Enabling Building Service Robots to Guide Blind People

Suleyman Balashzade **Alex Gbenimachor**

Abstract

Humanoid robots, including Softbank's Pepper, are mostly used as showpieces at events. Businesses would like to see these robots are used in more practical ways. They could work as receptionists or as tour guides. To complete these tasks, the robots must be able to move through a complex world.

This work aims to determine whether Pepper, is suitable to localize and navigate successfully in order to aid the blind or visually impaired people to find their way around the building.

This project proposed a Building Service Robot, located at the entrance of the building, to guide people to their desired destination or explain to them how to get there.

1 Introduction

The interest of utilizing humanoid robots as information and administration gadgets for people is developing with headways into pertinent innovations. For instance, for a financial organization, it would be recipient if a humanoid robot can interact with clients and offer information and responses to their inquiries rapidly. Within a customer office, a robot could welcome customers, answer initial questions and guide towards another location. Taking a gander at the business offices a robot would be handy to welcome guests. After dealing with identification it could guide a guest to the elevators or even to the location of the host. Another use case would manage little inside bundles, mail and such. Leaving these undertakings alone done by a robot is simply one more type of mechanization that could save employees time. While the automation of these tasks seems easy, most of them are still better handled by humans as the robots require to understand the intentions of the visiting humans. This can be hard for a computer and is an active subject of research. The tasks robots can handle more efficiently are mostly processing structured information with clear intent.

2 Related Work

Pepper is a human-shaped robot designed by Softbank Robotics. Originally created under the French company Aldebaran Robotics, widely known for their highly interactive robot NAO, it was acquired by Softbank Mobile Group in 2013, as the starting stone to build the Pepper project. The concept behind Pepper shows the desire to manufacture a friendly robot companion that is accessible for a wider range of customers. It's greatest features generate the ability to understand human emotions and react in a predefined manner to those emotions. Although most of Pepper's features are pre-programmed and show a much less conscious thinking process, its structural design and software capabilities offer a high predisposition for human interaction. This characteristic renders a perfect advantage to work on social navigation. From experience while developing this report, people are easily attracted to the robot's friendly design.

On the technical side, the robot incorporates a range of proximity and vision sensors that enable the development of tracking, localization and navigation algorithms. The wheel drive system is holonomic, allowing a broader range of movements that fit adequately into human related scenarios.

Softbank has created a coherent and well thought-out development environment for the NAO/Pepper robots, called Choreographe. It has the ability to read signals from the sensor, process these information and control the robot. However, experience indicates that for the implementation of a stable and more advanced application, the dedicated SDKs should be used.

In the paper "Enabling Service Robots to Guide Blind People" it is suggested, to use a smartphone to summon the robot for interaction, but in this project a face detection of a person that enters a building is used to start an interaction. Four different modes of interaction are available: Information kiosk, Escort Mode, Sighted guide mode, Holding the Robot Mode.

- **Information Kiosk**: The Robot provides directional Instruction, without escorting the User.
- **Escort Mode**: The Robot accompanies the User to the target destination
- **Sighted Guide**: In this mode, the robot moves side by side with the User
- **Hold the Robot Mode**: The User is holding the Robot.

This project has focused on the Information Kiosk and Escort modes.

3 Interaction

Interaction Control

Once the navigation system was implemented with optimal results from Pepper, it was necessary to investigate which functions from NAOqi were important to evaluate in the guiding behavior. Three (3) modes were needed to be designed for different routines within the control structure:

- Interactive Standing mode
- Interactive Guiding mode
- Show Room and Farewell behavior.

Interactive Standing

While Pepper waits for a visitor to trigger a guiding request, Pepper is set to Awareness mode. This feature from NAOqi enables Pepper to react to sound and face detection. The default settings for sound detection are not modified, but face detection is set to a maximum distance of four (4) meters.

It's important to mention that Pepper is easily attracted to sounds around its environment. This characteristic can create awkward reactions from Pepper, for example it could end up facing the wall if a sound is triggered next to this structure. In order to avoid this irrational behavior, the awareness mode was set to the head. In this mode Pepper will track sounds/people only with its head without any body displacements. This decision sets Pepper to a fixed location facing the most probable entrance for visitors.

