

Research Article

Open Access

H. M. Ravindu T. Bandara*, M. A. Viraj J. Muthugala, A. G. Buddhika P. Jayasekara, and D. P. Chandima

Understanding uncertain information in vocal description for creating virtual spatial maps

<https://doi.org/10.1515/pjbr-2019-0032>

Received March 31, 2019; accepted September 12, 2019

Abstract: Assistive robots are developed for supporting daily activities of elderly people to uplift the living standards. The assistive robots should be friendly, reliable, active, and comprehensible in order to satisfy the needs of elderly population. Human activities are frequently related to navigational tasks and human tend to use descriptions which include natural language phrases and uncertain terms such as “near”, “little”, “far”, “small”, “large”, “close” to describe about spatial information. Therefore assistive robots should be capable of analysing and understanding descriptions which contain natural language phrases with uncertain terms and creating a conceptual map for effective navigation. This paper proposes a method to understand spatial information in a description with uncertain terms and creates a conceptual map in a robot memory which can be linked with spatial map for purposeful, effective and human friendly navigation task. Human studies have been carried out to study different types of descriptions related to navigation tasks. The Virtual Spatial Data Identifier (VSDI) and Uncertain Term Identifier (UTI) modules have been introduced in order to evaluate the spatial information in description to create a virtual map. Results of the system have been compared with the results of a human study in order to evaluate performance of the proposed system.

Keywords: uncertain information, virtual spatial maps, assistive robotics

1 Introduction

The world’s growth rate of elderly population continues in an unrivalled scenario [1]. The ageing population faces physical, mental and intellectual impairments [2] and the support of caregivers is an essential matter in uplifting the living standard of the older people [3]. Furthermore, number of experienced human caretakers is far below the required number and the energy and time of the workforce for elderly care can be directed to the development of a country [4]. As a solution to the above-mentioned fact, assistive robots and devices can be used as a substitution for the human caregivers [5]. Assistive robots can provide a support in typical daily activities of elderly people such as support in finding items, navigational tasks, medical schedules, avoidal social issues [6, 7]. In order to perform assistive service tasks, the assistive robots should be capable of navigating effectively and purposively inside the human populated environments [8, 9]. For effective navigation, primitive low-level of motion predication abilities such as collision avoidance are required. However, solving of such primitive low-level navigation functionalities is no longer a focus problem in indoor robot navigation area since the availability of large number of low-level navigation controlling methods and software packages [10]. Nowadays, the main aim is to support for developing human-friendly navigation mechanism, in which robots can understand the human motions and intentions, and react as a companion of human [11]. Natural language is a flexible, intuitive medium that can enable such interactions, but language understanding requires robots to learn representations of their environments that are compatible with the conceptual models used by people [12–14].

Humans have the cognitive ability to create virtual maps about the environment [15, 16] based on the information received through natural language instructions without actually perceiving the environment. A situation where a deliveryman, who is new to an office, is sent to deliver an item to a completely strange location by his/her workmate (who is well aware of that location) can be considered as an example case for the explanation. In this

*Corresponding Author: H. M. Ravindu T. Bandara: University of Moratuwa; E-mail: ra-ravindu@uom.lk

M. A. Viraj J. Muthugala: Singapore University of Technology and Design; E-mail: muthugala@ieee.org

A. G. Buddhika P. Jayasekara: University of Moratuwa; E-mail: buddhikaj@uom.lk

D. P. Chandima: University of Moratuwa; E-mail: chandimadp@uom.lk

kind of a situation, typically the workmate provides informative instructions regarding the arrangement of the location where the item is needed to be kept. The information conveyed through these instructions could be used by the deliveryman to visualize the arrangement of the environment, where the item is needed to be kept, without actually perceiving it. This visualization in the mind of the person will contribute towards the effective navigation since it will help him in various ways such as to find an efficient path to go to the required location. Therefore, ability of an assistive robot to visualize an unknown environment without actually perceiving it will contribute towards enhancing the navigation capabilities of the robot. In order to possess such an ability, the robot should be capable of grasping the knowledge conveyed from the language instructions of the user. However, comprehending the knowledge transfer through the natural language discussion for creating a virtual map of a location is not an easy task since the natural voice instructions related to arrangements of environments often include uncertain terms such as “far”, “near” and “few”. Hence, those uncertainties must be correctly understood by the robot.

A method has been proposed in order to navigate mobile service robots using verbal instructions with uncertain terms such as “move near to the table” [17]. The Proposed system understands the uncertain information in user commands related to navigation such as “close”, “near” and “far” based on the environment and the previous experience of the robot. A robot experience model (REM) has been introduced. REM has the capability to understand the lexical representations in user commands and to adapt the perception of the robot on the uncertain information in heterogeneous domestic environments. The user commands are more similar to natural language phrases so the system is not bounded by a strict grammar model. This system does not use a representative map of an unknown environment though it uses previous experience of the robot which makes it difficult to operate in an unknown environment by only using voice instructions. The system has been tested on static environments that are previously known by the robot. However, the system lacks the ability of creating a virtual map about the environment based on descriptive instructions of the user. A method to create a conceptual spatial representation of indoor environments for mobile robots has been introduced [18, 19]. The system is capable of linking the knowledge acquired through conversations with the maps created from a laser scanner and a vision sensor. The problem of fusing information contained in natural language descriptions with the robot’s onboard sensors to construct spatial-semantic representations useful for in-

teracting with humans has been addressed in [14]. An architecture for performing efficient symbolic goal-directed exploration in previously unknown environments when provided with structured language phrases about the environment has been proposed in [20]. The robot creates a representative map called abstract map based on the topological structure and spatial layout of symbolically defined locations by organizing symbolic language description of the unseen environment. The abstract map is useful to reason about spaces beyond the robot’s known world. This system also uses the metric guidance provided by a spatial layout, and grounded observations of door labels for efficient navigation. In this abstract map exploring floor plan is efficient but it is lacking the capability of understanding positions of object based on voice description with uncertain informations. However, above proposed methods are lacking ability of interpreting uncertain information in language instructions related to the spatial descriptions. Moreover, the proposed systems are not effective when uncertain information is included. As an example, the methods are not capable of effectively extracting the language phrase such as, “table is far away from the refrigerator”. Furthermore, the system is not capable of creating a map of the environment solely from the information conveyed from the descriptive instructions of the user without fusing them with the sensory input of the robot used in order to perceive the environment (such as laser scanner).

Therefore, this paper proposes a novel method that can be utilized for a service robot to create virtual maps about previously unknown environments based on the descriptive language instructions with uncertain terms given by the users. These virtual maps would eventually contribute for efficient navigation of the robot. The overall functionality of the proposed system is explained in Section 2. The proposed concept to understand the description and uncertain information are explained in Section 3. Experimental results are presented and discussed in Section 4. Finally, the conclusion is presented in Section 5.

