



HAL
open science

Head-Body Motion Coordination for Human Aware Robot Navigation

Harmish Khambhaita, Jorge Rios-Martinez, Rachid Alami

► **To cite this version:**

Harmish Khambhaita, Jorge Rios-Martinez, Rachid Alami. Head-Body Motion Coordination for Human Aware Robot Navigation. 9th International workshop on Human-Friendly Robotics (HFR 2016), Sep 2016, Gênes, Italy. 8p. <hal-01568838>

HAL Id: hal-01568838

<https://laas.hal.science/hal-01568838v1>

Submitted on 25 Jul 2017

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.



HAL Authorization

Head-Body Motion Coordination for Human Aware Robot Navigation

Harmish Khambhaita¹, Jorge Rios-Martinez², and Rachid Alami¹

Abstract—Mobile robots equipped with a pan-tilt head need to use gaze direction to manifest its navigational intents for more acceptable human-robot interaction. We frame control of such gaze behavior as multi-criteria decision-making problem, and provide a solution to synchronize gaze control with robot’s navigation planner. This approach is useful in the context of robot navigation, where it may be inapt to display only a predefined gaze pattern due to the dynamic nature of the scene. By enabling two behaviors, *look-at-path* and *glance-at-human*, we demonstrate the effectiveness of our approach on a real robotic platform in a path crossing scenario. Furthermore, we discuss results of a video based user study conducted with 126 participants showing improved communication of robot’s navigational intentions with the proposed approach.

I. INTRODUCTION

Human-aware navigation planners already provide safe and socially acceptable motion of a robot [1]. Furthermore, directional cost functions [2] have shown to increase legibility of robot motions, where a robot attempts to solve spatial conflicts by adjusting velocity instead of path when possible. This approach is preferred by the human counterparts, however, it is not enough to reduce hesitation of humans and to ensure legibility and acceptability of the robot behavior. We use the modality of robot gaze to give explicit information about robot’s future plans and goals, and show how it enhances human-robot interaction during navigation. Specifically, our robot continuously looks ahead at its path to help humans anticipate its immediate navigation plans. It also performs a saccade like behavior towards a human partner to convey that the robot has seen her/him and it is going to avoid possible path conflicts. By taking navigation and gaze planner as a whole, our system can be used to produce more acceptable and legible robot behavior.

It is well known that humans predominantly use the modality of gaze as a cue for understanding intentions and mental states of others. “Even in the absence of any overtly executed action, observers can still read other people’s motor intentions, provided they can see a model’s face, in particular his or her gaze direction” is concluded by Castiello [3] in one of the most influential studies of human intention recognition. Besides, research in human behavioral psychology reveals that the constant alignment of the head with walking direction, as well as alignment of gaze with other humans and with static objects for the purpose of obstacle avoidance [4] are the prominent gaze behaviors associated with locomotion.

¹LAAS-CNRS, 7 avenue du Colonel Roche, 31400 Toulouse, France. {harmish.khambhaita,rachid.alami}@laas.fr

²Universidad Autónoma de Yucatán, Anillo Periférico Norte, Chuburna Hidalgo Inn, 13615, CP 97144, Mérida, Yucatán, Mxico. jorge.rios@correo.uady.mx

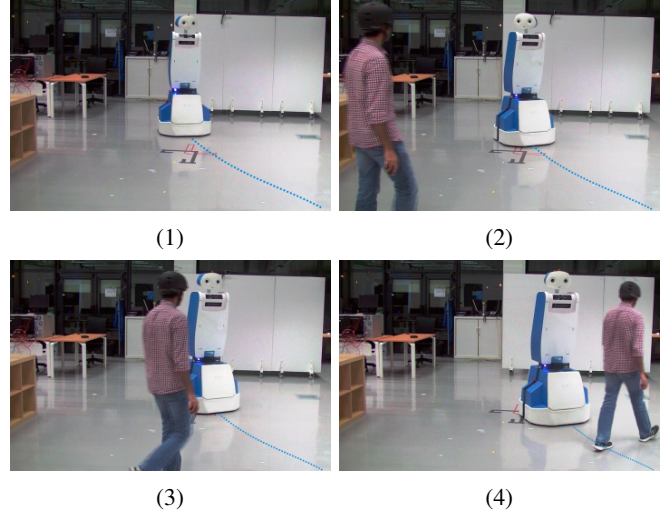


Fig. 1: A typical sequence of events during human-robot path crossing scenario. (1) The robot starts looking at the path. (2) A human enters in the area, and the robot detects a possibility of interference; the robot turns its head towards him. (3) After acknowledging the human, the robot turns its head back towards the path and slows down at the same time. (4) The human passes in front of the robot, the interaction episode ends.

How a robot with a pan-tilt head unit could simultaneously achieve such motions is still an open issue.

We propose the formulation of head motion control on a mobile-robot platform as a multi-criteria decision-making problem and solve it using the analytic hierarchy process, which is a novel contribution of this paper. We also provide implementation specifics for a synchronized head-behavior module that exhibit *look-at-path* and *glance-at-human* (see III-B) behaviors. A video-based pilot user study that evaluates our approach in real-world scenario is an additional key contribution.

Based on the behavioral psychology findings, on the onset of our investigation, we imagined a typical sequence of events involved in a human-robot path crossing situation as shown in Fig. 1. To achieve such interaction, we will now look into relevant parameters and social context, explored so far, associated with effective gazing behavior.

