Analysis of Social Navigation for a Robot

Andre Van den Berg

Analysis of Social Navigation for a Robot

Andre van den Berg
MOSIG Research Project - INRIA
Grenoble, France
andre.van-den-berg@inria.fr

Abstract

This article describes the different approaches currently followed in the field of social navigation for robots. A novel approach focusing on human activity called the Maisonasse attentional model is studied in more depth as well as its implementation in the robot operating system (ROS).

1 Introduction

1.1 Background

In the past, the field of robotics mainly comprised of robots within a controlled environment, for example a robot that assembles body parts of a vehicle. As technology advanced it became possible to expand the working environment of robotics to share the space inhabited by people. During these early days the focus of research was predominantly concerned with navigational safety or collision detection and prevention. However for smooth integration into human occupied areas, navigational safety and collision avoidance are not the only criteria to take into consideration. Moreover as described in [Reeves and Nass, 1996] humans expect a robot to interact in a way resembling how an actual human would behave. For this reason research has shifted towards enabling robots to act in a socially acceptable way and thereby improve the human robot interaction [HRI].

Modelling the movement and behaviour of a robot to accomplish this behaviour proves to be difficult. Humans have the ability to correctly interpret a scenario and interact accordingly conforming to certain social norms so as not to cause surprise, uneasiness or appear rude to other human beings. Also subtle differences in body posture, vocal levels etc, are interpreted correctly and their behaviour adjusted accordingly in a very short time span. This is not the case for robots. To achieve the goal of designing human-like robots, numerous challenges arise. The information gathered from the sensory models (be it visual, acoustic, map etc.) of robots allows only for an estimation of the scenario. For example visual sensors might not pick up transparent objects, dynamic objects can be incorrectly mapped as static objects and so forth. This leads to inconsistencies between the actual scenario and the perceived scenario. However even if it were possible to detect and describe the scenario one hundred percent accurately, one is still left with the difficulty of deciding on the correct behaviour and course of action to take based on the scenario.

1.2 Contribution

For a robot to navigate not only in a safe but also socially acceptable way, traditional path planning methods are not sufficient. [Maisonnasse *et al.*, 2006] illustrated that an attentional model can be used to detect interactions between people and objects. This paper builds on this idea and aims at describing how the attentional model can be used in a similar fashion for navigational purposes of a robot within a social context.

1.3 Outline

The paper is set out as follows: Section 2 contains a literature review focusing predominantly on the navigational aspect of human robotic interaction (HRI) and highlights the most common approaches followed. Section 3 expands on one of these approaches and describes in detail how attention between people and objects can be used in navigational planning. This is followed by the results obtained in section 4 and the discussion in section 5.

2 Literature review

To improve human robot interaction, three main areas of interest can be distinguished namely: *navigation, interaction, appearance*. Navigation focuses on the path and trajectory the robot should take to achieve its goal position, combining obstacle avoidance with social rules. Interaction is concerned with verbal communication and physical interaction with an actual human being, given that it is an objective. Lastly appearance is a criteria which includes among others, showing emotion, appearing human-like if the robot's goal is to interact with humans for example, or having an appearance that puts the human at ease and doesn't cause distress. Henceforth the focus will be on the navigational aspect of socially aware robots.

2.1 Navigational aspect of HRI

For a robot to reach a certain goal destination, the following three factors play a role in deciding which path and trajectory is most suitable: *safety, optimality in respect to time/energy/distance*, and *compliance to social/cultural rules*. Motion safety has been studied in detail and the three safety criteria such as set out in [Fraichard, 2007] underpins what is needed to avoid Inevitable Collision States (ICS) [Fraichard and Asama, 2003]. Route planning based on optimising the time/distance/energy to reach the target is achievable through methods such as A* [Hart *et al.*, 1968]. In general any criteria can be used to create a cost map that can be used as an input for the A* algorithm.

