User-Friendly Smartphone Interface to Share Knowledge in Human-Robot Collaborative Search Tasks

J. E. Domínguez-Vidal, Iván J. Torres-Rodríguez, Anaís Garrell and Alberto Sanfeliu

***Abstract*— Long-distance human-robot collaborative tasks require robust forms of knowledge-sharing among agents in order to optimize the performance of the task. In this paper, we propose to take advantage of the proliferation of mobile phones to use them as a reliable low-cost communication interface, as opposed to the use of specific gadgets or speech and gesture recognition techniques that are highly prone to failure in the presence of noise or occlusions. Our interface is focused on search tasks, and it allows the user to share with the other agents real-time information such as their position, their intention or even what they would like the other agents to do. To test its acceptability, a user study was conducted with 20 volunteers in a human-human scenario. A second round of experiments with other 30 volunteers was conducted to test different ways to encourage user interaction with our interface. Finally, real-life experiments were also conducted with a robot to apply skills learned to the desired scenario. We found a statistically significant improvement in the amount of information exchanged between agents.**

I. Introduction

Since the appearance of the iPhone in 2007, mobile phones have become more and more present in our lives, to the point of becoming small pocket computers equipped with multiple sensors. The same hardware development that has enabled this proliferation has also driven the development of robotics, allowing it to reach greater heights of cognition and sociability. Achievements in the fields, such as voice and gesture recognition as well as natural language processing, have enabled robots and humans not only to communicate, but also to interact or work together. However, for long-distance tasks, neither option allows reliable communication, due to ambient noise or lack of resolution.

In the specific case of human-robot collaborative search, the presence of obstacles or ambient noise makes it impossible to use the above alternatives to share information between both agents and thus optimize the search. Due to the above, in this paper, we present an user-friendly smartphone human-robot interaction (HRI) interface, which makes it unnecessary to use a specific gadget, and at the same time allows us to simplify the knowledge sharing between both agents in a way that is intuitive for the human and useful for the robot (see Fig. 1).

Work supported under the Spanish State Research Agency through the Maria de Maeztu Seal of Excellence to IRI (MDM-2016-0656) and ROCOTRANSP project (PID2019-106702RB-C21 / AEI / 10.13039/501100011033)

The authors are with the Institut de Robòtica i Informàtica Industrial (CSIC-UPC). Llorens Artigas 4-6, 08028 Barcelona, Spain. {jdominguez, itorres, agarrell, sanfeliu}@iri.upc.edu

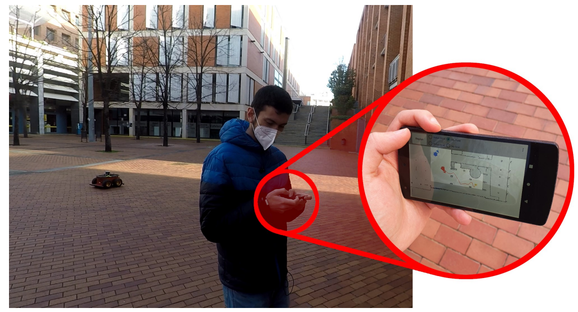

Fig. 1. **Participant using our interface to communicate with the robot:** It is not necessary to maintain visual contact with the robot at any time.

In this article, first, we check that our interface is acceptable by the user. To do this, we have designed a simplified version of our interface, and we have performed an user study on a search task between two humans in which one of them simulates the robot. We have also analyzed how to stimulate the use of the HRI interface in order to increase the user's collaboration in different aspects of the task. Finally, we present the full version of our interface, through which both the robot and the human can share all the information necessary for the correct development of this task, for example, their position and their intentionality.

In the reminder of the paper, we start describing the related works in Section II. In Section III we introduce the human-robot knowledge sharing in collaborative tasks. Section IV and V present the conducted experiments and the conclusions, respectively.

II. Related Work

This research is focused on developing a new interface to make robots capable of working cooperatively with people on map-based tasks.

The coordination among robots in collaborative tasks can be achieved with graphical model inference techniques as variational methods [1], [2] or other coordination algorithms [3] both for centralized [4] and decentralized collaborative scenarios [5], [6]. Examples can also be found where one or more team members are human rather than robots [7]. Regardless of the nature of the agents involved in a collaborative task, some communication channel needs to be established among them in order to optimize the performance of the task. This is especially important in human-robot teams as the two agents often represent in a different way the world associated with the task. For example, the TRADR project [8] focuses on enhancing human-robot teams collaboration in urban search and rescue (USAR) scenarios.

