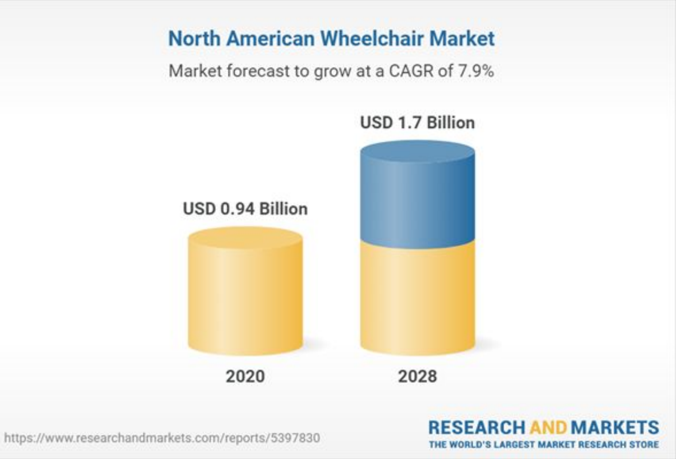
WeHelp: Wheelchair Helper

Team 3
Abulikemu Abuduweili
Alice Wu
Tianhao Wei
Weiye Zhao

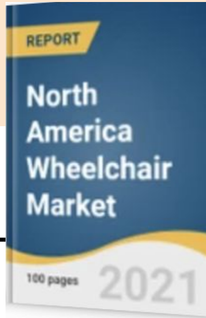
Problem: Wheelchair Users Need Help



In 2016, **3.3 million wheelchairs users,** 1.825 million are users aged 65+ in the U.S.

The North America wheelchair market size is expected to reach USD **1.7 billion** by 2028

Over **11 million people** reported a problem needed assistance with activities of daily living (ADLs)



Feedback from Stakeholder

- Our stakeholder: 70-years old, stroke, impair mobility, uses wheelchair on a daily basis

1) Needs assistance in small things (e.g. getting water, going to bathroom, taking shower);
2) Needs control by the wheelchair user;
3) Independence matters a lot

What we have achieved

1. <u>Human following</u>
   a. Robust human tracking
   b. Track from multiple positions
2. <u>Speech recognition</u>
   a. Recognize multiple commands
   b. Denoise the background
3. <u>Teleoperation</u>
   a. User takeover
   b. Robust pipeline (switch control)

Demo

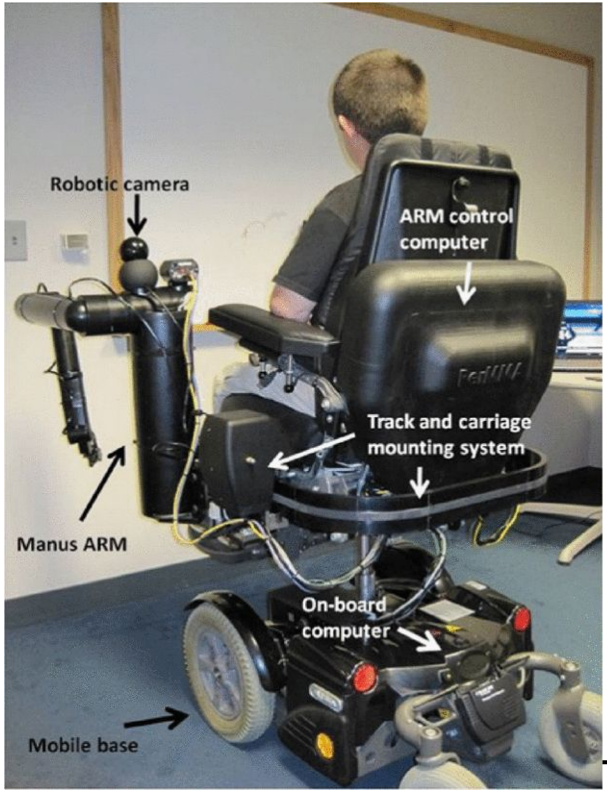
Team 3: WeHelp
Carnegie Mellon University
Contact
Alice Wu:
qichaow@andrew.cmu.edu
Tianhao Wei:
twei2@andrew.cmu.edu
Weiye Zhao:
weiyezha@andrew.cmu.edu
Abulikemu Abuduweili
abulikea@andrew.cmu.edu
drive

Feedback from Stakeholder

- Feedback based on the demo:

1) Take a long time to complete the task;
2) Difficult to control for people with arm lifting problem

State-of-the-art: robotic-caregiving system



PerMMA (Gen I): a personal assistive robot providing user and caregivers with enhanced manipulation (H. Wang *et al*., 2012).

