Preliminary Results on Deep Sensor Fusion for Endpoint Prediction in Transhumeral Prostheses

Eitan Gerber, Nili E. Krausz

Neuro-Robotics and Bionic Limbs (eNaBLe) Lab

Faculty of Mechanical Engineering, Technion

Introduction

- A key challenge in robotics is sim2real for robots working near/with people [1].
- Amputees are **unable** to coordinate multi-joint robotic prostheses, **amplifying the sim2real gap** by limiting available inputs.
- Previous work has shown correlation between gaze, sEMG and hand motion [2].
- We propose a **multi-modal sensor fusion model** as an endpoint predictor for a prosthetic arm controller.

Methods

- Gathered **healthy pilot data** and used it to design a proof-of-concept hybrid CNN/RNN model, and evaluated its prediction accuracy with **different sensor modalities**.
- Based on the preliminary results of the proof-of-concept predictor (Fig. 2), we chose the open-access Project Nymeria dataset [3].
- Trained both a **linear regressor (LR)** and a modified **UNet** using the Nymeria Data.
- Predictors: **Upper-arm coordinates**, **shoulder angles**, **gaze vector**, **gaze pitch&yaw angles**.
  - Including velocities and accelerations of arm, shoulder.
- Targets: **Hand coordinates.**
- Trained the UNet with a loss function the penalizes Euclidean distance and differences in velocity, and minimizes jerk.
- Compared the accuracy of the **LR** when using individual predictors and when using all predictors.

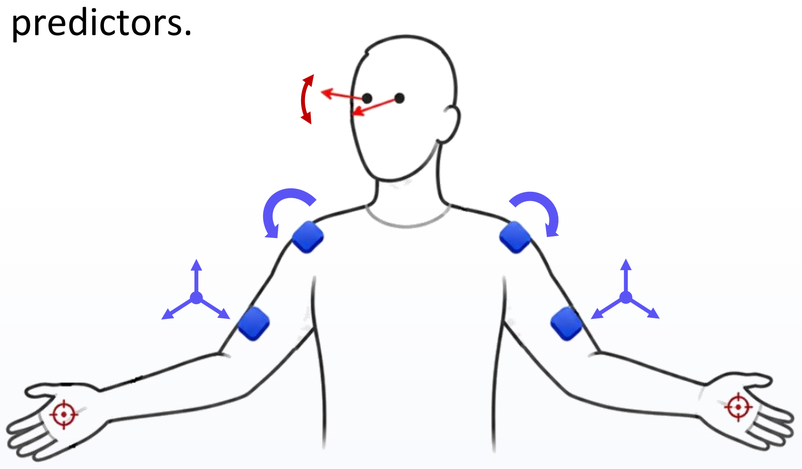


Fig 1. Illustration of the predictors and targets used.

Results

Average MAE When Using Different Sensors (pilot data)

All values in [mm], green=low MAE, red=high MAE

	Overall 3D	Palm			Wrist-In			Wrist-Out		
		Palm X	Palm Y	Palm Z	Wrist In X	Wrist In Y	Wrist In Z	Wrist Out X	Wrist Out Y	Wrist Out Z
Baseline (All Sensors)	65.53	45.46	31.63	22.76	38.07	30.61	25.62	40.14	33.78	23.78
sEMG Alone	99.78	75.49	45.54	27.73	62.67	41.60	27.72	67.21	45.40	28.20
Elevation Angles Alone	79.81	64.22	34.84	25.22	51.09	30.29	23.36	54.84	38.86	22.31
Optical Axes Alone	90.50	59.82	48.29	25.58	51.60	45.89	26.62	56.24	45.71	26.95
Eye Positions Alone	105.43	80.13	44.23	22.88	71.64	42.58	21.63	81.65	43.46	22.38
Head Accelerometer Alone	141.16	108.40	64.87	39.70	96.00	59.33	39.67	104.68	64.48	38.52
Head Gyro Alone	102.22	73.57	49.60	33.85	61.83	45.72	34.15	64.26	47.21	33.45
Control - No Sensors	108.03	79.99	50.84	41.76	66.08	45.82	39.94	67.70	50.84	39.30

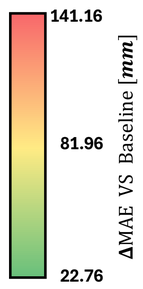


Fig 2. Average MAE of the predicted endpoint under baseline (all sensor modalities), individual sensor, and no sensors conditions using the proof-of-concept predictor with our pilot data. Elevation angles refer to the elevation angles of the upper arm according to Soechting & Flanders [3].