For the communications of robots and humans, several solutions have been stated. The two most commonly used methods for proximity tasks are based on speech recognition [9], [10] and gesture recognition [11]. Both methods can only be used in close proximity to the robot and are not very robust to the acoustic or illumination noise. Other solutions based on augmented reality or virtual reality can also be found [12], but these solutions are often based on specific hardware that is not available to everyone making them unsuitable for potential mass use.

In the case of mobile phones, despite being such a widespread gadget, their use in robotics has been limited to teleoperation tasks [13], [14]. In this work we will focus on the capabilities of mobile phones not only as teleoperation means but as graphical interfaces to share knowledge between humans and robots in a way that is effective regardless of the distance between the two agents and socially acceptable for the human. We also validate all our hypotheses through user studies, something that is usually absent in the mentioned works.

III. Human-Robot Knowledge Sharing in Collaborative Search Tasks

When a set of agents, humans, robots or mixed groups, must perform a task collaboratively, they should share not only information but knowledge. Focused on collaborative search tasks, it is necessary to know for example, the explored areas in order not to waste time re-exploring them. Moreover, due to the fact that we do not have access to the human mental knowledge about the task, neither its intentions, we need to design a HRI interface that facilitates getting information about the task for both the robot and the human and making a search plan agreement between them. The objective of the proposed HRI interface is that it can show the map, the explored areas, the human's and robot's pose, the human's intention, the shared human-robot plan and the possibility to modify the plan by the human agent.

It should be noted that collaborative search tasks can be carried out over arbitrarily large areas. Therefore, the communication between agents must be able to be achieved regardless of the distance, occlusions or obstacles between agents making voice or gestures inappropriate for the job.

A. Mobile Phone App as HRI

Fig. 2 depicts our proposed HRI diagram for performing a collaborative search task between a human and a robot. We propose to use something that almost all of us carry in our pockets, such as a smartphone, as the HRI interface. Based on the work from Kohler et al. [15], we have designed an interface that makes use of the touch screen and connectivity capabilities of any mobile phone. Through this interface, the human can see the same map on which both agents are working collaboratively and indicate on it their position and their intention. This interface is in charge of reading the human's clicks on the map and converting them into messages that the robot can understand. These messages are sent to the robot and used, together with the environmental information it receives from its own sensors, to calculate the route that optimizes the search. The specific planner and sensors used are irrelevant as they are transparent to the HRI interface. Finally, the robot sends its position and the routes it calculates to the interface, which translates them into visual information on the HRI map.

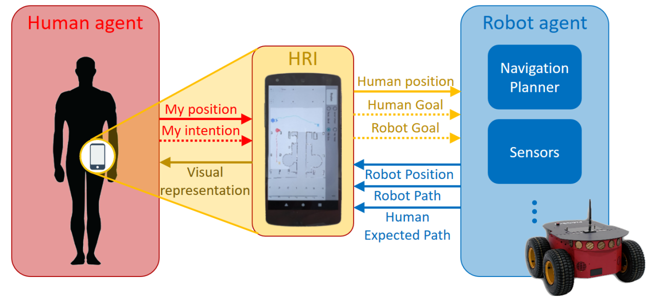

Fig. 2. **Information flow considered in our proposal.** Smartphone acts as an HRI interface converting graphic information understandable by the human into messages understandable by the robot and vice versa.

Fig. 3 depicts the interface's appearance. Human users can indicate their position using the *'My pose'* functionality and their intention using *'My goal'* and *'Robot goal'* options allowing them to also indicate to the robot what they would like it to do. The *'Replan'* button is used to indicate to the robot that the updated data is completed and that it can be used to redo its calculations. Note that it is not shown which areas have been explored and which have not for clarity and ease of use. However, the generality of this interface makes it possible to display this or any other information deemed necessary by simply modifying the layout of the main screen.

The communication can be done through a local wifi network or with a mobile network (3G or 4G) solving the problem of distance or occlusions that communication via voice or gestures could not solve. Regarding the human's pose accuracy in the map, it is not very good since it is done by the human by looking in the HRI map, but in our case it can be improved if the robot can detect the person using the its sensors.