Home Exploring Robot Butler, HERB

- Born in the robotics lab at the Carnegie Mellon University
- In 2010, Herb 2.0 was entirely redesigned: two arms and a head, custom electronics, cooling, and power to last up to **six hours**
- Can unload a dishwasher and even take apart an Oreo -> kitchen

PR2

Fetch drinks from a refrigerator - WILLOW GARAGE

- One of the most advanced research robots
- Became **available for purchase** in September 2010. (Cost: $400,000)
- Navigate autonomously and manipulate a wide range of objects:
  - Clean up tables, fold towels, and fetch drinks from the fridge

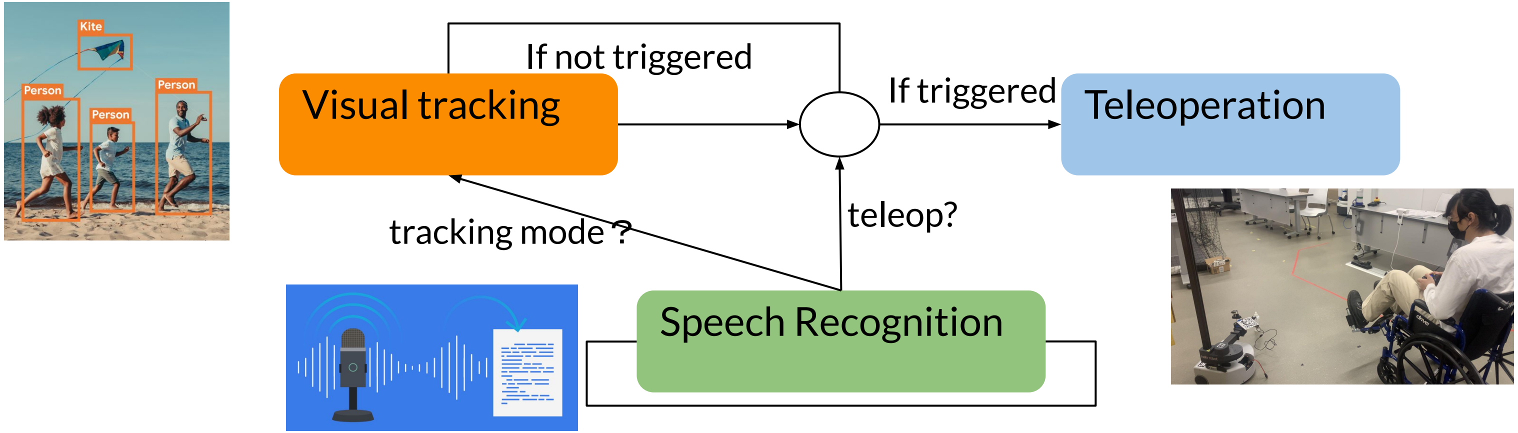
https://robots.ieee.org/robots/pr2/?gallery=video1