Interactive Guiding

The purpose of augmenting the cooperative action of guiding a person was addressed from the motion planning point of view. The aim was to increase this interaction by exploiting Pepper's humanoid design. There are 3 key features that were attended to:

- People prefer moving with a robot that shows human-like behaviors.
- There are different reactions of people to humanoid and non- humanoid robots.
- The distance between humans increases if there are few eye-contacts between people.

Partner Engagement

This point is addressed by using verbal interaction between Pepper and its Partner. This research was not devoted to the communication skills of robots, since this field of research is very complex enough on its own. Engaging in a normal conversation or small talk is far too complex to be implemented in this project. Therefore, this project limited to generating verbal feedback of the progress of the guiding task. It is clear that no reaction is expected from the Partner. Spoken dialogue systems are particularly suitable to this context. This project, concentrated or focused on implementing a task-oriented SDS. SDSs allow users to interact with machines by means of spoken dialogues in natural language. The general architecture of SDSs is comprised of six components, as shown in Figure 1. The speech input is first processed by a speech recognizer, which converts it to a written form. This is then passed to the language analyzer, which constructs a logical representation of the user's utterances. Using this representation, information on the previous discourse, and knowledge of the task to be performed, the dialogue manager may then decide to communicate with an external application or device, in some cases the robot's controller would convey a follow-up message to the user. In the latter case, a logical representation of the message is passed to response generator, which generates an appropriate response in written form and passes it to the speech synthesizer.

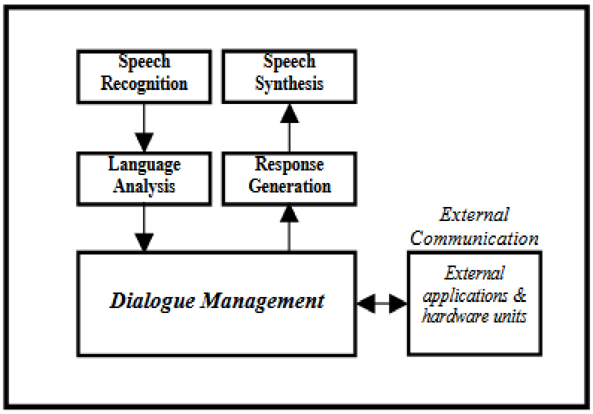


Figure 1: Architecture of spoken dialogue systems

Dialogue Management Techniques

The dialog manager controls turn-by-turn behavior. A simple dialog system may ask the user questions, then act on the response. Such directed dialog systems use a tree-like structure for control; frame- (or form-) based systems allow for some user initiative and accommodate different styles of interaction. More sophisticated dialog managers incorporate mechanisms for dealing with misunderstandings and clarification.

The most commonly used and simplest dialogue management techniques are state-based. These techniques represent the possible dialogues by a series of states, as shown in Figure 2. At each state, the system may ask the user for specific information, it may generate a response to the user, or it may access an external application. The structure of the dialogue is predefined, and at each state the user is expected to provide particular inputs. This makes the user's utterances easier to predict, leading to faster development and more robust systems at the expense of limited flexibility in the structure of the dialogues.

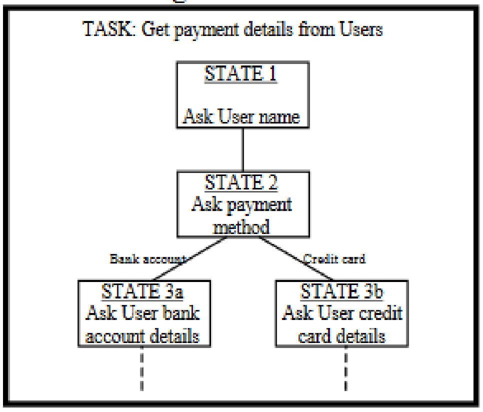


Figure 2: State-based dialogue

For simple tasks, state-based techniques are often the most practical solution.

The dialogue manager was designed to perform relatively short dialogues. The main goal, however, was to ensure that the user's input is interpreted correctly. There are confirmation sub-dialogues at key points to allow the users to check the robot's interpretation of their utterances, and to repeat them if necessary. Additional help and clarification messages are also available.

4 Resources Used For This Project

Google's Speech Recognition

Speech recognition is the inter-disciplinary sub-field of computational linguistics that develops methodologies and technologies that enables the recognition and translation of spoken language into text by computers. It is also known as automatic speech recognition (ASR), computer speech recognition or speech to text (STT). It incorporates knowledge and research in the linguistics, computer science, and electrical engineering fields.