B. Triggering Behaviours

The HRI performance is highly dependent on the users indicating their pose in the map. If the robot does not know the user's intention, it can still take control and tell the user which route it thinks they should follow to optimize the search, but it does need the user's position to do those calculations. In general, this problem is repeated in every situation and task, where the user's active participation is required.

Based on Fogg [16], we have designed three triggers or reminders that encourage the user to indicate their pose: *(1)* a toast (small message) on the screen to indicate to the user that more than a threshold time has elapsed since their last known position, *(2)* a pop-up window that blocks the usage of the interface until the user accepts and indicates their new position, and *(3)* a 'multimodal' trigger, i.e., a vibration of the device and a blink of the *'My pose'* option with the same time

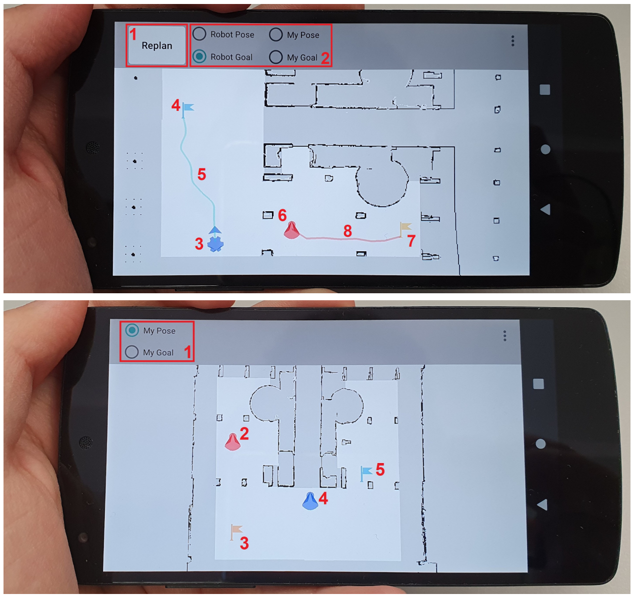

Fig. 3. **Mobile App main screen.** *Top:* Screenshot of an ongoing human-robot collaborative search where the human has used every available functionality. 1) Replan button. 2) Input data selection menu. 3) Robot's current position. 4) Human's desired goal for the robot. 5) Path calculated by the robot. 6) Human's current position. 7) Human's intended goal. 8) Path calculated for the human. *Bottom:* Execution using a simplified version of the interface for human-human collaborative search experiments. 1) Input data selection menu. 2) Human's current position. 3) Human's intended goal. 4) Human's partner position 5) Human's partner intended goal

threshold as the two previous triggers. Additionally, based on the results obtained with gamification techniques [17], a fourth trigger adds a counter to each available option to indicate to the user how many times they have indicated their positions or intentions[1].

Section IV-B shows more in detail the results obtained using the different triggers.

IV. EXPERIMENTS

It is necessary to carry out an user acceptability study to find out if the user feels comfortable using our interface, what is the added value and what is the cost to be paid. To do this, we have used a simplified version of the interface (see Fig. 3 - *Bottom*). This simplified interface has been tested with two persons, instead of a person and a robot to perform the search so that we can measure the effect of the interface on task performance by comparing the case of having it versus not having it. The same version has been used to test the effectiveness of the four triggers by stimulating the user to use the interface. The full version (Fig. 3 - *Top*), along with what is learned in these experiments, has been used for testing with the robot.

A. Experiments Setup and Methodology

The experiments were carried out in the Barcelona Robot Lab using an area of approximately 750 m^2, half of which corresponds to an open space and the other half to a covered area with multiple occlusions (walls and columns). (see Fig. 4).

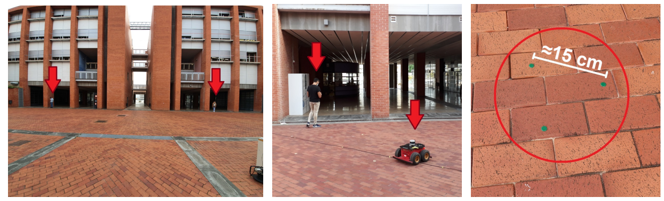

Fig. 4. **Experiment examples.** *Left:* Human-human pair collaboratively searching. *Center:* Human-robot pair doing the same task. *Right:* Group of three green Parcheesi tokens object of the search.