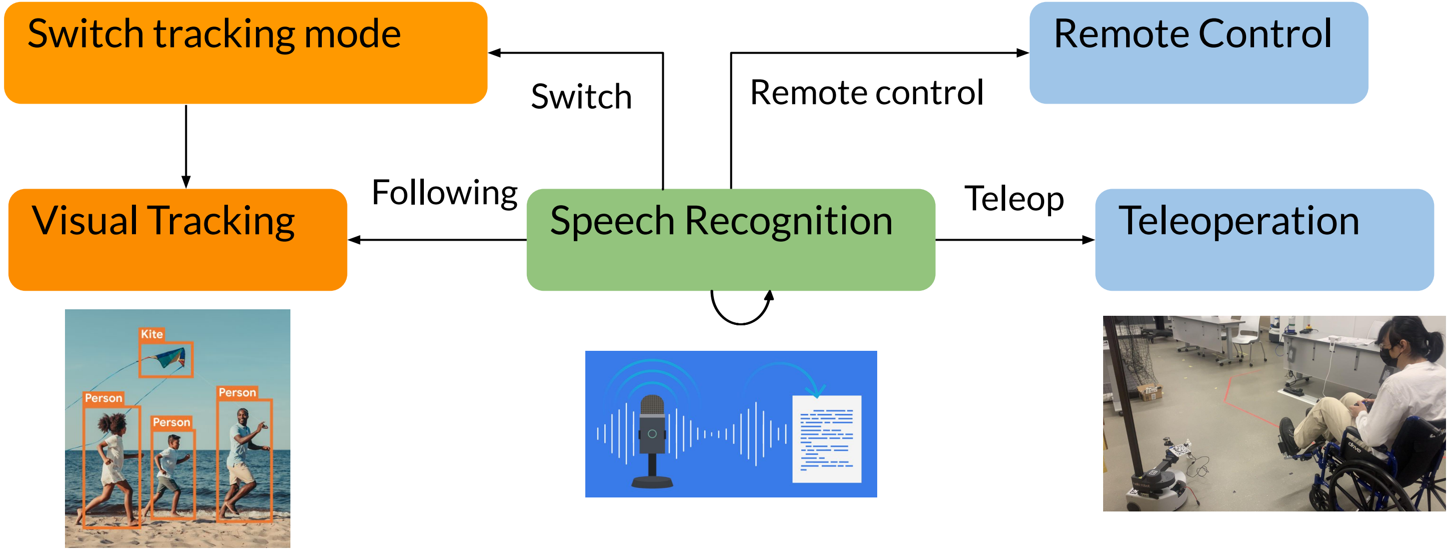
Toyota Human Support Robot (HSR)

The robot was able to fetch water and open doors for a quadriplegic in 2017

Task Decomposition and Methodology

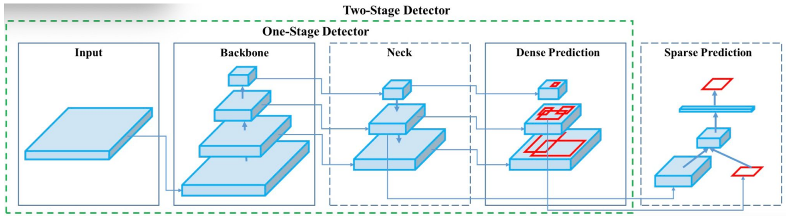
- The method is composed of 3 modules
  - Speech recognition: decide when the user needs remote help.
  - Visual tracking: follow the user at idle mode.
  - Teleoperation: the wheelchair user takes over the robot by a controller.



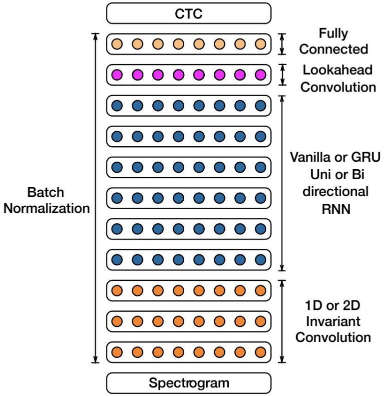

Switch tracking mode
Switch
Remote control
Remote Control
Visual Tracking
Following
Speech Recognition
Teleop
Teleoperation
Kite
Person
Person
Person

State-of-the-art: Methods

1. Visual Tracking
   a. Alexey Bochkovskiy, et al. "Yolov4: Optimal speed and accuracy of object detection." *arXiv preprint arXiv:2004.10934* (2020).
   b. Hao Zhang, et al. "DINO: DETR with Improved DeNoising Anchor Boxes for End-to-End Object Detection." *arXiv preprint arXiv:2203.03605* (2022).



