NRI: INT: COLLAB: Manufacturing USA: Intelligent Human-Robot Collaboration for Smart Factory

Robert X. Gao, Department of Mechanical & Aerospace Engineering, Case Western Reserve University

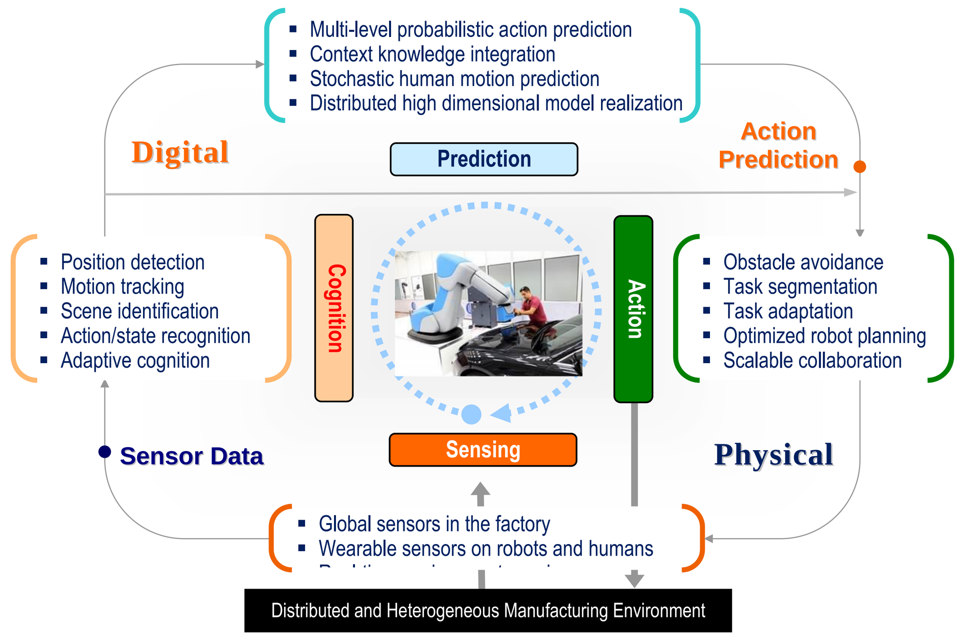
Ming Leu, Department of Mechanical & Aerospace Engineering, Missouri University of Science and Technology

Gloria Wiens, Department of Mechanical & Aerospace Engineering, University of Florida

Zhaozheng Yin, AI Institute, Department of Biomedical Informatics and Department of Computer Science, Stony Brook University

Research challenges in realizing human-robot collaboration in smart factories:

- Separating and highlighting relevant content from irrelevant/ ambiguous information in **sensing** data for human action, command and intention **recognition**.
- Accounting for uncertainty in human action and trajectory **prediction**.
- Estimating collision risk between robot and human trajectory during collaborative **action**.