We divided the experiments into two rounds. In the first round, 20 volunteers (age: μ=21.1, σ=3.46; most common ongoing or finished studies: B.Sc.) performed two collaborative search tasks, always teaming up with the same research assistant (so that the second component of all pairs always had the same effect on the outcome of the experiment). In the first task, they did not have the simplified interface to communicate, so they had to resort to voice, gestures and other instruments of human communication[2]. In the second task, they did have our interface running on two mobile phones, one for each member of the pair[3]. At the end of the second task, all volunteers were given a questionnaire to evaluate different aspects of interest using ANOVA tests.

In the second round, 30 different volunteers (age: μ=27.8, σ=5.21; most common ongoing or finished studies: M.Sc.) tested the effectiveness of the four triggers. As they are a different population sample, it is necessary to perform a first experiment with the interface and no triggers in order to standardize the results with those obtained in the previous round. After this, each volunteer performed two collaborative search tasks using in each one a different randomly selected trigger and again always with the same research assistant. At the end of all the tasks, another questionnaire was given in order to evaluate their subjective perception of the different aspects related to the triggers also using ANOVA tests.

In order to have objective data, the flow of messages sent through the two interfaces in all experiments was saved in .csv files. A total of 50 volunteers participated in up to 130 experiments. The interface was executed on two mobile phones, a Nexus 5 and a Galaxy S10, both with Android 10 (kernel 4.14, October 2020 compilation).

B. Acceptability and trigger experiments

Our first hypothesis is that the interface does not impair task performance, but can enhance it. The results of the questionnaires filled out by the volunteers after the first round of experiments (one search without the interface and one search with the interface) are depicted in Fig. 5.

As for the quickness, i.e., how much time the user needs to communicate relevant information, no statistically significant increase is shown (using the criterion of $p<0.05$): without interface μ=4.20, σ=1.54; with interface μ=5.00, σ=1.83;

[1]Four triggers: https://youtu.be/-xCuymEaoAY

[2]Episode without interface: https://youtu.be/G-w5QlThqu0

[3]Episode with interface: https://youtu.be/a205ZKZC_XA

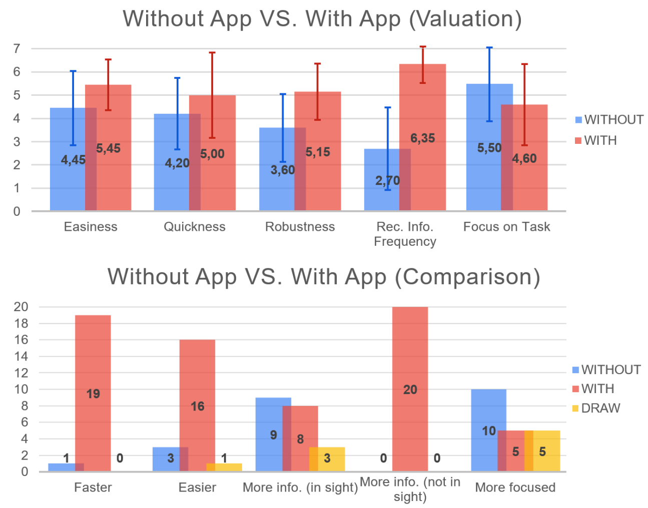

Fig. 5. **Main aspects of the acceptability user study.** *Top:* Valuation from 1 (low) to 7 (high) of the main aspects related to the use of the interface. *Bottom:* Number of times each option was selected in the comparison where the user had to chose among three options instead of valuate them numerically. As they are 20 volunteers, the maximum is 20 for each option.

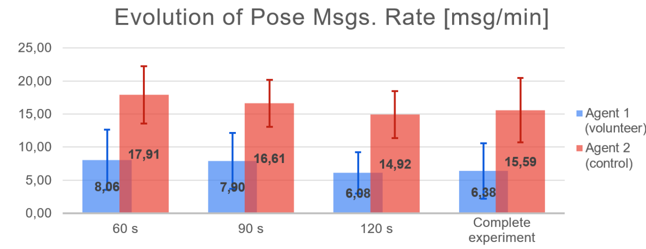

Fig. 6. **'My pose' messages rate throughout the experiment.** Evolution of the average number of messages per minute sent by both volunteers (blue) and the control user, always the same research assistant (red).