- The baseline model performs surprisingly well, with a 3D MAE of **6.5 cm**.
- The pilot data collected by us was limited in duration and biased in terms of contents.

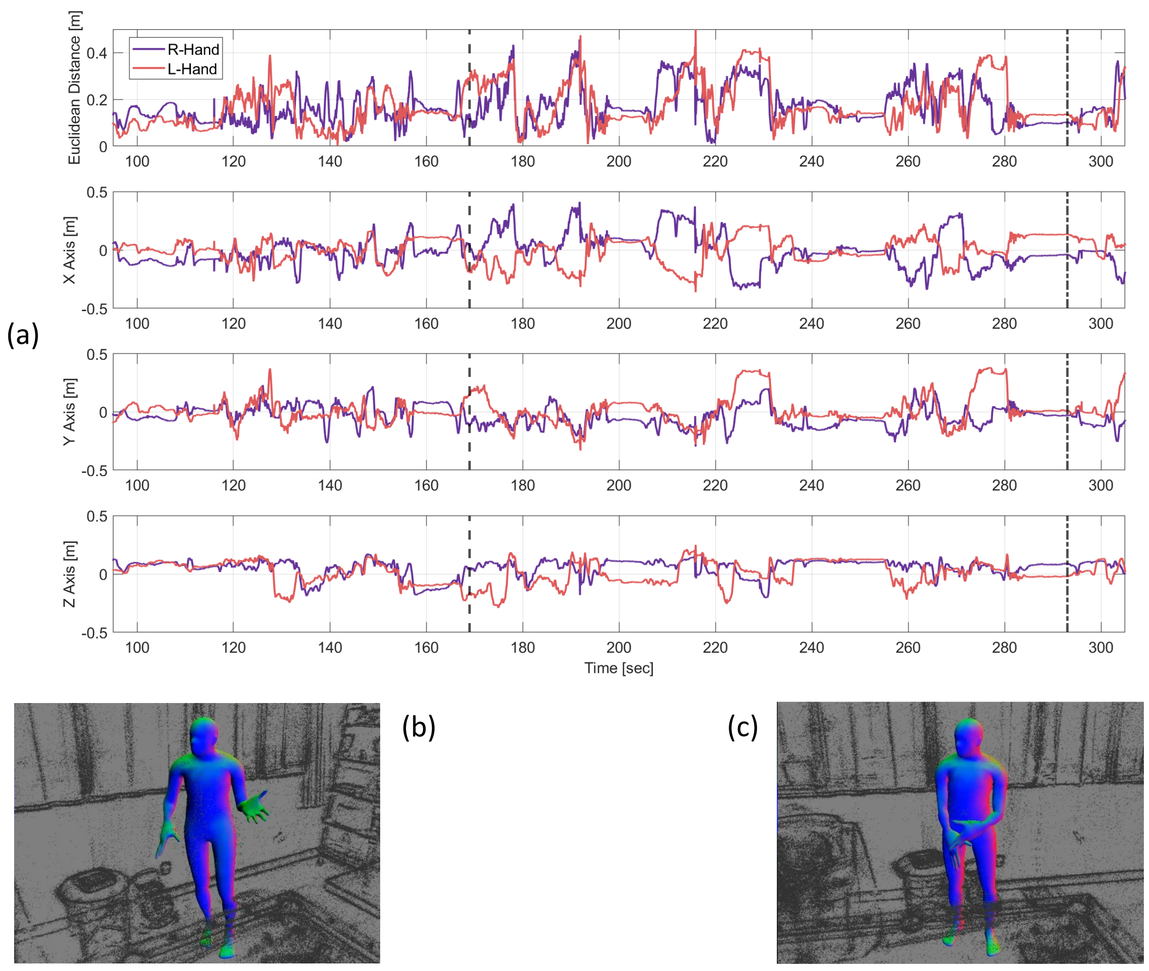


Fig 3. Distance graphs for a representative portion of the test data, specifically for 100-300s of one of the Project Nymeria recordings, showing segments with higher and lower LR prediction accuracy **(a)**, corresponding to the hands being used to gesture while walking **(b)** and the hands being used to hold a pool cue **(c)**.