2 System overview

Overall functionality of the proposed system is shown in Figure 1. This system can understand a description that narrates an unknown environment for robot given by the user and generate virtual maps using the data taken by the analysis of the description. The description may include phrases with uncertain terms that describe the position of objects such as “There is a chair in the far left corner of the room”. Voice description is converted into a

text and sent to the Instruction Identifier for analysis by voice recognition module. The Instruction Identifier understands the description and recognizes the sentences with the aid of the language memory which includes most commonly used linguistic data such as objects, uncertain terms and articles. Initially, the description is tokenized into sentences and then the extracted key information of each sentence is sent to the Instruction Processor (IP) one by one. For example, the above-mentioned example sentence will be sent to the IP as “is + chair +far left corner + room”.

The IP coordinates with three other modules namely, Uncertain Term Identifier (UTI), Virtual Spatial Data Identifier (VSDI) and Virtual Map Creator (VMC) in order to realize the creation of the virtual map. The UTI interprets the uncertain information in a sentence using data in the both VSDI and the VMC. The information interpreted by UTI is sent to the VMC. The VSDI processes the data gathered from UTI and map creator to interpret the relative position of an object. For every time a new object is encountered, IP creates a new ID (unique number linked with objects) in object memory so that every object can be uniquely accessible. Finally the IP amalgamates all the information related to the position of a previously unknown object and generates the virtual position of the object.

ViSMaLk (Virtual and Spatial Maps Linking) module is responsible for the linking of spatial map with virtual map. Spatial map is created by Map Data Processor (MDP) based on the sensory inputs received through the laser scanner. Process flow inside the MDP is shown in Figure 2.

3 Evaluation of description

As described in system overview Instruction identifier disassembles the description into sentences and filter out non-necessity terms. Then filtered sentences are analysed by the IP. human studies have been carried out for sentence categorisation and analysis of sentence structure.

3.1 Human study I

A human study was carried out with 25 participants. However, the participants were non-native English speakers. The participants were in the ages of 20 to 25 and fluent in English language. They have been asked to create a description for a given environment setup with 15 objects such as television, table, chair and fan. The written description should be comprehensible. Furthermore, written descriptions should provide a keen understanding about

the object setup for those who are unaware of the particular environment set up.

The results of the above experiment were analyzed and the sentence patterns were identified. Through that a model was created to extract the information from a description.

3.2 Description

A description is intermingled by sentences. As identified by human study I typically the first sentence determines about the focused environment. The rest of the sentences is furnished with the objects placed in the room such as tables, cupboards and chairs. In this scenario the supreme objective is to predefine the location of the objects related to the environment. As mentioned above the location is elucidated through the aid of objects. Exact values cannot be used in the initiation of virtual conceptual maps. Furthermore the sentence patterns used in describing the spatial locations are differentiated from the above. Description can be represented as

$$Description = \{s_1, s_2, s_3, \dots, s_i, \dots, s_n\}$$

s_i is the i^{th} sentence of the description. All the algorithms are developed based on following assumption which are made based on human study I.

- s_1 of the description describes the containing environment such as “This is a square shaped room”.
- s_i ($i \neq 1$) of the description is consisted with necessary information which includes a reference point or a reference object, uncertain terms which link the distance or direction relations are often used.
- One sentence has only one new object. But more than 1 reference object may use to explain relative reference.
- Description only consists of previously known lexical symbol of (included in language memory or object memory) uncertain terms and objects.
- Every new sentence introduces only one new object.
- When reference object of a sentence is a common name such as chair, table, it is identified as immediate previous object.

3.3 Sentence patterns

Any description can be expressed as a collection of sentences. As mentioned in the system overview description it is filtered by the Instruction Identifier and send to the IP. Those filtered sentences are configured under four cat-

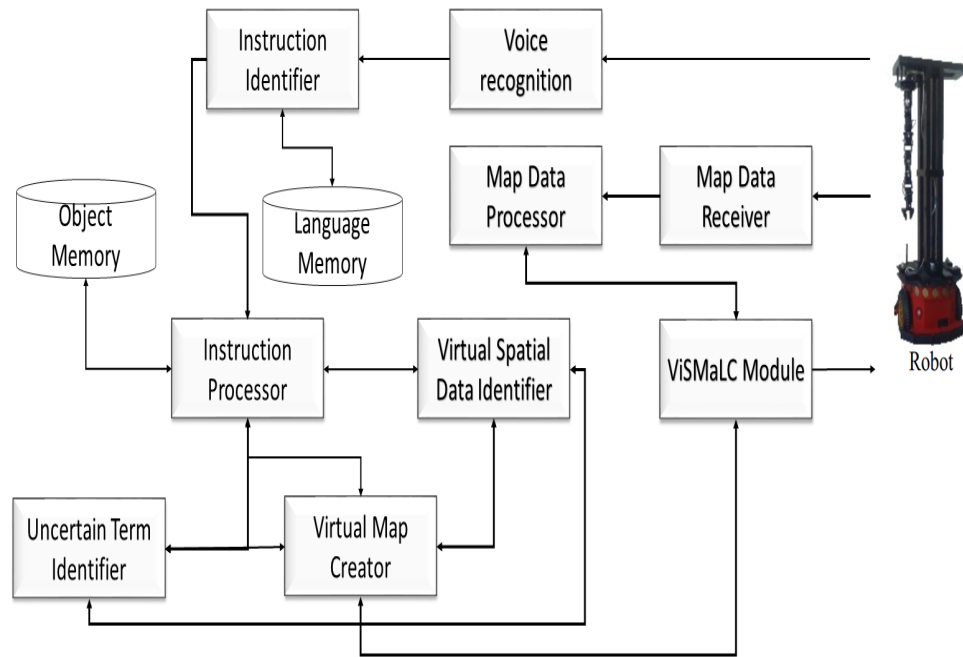


Figure 1: System overview.