The navigational part of social behaviour of a robot as a sub class of the whole HRI social behaviour, is interested solely in the path and trajectory of the robot. The social rules might include rules such as passing on the right/left, following distance, waiting in a queue, to name but a few. The senses with which a human perceives social behaviour is limited to sight and hearing. Touch is excluded, since for navigation the safety criteria will prevent touching and for interaction it falls within said category. A robot will at a certain point in time either try to attract attention or avoid distracting a humans attention. Which ever it may be, certain social behaviours need to be adhered to. The idea behind adhering to these social or cultural rules are to maximise human comfort and avoid rudeness or disturbance.

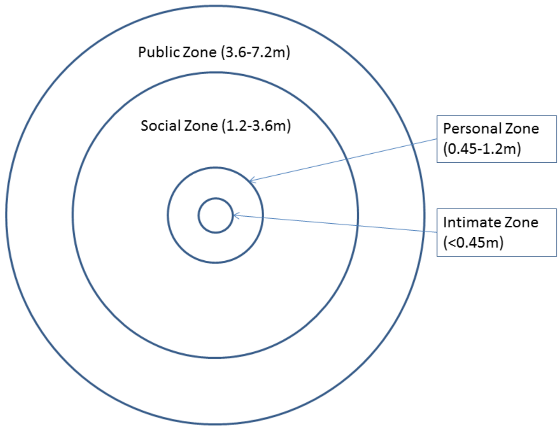

Figure 1: Representation of different zones according to Hall. (not to scale)

2.2 Consideration of social behaviour as related to Navigation

Proxemics

Different approaches have been suggested to try and achieve the goal of a robot that behaves socially acceptable, noteworthy is the field of *proxemics* first termed in 1966 by Edward T.Hall. [Hall, 1966] Proxemics acknowledges that a person feels comfortable with another person or robot up to a certain proximity within the space surrounding him or her. The distance that is deemed socially acceptable depends on the relationship between the persons as well as the type of interaction. Hall proposed four distinct circular zones around a person in increasing size: intimate zone, personal zone, social zone and public zone. Figure 1 gives an indication of these zones although the exact diameter depends heavily on culture, background and other factors.

The majority of research seems to focus heavily on proxemics as a criteria to achieve socially aware behaviour. In some cases the original model by Hall is slightly modified. Svenstrup et al.[Svenstrup *et al.*, 2010] effectively extend the size of the zones (intimate, personal, social, public) towards the area behind the person. This is done under the assumption that a person feels more uncomfortable with a person standing behind them. Others again adapt the model to cater for a certain social behaviour such as passing on the left. This is achieved for example by effectively increasing the zones to the right hand side of a person as done in [Kirby *et al.*, 2009] by means of an auxiliary cost function.

In all cases the robot's goal is to respect these zones depending on the task of the robot. As an example, if the robot needs to deliver an object to a person it will need to enter into the intimate zone. On the other hand if its task is only to convey a message, the robot needs to reside in either the personal zone or social zone depending on how acquainted the person is with the robot.

It is clear that navigation based on proxemics on its own lacks the capability of adequately dealing with the activity of humans. Two human beings might be in conversation within each others social zone and the robot would pass between them and deem it acceptable as it stayed outside their personal zone. In reality though most cultures would see this as unacceptable social behaviour. Therefore a different approach than solely proxemics is needed to overcome this problem.

Acoustics

Another social navigation approach focuses on *acoustics*. This approach can be subdivided into two scenarios. The first is for the situation where the robot interacts with a human and would thus want to be positioned such that the ambient noise is minimised. Martinson explains in [Martinson and Brock, 2007] how this can be achieved using a noise map as a cost function and letting the robot move towards the most tranquil place. This noise map is updated regularly using the a-priori knowledge as well as any newly detected sound sources. The second situation is applicable to when it is desirable for the robot not to be a distraction. In this case the ambient noise is used to mask the noise the robot generates itself [Martinson, 2007]. A motion path is then constructed, so as to minimize the detectable noise generated by the robot above the ambient noise.

Activity based approach

One of the key challenges to socially acceptable interaction is that a certain behaviour seen as acceptable at a point in time might be completely inappropriate at another time. This fact depends highly on the current *activity* with which a person or persons are busy. Take for example the scenario as mentioned in [Diego and Arras, 2011], where a noisy cleaning robot is used in a house environment. Cleaning within the television room while a person is watching a show is not socially acceptable. However if the person is just passing through the room, the robot need not interrupt its own cleaning task. Similarly greeting a person and seeking interaction while he is typing in front of a computer is acceptable, on the other hand if the person is busy with a video call on his computer it is best to wait until after the call before trying to engage with the person. This illustrates the need for determining human activity as a means of socially acting correctly.