II. RELATED WORK

The larger part of nonverbal human-robot interaction literature belongs to the studies of robot gaze and its utility for social signaling. Researchers have successfully used robot gaze to increase attention and thus engagement of

users [5], [6]. Moreover, sometimes deliberate delays [7] or premeditated motion patterns [8] are required to capture the attention of the partner. On the other hand, gaze direction can also communicate attentiveness and visual awareness of the robot [9]. Human-like head movements, adapted to the context of an interaction task, can help improve the fluency of the interaction. When well synchronized with rest of the body, gaze can even manifest forethought to improve robot's readability [10]. That being discussed, implementations of particular gaze behaviors lack generality and remain largely ad-hoc in nature.

A plurality of the robot gaze research is dedicated to situated interactions. It is only lately the community have begun to explore the effects of robot gaze in navigation context. Fiore et al. [11] have found that gaze affects the intensity of perceived emotional states when interacting with the cues of proxemics in a human-robot path crossing scenario. Nonetheless, their findings indicate that gaze of the robot is not as important a social signal for that particular scenario, meaning the gaze have no interactive effects on perceived social presence. This findings could be the consequence of the timing of particular gaze behavior execution, where the robot head was orienting towards human only in the beginning of an interaction episode. Our results, presented in Sec. IV-C, beg to differ in this regard. May et al. [12] have compared head orientation and visual light indicators for communicating turning signals, where a pre-scripted behavior controlled the head orientation. Their results indicate a positive impact on comfort felt by humans using both communication modalities, with participants favoring the indicators. It is, however, difficult to show the extent of turning direction with such indicators. Lu [13] have also tested a scripted *glance* behavior during the task of jointly navigating in a hallway, robot looking at passing humans for a certain amount of time and then looking back in front of the robot. However, they remain inconclusive whether the behavior was effectively giving the person acknowledgment that the robot saw them.

Search for the essential gaze control parameters leads us to follow the research on the gaze behavior associated with human locomotion. Humans initiate turning of head in the direction of travel before the body during locomotion. This anticipatory nature of the head orientation is known for a long time [14]. It is also evident that the head orientation control is initiated several meters before the turn [15]. With a series of experiments Bernardin et al. [16] have quantified the anticipation, they have observed that at normal walking speed the gaze shifts by 500 to 700 milliseconds in advance compared to the body. Besides that, the angle for gaze anticipation into the heading direction increases with the increasing curvature of the path. A recent user study by Unhelkar et al. [17] supports these findings and points out that in addition to *head orientation*, *body velocity* is also statistically important for anticipating a turning motion in humans. Kitazawa and Fujiyama [4] have investigated gaze patterns in a collision avoidance scenario with multiple pedestrians moving in a wide hallway shape area. Observed

participants were fixating their gaze on other pedestrian and static obstacles for avoiding a collision. Reported average distance of fixation was 3.97 m (SD = 0.54) for approaching participants and 1.9 m (SD = 0.71) for leading participants (average time for the fixation is not reported). Furthermore, during the experiment the participants fixated their gaze on other pedestrians when it was really necessary for collision avoidance. Put succinctly, the two most significant factors for the robot gaze are the direction of the path and other humans in the proximity.

Because there are multiple factors to take into account for controlling the robot's head direction during navigation, the decision for switching the head direction becomes complex. A behavioral framework developed by Srinivasan et al. [18], makes use of the social context for efficient generation of head behaviors. Authors have classified several types of social head gaze behaviors based on previous studies in human-robot interaction literature. We have adopted their proposed vocabulary in our implementation.

Zaraki et al. [6] have attempted a layered approach putting an attention layer between perception and gaze control layers. The attention layer selects the human that exhibits highest weighted sum of individual social features, and the robot turns its gaze toward the selected human. The specialized function for our *glance-at-human* task (see III-B) is inspired from this work. Only recently Yoo and Kim [19] have used a multi-criteria framework for gaze control, analogous to our proposal in Sec. III-A, however for a static robotic head. Their proposed algorithm considers all possible discrete gaze (pan and tilt) positions in the visual field as a candidate goal for the head controller. Seven of these discrete gaze positions are periodically selected as criteria, based on certain factors that affect the human gaze, e.g. faces, objects. Fuzzy measures reflecting user-defined weights of the factors are applied to the criteria points and the point with the maximum value is used for directing the head. Since these criteria points are of the same type as other candidate points, the solution process is comparable to that of a prioritization method, which is the critical difference from our proposed method. As one will see in Sec. III, we use several candidate gazing functions to select the most pertinent one depending on current perceptual inputs. Each function independently generates gazing points while respecting the timing constraints required by particular head behaviors. Our criteria to evaluate these candidate gazing functions are dynamically rated on perceived social context.

III. APPROACH

Different studies have shown that multiple, often competing, objectives are involved in human head gaze behavior. Selection of gaze direction depends on the prominent objective in the current *social context* that robot derives from its perceptual events. We make use of four social head gaze behaviors, summarized by Srinivasan et al. [18], which are:

- *Communicating Social Attention*, where robot shows interest by looking at a human.

- *Manifesting an Interaction*, where gaze turns towards objects relevant for current task, for instance obstacle-avoidance.
- *Projecting Mental State*, which includes expression of motion intent.
- *Establishing Agency*, where head gaze reinforces human-like aliveness.