TABLE I

PERCENTAGE OF 'MY POSE' TBM<10 S [%]

User	Experiment evolution			Complete
	60 s	90 s	120 s	Experiment
Control	94.79	94.36	94.25	90.95
Volunteer	81.82	81.76	70.59	72.94

$t(20)$=-1.45, p=0.162. The same does not occur with the ease of communicating with the partner: without μ=4.45, σ=1.60; with μ=5.45, σ=1.09; $t(20)$=-3.01, p=0.007. This helps us to confirm that the interface does not affect the speed of communication, but can make it easier. If we look at robustness, i.e., how difficult it is for the information to be misinterpreted, there is a significant improvement: without μ=3.60, σ=1.46; with μ=5.15, σ=1.22; $t(20)$=-4.01, p<0.001. This is even more noticeable when the amount of information received is analyzed: without μ=2.70, σ=1.78; with μ=6.35, σ=0.83; $t(20)$=-7.17, p<0.001. This shows that the interface can provide more and better information. The price to pay is a decrease in task concentration, although not statistically significant: without μ=5.50, σ=1.63; with μ=4.60, σ=$1,75$; $t(20)$=1.76, p=0.09.

In addition to rating the above aspects from 1 to 7, we also asked the volunteers to choose between the two possibilities (without interface and with), giving them also the draw as a valid option. The result is shown in Fig. 5 - *Bottom*. There is consensus that it is easier and faster to communicate through the interface as it is less prone to misunderstandings. There is unanimity that the interface provides more information when in line-of-sight with the partner. However, there is debate about which is the best way to obtain information when the partner is in view as well as which is easier to maintain concentration with.

Focusing on the exchange of *'My pose'* messages, which is the most important for the robot to know where its partner is when it cannot see it and thus optimize the search, there is a clear difference between the control user (someone with extensive knowledge of robotics and, therefore, perfectly aware of the needs of the robot for its optimal operation) and volunteer users (users with average or low knowledge of robotics and typical target of research in social robotics). Fig. 6 shows that the control user sends more than twice as many *'My pose'* messages as the average user, both if we look at the temporal evolution over the experiments or if we just look at the average of all experiments. This is only a problem if the frequency with which the average user refreshes their position is not high enough for the robot to interpolate its path. Considering that the search task involves searching for a small-sized object at ground level (see Fig. 4 - *Right*), the movement speed of the users ranges from 0.5 to 1.0 m/s. Therefore, we can set 10 s as a good threshold for the time between messages (TBM) of *'My pose'* type.

Table I shows the percentage of messages that are sent before exceeding this 10 s threshold with respect to the previous message. This percentage is calculated as the total number of messages sent in all experiments below this threshold divided by the total number of messages sent in all experiments. Roughly speaking, between 18% and 29% of the average user's messages do not meet this threshold, which can greatly complicate the robot's planning.

Our second hypothesis is that we can change user's behavior and encourage them to increase the number of messages they send within the recommended interval. This is also useful for other situations where it is necessary that the user performs some specific type of interaction. For this purpose, we have used the four triggers or reminders presented in Section III-B.

As mentioned in Section IV-A, volunteers in this second round of experiments correspond to a different population sample, so it is necessary to perform them a first experiment without any trigger in order to standardize the measurements. The results are shown in the first row of Table II. It was necessary to reduce the threshold from 10 s to 8.32 s to obtain similar percentages (76.77 vs. 76.82 on average). Therefore, this is the threshold we used to test the triggers. In other words, we repeated the search experiment two more times to each of the 30 volunteers in the second round using

TABLE II
PERCENTAGE OF 'MY POSE' TBM<8.32 S DEPENDING ON TRIGGER [%]

Trigger	Experiment evolution			Complete
	60 s	90 s	120 s	Experiment
None	77,98	77,91	76,26	75,12
Toast	71,57	77,07	78,67	74,69
Pop-up	**89,33**	86,86	**87,11**	80,01
Multimodal	89,22	**88,50**	86,19	**82,04**
Counters	84,46	84,77	81,31	80,21

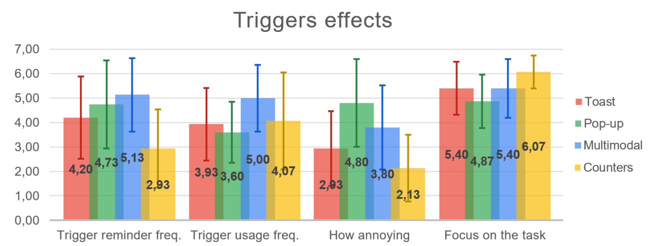

Fig. 7. **Main aspects of the triggers user study.** Valuation from 1 (low) to 7 (high) of the main aspects related to the effects of triggers.

a different trigger each time and tested the percentage of messages sent before 8.32 s. The result can be seen in the remaining rows of Table II.

