the area in front

of a human in order to obtain pointing places so that the robot can interact with him. Trafton et al. show in [15] and [16] a robot system that uses geometrical reasoning in perspective taking to take decisions about human reasoning. They prove how human-human interaction helps if it is applied to human-robot interaction. [17] and [18] describe in a 3D simulated environment using perspective taking to help a robot with the learning process, the authors establish some principles in perspective taking for HRI, and treat the problem on high level layer, but don't take into account robot motion to change its own perspective.

Nevertheless, to our knowledge mental rotation and perspective taking have not been used for a mobile robot who is in a close interaction with humans. In this article we present the importance of these characteristics in finding robot configurations to place itself.

### 3. PerSpective Placement (PSP) for HRI

As in every robot motion planning, the robot has to find a continuous path between its actual configuration (starting point) and its final configuration (destination point). Perspective Placement (PSP) is concerned by this last part of motion planning.

We define perspective placement as a set of robot tasks to find, validate and evaluate robot configurations in a human shared environment, performing perspective taking and mental rotation in order to interact and ease communication with humans.

#### 3.1. PSP for Humans

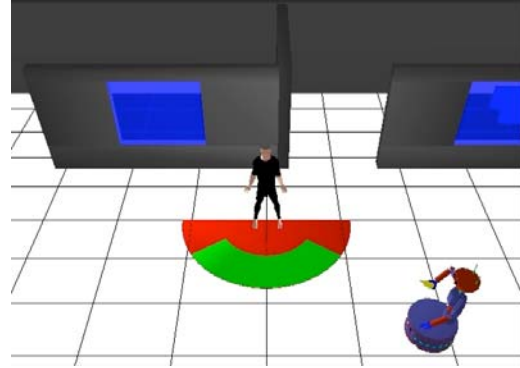
To interact with human, the robot has to find how to place itself in a configuration where it has direct contact with this person. This constraint helps to reduce search space to find such a point. This search can be subdivided in two phases: first, the robot has to find positions that belong to the human field of view represented as a semicircle area in front of human (with preference on the attentional field of view) and; second, it has to validate positions obtained for visual contact, preventing big visual obstructions from blocking robot perception.

The area in human attentional field of view, called here as "Interaction Area", is defined as the zone in front of the person by the angle and by the radius:

$$\alpha_{view} \mid 0^\circ \leq \alpha_{view} \leq 180^\circ$$

$$Rad \mid Rad_{min} \leq Rad \leq Rad_{max}$$

$\alpha$  and  $Rad$  and are defined by the characteristics of the interaction (e.g. proximity needed for just visual interaction or handing an object), by robot sensor capabilities and by human preferences.



**Figure 1. The interaction area is marked as the band in front of the human.**

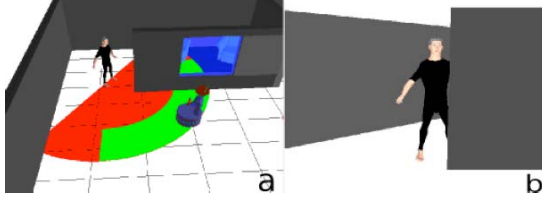
Once defined this area, illustrated in figure 1, a set of points around the person are generated and will be chosen in order to ensure these properties:

- **Collision Free:** Robot in this position must not be in collision neither with objects and persons nor with itself.
- **Sensor Oriented:** Selected sensors must be oriented towards the human in order to perceive it.
- **Without Visual Obstructions:** In sensor's acquisition, human has to be perceived depending on the interaction task.
- **Human-Like Positioning:** Based on user studies [19] of robot-human spatial placement, robot has to find spatial formation with the person in a human acceptable way.
- **Minimal Cost:** Robot should find a position that minimizes the cost based on human-robot distance and on human's comfort [20] [21].

To determine what is perceived by a camera, we use 2D perspective projection of the 3D environment. This projection is obtained from the sensors's position when the robot is placed in desired configuration. The projection obtained is a matrix  $MatP$  where the value of the position  $(x, y)$  represents one point in the object's projection image in sensor's field of view. In the figure 2, 2D projection is illustrated.

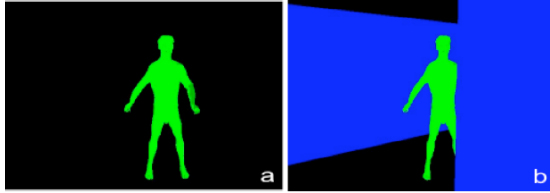
We define "Projection rate"  $Pr$  as the projection percentage of an element  $El$  (object, human or obstacle) on the environment represented in  $MatP$ .  $Pr$  is obtained by:  $Pr(El) = \sum MatP(x, y) \mid (x, y) \in El$ .