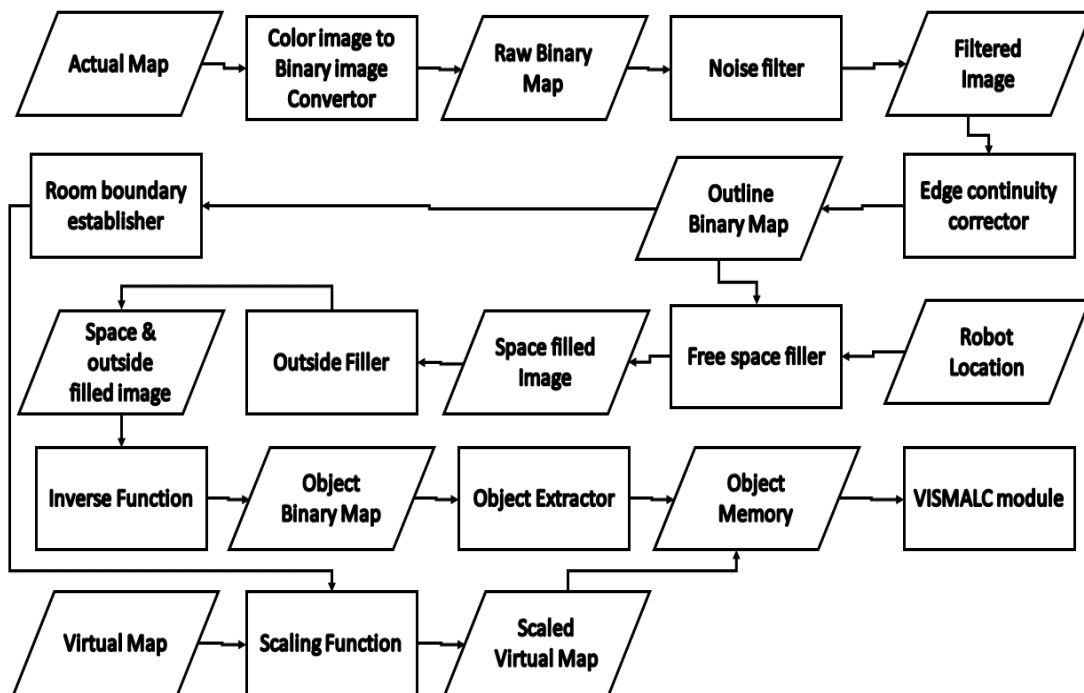


Figure 2: Process flow of a spatial map.

egories. After the evaluation of sentences IP identifies the new object described in the sentence.

$$Description = \{s_1, s_2, s_3, \dots, s_i, \dots, s_n\}$$

Therefore the s_i would be like

- Category 1 - This type of sentences is more often used in describing an environment. As an example "there is a chair in front of table". VSDI identifies "front + table" as reference object (ref. obj.) and use UTI to interpret "front". New object (new. obj.) is also identified by IP.

$$S_i \text{ category 1} = \langle \text{new obj.} \rangle \langle \text{ref. obj.} \rangle$$

- Category 2 - New objects are usually introduced relational to another object (ref. obj.) or position of environment combined with uncertain term such as "left corner", "right corner", "middle of left side". IP identifies those references as reference position (ref. pos.). Some sentences only have an uncertain term (e.g. there is a chair in the centre). For those sentences IP combines absolute position (room) as reference point by default.

$$S_i \text{ category 2} = \langle \text{new obj.} \rangle \langle \text{ref. pos.} \rangle$$

- Category 3 - There are also two procedures to express the same meaning as category 1 such as "near to the table there is a chair".

$$S_i \text{ category 3} = \langle \text{ref. obj.} \rangle \langle \text{new obj.} \rangle$$

- Category 4 - This category is used to explain relation of a new object to another two reference objects such as "There is a chair middle of the table and cupboard".

$$S_i \text{ category 4} = \langle \text{new obj.} \rangle \langle \text{ref. obj.} \rangle \langle \text{ref. obj.} \rangle$$

3.4 Uncertain terms

Uncertain terms are used in order to convey relative reference of objects with respect to the reference point or absolute position. In this design uncertain terms such as far, near, left, right, front, back and middle are used.

- Distance related

$$\langle UT_{Dis} \rangle = \langle \text{near} \mid \text{far} \mid \text{middle} \rangle$$

- Direction related

$$\langle UT_{Dir} \rangle = \langle \text{left} \mid \text{right} \mid \text{front} \mid \text{back} \rangle$$

- Reference point - Reference points are established in advance in an environment in order to express the position of objects in accordance to the known area of the environment. There are two reference points used in this system. Room is considered as an absolute position and all the other reference points are clarified based on the absolute position.

$$\langle \text{ref. point} \rangle = \langle \text{corner} \mid \text{centre} \rangle$$

- Reference position - Reference position is a position which is expressed relative to a reference point. Usually first few sentences have high probability to use reference position to mention position of an object. Reference position is a location which conveys about the reference point (eg; middle of the left side).

$$\langle \text{ref. pos.} \rangle = \langle UT_{dis. \text{ or } dir.} \rangle \langle \text{ref. point} \rangle$$

- Reference object - Reference object is used when user describes position of a new object relative to previously known object. As an example "There is a chair near to the table". In this sentence system identifies that the new object is "chair" and position of the chair is explained relative to the table. Here "table" is a previously known object. When previously known object combines with a uncertain term it will be considered as reference object (ref. obj.).

$$\langle \text{ref. obj.} \rangle = \langle UT_{dis. \text{ or } dir.} \rangle \langle \text{identified obj.} \rangle$$

3.5 Discussion - sentence patterns

1. These sentence patterns were derived from the human study for the purpose of creating a uncomplicated description model. The overall experiment scope is very broad. Therefore, the proposed model has no contribution over natural Language processing.
2. Important concern was to input a description and make it more manageable to the robot. Therefore the sentence patterns are not bounded by a strict grammar model.
3. A limited number of sentence patterns was selected. Therefore every sentence that may be used by human beings when they describe an environment, may not be included in defined patterns.
4. Human study was done in a domestic environment similar to Sri Lankan context created in a laboratory. Therefore square shaped rooms are considered. Model will be improved for other shapes in future works.

3.6 New object identification

Typical sentence in a description consisted with more than one object. Therefore it's important to identify which one is the new object before the analysis is done for the position. As an example typical sentence look likes “there is a table in front of chair”. There are two objects in this sentence. Among those objects the position of at least one object should be recognized. Since “chair” and “table” are common object names it's difficult to differentiate a new object from a reference object. To do that Spatial Category Interpretation algorithm has been developed. Example output of IP to VSDI is shown in Table 1. In order to identify the new object, IP searches entire sentence for uncertain terms. When the IP identifies the next object after the uncertain term, that object is named as the reference object (ref. obj.) with collaboration of object memory and sent it to the VSDI to interpret spatial information. Uncertain terms can represent directional uncertainty or distance uncertainty. UTI recognizes the category of identified term and use suitable algorithm to interpret spatial data.

3.7 Position identification

To identify the position of an object it is required to have spatial information about object. Spatial Category Interpretation algorithm has been developed to distinguish spatial data from a filtered sentence. As an example if the sentence is “There is a chair near to the table” instruction identifier module filters non-necessary articles and sends it to the instruction processing module and that sentence should be “chair + near + table”. After that the instruction processing module understands new object which is “chair” using object memory and Spatial Category Interpretation algorithm. Then the uncertain term identifier identifies the interpretation of uncertain term and virtual spatial data identifier generates virtual position of the new object considering positions of the other identified objects information collected from map creator. Finally the Instruction Processing Module collects calculated data from virtual spatial data identifier module and object memory to update map creator.