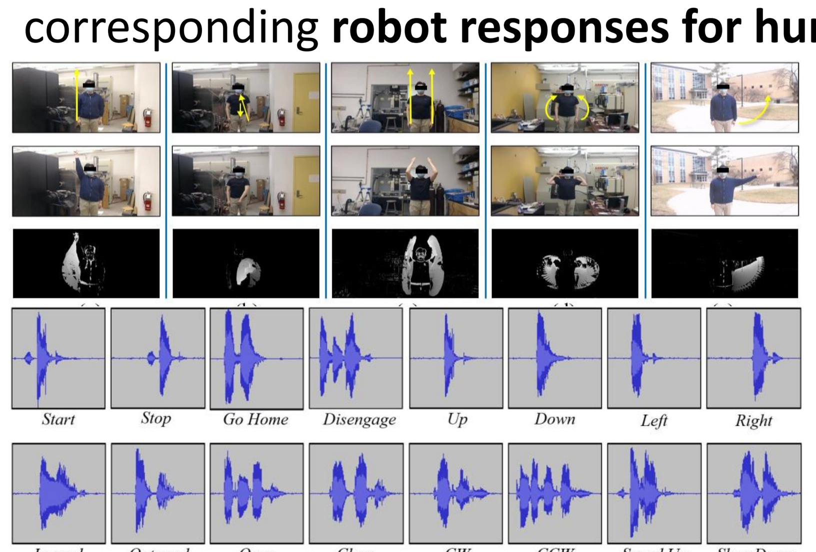


Sensing/ Cognition

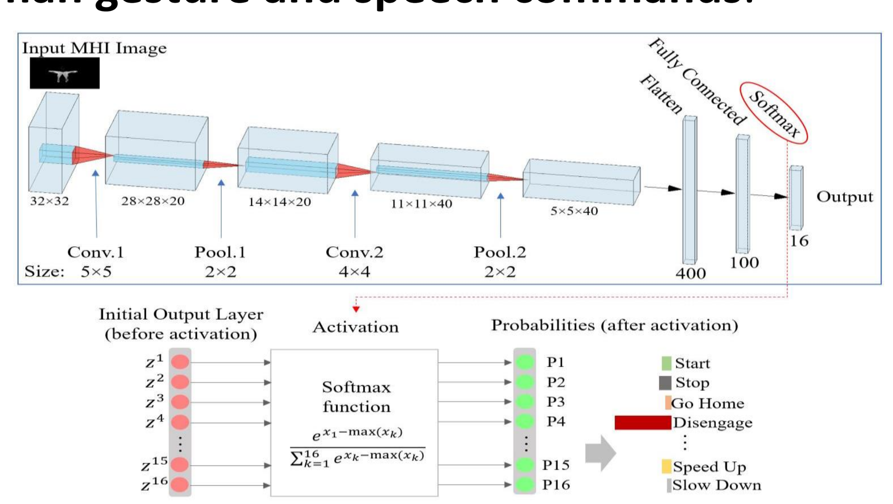
Real-Time Multi-modal Human-Robot Collaboration Using Gestures and Speech [1]

Objective: Design real-time human-robot collaboration system combining our designed dynamic gestures and natural speech commands.

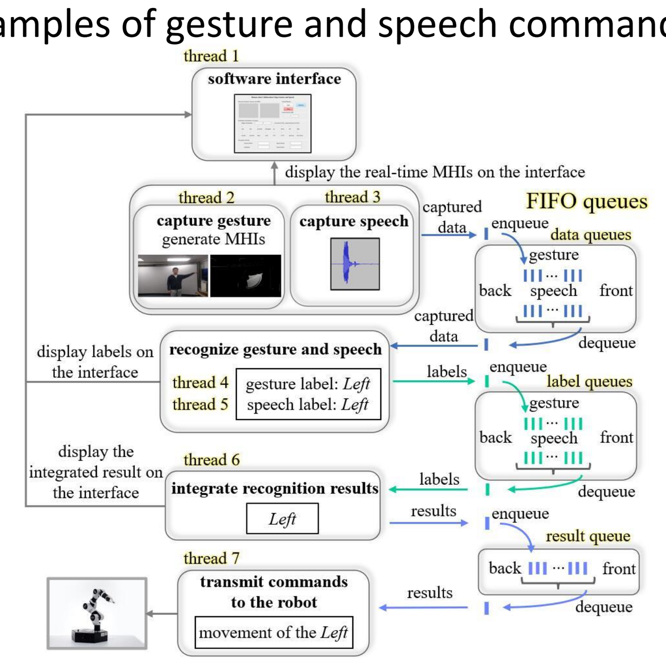
- Designed a real-time command recognizer for **designed dynamic gestures** in different backgrounds with **Motion History Image**-based feature extraction, and **natural speech commands** under ten different noisy backgrounds, achieving classification accuracy >95%.
- Developed a multi-threading system **controller** for parallel tasks execution and corresponding **robot responses for human gesture and speech commands**.