The projection rate of an element that is not projected  $Pr_{hidden}$  can be obtained with:  $Pr_{hidden}(El) = Pr_{desired}(El) - Pr_{visible}(El)$  where  $Pr_{visible}$  is the projection rate that considers visual obstructions (only visible



**Figure 2. a) Robot positioned in the interaction area. b) Computed robot perception (2D projection)**

projection). On the other hand,  $Pr_{desired}$  is the relative projection obtained without considering objects in the environment (as it should look without visual obstacles). In figure 3 we can observe the difference between desired and visible relative projections.



**Figure 3. Relative projections, here the person is the objective and differs from other elements on the environment. a) Desired relative projection b) Visible relative projection**

Objective's  $Ob$  visibility percentage,  $Watch$  is determined by:

$$Watch(Ob) = \frac{Pr_{visible}(Ob)}{Pr_{desired}(Ob)}$$

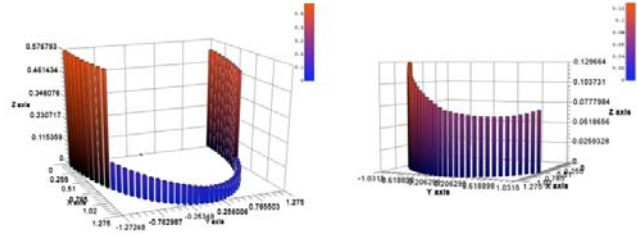
Finally, to know if a point is candidate for perspective placement by:  $Watch(Ob) \geq \mu$  where  $\mu$  is a threshold that corresponds to a desired percentage.

In the other hand, costs for each point are, as we have introduced before, based partially on the direct distance from robot's actual position  $Cost_{distance}$ , and the total cost is calculated as:

$$Cost_{total}(x, y) = (Cost_{distance}(x, y) * Gain_{distance}) + ((Cost_{frontal}(x, y) + Cost_{inarea}(x, y)) * Gain_{frontal}) + (Cost_{HAMP}(x, y) * Gain_{HAMP})$$

Where  $Cost_{frontal}$  is the distance from the front of the human gaze orientation, to incite a *face-to-face* human like position.  $Cost_{inarea}(x, y)$  is a fixed cost depending if the

point is in the interaction area (cost is higher otherwise).  $Cost_{HAMP}$  is the cost of the point calculated by the Human Aware Motion Planner (explained in detail on [20] and [21]) based on human safety and comfort.  $Gain_x$  is the weight given to this criterion. An example of calculated costs is shown in figure 4 for the scenario of figure 1.



**Figure 4. Computed costs of the points. a) Points in all around the field of view area, those points out of the interaction area have the highest cost. b) Points on the interaction area, lower cost are due to robot proximity and that are closer to front.**

Maximal and minimal ranges of interaction area are predefined based on Hall's interpersonal distances mentioned in [19] depending on desired interaction task.

### 3.2. PSP for Known Objects

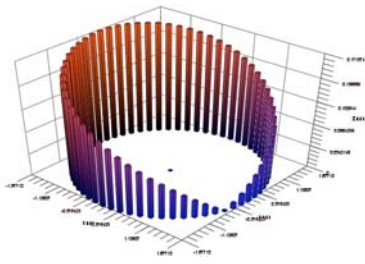
Objects in the environment like tables, desks, sofas, chairs, etc. All considered as fixed obstacles and are called here as "known objects". The robot knows object's model and its position in the environment.

Methodology of perspective placement for a known object, is similar to PSP for humans. Here the object's perspective is not taken into account, so the robot can place itself all around object's position. Nevertheless, robot has to take into account its own perspective to look for the best placement configuration.

In this case "Interaction area" is called "Approach area" (obviously because there is no interaction with unanimated objects). Maximal and minimal ranges are defined based only on collision avoidance and robot sensor limits. For a given point, costs are calculated with:

$$Cost_{total}(x, y) = (Cost_{distance}(x, y) * Gain_{distance}) + (Cost_{HAMP}(x, y) * Gain_{HAMP})$$

An example of costs function for a known object is illustrated in figure 5. Configuration search starts at the least expensive point that contains all the properties mentioned above.



**Figure 5. Costs of the points in modeling area around an object.**

### 3.3. PSP for Non-visible or Unknown Objects

Sometimes there are manipulable objects in the environment that are small. But they are not always at the same place and they can be hidden by another object. The object referenced by the human can be unperceivable to the robot (e.g. too far from robot sensors to create a model or know the exact position and orientation of the referenced object).