3.8 Human study II

Another human study has been conducted in relation to a human study to investigate the human behavior in the procedure of virtual placement of objects according to the

Input: Filtered S_i

Output: new_obj., UT_{dis} , ref. obj.

while New sentence found **do**

Search for the uncertain terms;

if number of uncertain terms < 2 **then**

Search for reference point;

if Found **then**

if UT_{dis} = centre **then**

Add absolute position next to
 UT_{dis} .

else

new_obj. = first object;
ref. point = ref. point from input S_i ;
return new_obj., UT_{dis} , ref. point;

end

else

Search for position related uncertain term;

if Found **then**

if UT_{dis} =

middle and no. of obj. == 3 **then**

new_obj. = first obj. of input;
textbf{return new_obj, UT_{dis} , ref.
obj. 1, ref. obj. 2;

else

ref. obj. = index of obj. next to
 UT_{dis} ;

end

if ref. obj. has common name
then

ref. obj. = previous obj. with
similar name;
new_obj. = remaining obj.;

else

go to next sentence

end

return new_obj, UT_{dis} , ref. obj.;

else

Go to next term

end

end

else

new obj = first obj. of sentence;
ref obj = final obj. of sentence;
return new_obj., UT_{dis} , ref. obj.;

end

end

Algorithm 1: Spatial category interpretation algorithm.

Table 1: Example sentences and outputs.

Sentence	Output
There is a chair near to the table	Ref. obj. = "table"
	New obj. = "chair"
	Uncertain term = "near"
	Uncertainty = distance
There is a cupboard in the left corner	Ref. point = "corner"
	New obj. = "cupboard"
	Uncertain term = "left"
	Uncertainty = direction
There is a chair in front of cupboard	Ref. obj. = "chair"
	New obj. = "cupboard"
	Uncertain term = "in front of"
	Uncertainty = direction

uncertain terms such as near, middle and far. A sketch of a picture as shown in Figure 3 was provided to 20 participants and the participants were requested to sketch the points according to the below mentioned sentences.

- “Place a chair near to the table”.
- “Place a chair middle of table and cupboard”.
- “Place a chair far from the table”.

According to the gathered data near, far and middle ranges are defined on the virtual scale based on coefficients δ_1 and δ_2 . δ_1 represents X axis and δ_2 represents Y axis in virtual map. Virtual position of object can be represented using δ_1 and δ_2 as shown in equations (1) and (2).

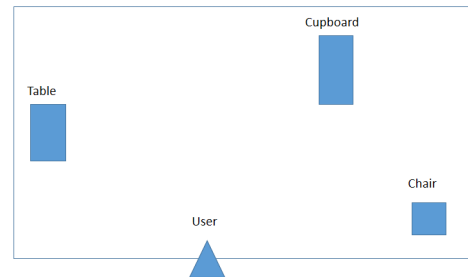
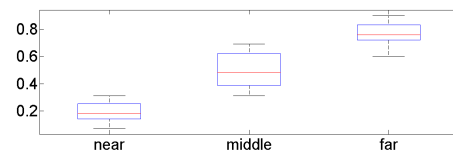
$$D = \{\delta_1, \delta_2\} \quad (1)$$

α and β are chosen as the 25th and 75th percentile of the distance range formulated by analysing dimensions of positioning new object by the participants. The chosen values for α and β given in Table 2. Experiment results are shown in Figure 4 as box-plots of better visualization of the data set.

$$0 < \delta_1, \delta_2 < 1 \quad (2)$$

Table 2: α and β values.

UT	α	β
near	0.14	0.25
middle	0.39	0.62
far	0.72	0.83

**Figure 3:** The map of the environment designed for human study explained in Section 3.1.**Figure 4:** Results of the human study explained in Section 3.1. used to find values for α and β .

3.9 Evaluation of uncertainty of position

In order to virtually locate an object in a conceptual map there are certain obstacles that have to be overcome before the interpretation of spatial information in a sentence into spatial data.

- When describing an environment, directions are depending on point of view of the describer. In this system VSDI considers all the directional information which are defined along the point of view of the robot.
- Objects that have fixed orientation such as sink, TV, cupboards have fixed front and back sides and objects such as tables desks front and back sides are depend

on user point of view. In here direction uncertain terms are identified by VSDI using direction module implemented with fuzzy direction inference system similar to [21].

Distance Uncertainty interpretation Algorithm has been developed to interpret spatial information in distance related uncertain term. VSDI recognizes the UTI and generates position for new object from the range calculated by Distance Uncertainty interpretation Algorithm. Output of the Distance Uncertainty interpretation Algorithm has random values based on the uncertain information. Therefore, output for the same uncertain term may different for different executions of the algorithm. Virtual Map Scaling Algorithm has been developed in order to scale virtual map into actual map and "Virtual and Spatial Maps Linking Algorithm" has been developed to establish links between objects in actual map and virtual map.

3.9.1 Distance Uncertainty interpretation Algorithm

Result: X_0, Y_0

n = number of identified objects;

$v = \{\text{Min}, \text{Max}\}$;

if $n \neq 0$ **then**

if $UI == \text{"middle"}$ **then**

$d_i = \{\text{distance between ref. obj.}\} / 2$

else

$d_i = \{v\{d_1, d_2, d_3, \dots, d_n\}\}$;

end

else

$d_i = \{\text{distance between ref. point and user}\} / 2$;

end

$D_1 = \alpha d_i$;

$D_1 = \alpha d_i$;

$D_2 = \beta d_i$;

$X_1 = x_{D1}$;

$X_2 = x_{D2}$;

$Y_1 = y_{D1}$;

$Y_1 = y_{D2}$;

$X_1 < X_0 < X_2$ and $Y_1 < Y_0 < Y_2$;

Algorithm 2: Distance uncertainty interpretation.

Result: Object Map

P - Pixels in binary image;

Avg_left_boundary ;

$_points = 0$;

for y in $0, \text{bitmap_height} - 1$ **do**

for condition do

if $P_{xy} == \text{Filled}$ **then**

if $_points == 0$ **then**

$_left_boundary = x$;

$_points = 1$;

else

Avg_left_boundary

$= (\text{Avg_left_boundary} \times _points + x) \div (_points + 1)$;

$_points = _points + 1$;

end

else

 skip;

end

end

end

Algorithm 3: Virtual map scaling algorithm.

Result: Object Map

$C = C_n \in$ Centroids of the objects in actual map;

Q = Copy of C ;

$V = V_m \in$ Points of the scaled virtual map;

$\text{Min_Distance} = \text{Virtual_to_actual_tolerance}$;

$C_{n_0} = \text{empty}$;

for V_m in V **do**

for C_n in Q **do**

$\text{Diff} = ((Q_n.x - V_m.x)^2 + (Q_n.y - V_m.y)^2)^{0.5}$;

if $\text{Diff} < \text{Min_Distance}$ **then**

 Link V_m to C_{n_0} ;

 Remove C_{n_0} from Q // Remove the linked one from the list;

else

V_m Link = empty // unable to find a object in actual map;

end

end

end

Algorithm 4: Linking virtual and spatial maps.