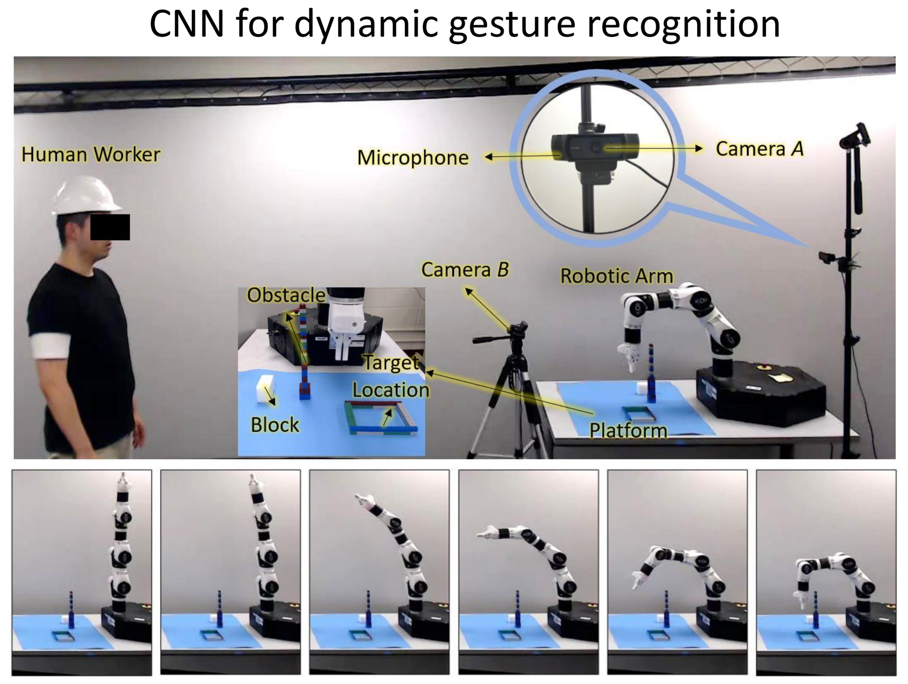
Samples of gesture and speech commands



CNN for dynamic gesture recognition



Multi-threading architecture of control system



Overview of the HRC system

Sensing/ Cognition

Wearable Motion Capture: Reconstructing and Predicting 3D Human Poses from Wearable Sensors [2]

Objective: Reconstruct and predict 3D human poses from the wearable IMU sensors and the wearable cameras.

- Proposed the **wearable motion capture problem** of reconstructing and predicting **3D human poses** from the wearable IMU sensors and wearable cameras.
- Developed a novel **Attention-Oriented Recurrent Neural Network (AttRNet)** that contains a sensor-wise attention-oriented recurrent encoder, a reconstruction module, and a dynamic temporal attention-oriented recurrent decoder, to **reconstruct the current pose and predict the future poses**.