Here, our approach to this problem is to create a simple geometrical object, called “search sphere”, that could be set on possible places like the one referenced by human or on surfaces of known objects (e.g. on the top of a table). Discovering the referred object can be interpreted as seeing the search sphere.

Due to the search sphere’s shape and position, its approaching area and its costs functions are placed around the sphere as it is in for known objects. Maximal and minimal approaching ranges are defined by the specifications of the task.

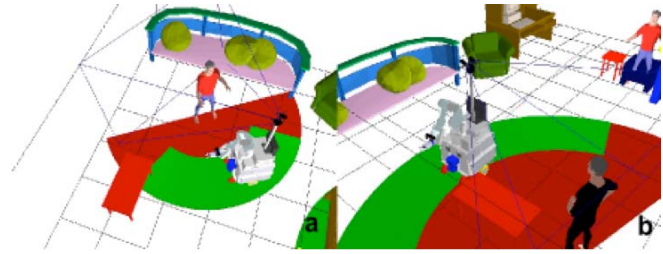
## 4. Simulation Results

Our perspective taking system is implemented in C and integrated and tested within the Move3D [23] software platform developed at LAAS-CNRS.

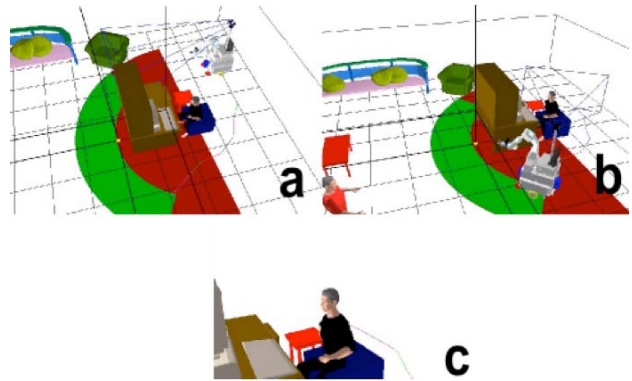
Fig. 6 illustrates two similar situations where the robot has to interact with a person, in the first scenario (fig. 6-a) there is no obstacles between human and robot, and on the second one (fig. 6-b) we can see an obstacle between them but it doesn’t cause a visual obstruction for the robot, the interaction zone is also further from human.

In the scenario illustrated in figure 7, a desk prevents the robot to see the human that is seated behind this visual obstruction. Robot finds a position where it can see the person whom it wants to interact.

Figure 8 gives examples of how PSP can find a configuration for known object (a furniture) in two different situations. In the first case with a person is blocking the obstacle



**Figure 6. Simple case: Results of PSP for Humans. Robot have found a point near the human where it can see him**



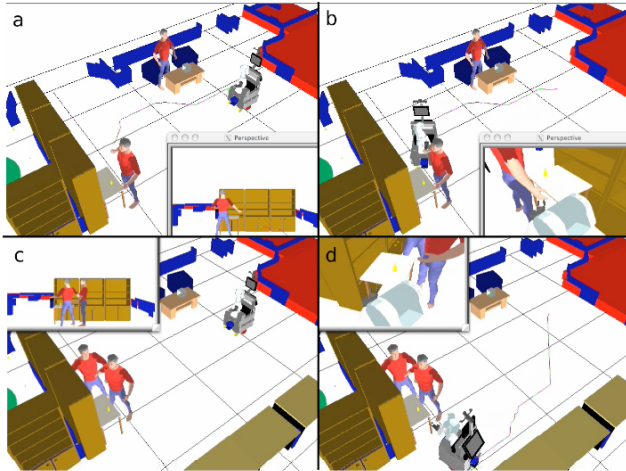
**Figure 7. Results of PSP for Humans with visual obstacles. a) Initial scenario with a human is partially hidden by a desk. b) Robot final configuration where it can see de human. c) Robot’s perspective in its computed configuration**

and causing a visual obstruction, here the robot finds a position on the right side of the person, by avoiding it in order to perceive the table. Note that the robot computes a final configuration where it avoids not only collisions with human or obstacles but also with visual obstacles.

On the second situation an extra person is occluding the table from the position chosen on the first case, here the robot computes a configuration on the other side of table where none of the persons is present on the environment are visually blocking this table.

## 5. Robot Implementation and results

The PSP system is implemented into OpenGenom [24] with the help of other modules of supervision, task planning, human detection and tracking, robot self localization, path and motion planning, etc. that forms the same architecture as described in detail on [25].