Figure 5: Explaining descriptions to the robot.

4 Results and Discussion

4.1 Experimental setup

The proposed system has been implemented on MIRob [17] platform. MIRob during experiments session is shown in Figure 5. First the system was updated with object memory and language memory based on typical environment. Two descriptions were made about same environment setup from different order of explaining structure. Experiment was carried out with 10 participants.

4.2 Experimental methodology

The descriptions given below were given to the participants and asked them to sketch plan of visualized map in their mind in a paper. It has been made sure that described environment was not familiar with any participant. After same descriptions were given to the robot and compared the output data with result of the human study.

- 1 “This is a square shaped room. In the left corner there is a table . There is a cupboard near to the table. In the far right corner there is a sofa. There is a table fan in the left side of the sofa. There is a photocopy machine in the centre of the room. In the far left corner there is a paper rack”.
- 2 “This is a square shaped room. In the far right corner there is a table In the centre of the room there is a conference table. There is a stool near to the table. There is a chair near to the conference table. In the left corner there is a filter. There is a shoe rack near to the filter”

4.3 Experimental results

4.3.1 Description 1

Description 1 was given to the system and output data was gathered in several steps during the process such as after being processed by Instruction Identifier, VSDI and IP. gathered data are shown in Tables 3 and 4.

Table 3 shows the sentences received by IP for further processing. As shown here objects and other necessary words should be included in object memory in order to identified by IP. After this each sentence was sent to the VSDI and UTI for position recognition. In Table 4 all the data input into VSDI by IP and UTI are given. In system implementation unique ID for each object is used.

In Table 7 all the final positioning data output by IP to the VMC are shown.

Table 3: Output of instruction identifier.

<i>i</i>	Filtered sentence
1	"left" + "corner" + "table"
2	"cupboard" + "near" + "table"
3	"far" + "right" + "corner" + "sofa"
4	"table fan" + "left" + "sofa"
5	"photocopy machine" + "centre"
6	"far" + " right" + " corner" + "paper rack"

Table 4: Input of VSDI.

<i>i</i>	New object	ID	UI_{dis}	UT_{dir}	Ref. point or obj.
1	table	1	-	left	corner
2	cupboard	2	near	-	1
3	sofa	3	far	right	corner
4	table fan	4	-	left	3
5	photocopy machine	5	-	-	centre
6	paper rack	6	far	right	corner

Table 5: Output of IP for VMC.

ID	Object	Position
1	table	0.16, 0.16
2	cupboard	0.18, 0.34
3	sofa	0.83, 0.83
4	table fan	0.77, 0.83
5	photocopy machine	0.5, 0.5
6	paper rack	0.16, 0.83

Table 6: Output of system and human study.

ID	Object	System output
1	table	0.16,0.16
2	cupboard	0.18,0.34
3	sofa	0.83,0.83
4	table fan	0.77,0.83
5	photocopy machine	0.5,0.5
6	paper rack	0.16,0.83

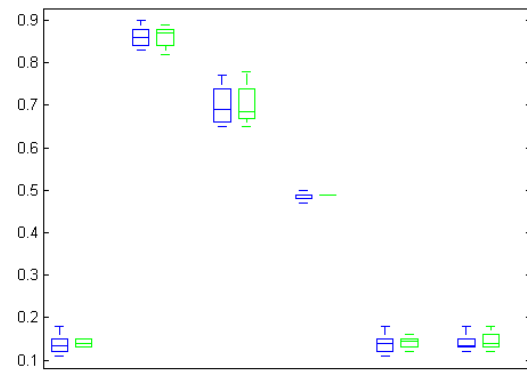
4.3.2 Comparison between the result of system and human study - Description 1

In consideration to the collected data through the system output and through the derivations of the human study the arithmetic mean and the standard deviation of “x” and “y” variables has been measured. The summarized values are included in Table 7 relevant to Description 1. The variations of X and Y coordinates of each object are given in Figures 6(a) and 6(b) as box plot for better visualization of the data. Furthermore these data have been subjected to a t-test which calculates the statistical significance related to the average values obtained from the human study and the system. Accordingly (p) value obtained through the t-test results is more than 0.05 for X and Y coordinates of the all the objects included in the description. Moreover, the difference between the human study and the output obtained through the system is not statistically significant. Furthermore, the locations of objects obtained from the human study and the robot are shown in Figure 7(a) and example map drawn by participant are shown in Figure 7(b).

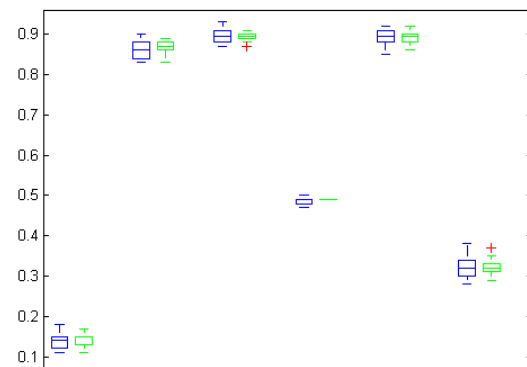
4.3.3 Description 2

Description 2 was also given to the system and output data was gathered in several steps during the process such as after being processed by Instruction Identifier, VSDI and IP. Output data of Instruction Identifier and input data of VSDI are shown in Tables 8 and 9. Procedure and contribution sequence of the system units are same as above explanation in Description 1.

Results calculated from human study and sketch plans of visual maps created by participants are also shown in Table 10. As mentioned before 10 participants were participated in experiment. Data shown in Table 10



(a)



(b)

Figure 6: (a) Boxplot for X axis data - Description 1, (b) Boxplot for Y axis data - Description 1.

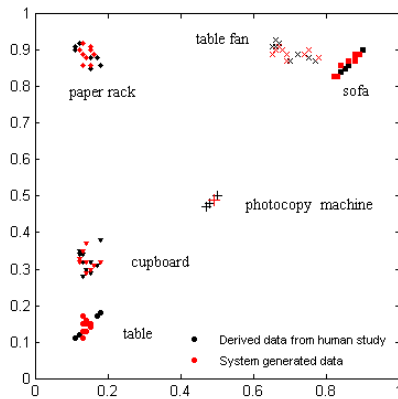
are calculated by taking position data of sketches of 10 participants.