These social behaviors are the criteria upon which we evaluate each of the available alternative functions for pointing the robot gaze. Process of updating the most relevant *social behavior* is usually continuous and rapid, especially for mobile robots as they navigate through highly dynamic environments like shopping malls, museums or airports. Notably, previous work appears to neglect this high dynamic nature of robot gaze behavior.

A. Framework

We see generation of overall *robot head behavior* during navigation task as a multi-criteria decision-making problem. One can think of a robot having multiple choices for pointing its head, e.g. looking into the direction where it is going to move the next moment; or looking at an object to act upon; or simply moving the head down to express an emotion of sadness etc. The choice of the best among several alternatives requires evaluating each alternative against a set of criteria indicative of the social gaze behaviors. These criteria are sometimes common among multiple situations. However the importance of these criteria may change as the situation, goal or task changes. This importance is quantified with weights (a list of user-defined scalar values).

Distinctive choices for gazing points usually originate from different types or sources of information. For example, human detection module provides the position of a human to look at. Similarly, manipulation planning module can provide information about the next object to grasp. However, often there are more than one alternatives that arise from the same source of information. Consider a situation where there are multiple humans in front of the robot, who should the robot look at? Selection within these alternatives is subjected to dedicated computation methods. Thus, we propose a scheme where information from distinct sources is processed using definite *behavioral functions*. Each of the behavioral function takes a set of candidate 3D points (C) belonging to the same source and computes a single candidate point (p_f) for directing the robot gaze. Therefore, these functions are of the following signature,

$$f : C \subset \mathbb{R}^3 \times \mathcal{D} \rightarrow \mathbb{R}^3 \times \mathbb{R}^k$$

where, \mathcal{D} is set of domain-dependent parameters and k is the number of criteria under consideration.

The analytic hierarchy process (AHP), introduced by Saaty [20], finds a solution for multi-criteria decision-making problem by assessing and prioritizing the options. It uses a multi-level hierarchical structure of objectives, criteria, sub-criteria, and alternatives. Solution using AHP involves paired comparison of involved criteria, which gives us the weight vector (w_j). At present, however, we only consider

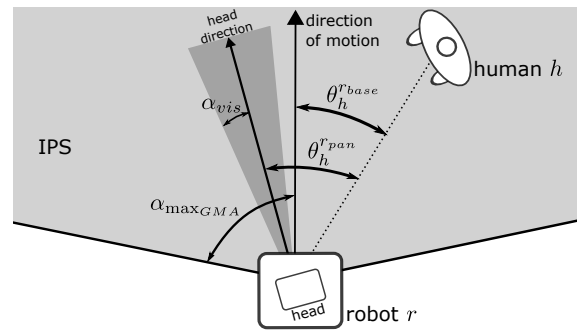


Fig. 2: Illustration of angles involved in computation of gazing point for the robot head. $\alpha_{\max_{GMA}}$: maximum gaze-movement angle, α_{vis} : visibility angle, $\theta_{h_L}^{base}$: robot base to human angle, $\theta_{h_L}^{pan}$: robot head-pan axis to human angle, IPS: information process space (area in light grey color).

one level of criteria for our purpose. Once the weight vector is computed, an alternative is chosen that receives the maximum score,

$$P_{head} = \arg \max_{p_f} \sum_{j=1}^k v_j(p_f) \cdot w_j$$

where $v_j(p_f)$ is a value vector assigned by the behavioral functions, it represents importance of the candidate point p_f with respect to each criteria.

B. Behavioral Functions

A well-designed implementation is needed to demonstrate capabilities of the framework. To this end, we have implemented two gaze behaviors that we found most pertinent for social navigation:

- *look-at-path*: looking at the planned navigational path.
- *glance-at-human*: acknowledging humans with a short glance.

We will now detail the associated behavioral functions along with the underlying communication architecture.

1) *The look-at-path behavior*: Standard practice in navigation planning is to differentiate between generation of a geometric path with a static map of the environment (*global-planning*), and computation of the motor commands avoiding dynamic obstacles (*local-planning*). For *look-at-path* functionality, we use output trajectory from the *local-planning* module, as it gives us the best estimate of where the robot is going in the immediate future. This functionality corresponds to two social gaze behaviors, *Projecting Mental State* and the *Establishing Agency*. Therefore, the alternative point provided by this functionality has a higher score (1.0 during the experiments) for the corresponding attributes.

Angle between the movement direction and the gaze direction is defined as *gaze-movement angle* (GMA) by Park et al. [21], illustrated in Fig. 2. Limits on this GMA value define an *information process space* (IPS), which is a conceptual area that determines the spatial boundary for observing humans. IPS is the area which humans use for

avoiding potential collisions and shown to depend on the current velocity. We likewise enforce limits on robot GMA to enforce naturalness of head motions. When the path-planner gives a trajectory in which the last point of the trajectory is outside this maximum allowed GMA ($\alpha_{\max_{GMA}}$), we limit the robot pan-angle to remain within the IPS.

2) *The glance-at-human behavior*: When the robot detects humans in its environment, a second function is activated for calculating an alternative gazing point, that leads the robot to perform a saccade like behavior towards the human. Each of the behavioral functions provides only one alternative for the decision process. Hence, when multiple humans are detected, the robot favors the one that requires the most urgent attention. Both relative position and velocity of detected humans are taken into account for ranking which human to look at first, step-5 in algorithm Alg. 1. This ranking scheme is inspired from the experiments of Kitazawa and Fujiyama [4], which shows that it is preferable to look at the human which is coming towards the robot compared to the one who is going away. The \vec{r} and the \vec{h} represent the position vectors for the robot and the human respectively.