- The **LR** predictor achieved a 3D MAE of **17.3 cm** over 8 subject and 16 hours of data.
- The **UNet** predictor achieved a 3D MAE of **11.8 cm** over 8 subject and 16 hours of data.
- Both predictors showed strong generalizability between subjects and tasks.

Deep sensor fusion eNaBLes prediction of bilateral hand kinematics

Conclusions and Future Work

- The full results were submitted in a paper to IEEE BSN 2026.
- Acquire full dataset combining MoCap, IMU, sEMG, and gaze-tracking data recorded during simulated ADLs.
- Identify relevant **synergies** between sensor modalities.
- Create an efficient **real-time endpoint predictor and controller**.
- Control a robotic hand and FRANKA arm to demonstrate real-time ***prediction of user intent***.
- Test controller with transhumeral amputees user and custom prosthesis.

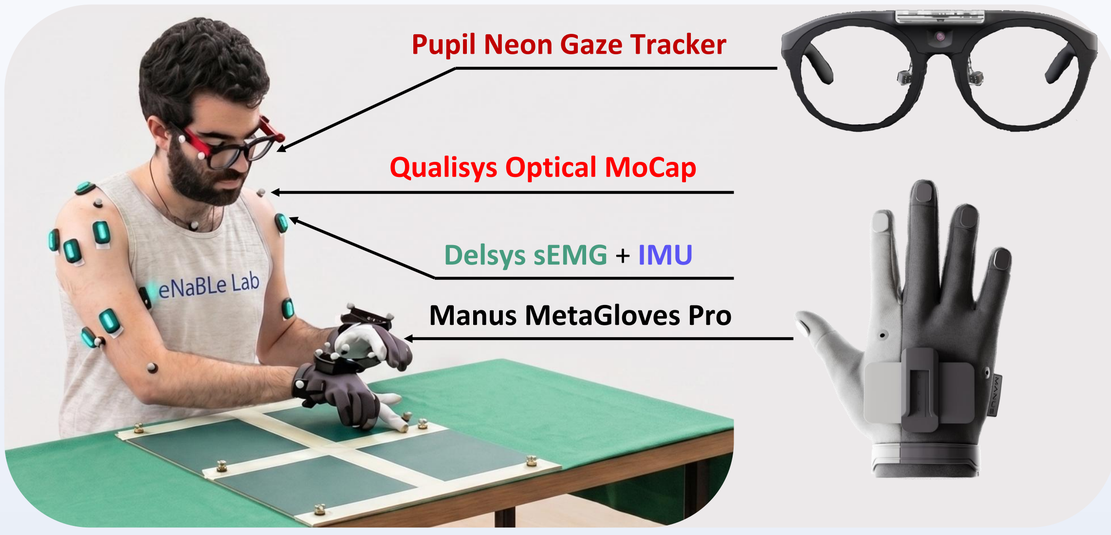


Fig 5. Illustration of the experimental setup and sensors

[1] S. Höfer et al., "Sim2real in robotics and automation: Applications and challenges," IEEE Trans. Autom. Sci. Eng., vol. 18, pp. 398–400, Apr. 2021.

[2] N. E. Krausz, D. Lamotte, I. Batzianoulis, L. J. Hargrove, S. Micera, and A. Billard, "Intent Prediction Based on Biomechanical Coordination of EMG and Vision-Filtered Gaze for End-Point Control of an Arm Prosthesis," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 28, no. 6, pp. 1471–1480, Jun. 2020, doi: 10.1109/TNSRE.2020.2992885.

[3] L. Ma et al., "Nymeria: A massive collection of multimodal egocentric daily motion in the wild," in Proc. ECCV, 2025, pp. 445–465.

[4]] J. F. Soechting, M. Flanders, "Errors in Pointing Are Due to Approximations in Sensorimotor Transformations", Journal of Neurophysiology, vol. 62, issue 2, Aug. 1989, doi: 10.1152/jn.1989.62.2.595.

This research is funded by the Ministry of Innovation, Science and Technology Grant #0007458