4.3.4 Comparison between the result of system and human study - Description 2

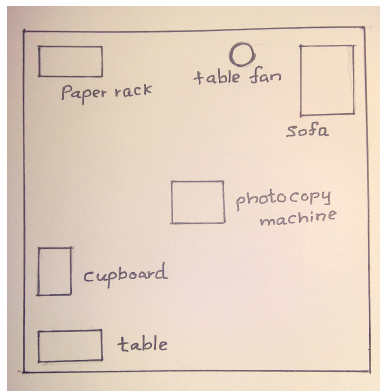
In consideration to the collected data through the system output and through the derivations of the human study the arithmetic mean and the standard deviation of “x” and “y” variables have been measured. The summarized values are included in Table 10 relevant to Description 1. The variations of X and Y coordinates of each object are given in Figures 9(a) and 9(b) as box plot for better visualization of the data. Furthermore these data have been subjected to a t test which calculates the statistical significance related

Table 7: Output of system and human study - Description 1.

ID	Object	System output				Derived from human study			
		X		Y		X		Y	
		mean	S.D	mean	S.D	mean	S.D	mean	S.D
1	table	0.139	0.008	0.14	0.016	0.139	0.022	0.144	0.022
2	cupboard	0.142	0.018	0.323	0.022	0.322	0.0193	0.322	0.028
3	sofa	0.864	0.024	0.864	0.020	0.866	0.023	0.866	0.023
4	table fan	0.7	0.040	0.897	0.011	0.893	0.040	0.893	0.019
5	photocopy machine	0.488	1.110E-16	0.488	1.110E-16	0.49	0.008	0.49	0.008
6	paper rack	0.14	0.012	0.892	0.018	0.891	0.022	0.891	0.023



(a)



(b)

Figure 7: (a) The locations of objects obtained from the human study and the robot - Description 1, (b) A map drawn by a participant for Description 1.**Table 8:** Output of instruction identifier - Description 2.

<i>i</i>	Filtered sentence
1	"far"+"right"+"corner"+"table"
2	"centre"+"conference table"
3	"stool"+"near"+"table"
4	"chair"+"near"+"conference table"
5	"left"+"corner"+"filter"
6	"shoe rack"+"near"+"filter"

Table 9: Input of VSDI - Description 2.

<i>i</i>	New object	ID	UT	UT	Ref. point of obj.
1	table	1	far	right	corner
2	conference table	2	-	-	centre
3	stool	3	near	-	1
4	chair	4	near	-	2
5	filter	5		left	corner
6	shoe rack	6	near	-	5

to the average values obtained from the human study and the system. Accordingly p value obtained through the t-test results is more than 0.05 for X and Y coordinates of all the objects included in the description. Moreover, the difference between the human study and the output obtained through the system is not statistically significant. Furthermore the locations of objects obtained from the human study and the robot are shown in Figure 10(a) and example map drawn by participant are shown in Figure 10(b).

Spatial maps for Description 1 and Description 2 are shown in Figures 8 and 11. Map preprocessing algorithms and methods have been explained in [22].

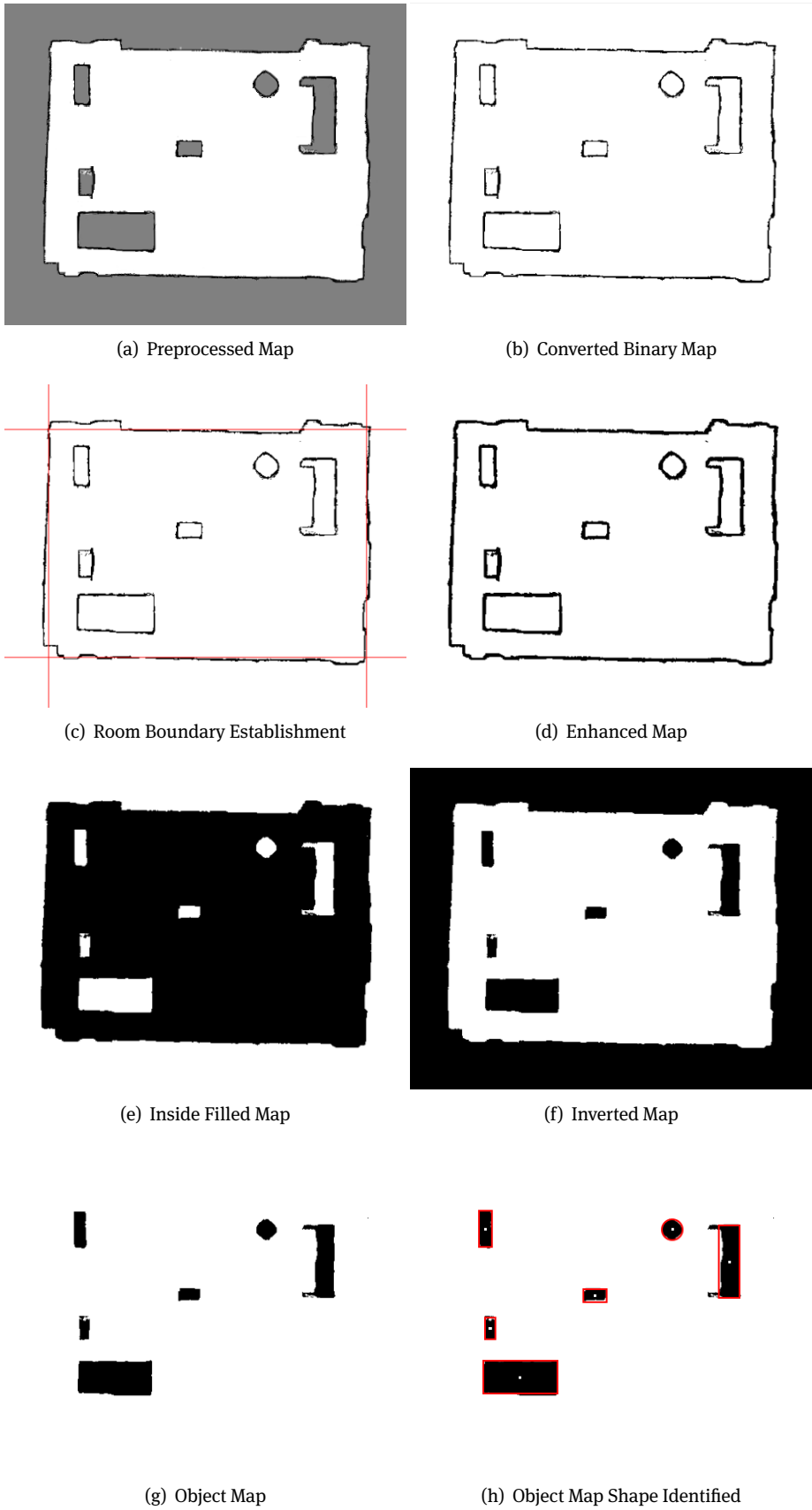
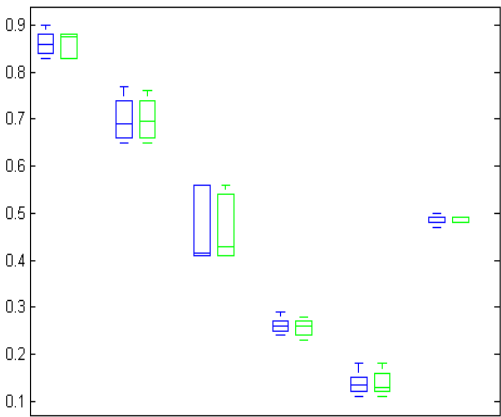