Furthermore, we need to keep track of humans that are already acknowledged by the robot (in set \mathcal{H}_L), to avoid triggering multiple saccade behaviors on new position updates. We have defined a visibility angle α_{vis} , the human is considered acknowledged once found within this angle and *tracking-id* of the human is added to the set \mathcal{H}_L . Triggering of further saccade behavior is suspended during an active saccade. This relatively straightforward procedure for dealing with human tracking updates is written down in Alg. 1. The procedure `HumansUpdate` receives a list of position of currently tracked humans (\mathcal{H}), current robot position (\vec{r}) and positions of robot joints (\mathcal{J}). The algorithm returns single human (h_L) position for the robot to perform a saccade behavior towards that human.

Algorithm 1 HUMANSUPDATE(\mathcal{H} , \vec{r} , \mathcal{J})

- 1: $[\theta_{h_L}^{pan}, \theta_{h_L}^{base}] = \text{COMPUTE_ANGLES}(\mathcal{J}, h_L)$
 - 2: **if** $\theta_{h_L}^{pan} < \alpha_{vis}$ **or** $\theta_{h_L}^{base} > \alpha_{\max_{GMA}}$ **then**
 - 3: $\mathcal{H}_L \leftarrow \mathcal{H}_L \cup h_L$
 - 4: **else**
 - 5: $h_L = \arg \max_{h \in \mathcal{H} \setminus \mathcal{H}_L} \frac{(\vec{r} - \vec{h}) \cdot (\dot{\vec{r}} - \dot{\vec{h}})}{\|\vec{r} - \vec{h}\|^2}$
 - 6: **end if**
 - 7: **return** h_L
-

Three of the social gaze behaviors, *Establishing Agency*, *Communicating Social Attention* and *Manifesting an Interaction* afford this *glance-at-human* functionality, therefore, respective attributes are set to a higher value (1.0 during the experiments). We do not prune the head-tilt angle (β) in this function, the limits of which are enforced by the head-tilt motor controller. On each point-head request the head motor controllers compute the pan- and the tilt-joint velocities with a hyperbolic function that is proportional to the difference of current (\mathcal{J}) and required joint states.

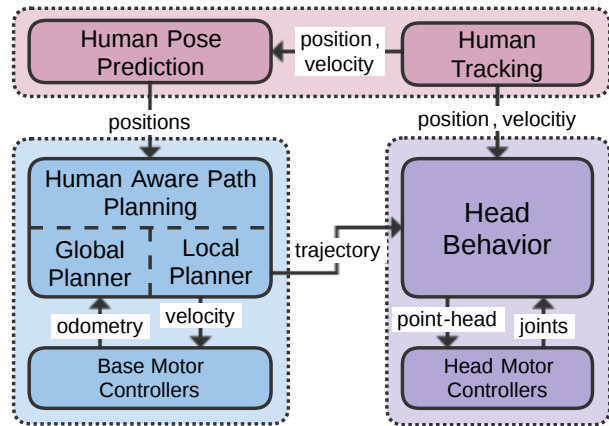


Fig. 3: Architecture of the proposed head behavior scheme.

C. Implementation

For path planning, we have adopted the widely used and tested “move_base”¹ navigation framework on our robot. As a local-planner we have used a modified version of directional-cost based algorithm developed by Kruse et al. [2]. This human-aware path planning algorithm makes the robot slow down when a collision with a human is predicted, avoiding abrupt changes in the path direction. Since the *look-at-path* behavior is tightly coupled with the local-planning module, sudden changes in the path direction would result in an unwanted trembling head motion. The directional cost-based algorithm avoids such abrupt change in motion direction and, therefore, it is particularly useful for our application.

We employed OptiTrack² motion capture system for human tracking, which publishes positions and velocities of detected humans (denoted by set \mathcal{H}) at a certain frequency (10 Hz during our experiments³). We have also developed a simple velocity-obstacle based human pose prediction module to support the human-aware path planning. Overall architecture schema for joint control of head and base of the robot is illustrated in Fig. 3. The *Head Behavior* module is developed as original contribution for this work. Other modules of the framework are comprised of our previous work and widely used open-source ROS modules.

As the local-planner is responsible for dynamic obstacle avoidance. At every iteration it generates a trajectory (a series of 2D points) and computes velocity commands for the robot base controller. The *Head Behavior* receives this planned trajectory, and selects the endpoint of the trajectory to compute the next gazing point for the *look-at-path* behavior. We add a scalar constant value to the Z-coordinate to adjust the height of the gazing point, and thus the head-tilt axis.

The *Head Behavior* software module is well integrated

¹http://wiki.ros.org/move_base

²<http://www.optitrack.com>

³Although the motion capture system delivers data at higher frequency (about 100 Hz), we apply a moving average filter and re-sample the filtered data at 10 Hz to have better estimate of velocities.

with the ROS⁴ framework and tested on a service robot that was deployed for duration of three weeks at Amsterdam-Schiphol airport for guiding passengers [22]. Nevertheless, the module is not limited to that particular robot. In fact, any robot with a *head-pan* and *head-tilt* joints would benefit from it⁵.

D. Synchronization

Synchronization between head- and body-joints of the robot, as well as between robot joints and tracked human positions is very important to achieve a desired and meaningful behavior, which requires substantial engineering and programming efforts to accomplish. Perhaps this is one of the reasons, why we see a big disproportion in the implementations of gazing behaviors among static and dynamic situations.