In this project, Google's speech recognition was used to identify and process users voice. It receives a spoken sentence, calls the ASR to get the corresponding text.

GTTS (Google's Text To Speech)

This is the Library which was used to convert pepper response to speech.

Choregraphe

This is a multi-platform desktop application, that allows one to:

- Create animations, behaviors and dialogs,
- Test them on a simulated robot, or directly on a real one,
- Monitor and control the robot

Choregraphe allows to create applications containing Dialogs, services and powerful behaviors, such as interaction with people, dance, e-mails sending, without writing a single line of code.

It was used in this project to simulate the interaction between Pepper and the User

V-REP

The Virtual Robot Experimentation Platform - is a 3D robot simulation software, with integrated development environment, that allows one to model, edit, program and simulate any robot or robotic system (e.g. sensors, mechanisms, etc.).

It has been used to design mock building setting and implement guiding behavior.

OpenCV

The OpenCV (Open Source Computer Vision Library) is a library of programming functions mainly aimed at real-time computer vision to detect face, to initiate the contact with user and start cooperation.

5 Beginning Of Task

As soon as Pepper confirms the objective location, Pepper pronounces phrases:

• Please follow me to the <location depending on users input>.

Farewell

Once Pepper and the Partner reach the target location, Pepper stops in front of the room and informs the Partner about it. This is designed to give a clear indication that the task is completed and where the Partner should direct after Pepper leaves toward the home location. This is accompanied by the following expression:

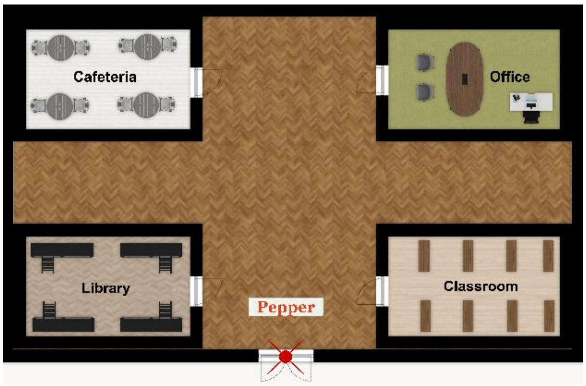
• We're now in front of the <location depending on users input>. Is there anything else I can do for you?

If no help further needed after a couple of seconds, Pepper gives a warm farewell by saying:

• Ok, I am glad I could help. Bye now.

6 Experiment

The experiment was carried out at the virtual environment in this case, at the entrance of the building inside of Sapienza. The door, marked with red sign in Figure 3, is the main point of entry of people to the building. As a result, Pepper's initial position was defined to be in front of the door, highlighted with a label "Pepper". This specific location was referred to as Home throughout the test.



In order to describe the set of steps that are followed on every trial, there is need to define the two subjects that form part of the test:

- Participant: Assumes the role of the person being guided by the robot.
- Pepper: Humanoid robot programmed as a guide.

Face Learning/Recognition

Recall that Pepper uses face recognition to start the guiding behavior. Face recognition triggers Pepper to perform greeting and read the list of operations that it can offer.

This was done by using OpenCV and Haar Cascade file. The Haar Cascade file is part of the OpenCV library, which is used for extracting features from objects. The features, that are extracted using Haar Cascades files, includes face features, nose features, eyes etc. For this project, the face features extraction file was used. Upon seeing a person at the entrance of the building, Pepper steps forward to this person. The file known as peppervision, which controls this action, contains the detectfaces() function. Within the function, the head of the robot was tilted using the head-joint variables and below can be seen the initialization of Haar Cascades file, required for our face detection.

def DetectFaces(motionService):

HeadJoint = ["HeadPitch", "HeadYaw"]

HeadAngles = [-5,6, 0]

*HeadAngles= [x * motion.TO_RAD for x in HeadAngles]*

motionService.angleInterpolationWithSpeed(HeadJoint, HeadAngles, 0.6)

cascadePath = "haarcascade_frontalface_default.xml"

faceCascade = cv2.CascadeClassifier(cascadePath)

vidobj = cv2.VideoCapture(0)

Accompany the User to the target room

The Accompany Solution, is performed using the V-REP simulator, by performing a path planning that moves the robot to the target Area. During the movement, Pepper gives information on what actions it performs. For instance: “Pepper informs the user that it is performing action to turn left, during this process”.