2. Speech Recognition
   a. Dario Amodei, et al. "Deep speech 2: End-to-end speech recognition in english and mandarin." *ICML*, 2016. (DeepSpeech 0.9.3, 2020)
   b. Yu- An Chung, et al. "W2v-bert: Combining contrastive learning and masked language modeling for self-supervised speech pre-training." *arXiv preprint arXiv:2108.06209* (2021).

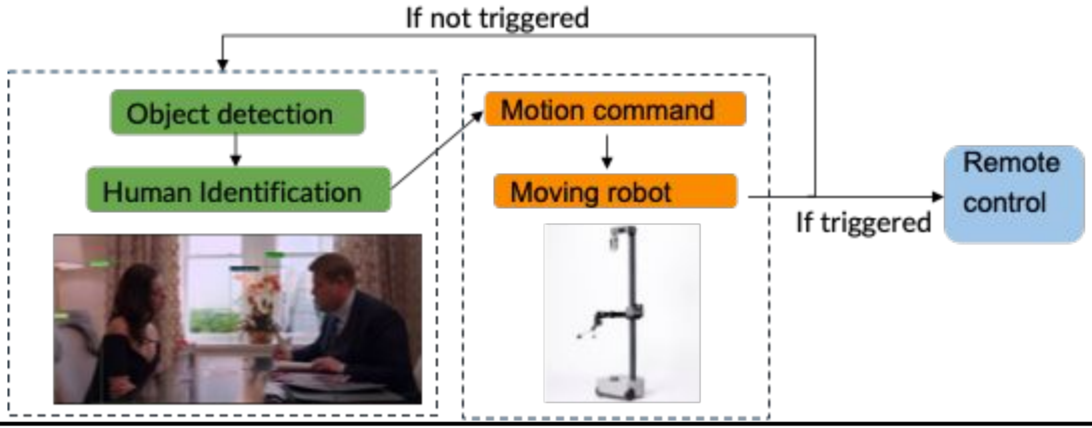


Human Following (Visual tracking)

Object detection: Based on YOLO-v4.

Human Following: The Robot follows the wheelchair user and moves itself to get the wheelchair user in the center of the frame.

1) Tracking from behind 2) Following from accompany mode (on left/ right)



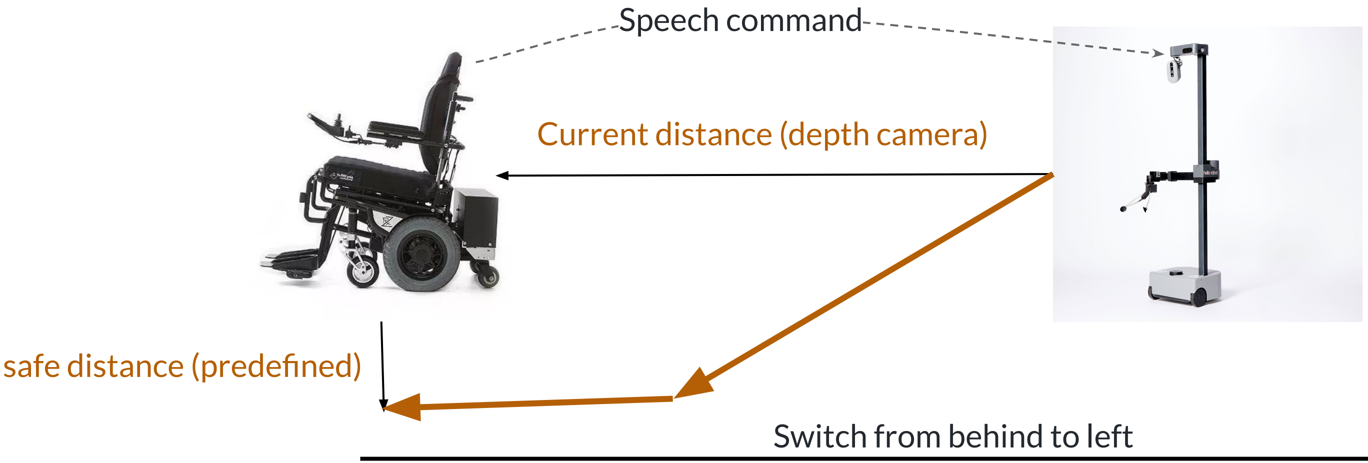
Rotation -> deviation (RGB camera)

Transition -> distance (depth camera)

Human Following (different tracking modes)

Allow the robot to follow in a more natural way.