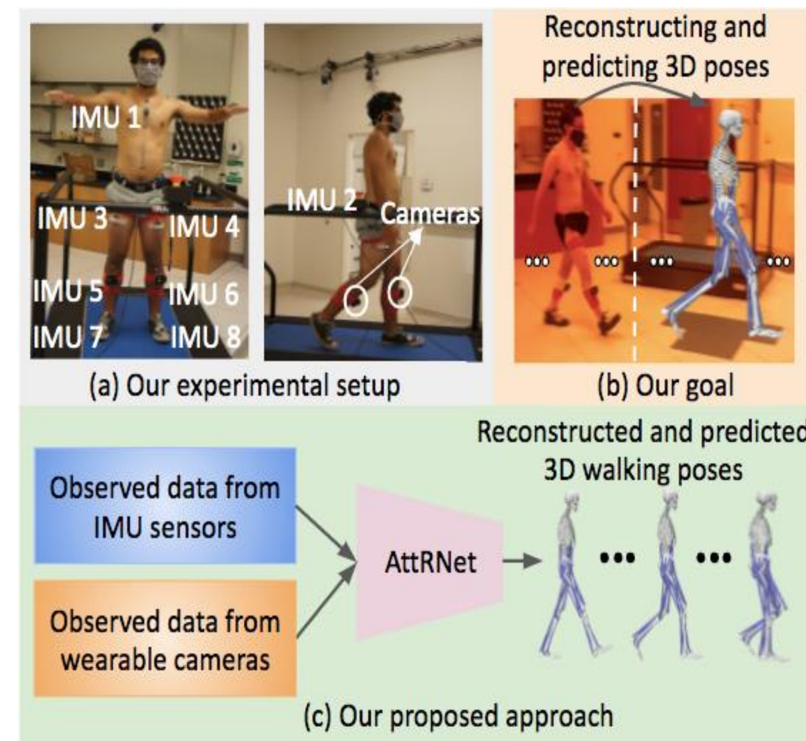


Illustration of our proposed approach.

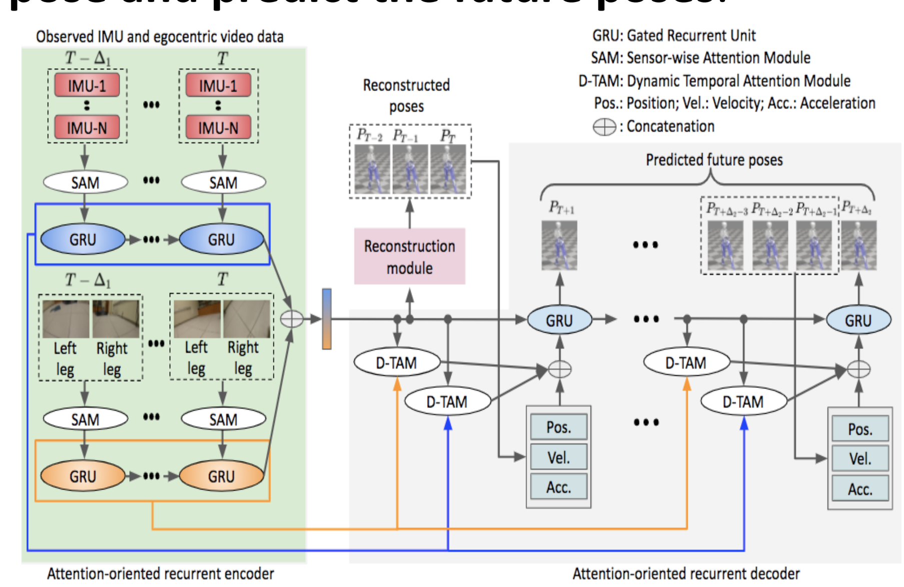
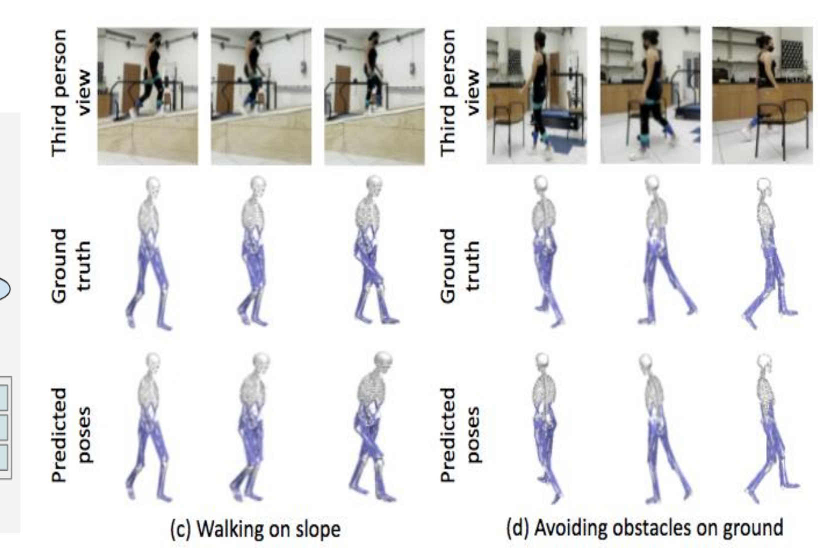


Illustration of our proposed AttRNet.