**Figure 8. Approaching to the table, a) Initial position, b) Final configuration looking at the table c), d) initial and final configuration with two persons hiding the table.**

The whole system has been carried into our robot *Jido* equipped of three Pentium IV processors, one Laptop with a Intel core-duo processor, two 2D SICK laser scanner (front and rear), a six DOF Mitsubishi PA-10 manipulator, 1 tilt and pan stereo camera, 1 color stereo cameras on the end axis of the manipulator, 8 sonar sensors, and three finger tactile hand.

In figure 9, a scenario is shown with a person interacting with the robot. The person indicates the robot to go to pick up an object on the table and bring it back to him (“bring me the yellow bottle”). The robot makes the respective plan with the sequence of different tasks. PSP module finds a valid configuration to execute each task that implies navigation motion of the robot.

The first configuration is placed to see the table where the bottle is located, maximum and minimum distances are set according to arm capabilities for grasping the bottle. Second configuration is found to get close to human and give him the bottle.

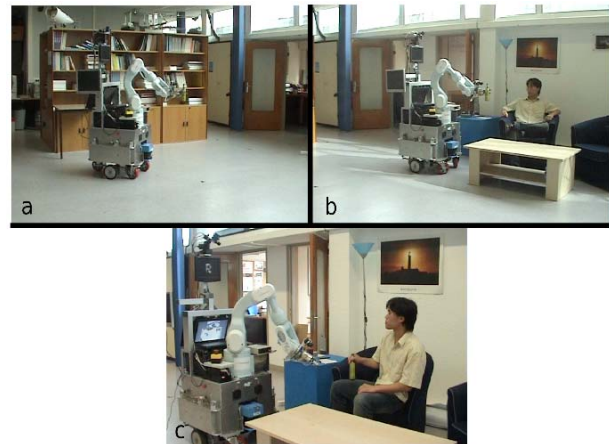
Another example is illustrated in 10, where an obstacle (a table) prevents the robot from several positions close to the human to bring him the bottle.

## 6. Conclusions and future work

In this work, we have shown how using perspective taking concepts added to motion planning techniques can be useful on human-robot interaction. Furthermore, we have presented different algorithms that take into account human and robot perspectives to determine different robot configurations, where the robot can perceive its objective (human or



**Figure 9. First PSP for known objects, approaching to the table, a) Initial position, b) Final configuration looking at the table. Second, PSP for humans, c)-d) Initial and final configuration looking at human**



**Figure 10. PSP to hand an object to a human sat in front of a table a) Initial position, b)-c) Final configuration looking at the human. The robot is perceiving human by cameras on top. The robot finds its final configuration to complete the task.**

objects) without any collision and visual obstructions. This allows robot to plan its trajectory to a position where both, the human and the robot, can see each other and interact.

A possible extension of this approach is a fusion of a task planner with these geometrical tools. This could increase the precision for obtaining viable tasks. Another extension of PSP, could be on the close human-robot interaction where the robot has to reason about manipulating objects, deciding

positions on a shared space with human.

Also we plan to integrate these systems to obtain a general solver which will produce friendly and socially acceptable robot positions.

## Acknowledgements

Special thanks to Assia and Juan Pablo for their comments and suggestions.