With our implementation, we were able to achieve the desired collaboration in human-robot path crossing scenario as shown in Fig. 1. Time-series plot of commanded and actual pan-angles is presented in Fig. 4, for a scenario in which human moves across the robot’s path in front of the robot. The robot starts with moving and looking towards its navigational plan; it executes the *glance-at-human* behavior when the human appears on the scene, and after finishing the glance behavior the robot again turns its head towards the plan until it reaches the goal position.

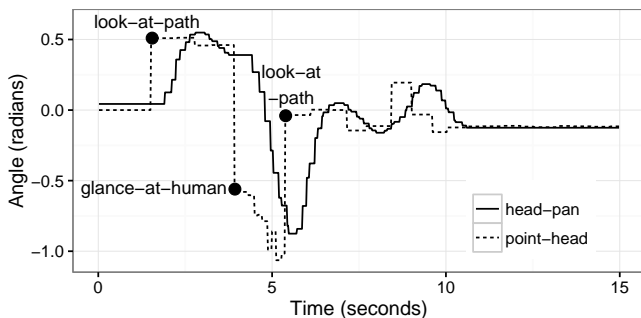


Fig. 4: An example time-series plot of commanded gaze points (point-head) and the angle between the front position of robot base and head-pan axis (head-pan). Marker-points indicate change of behavioral functions.

IV. EVALUATION

Based on the design of head behaviors, we anticipate that human’s perception of the robot motion will improve both in interactive and non-interactive situations. We set-up a video based user study to determine whether it is true that the proposed head movements visibly improves the robot motion quality. Hypotheses for the outcome of the study were following:

- *Hypothesis 1*: Anticipatory head movements during navigation will positively affect the perception of the robot’s navigation intents.

⁴<http://www.ros.org/>

⁵Source code for the module is available at https://github.com/harmishhk/hanp_head_behavior

- *Hypothesis 2*: Head behavior with both of the *look-at-path* and *glance-at-human* functionalities will be evaluated as more favorable over no head movements.

Hypothesis 1 is related to objective improvement in robot motion, whereas the *Hypothesis 2* is linked to the subjective improvement since it concerns participants’ subjective choice about the robot behavior.

A. User Study Design

There is a precedent for video based studies performed with real [23] as well as a simulated [10] robot for human-robot interaction. Results of the study by Syrdal et al. [23] suggest that video prototyping is an excellent source to gain insights regarding user experiences related to the assessment of the human-robot interactions. Furthermore, the video prototyping allowed us to engage geographically diverse set of participants while providing a consistent experience for each of them. This study was conducted in four different languages and with participants from three different countries.

We deployed the *Head Behavior* modules on a custom designed service robot platform, equipped with a two degrees-of-freedom head. Head pan joint can fully turn towards the back of the robot, while tilt joint is physically limited to move up to 30° up and down from the straight looking position. Head of the robot has two passive eyes that can only move horizontally to the left and the right, however, they were not used in our analysis. The robot is also equipped with a display in the back, which was switched off during the recording.

We recorded seven videos in an experimental area of about 5.5 by 9 meters. Surrounding place to the experimental area was partially covered with large wooden boards to reduce any distraction caused by it. Lighting conditions were maintained same throughout the recording. And the audio was removed from the video before it was shown to the participants. During experiments the weight for *Establishing Agency* criteria was assigned to $w = 0.9$, remaining weights were set to value 1.0. We mentioned that the total distance where the robot looks depends on the immediate path of the robot, in these experiments the robot was planning for a time-window of 4 seconds and thus watching the point where it will reach 4 seconds in the future. These values were empirically found to be most suitable for our task, however, they can be learned and adapted on-line.

The first four of these videos did not involve any human. Here the task of the robot was to move from the start position as shown in Fig. 5a to one of the end positions, either going towards the door or the corridor. For both of the goal positions, we recorded two videos, one with the *look-at-path* behavior and other without (head position fixed, looking straight ahead). After recording, all four videos were cropped until the frame where the body of the robot starts moving towards the goal direction. Only when the *look-at-path* behavior was enabled, the robot head starts moving at the marked position in Fig. 5a.

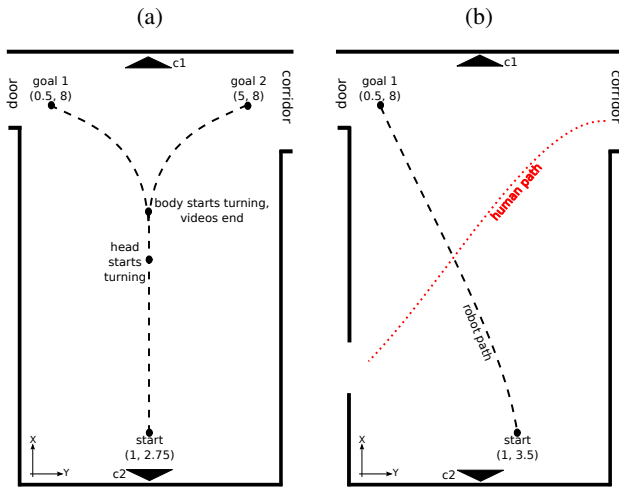


Fig. 5: Illustrative layout of the experiment space, (a) for recording videos without human and (b) for videos with human.