Direction Information

If a person requires an information on how to get to their destination, Pepper explains to him how to get there, additionally performing directional gestures. Accompanying is not done by this action.

Path Planning

It is defined as planning problem, that involves the computation of sequence X,Y points for finding the location from the start location to the end location.

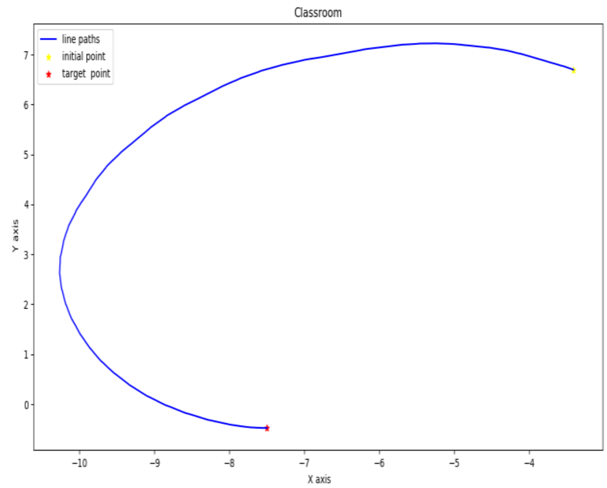
Deviation cost has been computed., If Pepper perceives an obstacle in front of it, it avoids it. Below is the deviation formula.

$$dev = robotRotation - devAngle$$

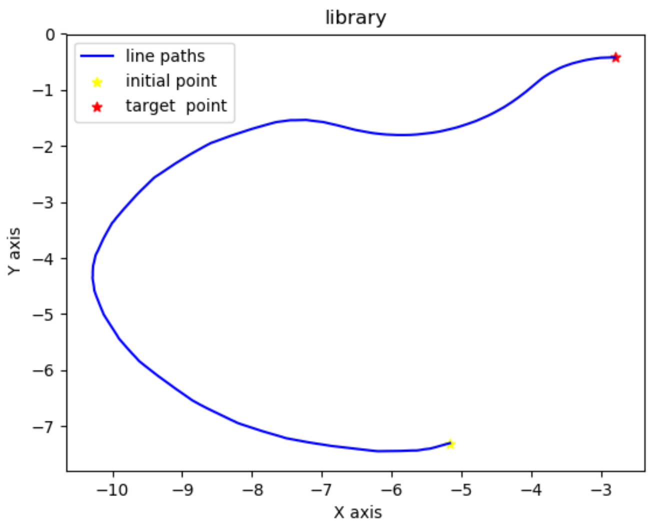
7 Results

The Results show the Robot's Navigation, between each building in the scenarios

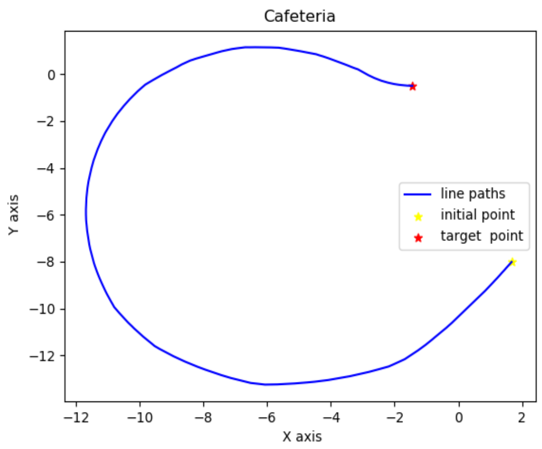
The Classroom Navigation:



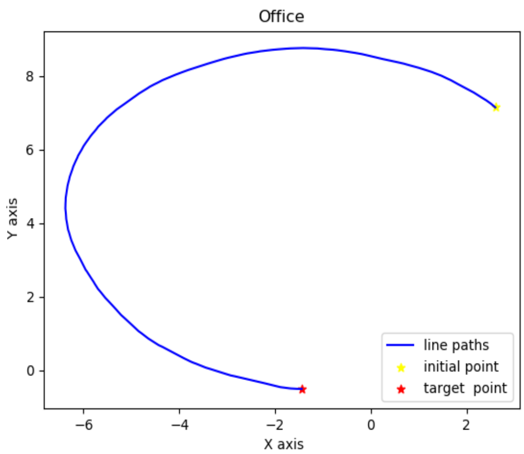
The Library Navigation:



The Cafeteria Navigation:



The Office Navigation:



8 Conclusion

This paper has presented a model, which was developed to guide people using Pepper that work cooperatively. Unfortunately, no participants were involved in this project. We believe that human feedback is a key point for the development of robots whose purpose is interacting with people.