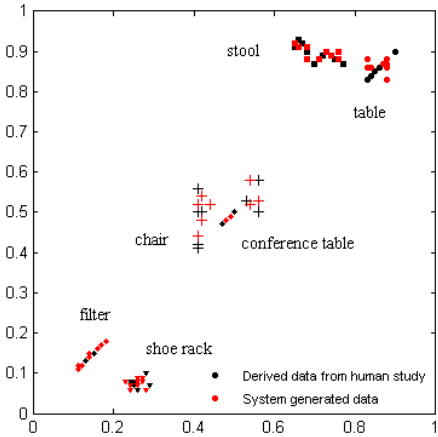
Figure 8: Spatial map processing steps for Description 1.

Table 10: Output of system and human study - Description 2.

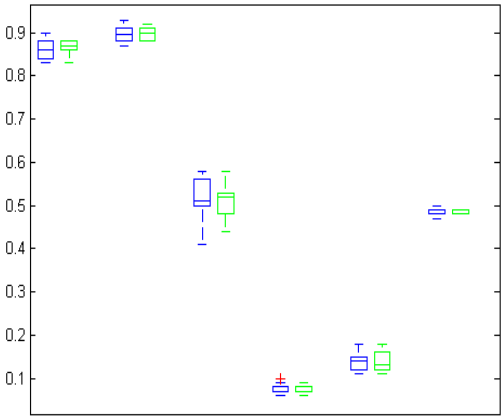
ID	Object	System output				Derived from human study			
		X		Y		X		Y	
		mean	S.D	mean	S.D	mean	S.D	mean	S.D
1	table	0.86	0.022	0.867	0.014	0.864	0.023	0.864	0.023
2	conference table	0.487	0.004	0.487	0.004	0.488	0.008	0.488	0.008
3	stool	0.703	0.040	0.898	0.014	0.7	0.040	0.897	0.019
4	chair	0.469	0.062	0.509	0.041	0.468	0.069	0.51	0.055
5	filter	0.137	0.024	0.139	0.023	0.139	0.022	0.14	0.022
6	shoe rack	0.258	0.015	0.076	0.010	0.26	0.015	0.078	0.010



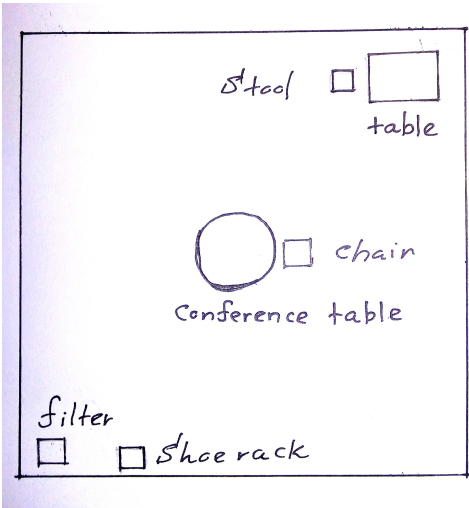
(a)



(a)



(b)



(b)

Figure 9: (a) Boxplot for X axis data - Description 2, (b) Boxplot for Y axis data - Description 2.

Figure 10: (a) The locations of objects obtained from the human study and the robot - Description 2, (b) A map drawn by a participant for Description 2.

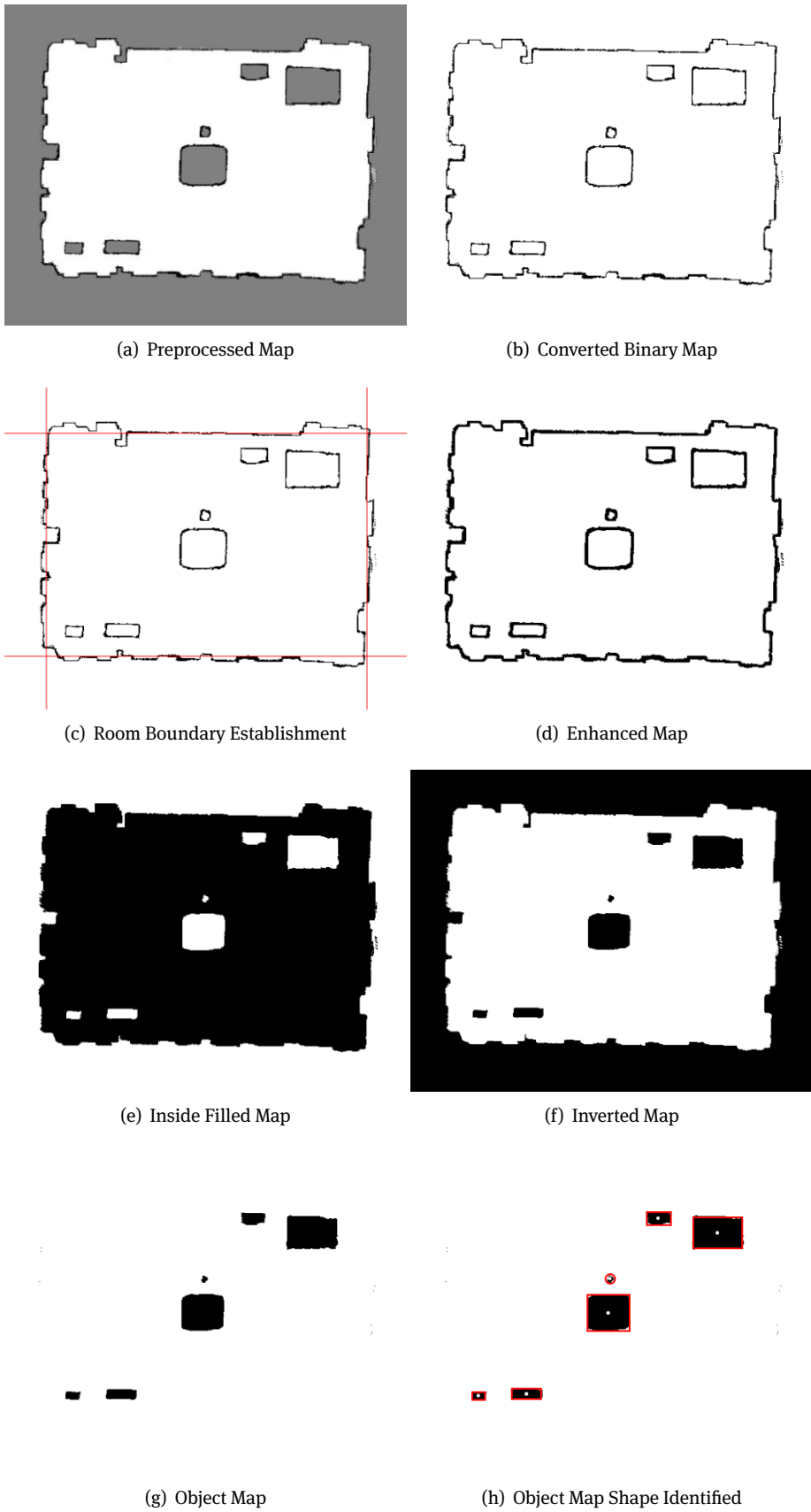


Figure 11: Spatial map processing steps for Description 2.