Last three videos were recorded where a human played the role of an interferer who crosses the robot path as shown in the Fig. 5b. The human was tracked with four passive reflective markers placed on a light-weight helmet, that the human wore. The three videos were recorded with following three conditions:

- (A) All of the head-behavior functions disabled.
- (B) Only the *look-at-path* function enabled.
- (C) Both the *look-at-path* and *gaze-at-human* enabled.

All of the videos in the study were recorded with two viewpoints⁶, one in the front of the robot and another in the back, denoted by c1 and c2 respectively in Fig. 5. Each experiment was shown to the participants from both viewpoints. These videos were embedded in a web-page⁷, after watching each video, the participants were requested to answer a question on the same web-page. A short introductory message about the study procedure was displayed before showing the videos.

B. Procedure

The study procedure was divided into two parts.

- Part 1, where we only manipulated the *look-at-path* behavioral condition.
- Part 2, where we manipulate both the *look-at-path* and *glance-at-human* behavioral conditions.

For the first part, participants were randomly assigned to watch videos where the actual navigation goal was one of the two positions shown in Fig. 5a. This way we remove any effect on the goal direction caused by the experimental setup or the robot appearance. Once assigned to an option, participants watched one video where movements of the robot head were disabled and another video of robot exhibiting the *look-at-path* behavior. The order of the videos

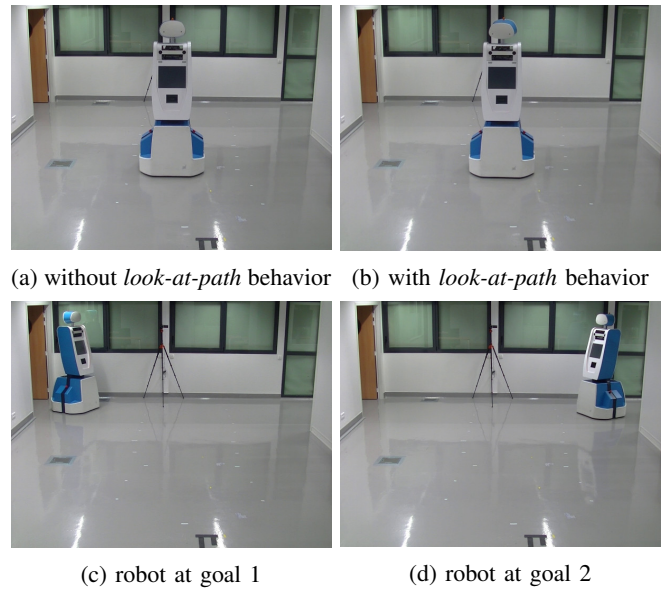


Fig. 6: Condition and goals *look-at-path* behavior.

was counterbalanced across participants. After watching each video, they were posed with the following question.

- Where is the robot going?

To respond to this question, participants were asked to choose one of the two goal positions, as shown in Fig. 6c and 6d.

As discussed earlier, to remove the influence of robot’s body-movements on participants’ answers, we stopped the video just at the moment where the body starts turning towards the goal position, Fig. 6a and 6b show the snapshots of these moments. Participants’ response to these videos will be used to test the *Hypothesis 1*. It should be noted that we have added a couple of virtual obstacles in the planning cost-map to “force” the path planning system to produce a path that makes the robot move straight forward for a certain distance before initiating the goal-directed turning motion.

The second part of the procedure concerns the *Hypothesis 2*. Participants were shown three videos with the conditions A, B, and C, as listed before. Again, the order of the videos was counterbalanced across participants. After watching all three videos, the participants were asked to choose which behavior of the robot they like the most and the least. Once participants register their liking, we requested them to see the videos with condition A (no head behavior) and condition C (both *look-at-path* and *glance-at-human* behaviors enabled) once more, in counterbalanced orders. After watching each video, participants were asked to rate their responses to the following statements on a five-level Likert scale ranging from *strongly disagree* (= 1) to *strongly agree* (= 5):

- Q1: The robot noticed the person.
- Q2: Robot’s actions were clear for the person.
- Q3: Robot’s behavior was often in direct response to person’s behavior.
- Q4: Person did not receive robot’s attention.

These statements are adapted from the social presence measure introduced by Harms and Bioocca [24]. Since our

⁶Videos used for the user study: http://bit.ly/hbmc_video

⁷An example of on-line form: http://bit.ly/hbmc_form

study captures a third person’s perspective, we have selected only four (of total six) dimensions, removing the dimensions that measure affective and emotional aspects of the interaction episode. The four selected dimensions are *co-presence* (Q1), *perceived message understanding* (Q2), *perceived behavioral interdependence* (Q3), and *attentional allocation* (Q4).

C. Results

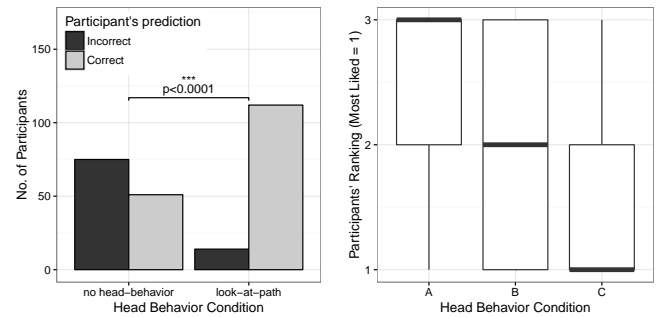
We collected data from $N = 126$ participants between age 17 and 59 ($Mean = 27.82, SD = 6.44$) who volunteered for the video based user study. Out of 126 participants, 67 choose to answer in English, 36 in French, 20 in Spanish and 3 in German.