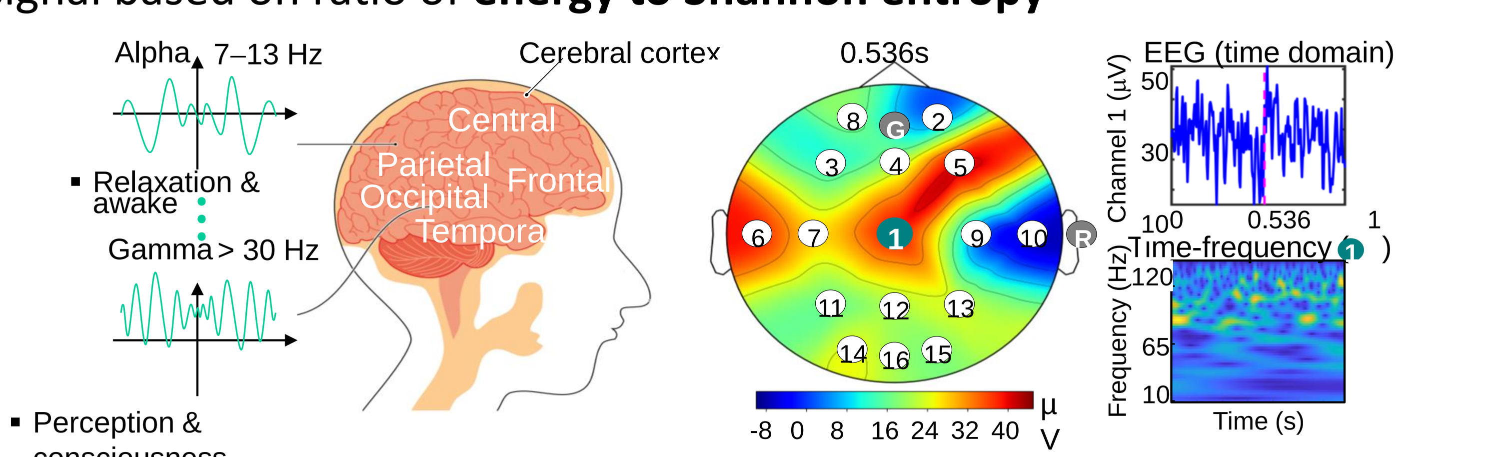
Visualization of our human pose prediction over time.

Prediction

Future Action Prediction driven by Brainwaves [3]

