IRRMA: An Image Recommender Robot Meeting Assistant

Benoît Alcaraz[1][0000−0002−7507−5328], Nina Hosseini-Kivanani[1][0000−0002−0821−9125], and Amro Najjar[2][0000−0001−7784−6176]

[1] University of Luxembourg, Luxembourg
{benoit.alcaraz,nina.hosseinikivanani}@uni.lu
[2] Luxembourg Institute of Science and Technology (LIST)
amro.najjar@list.lu

Abstract. The number of people who attend virtual meetings has increased as a result of COVID-19. In this paper, we present a system that consists of an expressive humanoid social robot called QTRobot, and a recommender system that employs natural language processing techniques to recommend images related to the content of the presenter's speech to the audience in real time. This is achieved utilising the QTRobot's platform capabilities (microphone, computation power, and Wi-Fi).

Keywords: Robot Assistant · Meeting Assistant · Recommender System.

1 Introduction

Presentations and meetings are recurrent tasks in today's workplace. Furthermore, the COVID-19 situation has shifted meetings modes and forced users to increase the number of meetings they attend [2]. As a result, technologies that improve the productivity of meetings are highly valuable. Some meeting assistant technologies have been developed in recent decades to help with this. Recommender systems (RSs) [1] are used to provide information for users and facilitate smooth decision-making when participants are overloaded with information. This issue has been addressed in the literature of smart assistants and intelligent personal assistants for decades. For instance, Augmented Multi-party Interaction (AMI) [3], a recommender system based on what the speaker says, provides participants with textual documents and resources from previous meetings in real time. However, we can identify two problems with this work.

First, the recommendations are often limited to those of previous meetings or what the speaker has provided. This makes the system unsuitable for meetings where the topic under consideration has not been previously discussed. It also requires additional work on the part of the speaker, who needs to provide resources to fill the database. Secondly, even though textual documents provide a lot of valuable information, they cannot be viewed during a presentation without the participants missing information from the speaker. Furthermore, reading

documents is time-consuming. If the number of documents recommended during a meeting is too high, people may simply ignore them due to lack of time, and so the system is no longer valuable.

Addressing these issues requires expanding the recommendation database while also improving its fit for the meeting context. The database can be extended with data available on the Internet. It is possible to create a whitelist of documents from which appropriate recommendations can be taken. Websites like Wikipedia Commons[3], the Wikipedia image database, can be added to the whitelist. Furthermore, a recommendation fits a meeting context if it can be viewed without the audience being distracted from what the speaker is saying. Images fit such contexts perfectly as they can be viewed while listening to the speaker. A photo, a schema, or a statistical graph are all examples of recommended images.

This paper is a demonstration for PAAMS 2022 of our Image Recommender Robot Meeting Assistant (IRRMA). The information is displayed through the interface of an expressive humanoid social robot called QTRobot standing on the meeting table in front of the participants. Figure 1 shows IRRMA during a meeting. Figure 2 shows the architecture of our approach.

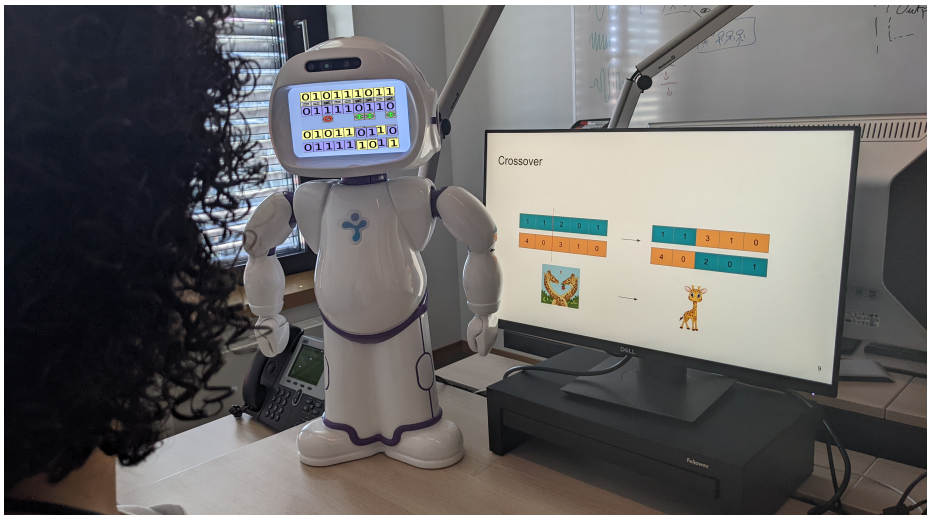

Fig. 1. QTRobot using IRRMA next to the main presentation monitor.

In meeting situations, IRRMA can help users better retrieve information from the presenter, and it can fine-tune the assistance during usage. IRRMA will provide as information in an application domain (i.e. providing "live" recommendations during meeting sessions).

[3] https://commons.wikimedia.org

More specifically, the advantages of IRRMA are threefold: (i) it provides information that can be consulted during the meeting, (ii) it can catch the attention of the audience more effectively than a monitor showing recommendations, (iii) it can search for recommendations based on the contents of a large database without requiring too much preparation time from the speaker.