Field of view is commonly used to determine a persons attention and infer the activity they are currently busy with. [Rios-Martinez *et al.*, 2012] refers to this as an information process space (IPS). This approach does however have its limitations as not all objects in a persons field of view necessarily holds his or her attention. Furthermore acoustic attention is completely ignored in this case, in other words a humans attention regarding acoustics is normally omnidirectional and not limited to his or her field of view. Maisonasse et al.[Maisonnasse *et al.*, 2006] proposes an *attentional model* to determine which amount of a person's attention lies with an object. The information gathered can then be used to determine the current activity of the person. The advantages of this approach is that it takes into account the salience of all objects related to the person, which can range from a computer, telephone, screen or another human being.

3 Attention-Based Navigation

In order to detect the interactions that exist between people and objects within their environment, it is necessary to take into account all possible interactions between people as well as interactions between people and objects. This can only be achieved with an analysis on scene level.

The attentional model acknowledges that a human possesses a certain amount of attention and that this attention is divided between different objects. The attentional model aims to model this attention span of a human and provide this model to a robot which can then use this information to plan its motion given a certain goal.

3.1 Details of the attentional model

As stated in [Fischer, 2012], all objects that are of interest (which draw attention) are defined as objects in the workspace. These objects are parametrised by their position, velocity and an attentional component. The attentional component consists of the following variables: (O, T, S, α_{ini}, α_{end}). O is defined as the orientation of the object, T can take one of two values Visual or Audio and indicates the kind of disturbances. S is the salience, level of distraction an object possesses, note that this variable can be time-dependant. Lastly the angles α_{ini}, α_{end} define the angle within which the object possesses its salience. Humans form a subclass of type object and is named Agents.

Each agent possesses a certain amount of attention that it can distribute between different objects. The first step to model this attention as described in [Fischer, 2012] is to define an average attention vector consisting of a size and a direction in the R^2 space. This average attention is itself broken up into two vectors namely intention and distraction. The intention is related to the current activity with which the agent is busy, while the distraction is a vector acquired by summing over all the possible objects (j) that can distract the Agent (i) from his current activity. The exact detail of how the distraction vectors are calculated is omitted for conciseness but can be found in [Maisonnasse *et al.*, 2006]. The resulting distraction vector is then as follows:

$$\vec{Distraction_i}(t) = \sum_{j,j\neq i} \vec{Distraction_{ij}}(t) \tag{1}$$

and finally the average attention vector is obtained by combining the two vectors after applying specific weights to them.

$$\begin{aligned} \vec{Attention_i}(t) = \lambda * \vec{Intention_i}(t) + \\ \mu * e^{\frac{-2*||Intention_i(t)||}{Intention_{max}}} \vec{Distraction_i}(t) \end{aligned} \tag{2}$$

An example of this attention vector for an agent called object 1 is shown in figure 2 from [Maisonnasse *et al.*, 2006].

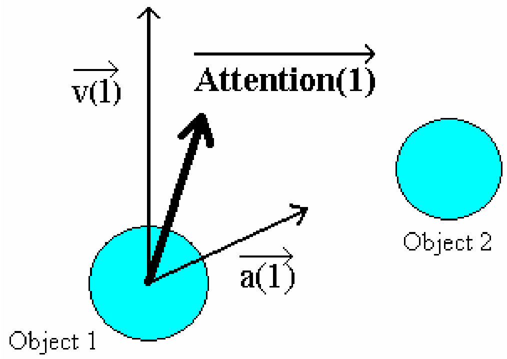

Figure 2: Attention vector of object 1 (An agent)

The exponential term in the weighting of the distraction vector allows a decrease in distraction given a very high value of intention. This corresponds to a scenario where a person is highly focused on a task and distracting him or her would require a high amount of effort.