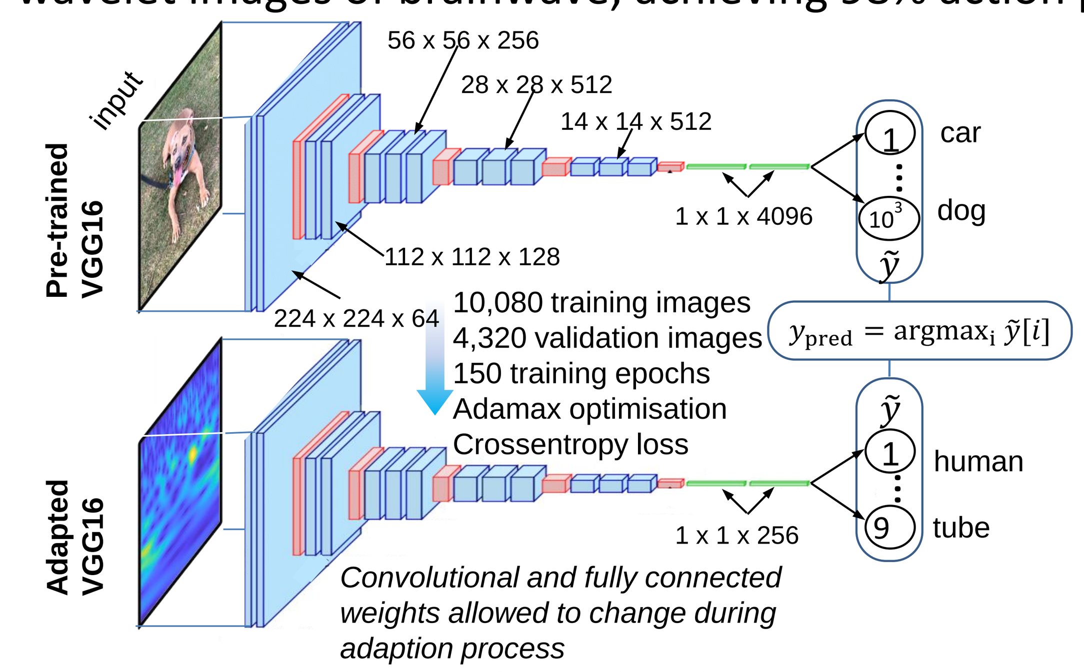
Objective: Accurate extraction and interpretation of stimulus-free brain EEG signals from human worker for future action prediction, to complement standard sensing modalities such as voice, gesture and haptic

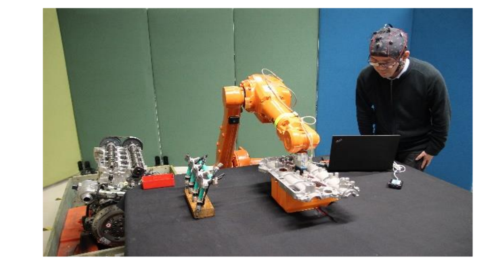
- Optimized base wavelet selection to maximize information extraction from brain EEG signal based on ratio of **energy to Shannon entropy**



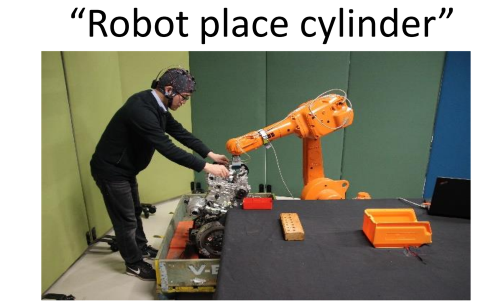
Structure and waves of brain and wavelet transform of brainwave to time-frequency images

- **Adapted** a pre-trained VGG16 convolutional neural network and **fine-tuned** classification layers for predicting future robot action based on time-frequency wavelet images of brainwave, achieving 98% action prediction accuracy





"Robot place cylinder"

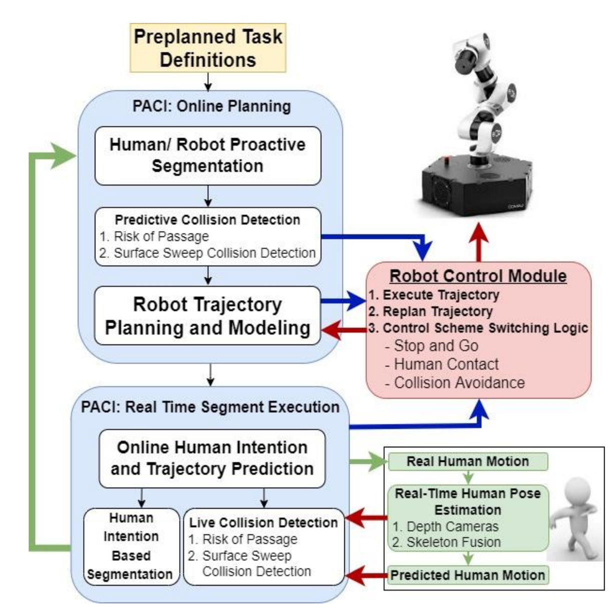


"Robot insert tube"