## References

- [1] B. Tversky, P. Lee, and S. Main Waring, "Why do Speakers Mix perspectives", *Spatial Cogn. Computat.*, vol. 1, 1999, pp 312-399.
- [2] Lee, P. U. and Tversky, B. "Costs of Switching Perspectives in Route and Survey Description", *Proceedings of the twenty-third Annual Conference of the Cognitive Science Society*, Edinburgh, Scotland 2001.
- [3] Edith K. Ackerman, "Perspective-Taking and Object Construction: Two keys to learning", *Constructionism in Practice: Designing, Thinking, and Learning in a Digital World*, Mahwah, New Jersey, 1996.
- [4] John P. Costella, "A Beginner's Guide to the Human Field of View", School of Physics, The University of Melbourne, November 1995.
- [5] Hansong Zhang, Dinesh Manocha, Tom Hudson and Kenny Hoff, "Visibility Culling Using Hierarchical Occlusion Map", *Proceedings of SIGGRAPH*, 1997.
- [6] Antonio Gomes de Sá and Joachim Rix, "Virtual Prototyping - The Integration of Design and Virtual Reality", *CAD Tools and Algorithms for Product Design*, 1998, 128-150.
- [7] Pascal Mueller, Peter Wonka, Simon Haegler, Andreas Ulmer, Luc Van Gool, "Procedural Modeling of Buildings", *ACM Transactions on Graphics*, volume 25, number 3, pages 614-623. 2006.
- [8] M. Alejandra Menchaca-Brandan and Andrew M. Liu and Charles M. Oman and Alan Natapoff, "Influence of perspective-taking and mental rotation abilities in space teleoperation", *HRI '07: Proceeding of the ACM/IEEE international conference on Human-robot interaction*, 2007, Arlington, Virginia, USA.
- [9] S.W. Hsu and T.Y. Li, "Third-Person Interactive Control of Humanoid with Real-Time Motion Planning Algorithm", *Proc. of IEEE International Conf. on Intelligent Robots and Systems (IROS'06)*, 2006, Beijing, China.
- [10] Josh Faust, Cheryl Simon, William D. Smart, "A Video Game-based Mobile Robot Simulation Environment", *Proc. of IEEE International Conf. on Intelligent Robots and Systems (IROS'06)*, 2006, Beijing, China.
- [11] Andrea Bottino and Aldo Laurentini, "What's NEXT? An interactive next best view approach", *Journ. of Pattern Recognition*, 2006, pages 126-132.
- [12] Y.F. Li, B. He and Paul Bao, "Automatic view planning with self-termination in 3D object reconstructions", *Journ. on Sensors and Actuators 2005*, Pages 335-344.
- [13] Richard Pito, "A Solution to the Next Best View Problem for Automated Surface Acquisition", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1999.
- [14] Richarz, J. Martin, C. Scheidig, A. Gross, H.M., "There You Go! - Estimating Pointing Gestures In Monocular Images For Mobile Robot Instruction", *Robot and Human Interactive Communication ROMAN 2006.*, Univ. of Hertfordshire, Hatfield, UK, .
- [15] J. Gregory Trafton, Nicholas L. Cassimatis, Magdalena D. Bugajska, Derek P. Brock, Farilee Mintz, Alan C. Schultz, "Enabling effective human-robot interaction using perspective-taking in robots", *IEEE Transactions on Systems, Man, and Cybernetics, Part A*, 2005, 460-470.
- [16] Trafton, J.G.; Schultz, A.C.; Bugajska, M.; Mintz, F., "Perspective-taking with robots: experiments and models", *Robot and Human Interactive Communication ROMAN 2005*, Page(s): 580 - 584.
- [17] Cynthia Breazeal, Matt Berlin, Andrew Brooks, Jesse Gray, Andrea L. Thomaz, "Using perspective taking to learn from ambiguous demonstrations", *Robotics and Autonomous Systems*, 2006, pgs: 385-393.
- [18] Matt Berlin, Jesse Gray, Andrea L. Thomaz, Cynthia Breazeal, "Perspective Taking: An Organizing Principle for Learning in Human-Robot Interaction.", *AAAI*, 2006, Boston, Mt.
- [19] Helge Huettenrauch, Kerstin Severson Eklundth, Anders Green, Elin A. Topp, *Investigating spatial relationships in human-robot interaction*, Proc in (IEEE/RSJ) International Conference on Intelligent Robots and Systems, 2006. Beijing, China
- [20] E. Akin Sisbot and Luis F. Marin and Rachid Alami and Thierry Siméon, *A mobile robot that performs human acceptable motion*, Proc in (IEEE/RSJ) Interna-

tional Conference on Intelligent Robots and Systems, 2006. Beijing, China

- [21] E. Akin Sisbot and Luis F. Marin-Urias and Rachid Alami and Thierry Siméon, *A Human Aware mobile robot motion planner* IEEE transactions on Robotics, vol. 23 No. 5. Oct.2007.
- [22] E. Akin Sisbot and Luis F. Marin-Urias and Rachid Alami, *Spatial reasoning for human robot interaction* Proc in (IEEE/RSJ) International Conference on Intelligent Robots and Systems, 2007. San Diego, CA. USA.
- [23] T. Siméon, JP. Laumond, F. Lamiroux, *Move3D: a generic platform for motion planning*, 4th International Symposium on Assembly and Task Planning, Japan,2001.
- [24] S. Fleury and M. Herrb and R. Chatila *Genom: a tool for the specification and the implementation of operating modules in a distributed robot architecture* IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 1997, Grenoble, FR.
- [25] E. Akin Sisbot and Maxime Ransan and Aurelie Clodic and Rachid Alami, *Supervision and Motion Planning for a Mobile Manipulator Interacting with Humans* Proc in (ACM/IEEE) International Conference on Human Robot Interaction, 2008. Amsterdam, HL.