University POLITEHNICA Bucharest

Faculty of Automatic Control and Computers

ZAVOD IZRIIS
Informacijski raziskovalni inštitut za izobraževanje in svetovanje

A PLATFORM THAT AIMS TO HELP PEOPLE TO LEARN HOW TO INTERACT WITH ROBOTIC PLATFORMS

A.C. Popescu[1], I.A. Awada[1], I. Mocanu[1], O. Cramariuc[2], N. Samar Brenčič[3]

[1] University Politehnica of Bucharest (ROMANIA)

[2] IT Center for Science and Technology (ROMANIA)

[3] M.A., Izriis Institute, Ljubljana (SLOVENIA)

EDULEARN21

Introduction

Context, motivation:

- Advances in robotics
- Large social group
- High healthcare costs

Source: Boston Dynamics

Solution:

- Human–robot interaction plays a crucial role in the booming market for intelligent personal-service.

Introduction (2)

- The number of older persons is growing faster than the numbers of people in any other age groups (a study of the Department of Economic and Social Affairs of the United Nations)
- It is estimated that by 2030, people over the age of 60 will represent more than 25% of Europe's population.
- Developing solutions for improving the quality of life of older adults is justifiable

Introduction (3)

- There are different systems that assist people in need
- Robotic platforms are being used as assistive or social robots
- Human-robot interaction is needed:
  - **Voice commands**
  - **Hand gestures**

Gesture Recognition

- Gesture interactions allow the user to interact with a robotic platform through hand and/or body gestures.
- We used only hand gestures.
- A simple CNN architecture that contains 7 hidden layers (2D convolutions followed by 2D max pooling operations) with ReLU as activation function, and dropout and one fully connected layer with softmax activation was used.
- The optimizer is Adam and the loss function is categorical cross-entropy.
- Size of the input images is 100×100 pixels representing shapes of the hand.

Gesture Recognition (2)

- Size of the input images is 100×100 pixels representing shapes of the hand



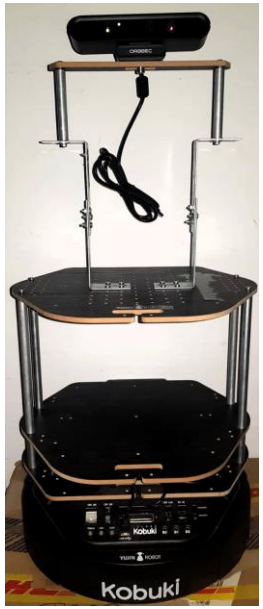

Voice Commands

Voice commands are analysed through the following components:

- Audio Preprocessing: ensures the best functionality of the voice module.
- Automatic Speech Recognition (ASR): converts the spoken words of the user into written text.
- Natural Language Understanding (NLU): converts the written text into a machine-reading representation.
- Dialog Management (DM): maintains the history of the conversation, decides the system answer and it is responsible of the state and flow of the conversation.
- Text-to-Speech (TTS) Synthesis: converts the text into artificial human speech.

Evaluation

- 30 elderly people have evaluated the human-robot interaction through gesture and voice commands on a Turtlebot robot and a TIAGo robot.



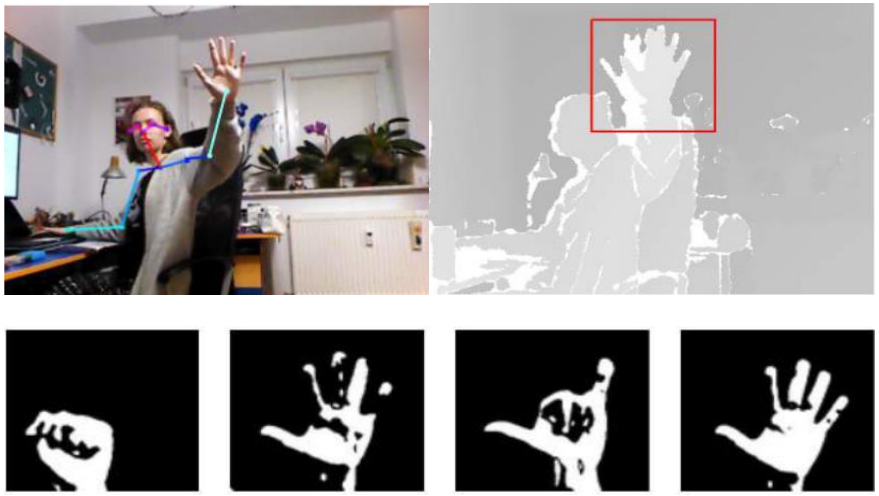
Kobuki Turtlebot with ORBBEC Astra Pro

Name	Dimensions	Autonomy	Computing	Cost
TIAGo	110-145 cm x 54 cm	9 hours	Intel Core i7 processor	25-59.000 $
Turtlebot	35.5 x3 5.4 cm	7 hours	None: an attach an external laptop/PC (e.g., Intel NUC)	780 $