Toast is counterproductive. The user relies on something to remind them to use the interface, but toast is not effective enough to do this. The Pop-up and Multimodal have similar effectiveness due to different reasons. Pop-up is based on something people want to avoid, while Multimodal is a real reminder acting only in case it is needed. Counters is proof that gamification techniques can give good results (an increase of 6% in this case) with very little computational burden and annoyance for the user.

Fig. 7 summarizes the main aspects evaluated in the post-task questionnaire. As for the subjective frequency with which the trigger reminds the user to use the remembered option, there is statistical significance in favor of Multimodal (M) over Counters (C): M μ=5.13, σ=1.37; C μ=2.93, σ=1.65; $t(15)$=2.79, p=0.014. This means that users feel that Counters is much more sibylline. Pop-up (P) and Muntimodal tie in this aspect, but not about the trigger usage frequency, i.e., how frequently users use an option because the trigger remembered them to do it: P μ=3.60, σ=1.25; M μ=5.00, σ=1.37; $t(15)$=-2.19, p=0.046. This means that users recognize that Multimodal is more effective in modifying their behavior even though they consider both Pop-up and Multimodal to be equally frequent. Regarding annoyance, Multimodal is considered to be more annoying than Counters: M μ=3.80, σ=1.72; C μ=2.13, σ=1.36; $t(15)$=2.45, p=0.028. Pop-up has even worse valuation. Finally, Counters also shows a significant increase over Pop-up in terms of task concentration, but not over Multimodal: P μ=4,87, σ=1.02; C μ=6.07, σ=0.68; $t(15)$=-2.17, p=0.048.

As we did in the acceptability study, we also asked the volunteers to choose between the two triggers they

More efective?	Trigger T	Trigger P	Trigger M	Trigger C
Trigger T		1,5 (3)	0 (0)	1 (0)
Trigger P	3,5 (3)		2 (0)	3 (0)
Trigger M	5 (0)	3 (0)		3 (0)
Trigger C	4 (0)	2 (0)	2 (0)	

More annoying?	Trigger T	Trigger P	Trigger M	Trigger C
Trigger T		0 (0)	2 (2)	3 (0)
Trigger P	5 (0)		5 (0)	5 (0)
Trigger M	3 (2)	0 (0)		4 (2)
Trigger C	2 (0)	0 (0)	1 (2)	

More intrusive?	Trigger T	Trigger P	Trigger M	Trigger C
Trigger T		0 (0)	1 (2)	4 (0)
Trigger P	5 (0)		2 (0)	5 (0)
Trigger M	4 (2)	3 (0)		5 (0)
Trigger C	1 (0)	0 (0)	0 (0)	

Easier to focus?	Trigger T	Trigger P	Trigger M	Trigger C
Trigger T		5 (0)	2 (0)	1,5 (1)
Trigger P	0 (0)		1 (0)	0,5 (1)
Trigger M	3 (0)	4 (0)		3 (2)
Trigger C	3,5 (1)	4,5 (1)	2 (2)	

Fig. 8. **Comparison of triggers.** T=Toast, P=Pop-up, M=Multimodal, C=Counters. One point each time a trigger is selected for each aspect. 0.5 for both triggers in case of draw. Number of draws in parentheses. Each row indicates the result of comparing the selected trigger with all other triggers. The maximum score is 5.0 since each of the 30 volunteers performed only 1 of the 6 possible combinations of 2 triggers.

had used (or draw) with respect to the above parameters. Fig. 8 shows user preferences. In general, they consider that the most effective trigger is Multimodal, that Pop-up and Multimodal are equally intrusive and that Pop-up is the most annoying. This indicates that intrusiveness is seen as a positive quality in the case of Multimodal. Finally, Counters is also considered the best to facilitate keeping focused on the task.