Switch between three modes: 1) Following from behind 2) Following on right 3) Following on left.



Evaluation: Human Following

Human following: success rate (three following modes)

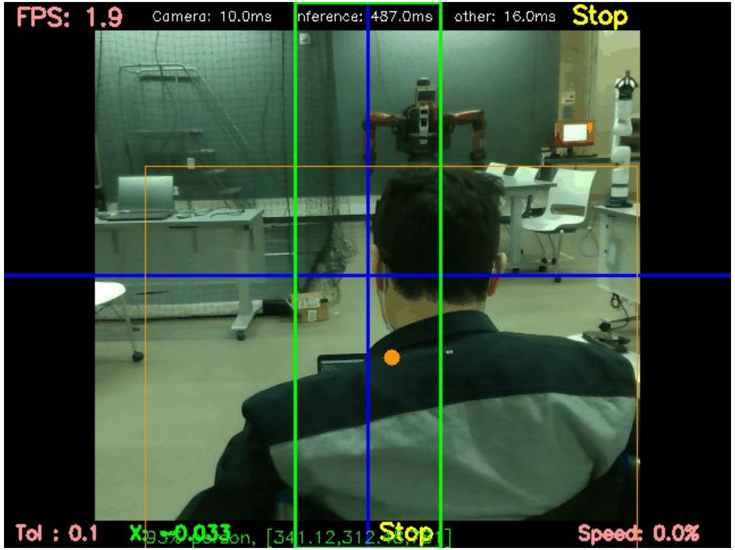
speed\mode	behind	on right	on left
0.1 m/s	100%	100%	100%
0.2 m/s	100%	60%	85%
0.3 m/s	100%	20%	30%
1.0 m/s	100%	0%	0%

- The success rate of human following relies on the speed of the wheelchair user. But more importantly, following from behind have much higher performance than following alongside. Because tracking from behind has 1) better view (camera vision), and 2) flexible moving (base motion).

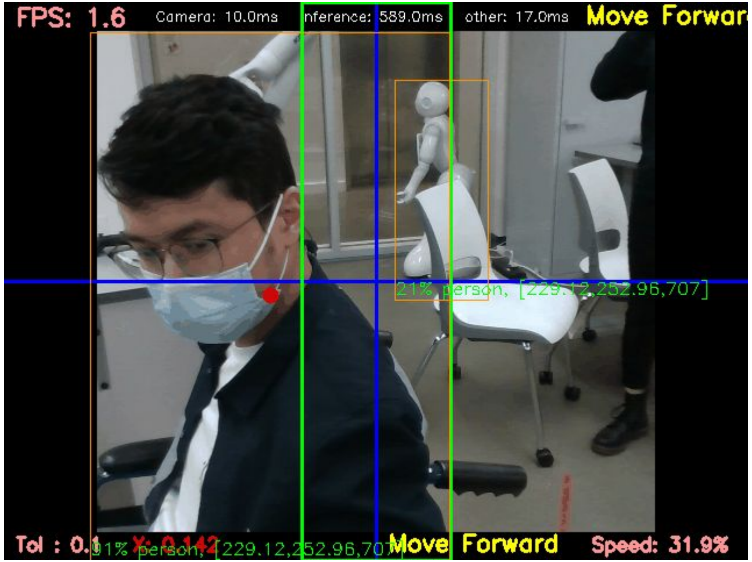
Human Following - Demo

Following from behind
then switch to left mode

Following from behind (robot's view)