For the first two videos we manipulated the behavioral condition *look-at-path*, our aim was to get an objective measure of improvement in robot motions. Results are plotted in Fig. 7a, where we see an increase in the ratio of participants who inferred the correct goal when *look-at-path* condition was *true*. The difference in participants’ predictions of the robot goal is significant according to McNemar’s chi-squared test ($\chi^2(DF = 1, N = 126) = 50.704, p < 0.0001$). Since the length of videos with and without *look-at-path* behavior was same, these results show that, in given amount of time, the *look-at-path* behavior results in a more accurate perception of robot’s navigational goal. Therefore, we found support for *Hypothesis 1* concerning accuracy. The results also confirm that in absence of any cue the probability of successful prediction of the robot goal is equivalent to what is expected due to chance.

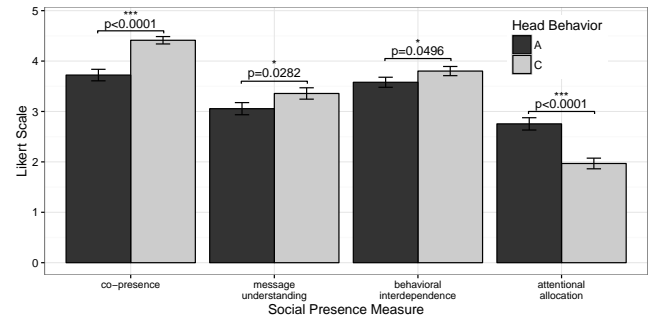
For the second part of the user study involving a human-robot path crossing scenario, the box plot of participants’ liking about above listed three behaviors, A (the robot head looking straight), B (the robot head looking at its path) and C (the robot head looking at path and throwing a glance at the human) are plotted in Fig. 7b. The Friedman rank sum test reveals significant differences in participants’ liking about the three types of robot motions, ($\chi^2(DF = 2, N = 126) = 50.683, p < 0.0001$). Moreover, the post hoc analysis with Wilcoxon signed rank test shown significant differences between all three pairs of conditions as well. That means, the participants significantly preferred the robot motions condition-C over the other two variants.

Fig. 7c shows participants’ responses to the social-presence-measure questionnaire among four selected dimensions⁸. We see significant improvement along the *co-presence* dimension with C-condition. Complementary to this, we also see an increase in perceived attention given to the human by the robot (note the negative nature of Q4 for *attention allocation*). Both of the above mentioned results indicate an increase in robot’s awareness towards the human. Participants’ response towards *Perceived message understanding* measure hints at overall enhancement in fluency of the interaction episode. The dimension of *perceived*

⁸The comparisons are annotated with significance levels using “*”, *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$



(a) Bar chart showing success-ratio of participants’ prediction of robot goal; indicating significant improvement with *look-at-path* behavior. (b) Box plot of participants ranking over three head behavior conditions A, B and C; where C turn out to be the most liked behavior condition.



(c) Bar charts of participants’ responses to social-presence-measure questionnaire.

Fig. 7: Results of the user study.

behavioral interdependence also shows little improvement, meaning the robot being more responsive to the human presence. These results in combination with participants’ ranking partially support the *Hypothesis 2*.

D. Limitations

There are several limitations present in the design of this study. It gives us a third-person’s perspective on human-robot interactions, which is not sufficient to adequately assess the quality of the interaction episode with all dimensions of the social-presence-measure test. The results are nonetheless meaningful, as suggested by Syrdal et al. [23] it gives valuable insights regarding user experiences related to the assessment of human and robot interactions.

A comprehensive study should also include auditory clues and multiple humans, which could directly affect the perceived safety of individual participants. The context of the experiment, a path crossing scenario within a relatively small area, limits the generalizability of our results. Subsequent experiments would need to involve other shared navigation situations to further elaborate the relationships between a robots head expressions and how these are interpreted as cues for its navigational intents.

V. CONCLUSION

We reviewed literature on gaze behaviors from multiple disciplines including human-robot interaction, psychology,

neurobiology, and computer simulation. Based on our review, we put forward a plausible framework to determine dynamically an optimal gazing point, in which we formulated the robot gaze selection as a multi-criteria decision-making problem and solved it with the use of an analytic hierarchy process based algorithm. The proposed algorithm first employs dedicated functions to provide gazing points based on different sources of information. Each candidate functions are then globally evaluated on account of perceived social context. Along with the framework, we provided implementation details that testify our proposal. Lastly, to assess the resulting gazing behavior of the robot, we carried out a video based user study that supports our hypotheses and suggests improvements in participants' perception of the robot motions both objectively and subjectively.

We plan to explore other behavioral functions, e.g. robot acknowledging a group of people, showing certain emotions while navigating, and functions that explicitly take affordances of the environment into account. A machine learning based approach for dynamically adapting the criteria weight vector coupled with higher level perception modules is also an exiting avenue for the future work. In the user study, we did not include any *placebo* behavior (for example robot looking at random points), testing such behavior against the proposed approach and with a first-person user study will also be considered for the future work.

ACKNOWLEDGMENT

This work has partly been supported by the European Commission under grant agreement No 688147 (MuMMER).