In general, the conclusion is that if it is needed to make sure that the user will behave in a specific way at a specific time, the most effective way is to use a multimodal trigger, that is, something that interacts in multiple ways with the user. However, care must be taken to ensure that the user does not become fatigued or the result may be the opposite of what is expected. If sustained behavior is needed, gamification techniques seem to show more promising results.

C. Human-robot collaborative search experiments

It is time to use the full version of the interface shown in Fig. 3 - *Top* applying the knowledge gained in the previous experiments to perform the same task, but between a human and a robot as depicted in Fig. 4 - *Center*[4].

We added to the previous version the ability to tell the robot where it is expected to go with a *'Robot goal'* option. We also created a *'Replan'* button to request two new plans, one for the robot and one for the human, once the user has updated all the data they wish to use. The possibility of including the person's intention through the HRI interface (with both *'My goal'* and/or *'Robot goal'* messages) allows the emergence of different types of relationships: master-slave with the person as master if they tell the robot what to do or as slave if they allow the robot to make the decisions to optimize the search. The peer-to-peer relationships can also appear if the person indicates their own goals and lets the robot adapt as it wishes. More details about the robot planner can be found in [18].

The robot includes a Lidar and a depth camera for navigation purposes. In addition, the Lidar is used to track the user whenever the user is in sight and within 20 m of the robot [19]. This reduces the inconvenience on the user of continually indicating its pose. However, when the tracker

[4]Human-Robot search: `https://youtu.be/nuEECzNKcEI`

TABLE III

HUMAN-ROBOT EXPERIMENTAL RESULTS RELATED WITH HRI USAGE

Parameter	Mean	Std. Dev.
Tracked time [%]	49,16	14.52
No. tracker re-coupling	5,538	2,578
'My pose' TBM <10 s [%]	77.14	–
'My pose' TBM <11 s [%]	85.71	–

loses the user (due to occlusions), the user will have to start indicating its position until it is in sight of the robot again. To achieve this, we have implemented a double trigger. Firstly, the color of the icon changes from green to red to visually indicate that the robot is not tracking the user (similar to Toast) and, secondly, the same Multimodal trigger explained above vibrates the mobile phone if more than $10\ s$ have passed since the last update of the user's pose.

Table III shows that with these modifications we have achieved that the robot can track the user up to almost half of the time. This is because it manages to relocate the user more than 5 times per experiment thanks to the user's indication of where it is when the robot is not watching it. As for the TBM when the user is not being tracked, 77% of the messages have been sent before the $10\ s$ threshold. This is higher than shown in Table I for complete experiments, but lower than the same measurement in Table II for Multimodal trigger. The explanation is that in this case, people do not know when the robot is going to lose their track so they tend to focus on the task and use the vibration to know when it is necessary to start sending their position. Because of this, if we look at the effectiveness of the trigger one second after it is activated, it goes up to almost 86%. In other words, we could move the trigger threshold to $9\ s$ to achieve the desired effect at the end of the $10\ s$ we had set as a limit.

V. CONCLUSIONS

We have developed an interface that can function as an HRI interface using a gadget that almost all of us carry in our pocket such as a mobile phone. We have conducted a user study to verify that this interface is acceptable to the user and that it can improve communication in collaborative search tasks. Anticipating the need for the user to use our interface in a specific way, we have developed and tested different methods to encourage the desired behavior. Finally, we have applied all of the above to improve the performance of a robot in human-robot collaborative search tasks.

One possibility for future work, is the creation of a web version of our interface. This would allow it not to depend on the specific hardware or the Operating System of each device, making it universally available. The disadvantage of this version would be that it would not have access to the sensors built into mobile phones (accelerometers, gyroscopes, cameras...). The use of these sensors also opens up a myriad of future possibilities to, for example in our case, automate the collection of the user's location.

ACKNOWLEDGMENTS

The authors want to express their gratitude to the fifty volunteers who made this work possible. Specially to Nicolás Rodríguez and David Henry for their infinite patience.

REFERENCES

[1] D. Koller and N. Friedman, *Probabilistic Graphical Models. Principles and Techniques*. MIT Press, 2009.