5 Conclusion

A method has been introduced to identify spatial information of given description with uncertain terms. This makes user to work with the robot in a more comfortable way. It will be very useful in navigation task when robot doesn't process awareness of the environment. Natural voice instruction can be used to make sure that robot can acknowledge about object in a unknown environment.

Identification of natural language phrases has been improved with Instruction Identifier and object memory. IP with VSDI was introduced to understand the relation between object references in virtual environment. UI has been improved to identify uncertain information in description. Novel method to quantify uncertain distance terms in virtual map has been developed.

The proposed system is capable of creating a virtual map or imagine a representative map of an unknown environment before the robot actually perceive visual sensory information from the environment.

References

- [1] World population ageing 2015, Population Division, Department of Economic and Social Affairs, United Nations, ST/ESA/SER.A/390, 2015
- [2] T. Heller, H. M. Gibbons, D. Fisher, Caregiving and family support interventions: Crossing networks of aging and developmental disabilities, *Intellectual and Developmental Disabilities*, 2015, 53(5), 329–345, DOI: 10.1352/1934-9556-53.5.329
- [3] U. Schneider, J. Kleindienst, Monetising the provision of informal long-term care by elderly people: estimates for european out-of-home caregivers based on the well-being valuation method, *Health & Social Care in the Community*, 2016, 24(5), e81-91, DOI: 10.1111/hsc.12250
- [4] A. Squires, et al., A systematic survey instrument translation process for multi-country, comparative health workforce studies, *International Journal of Nursing Studies*, 2013, 50(2), 264–273, DOI: 10.1016/j.ijnurstu.2012.02.015
- [5] D. Feil-Seifer, M. J. Mataric, Defining socially assistive robotics, In: *Proceeding of the 9th International Conference on Rehabilitation Robotics (ICORR 2005)*, IEEE, 2005, 465–468
- [6] A. M. Okamura, M. J. Mataric, H. I. Christensen, Medical and health-care robotics, *IEEE Robotics & Automation Magazine*, 2010, 17(3), 26–37, DOI: 10.1109/MRA.2010.937861
- [7] D. Feil-Seifer, M. J. Mataric, Ethical Principles for Socially Assistive Robotics, *IEEE Robotics & Automation Magazine*, 2011, 18(1), 24–31, DOI: 10.1109/MRA.2010.940150
- [8] A. Tapus, M. J. Mataric, B. Scassellati, Socially assistive robotics [grand challenges of robotics], *IEEE Robotics & Automation Magazine*, 2007, 14(1), 35–42, DOI: 10.1109/MRA.2007.339605
- [9] A. Tapus, M. Maja, B. Scassellati, The grand challenges in socially assistive robotics, *IEEE Robotics and Automation Magazine*, 2007, 14(1)
- [10] V. Nguyen, C. Jayawardena, A technical review of motion prediction methods for indoor robot navigation, Technical Report, 2015
- [11] A. Sharkey, Robots and human dignity: a consideration of the effects of robot care on the dignity of older people, *Ethics and Information Technology*, 2014, 16(1), 63–75, DOI: 10.1007/s10676-014-9338-5
- [12] S. Frennert, B. Östlund, Review: Seven matters of concern of social robots and older people, *International Journal of Social Robotics*, 2014, 6(2), 299–310, DOI: 10.1007/s12369-013-0225-8
- [13] E. A. Sisbot, L. F. Marin-Urias, X. Broquere, D. Sidobre, R. Alami, Synthesizing robot motions adapted to human presence, *International Journal of Social Robotics*, 2010, 2(3), 329–343, DOI: 10.1007/s12369-010-0059-6
- [14] S. Hemachandra, F. Duvallet, T. M. Howard, N. Roy, A. Stentz, M. R. Walter, Learning models for following natural language directions in unknown environments, In: *2015 IEEE International Conference on Robotics and Automation (ICRA)*, IEEE, 2015, 5608–5615, DOI: 10.1109/ICRA.2015.7139984
- [15] C. Ware, *Information Visualization: Perception for Design*, Morgan Kaufmann, 2012
- [16] T. Wolbers, M. Hegarty, What determines our navigational abilities?, *Trends in Cognitive Sciences*, 2010, 14(3), 138–146, DOI: 10.1016/j.tics.2010.01.001
- [17] M. V. J. Muthugala, A. B. P. Jayasekara, Enhancing human-robot interaction by interpreting uncertain information in navigational commands based on experience and environment, In: *2016 IEEE International Conference on Robotics and Automation (ICRA)*, IEEE, 2016, 915–2921, DOI: 10.1109/ICRA.2016.7487456
- [18] H. Zender, O. M. Mozos, P. Jensfelt, G.-J. Kruijff, W. Burgard, Conceptual spatial representations for indoor mobile robots, *Robotics and Autonomous Systems*, 2008, 56(6), 493–502, DOI: 10.1016/j.robot.2008.03.007
- [19] H. M. R. T. Bandara, M. A. V. J. Muthugala, A. B. P. Jayasekara, D. Chandima, Cognitive spatial representative map for interactive conversational model of service robot, *2018 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, IEEE, 2018, 686–691, DOI: 10.1109/RO-MAN.2018.8525778
- [20] B. Talbot, O. Lam, R. Schulz, F. Dayoub, B. Upcroft, G. Wyeth, Find my office: Navigating real space from semantic descriptions, *2016 IEEE International Conference on Robotics and Automation (ICRA)*, IEEE, 2016, 5782–5787, DOI: 10.1109/ICRA.2016.7487802
- [21] M. A. V. J. Muthugala, A. G. B. P. Jayasekara, Interpreting uncertain information related to relative references for improved navigational command understanding of service

robots, 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2017, 6567–6574, DOI: 10.1109/IROS.2017.8206568

- [22] H. M. R. T. Bandara, B. M. S. S. Basnayake, A. G. B. P. Jayasekara, D. P. Chandima, Enhancing conceptual spatial map by amalgamating spatial and virtual cognitive maps for domestic service robots, 2018 2nd International Conference on Electrical Engineering (EECon), IEEE, 2018, 150–155, DOI: 10.1109/EECon.2018.8541010