To specify the final attention between an agent and an object given the attention vector as defined in equation 2 the first step is to normalise this value to a value between 0 and 1 as described in [Fischer, 2012]. Intuitively the attention of an agent is not limited to one single specific direction as suggested by the attention vector. To take into account that an agent will focus his attention within a certain azimuth centred around the attention vector, a gaussian distribution based on the attention vector of each agent is defined as $F_c(\alpha)$ where c is the normalized attention value related to the size of the attention vector. (see [Fischer, 2012]). Figure 3 adapted from [Maisonnasse, 2007] shows how the angle between the vector $\vec{Attention_i}(t)$ and the object has an impact on the value of $F_c(\alpha)$. The dotted lines show that for an angular displacement of -0.5 radians, the attentional distribution changes from approximately 0.48 for the case where the object is in the exact same direction as the vector $\vec{Attention_i}(t)$ to 0.3.

The final parameter called $resource_{ij}$ which describes the attention agent i pays to object j is then obtained as follows:

$$resource_{ij} = F_{c_i}(\alpha_{ij}) * \frac{||Distraction_{ij}(t)||}{||Distraction_i(t)||} \tag{3}$$

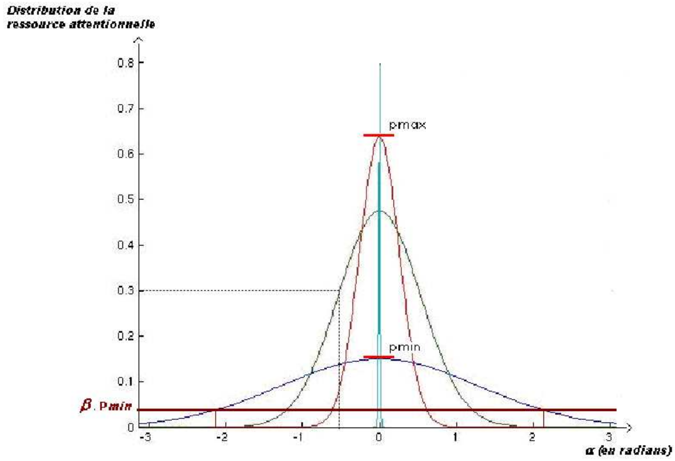

Figure 3: Function $F_c(\alpha)$ for different values of c

Computation of resources

The computation of the parameters $resource_{ij}$ between agent i and object j for all agents and objects is calculated using the following algorithm adapted from [Fischer, 2012].

Algorithm *ComputeAttentionMatrix(S)*
input: The set of agents and other objects S
constants: λ = 60, μ = 40, intentionmax = ?, K = ?
output: The attention resource distribution matrix M
for all i $\in$ Agents(S) **do**
  for all j $\in$ Objects(S) **do**
  distance $\leftarrow$ Position(j) - Position(i)
  distraction[i][j] $\leftarrow$ Salience(j, i) $*distance/||distance||^3$
  distraction sum[i] $\leftarrow$ distraction sum[i] + distraction[i][j]
  end for
end for
for all i $\in$ Agents(S) **do**
  dedication $\leftarrow \mu * e^{(2*intention[i]/intention_{max})}$
  attention[i] $\leftarrow\lambda*$ intention(i) +
    dedication $*$ distraction sum[i]
  c $\leftarrow tanh(K * ||attention[i]||)$
  for all j $\in$ Objects(S) **do**
  contribution[i][j] $\leftarrow$ distraction[i][j]/distraction sum[i]
  M[i][j] $\leftarrow$ F (c, Angle(i, j)) $*$ contribution[i][j]
  end for
end for

4 Results

4.1 Scenario

For a scenario consisting of four agents (A1-A4) and three objects (O1-O3), the final resource allocation is represented in table 1 (adapted from [Maisonnasse, 2007]). The information in the table can be interpreted as follows: a row entry shows how the focus of the agent corresponding to the row is distributed amongst the different objects in the workspace (per column). A column entry shows how the specific object corresponding to the column holds the attention of all the agents. As mentioned previously an agent is a subclass of object and is therefore included in the columns. As an example, assume objects 1 and 2 are both robots capable of interacting with a human/agent and that table 1 is a given configuration in the state time space. Let the task be to interact with agent 2, the table then indicates that the best option would be to use object 1 as it holds the highest relative attention (0.35) of agent 2. Let us say that the task changes slightly in that a robot still has to interact with agent 2, however it should not divert the attention of agent 1. In this case the robot corresponding to object 2 will be the preferred option. The reason being that even though it does not hold the highest relative attention of agent 2, it holds no attention of agent 1 while still holding a relatively high attention of agent 2.