2 Main Purpose

IRRMA's aim is to suggest images that are related to what the speaker is saying. The goal is to add more information and references to a presentation in order to improve its quality. A statistical graph, for example, can be displayed to evaluate what the speaker is currently saying, or a photo can be used to illustrate the speech. A whitelist of websites can be used as a database to find images. This provides access to an enormous number of suitable images. Also, if the topic is known prior to the presentation, a whitelist can be created of websites strictly related to the topic. This would allow recommendations of a higher quality. We have chosen to use a QTRobot to represent IRRMA because it can grab the attention of participants when a recommendation is made. It also better fits the role of "assistant" in that way. We intend to design and develop an interactive user support system. Multiple challenges have been identified, including (i) speech processing and keyword extraction, (ii) developing the recommendation system, and (iii) how to share the recommendation with the audience. We propose an early stage demonstration of IRRMA in a meeting situation, with a focus on the recommendations and the way we present them.

3 Demonstration

Figure 2 depicts the architecture of our system. Our procedure can be broken down into several steps. To begin with, IRRMA has to extract the keywords of the speech. It uses the QTRobot microphone to listen to the speaker. It then uses speech-to-text technologies to convert the speech to text. It then extracts the keywords from the transcript using a custom-made algorithm. Next, it must choose a recommendation based on the keywords obtained. To do so, it uses the keywords to iterate and search through the images in its whitelist. If a suitable image is found, it is chosen as the recommendation.

Finally, the recommendation has to be presented to the audience. This step largely depends on the type of meeting (i.e. online, offline or hybrid). The QTRobot can be connected as a user of the remote meeting application, which then shares its screen with the online participants. It is placed on the table next to the presenter and faces the audience. The recommendation can then be displayed on the tablet held by the QTRobot. The QTRobot makes a movement with its arms when displaying the recommendation to attract the audience's attention. Although the speaker's monitor is present in our architecture, it is possible that the speaker has nothing to show.

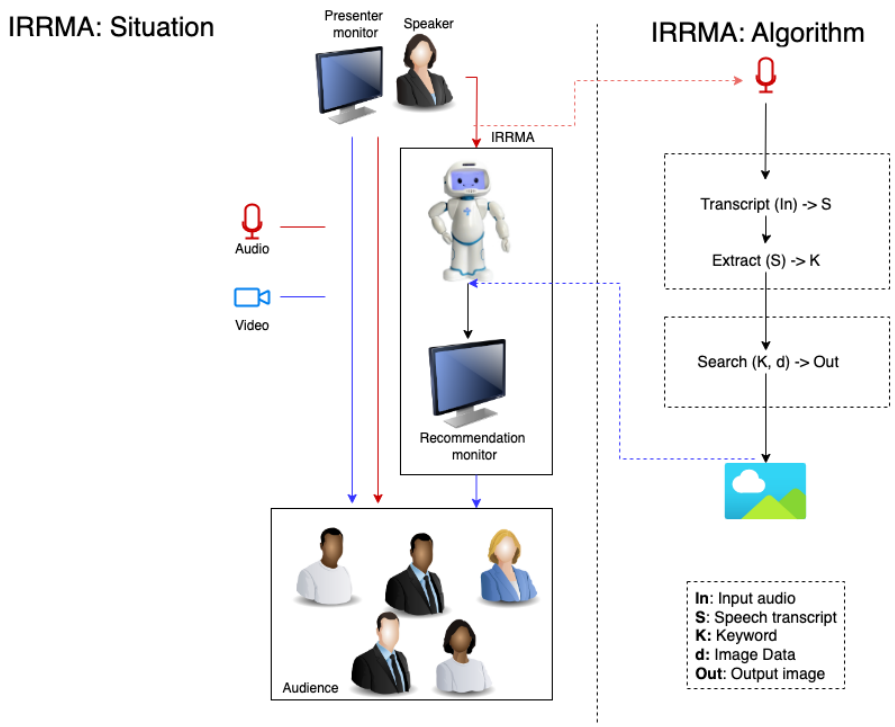

Fig. 2. Architecture of IRRMA in situ and its algorithm.

4 Conclusion

Our system, an Image Recommendation Robot Meeting Assistant, aims to recommend relevant images in real time based on the contents of a presentation. Natural language processing technologies are used to accomplish this. We also use the possibilities that this embodiment provides to improve the trust among audience's into the system [4].

Our contribution is relevant for meetings where the speaker is giving a presentation, as well as for other situations where the speaker dose not necessarily have slides to show. We plan to improve the system in the future so that it can handle multimedia recommendations and follow a meeting with multiple speakers.

Acknowledgments

This work has been funded by the Luxembourg National Research Fund under Grant No. IS/14717072.

This work is partially supported by the Chist-Era grant CHIST-ERA19-XAI-005, and by *(i)* the Swiss National Science Foundation (G.A. 20CH21_195530), *(ii)* the Italian Ministry for Universities and Research, *(iii)* the Luxembourg National Research Fund (G.A. INTER/CHIST/19/14589586), *(iv)* the Scientific and Research Council of Turkey (TÜBİTAK, G.A. 120N680).

References

1. Bobadilla, J., Ortega, F., Hernando, A., Gutiérrez, A.: Recommender systems survey. Knowledge-based systems **46**, 109–132 (2013)
2. DeFilippis, E., Impink, S.M., Singell, M., Polzer, J.T., Sadun, R.: Collaborating during coronavirus: The impact of covid-19 on the nature of work. Tech. rep., National Bureau of Economic Research (2020)
3. Ehnes, J.: A tangible interface for the ami content linking device—the automated meeting assistant. In: 2009 2nd Conference on Human System Interactions. pp. 306–313. IEEE (2009)
4. Stroessner, S.J., Benitez, J.: The social perception of humanoid and non-humanoid robots: Effects of gendered and machinelike features. International Journal of Social Robotics **11**(2), 305–315 (2019)