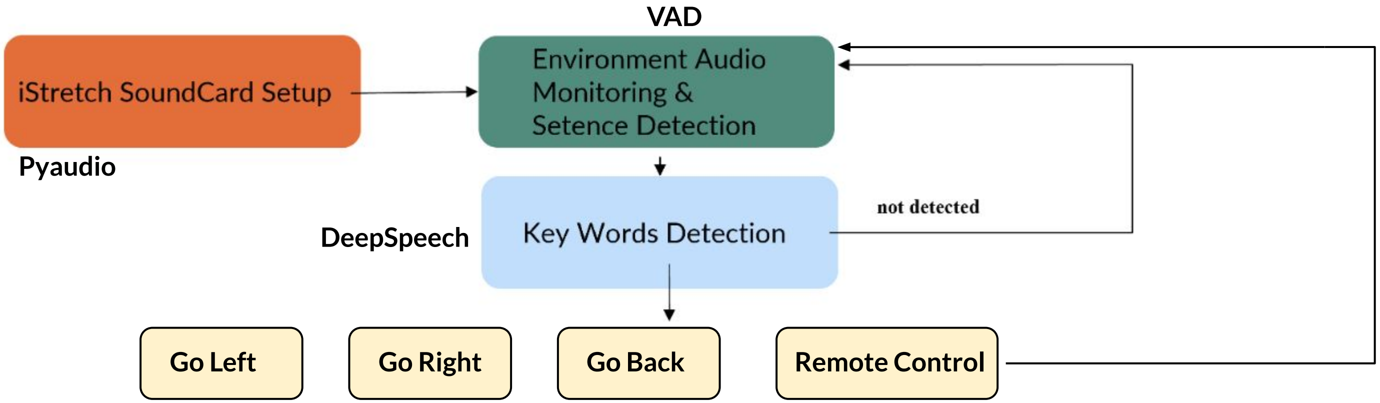


Following on left (robot's view)



Speech Recognition

- Pyaudio, which provide the Python bindings for PortAudio, the cross-platform audio I/O library. Device assignment and streaming binding
- Voice Activity Detection. 1) concatenate voiced frames, 2) complete sentence detection
- DeepSpeech for audio recognition and key words detection



Evaluation: Speech Recognition

Speech Recognition: success rate (4 subjects, each keywords 20 repetition)

Noise\keywords	“Left”	“Right”	“Back”	“Remote”	Average
50 db	57.5%	85%	85%	37.5%	66.2%
60 db	50%	61.3%	83.7%	30%	56.2%

50 db

Subjects	“Back”	“Left”	“Right”	“Remote”
Subject A	95%	5%	50%	15%
Subject B	95%	95%	100%	60%

- Recognition success rate heavily relies on accent of speaker
- On average, twice repetition will successfully trigger subsequent control

Teleoperation

We decided to let the user to control the robot instead of the remote assistant. There are several advantages:

- Independence
- Privacy
- Does not rely on internet connection
- Full view of the environment

Evaluation: Teleoperation

Daily tasks and consuming time.

Task	First timer	Fluent user
Retrieve objects	~60s	~40s
open doors	~90s	~60s

Next steps

- **Speech recognition**: bluetooth microphone, other possibility like BMI, gesture control etc.
- **Visual tracking**: combine with trajectory prediction, trajectory planning, collision avoidance, SLAM.
- **Teleoperation**: Inspect which UI is the most intuitive. Command memorization
- **Task design**: more complex tasks such as opening doors, wearing shoes, making phone call/zoom

Conclusions

The pipeline is generally robust and practical, it can benefit the wheelchair users in daily life by increasing their independence. But there are still some challenges:

- The robot moves too slow.
- The robot may lose track of the user due to occlusion.
- Environment noise affects the speech recognition if the microphone is too far.

Live Demo

References

J.M. McNeil. "Americans with disabilities: 2005, household economic studies, current population reports," pp. 70-117. 2008

H. Wang *et al.*, "The Personal Mobility and Manipulation Appliance (PerMMA): A robotic wheelchair with advanced mobility and manipulation," *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2012, pp. 3324-3327, doi: 10.1109/EMBC.2012.6346676.

Hannun, Awni, et al. "Deep speech: Scaling up end-to-end speech recognition." *arXiv preprint arXiv:1412.5567* (2014).

Redmon, Joseph, et al. "You only look once: Unified, real-time object detection." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.

Redmon, Joseph, and Ali Farhadi. "Yolov3: An incremental improvement." *arXiv preprint arXiv:1804.02767* (2018).