REFERENCES

- [1] 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, no. 5, pp. 874–883, 2007.
- [2] T. Kruse, A. Kirsch, H. Khambhaita, and R. Alami, "Evaluating Directional Cost Models in Navigation," in *Proc. ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 2014, pp. 350–357.
- [3] U. Castiello, "Understanding other people's actions: intention and attention," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 29, no. 2, pp. 416–430, 2003.
- [4] K. Kitazawa and T. Fujiyama, "Pedestrian Vision and Collision Avoidance Behavior: Investigation of the Information Process Space of Pedestrians Using an Eye Tracker," in *Pedestrian and Evacuation Dynamics 2008*, W. W. F. Klingsch, C. Rogsch, A. Schadschneider, and M. Schreckenberg, Eds., 2010, pp. 95–108.
- [5] B. Mutlu, J. Forlizzi, and J. Hodgins, "A Storytelling Robot: Modeling and Evaluation of Human-like Gaze Behavior," in *Proc. IEEE-RAS International Conference on Humanoid Robots*, 2006, pp. 518–523.
- [6] A. Zaraki, D. Mazzei, M. Giuliani, and D. De Rossi, "Designing and Evaluating a Social Gaze-Control System for a Humanoid Robot," *IEEE Transactions on Human-Machine Systems*, vol. 44, no. 2, pp. 157–168, 2014.
- [7] H. Admoni, A. Dragan, S. S. Srinivasa, and B. Scassellati, "Deliberate Delays During Robot-to-human Handovers Improve Compliance with Gaze Communication," in *Proc. ACM/IEEE International Conference on Human-Robot Interaction*, 2014, pp. 49–56.
- [8] M. Gharbi, P.-V. Paubel, A. Clodic, O. Carreras, R. Alami, and J.-M. Cellier, "Toward a better understanding of the communication cues involved in a human-robot object transfer," in *Proc. IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, 2015, pp. 319–324.
- [9] C. Breazeal, C. Kidd, A. Thomaz, G. Hoffman, and M. Berlin, "Effects of nonverbal communication on efficiency and robustness in human-robot teamwork," in *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2005, pp. 708–713.
- [10] L. Takayama, D. Dooley, and W. Ju, "Expressing Thought: Improving Robot Readability with Animation Principles," in *Proc. ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 2011, pp. 69–76.
- [11] S. M. Fiore, T. J. Wiltshire, E. J. C. Lobato, F. G. Jentsch, W. H. Huang, and B. Axelrod, "Toward understanding social cues and signals in human-robot interaction: Effects of robot gaze and proxemic behavior," *Frontiers in Psychology*, vol. 4, no. 859, 2013.
- [12] A. May, C. Dondrup, and M. Hanheide, "Show me your moves! Conveying navigation intention of a mobile robot to humans," in *Proc. European Conference on Mobile Robots (ECMR)*, 2015, pp. 1–6.
- [13] D. V. Lu, "Contextualized Robot Navigation," Ph.D. dissertation, Washington University in St. Louis, 2014.
- [14] A. E. Patla, A. Adkin, and T. Ballard, "Online steering: coordination and control of body center of mass, head and body reorientation," *Experimental Brain Research*, vol. 129, no. 4, pp. 629–634, 1999.
- [15] H. Hicheur, S. Vieilledent, and A. Berthoz, "Head motion in humans alternating between straight and curved walking path: Combination of stabilizing and anticipatory orienting mechanisms," *Neuroscience Letters*, vol. 383, no. 1–2, pp. 87–92, 2005.
- [16] D. Bernardin, H. Kadone, D. Bennequin, T. Sugar, M. Zaoui, and A. Berthoz, "Gaze anticipation during human locomotion," *Experimental Brain Research*, vol. 223, no. 1, pp. 65–78, 2012.
- [17] V. V. Unhelkar, C. Pérez-D'Arpino, L. Stirling, and J. A. Shah, "Human-Robot Co-Navigation Using Anticipatory Indicators of Human Walking Motion," in *Proc. IEEE International Conference on Robotics and Automation (ICRA)*, 2015, pp. 6183–6190.
- [18] V. Srinivasan, R. R. Murphy, and C. L. Bethel, "A Reference Architecture for Social Head Gaze Generation in Social Robotics," *International Journal of Social Robotics*, vol. 7, no. 5, pp. 601–616, 2015.
- [19] B.-S. Yoo and J.-H. Kim, "Fuzzy Integral-Based Gaze Control of a Robotic Head for Human Robot Interaction," *IEEE Transactions on Cybernetics*, vol. 45, no. 9, pp. 1769–1783, 2015.
- [20] T. L. Saaty, "How to make a decision: The analytic hierarchy process," *European Journal of Operational Research*, vol. 48, no. 1, pp. 9–26, 1990.
- [21] J. H. Park, F. A. Rojas, and H. S. Yang, "A collision avoidance behavior model for crowd simulation based on psychological findings," *Computer Animation and Virtual Worlds*, vol. 24, no. 3-4, pp. 173–183, 2013.
- [22] R. Triebel et al., "SPENCER: A Socially Aware Service Robot for Passenger Guidance and Help in Busy Airports," in *Proc. Conference on Field and Service Robotics (FSR)*, 2015, pp. 21–34.
- [23] D. S. Syrdal, N. Otero, and K. Dautenhahn, "Video prototyping in human-robot interaction: results from a qualitative study," in *Proc. European Conference on Cognitive Ergonomics (ECCE)*, 2008.
- [24] P. C. Harms and P. F. Biocca, "Internal Consistency and Reliability of the Networked Minds Measure of Social Presence," in *Proc. 7th Annual International Workshop on Presence*, M. Alcaniz and B. Rey, Eds., 2004.