Table 1: Resource table showing Objects O and Agents A

	A1	A2	A3	A4	O1	O2	O3
A1	-	0.15	0.02	0.02	0.34	0.0	0.0
A2	0.5	-	0.11	0.04	0.35	0.20	0.0
A3	0.04	0.2	-	0.12	0.54	0.07	0.08
A4	0.04	0.3	0.17	-	0.12	0.2	0.4

Figure 4 shows a partial representation of the scenario from table 1 with the relevant resources shown for the scenario.

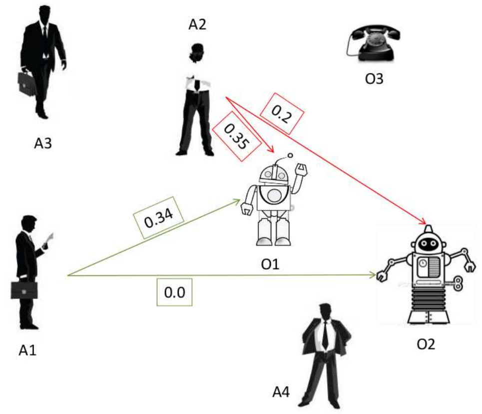

Figure 4: Representation of the scenario

4.2 Cost function

As discussed in the previous section the attentional model describes the amount of resources each agent allocates to each object respectively. The goal is to develop a cost function, based on the attentional model, that can be used in robot navigation. As a robot is itself one of the objects in the workspace, the table resulting from the attentional model gives an indication of the amount of attention the robot object holds for each agent. The cost map can then be created by evaluating the resource relationships between agents and the robot for different positions of the robot at different times.

Similar to the acoustic approaches by [Martinson and Brock, 2007] and [Martinson, 2007] there exists two distinct situations, the first is when the goal is to interact or disturb an agent and the second when it is desirable to cause as little disturbance as possible for the agent. Depending on the scenario the goal would be to either maximise or minimise the cost function. More complex situations might also exist where one agents attention should be attracted while another agents attention should not be diverted.

It is necessary that the robot also respects the human personal space based on proxemics as described by [Hall, 1966]. To ensure this, the cost function derived from the attentional model is linearly combined with a dynamic personal space cost function for each agent. This dynamic personal space cost function is as described in [Scandolo and Fraichard, 2011] and displayed in figure 5. The higher cost in red corresponds to the area behind a person.

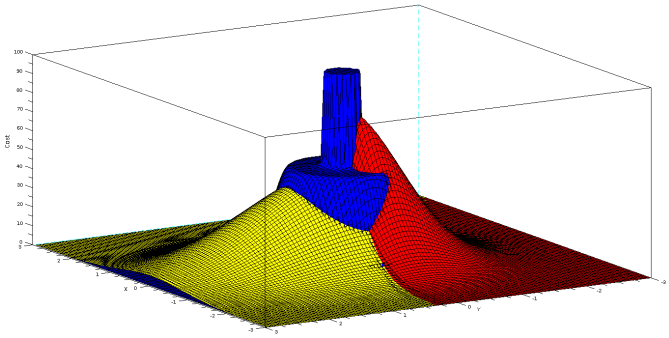

Figure 5: Illustration of the dynamic personal space cost function surrounding a person

4.3 Future work

The attentional model and cost map generation has been implemented in code (C++). Unfortunately due to time constraints the final integration within the ROS framework and simulation in the Morse environment has not been completed and is still in progress.