The validation of the model has been demonstrated throughout an extensive set of simulations. Our method can handle realistic situations, such as dealing with indoor environments and obstacles within. For that reason, this work can be applied to certain specific real robot applications, for instance, guiding tourists or accompanying professional visitors.

In future work, we aim to obtain more sophisticated robot behavior, by performing user experiments with our system, in order to study their reactions and eventually modify the system. Another thing that needs to be built is an ability of Pepper to understand that its Partner has deviated from the guiding task so that it could stop and face the side where the person was last tracked to search for him..

References

[1] Aethon: Autonomous Mobile Robots and Tracking Solutions:

http://www.aethon.com/. Accessed: 2015-10-05.

[2] Aigner, P. and McCarragher, B. 1999. Shared control framework applied to a robotic aid for the blind. IEEE Control Systems Magazine. 19, 2 (Apr.1999), 40–46.

[3] Alkhanifer, A. and Ludi, S. 2014. Towards a situation awareness design to improve visually impaired orientation in unfamiliar buildings: Requirements elicitation study. 2014 IEEE 22nd International Requirements Engineering Conference (RE) (Aug. 2014), 23–32.

[4] Alkhanifer, A.A. and Ludi, S. 2014. Visually impaired orientation techniques in unfamiliar indoor environments. Proceedings of the 16th international ACM SIGACCESS conference on Computers & accessibility - ASSETS '14 (New York, New York, USA, Oct. 2014), 283–284.

[5] Azenkot, S. et al. 2011. Enhancing Independence and Safety for Blind and Deaf-Blind Public Transit Riders. Proceedings of the 2011 annual conference on Human factors in computing systems - CHI '11 (New York, New York, USA, May 2011), 3247–3256.

[6] Banovic, N. et al. 2013. Uncovering information needs for independent spatial learning for users who are visually impaired. Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility - ASSETS '13 (New York, New York, USA, Oct. 2013), 1–8.

[7] Being a Sighted Guide: http://www.afb.org/info/friends-andfamily/etiquette/being-a-sighted-guide/235. Accessed: 2014-09-07.

[8] Butz, A. et al. 2001. A hybrid indoor navigation system. Proceedings of the 6th international

conference on Intelligent user interfaces - IUI ’01 (2001), 25–32.

[9] Chumkamon, S. et al. 2008. A blind navigation system using RFID for indoor environments. 2008 5th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (May 2008), 765–768.

[10] Coroama, V. 2003. The Chatty Environment - A World Explorer for the Visually Impaired. Adjunct Proceedings of Ubicomp 2003. (2003).

[11] Atri, E. et al. 2007. A system to aid blind people in the mobility: A usability test and its results. 2nd International Conference on Systems, ICONS 2007 (Apr. 2007), 35–35.

[12] Ellis, R.D. and Kurniawan, S.H. 2000. Increasing the Usability of Online Information for Older Users: A Case Study in Participatory Design. International Journal of Human-Computer Interaction. 12, 2 (2000), 263–276.

[13] Faria, J. et al. 2010. Electronic white cane for blind people navigation assistance. World Automation Congress (WAC), 2010.

[14] Feng, C. et al. 2015. Designing a Robot Guide for Blind People in Indoor Environments. Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction Extended Abstracts - HRI’15 Extended Abstracts (New York, New York, USA, Mar. 2015), 107–108.

[15] Golding, A.R. and Lesh, N. 1999. Indoor navigation using a diverse set of cheap, wearable sensors. Digest of Papers. Third International Symposium on Wearable Computers (1999), 29–36.

[16] Shiri Azenkot, Catherine Feng, and Maya Cakmak Enabling Building Service Robots to Guide Blind People A Participatory Design Approach