[2] I. Rezek, D. S. Leslie, S. Reece, S. J. Roberts, A. Rogers, R. K. Dash, and N. R. Jennings, "On similarities between inference in game theory and machine learning," *Journal of Artificial Intelligence Research*, vol. 33, pp. 259—-283, 2008.

[3] A. Kumar, S. Zilberstein, and M. Toussaint, "Probabilistic inference techniques for scalable multiagent decision making," *Journal of Artificial Intelligence Research*, vol. 53, pp. 223—-270, 2015.

[4] R. Luna and K. E. Bekris, "Efficient and complete centralized multi-robot path planning," in *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2011, pp. 3268–3275.

[5] T. Regev and V. Indelman, "Multi-robot decentralized belief space planning in unknown environments via efficient re-evaluation of impacted paths," in *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2016, pp. 5591–5598.

[6] G. Best, O. M. Cliff, T. Patten, R. R. Mettu, and R. Fitch, "Dec-mcts: Decentralized planning for multi-robot active perception," *The International Journal of Robotics Research*, vol. 38, no. 2-3, pp. 316–337, 2019.

[7] V. Govindarajan, S. Bhattacharya, and V. Kumar, "Human-robot collaborative topological exploration for search and rescue applications," in *Distributed Autonomous Robotic Systems*, vol. 112. Springer Japan, 2016, pp. 17–32.

[8] I. Kruijff-Korbayová, F. Colas, M. Gianni, F. Pirri, J. de Greeff, K. Hindriks, M. Neerincx, P. Ögren, T. Svoboda, and R. Worst, "Tradr project: Long-term human-robot teaming for robot assisted disaster response," *KI-Künstliche Intelligenz*, vol. 29, no. 2, pp. 193–201, 2015.

[9] R. K. Megalingam, M. Manaswini, J. Yannam, V. S. Naick, and G. N. Chowdary, "Human robot interaction on navigation platform using robot operating system," in *2020 Fourth Int. Conf. on Inventive Systems and Control (ICISC)*, 2020, pp. 898–905.

[10] Y. Zhang, Z. Lu, C. Wang, C. Liu, and Y. Wang, "Voice control dual arm robot based on ros system," in *2018 IEEE Int. Conf. on Intelligence and Safety for Robotics (ISR)*, 2018, pp. 232–237.

[11] P. Neto, M. Simão, N. Mendes, and M. Safeea, "Gesture-based human-robot interaction for human assistance in manufacturing," *The International Journal of Advanced Manufacturing Technology*, vol. 101, 03 2019.

[12] P. Marti and I. Iacono, "Social and empathic behaviours: Novel interfaces and interaction modalities," in *2015 24th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, 10 2015, pp. 217–222.

[13] A. Speers, P. M. Forooshani, M. Dicke, and M. Jenkin, "Lightweight tablet devices for command and control of ros-enabled robots," in *2013 16th Int. Conf. on Advanced Robotics (ICAR)*, 2013, pp. 1–6.

[14] R. Codd-Downey and M. Jenkin, "Rcon: Dynamic mobile interfaces for command and control of ros-enabled robots," in *2015 12th Int. Conf. on Informatics in Control, Automation and Robotics (ICINCO)*, 2015, pp. 66–73.

[15] D. Kohler and K. Conley, "Rosjava - an implementation of ros in pure java with android support," 2011. [Online]. Available: https://github.com/rosjava/rosjava_core

[16] B. Fogg, "A behavior model for persuasive design," in *Proceedings of the 4th Int. Conf. on Persuasive Technology*, 2009, p. 40.

[17] M. Michael Sailer, J. U. Hense, S. K. Mayr, and H. Mandl, "How gamification motivates: An experimental study of the effects of specific game design elements on psychological need satisfaction," in *Computers in Human Behavior*, no. 69, 2017, pp. 371–380.

[18] M. Dalmasso, A. Garrell, J. E. Domínguez-Vidal, P. Jiménez, and A. Sanfeliu, "Human-robot collaborative multi-agent path planning using monte carlo tree search and social reward sources," in *IEEE Int. Conf. on Robotics and Automation (ICRA)*, 2021.

[19] V. Vaquero, E. Repiso, and A. Sanfeliu, "Robust and real-time detection and tracking of moving objects with minimum 2d lidar information to advance autonomous cargo handling in ports," *Sensors*, vol. 19, no. 1, p. 107, 2019.