TIAGo

Gesture Recognition Evaluation

- Hand area detection based on skeleton and depth information



95.4% accuracy on own dataset with 1000 samples per class.
Confusion matrix:

	fist	palm	swing	peace	index
fist	97				3
palm		99		1	
swing		1	96	2	1
peace	2	1		93	4
index			3	5	92

Voice Commands Evaluation

- Voice commands were implemented in English and Romanian languages.
- Only for the TIAGo robot the users evaluated the full set of commands while for the Turtlebot robot, the users evaluated only the navigation commands.
- In total 2160 interactions were collected:

Grouped by language
- 1080 in English: 900 for TIAGo, 180 for Turtlebot,
- 1080 in Romanian: 900 for TIAGo, 180 for Turtlebot.

Grouped by robot
- 1800 for TIAGo,
- 360 for Turtlebot.

Voice Commands Evaluation (2)

Language	User's Command	Recognition Percentage		
		TIAGO	Turtlebot	Average
English	Display my agenda.	100.00%	N/A	100.00%
	Navigate toward the fridge.	96.67%	90.00%	93.33%
	How will the weather be tomorrow?	96.67%	N/A	96.67%
	Go to the table.	100.00%	96.67%	98.33%
	Display my blood pressure.	100.00%	N/A	100.00%
Romanian	Afișează-mi agenda	100.00%	N/A	100.00%
	Du-te spre frigider	93.33%	83.33%	88.33%
	Cum va fi vremea mâine?	100.00%	N/A	100.00%
	Mergi la masă.	100.00%	93.33%	96.67%
	Afișează cât îmi este tensiunea.	100.00%	N/A	100.00%

A set of samples from the ASR evaluation results

Voice Commands Evaluation (3)

Language	User's Command	Intent	Entity(ies)	Output
English	Display my agenda.	get_calendar	calendar_entity: agenda output_entity: display	Displaying calendar. (Visual) N/A
	Navigate toward the fridge.	nav_to	target_entity: fridge	(Phonetic & Action) 100%
	How will the weather be tomorrow?	get_weather	dateandtime_entity: tomorrow location_entity: none (GPS_current_location) output_entity: none (default)	The weather will be sunny in Bucharest. (Visual & Phonetic - Default Option) 100%
	Go to the table.	nav_to	target_entity: table	I am going to the table. (Phonetic & Action) 100%
	Display my blood pressure.	get_health	health_entity: blood pressure output_entity: display	Your blood pressure is 134/81 mmHg. (Visual) N/A
Romanian	Afișează-mi agenda	get_calendar	calendar_entity: agenda output_entity: afișează-mi	Displaying calendar. (Visual) N/A
	Du-te spre frigider	nav_to	target_entity: frigider	Mă duc spre frigider. (Phonetic & Action) 100%
	Cum va fi vremea mâine?	get_weather	dateandtime_entity: tomorrow location_entity: none (GPS_current_location) output_entity: none (default)	Mâine, vremea va fi însorită în București. (Visual & Phonetic - Default Option) 98.33%
	Mergi la masă.	nav_to	target_entity: masă	Mă duc spre masă. (Phonetic & Action) 100%
	Afișează cât îmi este tensiunea.	get_health	health_entity: tensiunea output_entity: afișează	Tensiunea dumneavoastră este 134/81 mmHg. (Visual) N/A

A set of samples from the voice module evaluation results

Conclusions and Future Work

Conclusions:

- Two types of human-robot interaction: using gesture and voice commands.
- These interactions were tested on two robots with different capabilities: Turtlebot and TIAGo.
- Proven to be easily deployable on a robotic platform.

Future work:

- Specialize/improve both gesture and voice commands recognition.
- Enrich the set of gestures and voice commands.
- Perform tests with users of different age and gender.

Acknowledgement

This work was supported by a grant of the CCDI--UEFISCDI and of the AAL Programme with co-funding from the European Union's Horizon 2020 research and innovation programme project "**INCARE -- Integrated Solution for Innovative Elderly Care**", project number AAL-2017-059-INCARE.

http://www.aal-incare.eu/

Thank you!

irina.mocanu@upb.ro