5 Discussion

The various approaches mentioned all have their advantageous and disadvantageous related to socially acceptable behaviour of a robot. A purely proxemics approach for example does not take into account the current activity that a user is engaged with. A purely activity based approach might on the other hand neglect to take into account the discomfort a user feels based on proxemics. Depending on the scenario one or the other approach might seem more suited. The social importance deemed most desirable such as silent movement, respecting personal space, non-intimidating speeds, considering human activity will in the end determine which approach is followed. The attentional model holds promise as each interaction between object and agent can be modelled individually whether it holds acoustic or visual attention. The model is also not limited to a certain field of view as parameters α_{ini}, α_{end} describing the azimuth of the objects salience can be changed independently of the objects orientation. The major challenges in the attentional model is how to choose relevant values for the models parameters as well as taking into consideration that these values can be time depended as well.

References

[Diego and Arras, 2011] Gian Diego and K.O. Arras. Please do not disturb! minimum interference coverage for social robots. In *Intelligent Robots and Systems (IROS), 2011 IEEE/RSJ International Conference on*, pages 1968–1973, 2011.

[Fischer, 2012] Thomas Fischer. Implementation of the maisonnasse model. Internal document, July 2012.

[Fraichard and Asama, 2003] T. Fraichard and H. Asama. Inevitable collision states. a step towards safer robots? In *Intelligent Robots and Systems, 2003. (IROS 2003). Proceedings. 2003 IEEE/RSJ International Conference on*, volume 1, pages 388–393 vol.1, 2003.

[Fraichard, 2007] T. Fraichard. A short paper about motion safety. In *Robotics and Automation, 2007 IEEE International Conference on*, pages 1140–1145, 2007.

[Hall, 1966] E.T. Hall. *The hidden dimension*. A Doubleday anchor book. Doubleday, 1966.

[Hart *et al.*, 1968] Peter E Hart, Nils J Nilsson, and Bertram Raphael. A formal basis for the heuristic determination of minimum cost paths. *Systems Science and Cybernetics, IEEE Transactions on*, 4(2):100–107, 1968.

[Kirby *et al.*, 2009] R. Kirby, R. Simmons, and J. Forlizzi. Companion: A constraint-optimizing method for person-acceptable navigation. In *Robot and Human Interactive Communication, 2009. RO-MAN 2009. The 18th IEEE International Symposium on*, pages 607–612, 2009.

[Maisonnasse *et al.*, 2006] Jrme Maisonnasse, Nicolas Gourier, Oliver Brdiczka, and Patrick Reignier. Attentional model for perceiving social context in intelligent environments. In Ilias Maglogiannis, Kostas Karpouzis, and Max Bramer, editors, *Artificial Intelligence Applications and Innovations*, volume 204 of *IFIP International Federation for Information Processing*, pages 171–178. Springer US, 2006.

[Maisonnasse, 2007] J. Maisonnasse. *Estimation des relations attentionnelles dans un environnement intelligent*. 2007.

[Martinson and Brock, 2007] E. Martinson and D. Brock. Improving human-robot interaction through adaptation to the auditory scene. In *Human-Robot Interaction (HRI), 2007 2nd ACM/IEEE International Conference on*, pages 113–120, 2007.

[Martinson, 2007] E. Martinson. Hiding the acoustic signature of a mobile robot. In *Intelligent Robots and Systems, 2007. IROS 2007. IEEE/RSJ International Conference on*, pages 985–990, 2007.

[Reeves and Nass, 1996] Byron Reeves and Clifford Nass. *How people treat computers, television, and new media* *like real people and places*. CSLI Publications and Cambridge university press, 1996.

[Rios-Martinez *et al.*, 2012] J. Rios-Martinez, A. Renzaglia, A. Spalanzani, A. Martinelli, and C. Laugier. Navigating between people: A stochastic optimization approach. In *Robotics and Automation (ICRA), 2012 IEEE International Conference on*, pages 2880–2885, 2012.

[Scandolo and Fraichard, 2011] Leonardo Scandolo and Thierry Fraichard. An anthropomorphic navigation scheme for dynamic scenarios. In *Robotics and Automation (ICRA), 2011 IEEE International Conference on*, pages 809–814. IEEE, 2011.

[Svenstrup *et al.*, 2010] M. Svenstrup, T. Bak, and H.J. Andersen. Trajectory planning for robots in dynamic human environments. In *Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on*, pages 4293–4298, 2010.