VGG16 structure and adaptation for EEG-based action prediction and example predicted actions

Action

Human-Robot Proactive-n-Reactive Behavior Intelligence [4,5]

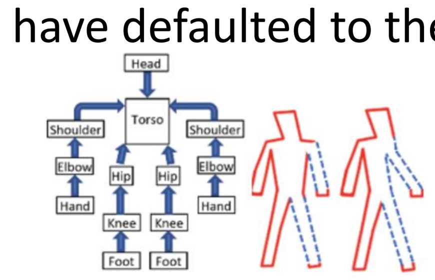


PACI Control Architecture

Objective: Seamless integration of sensing, cognition, and prediction into robot controller yielding efficient Proactive Adaptive Collaboration Intelligence (PACI) to ensure safe interactions with humans and mitigate production disruptions.

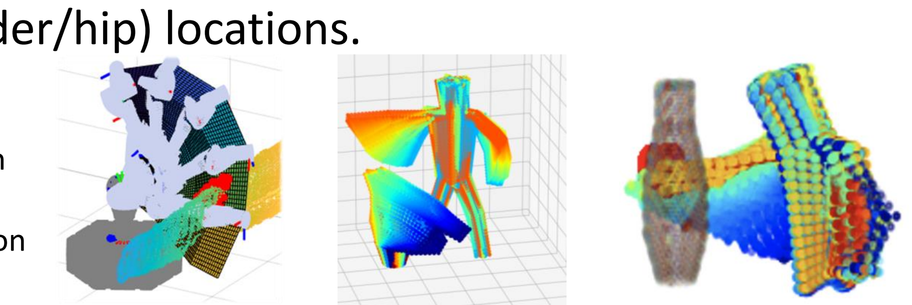
Segmentation and Augmented Intelligence: Control provides 'proactive-n-reactive' robot response using a segmentation framework of predicted, real time, and behavioral classifications to allow for modularity and flexibility within the control architecture.

Predictive Collision Detection: Developed a parent-child human kinematics algorithm to correct vision tracker data of joint locations. The human outlines show repairing the skeleton if the elbow/knee (dashed lines) are inaccurate, where they have defaulted to their parent (shoulder/hip) locations.



Parent/Child Kinematic Model

Improved Temporal Point Cloud generation for predicted human-robot collision detection



Skeleton Fusion w/ Multiple Depth Cameras: Using multiple depth cameras to mitigate occlusion errors, developed a method that decomposes multiple Deep Convolutional Neural Newtork (DCNN) skeleton pose estimates into quaternion representations, uses filtering/state-estimation; stitching (fusing) 'best' quaternions into the fused skeleton for tracking human movement. DCNN human skeleton trackers compared were 2D/3D Media Pipe and OpenPose.

Tracked Human Skeleton with Robot as Occlusion

2022 NRI & FRR Principal Investigators' Meeting

April 19-21, 2022

NSF/NRI 2.0: ENG/CMMI and CISE/IIS

Award ID#: 1954548/1830479/1830383/1830295

Related Publications

[1] H. Chen et al., "Real-Time Multi-modal Human-Robot Collaboration Using Gestures and Speech", ASME JMSE, 2022 in review

[2] M. Moniruzzaman *et al.*, "Human action recognition by discriminative feature pooling and video segment attention model," IEEE TMM, 2021.

[3] L. Wang *et al.*, "Function block-based human-robot collaborative assembly driven by brainwaves," CIRP Annals, 2021.

[4] G. Streitmatter, and G. Wiens, "High Fidelity Human Modeling via Integration of Skeleton Tracking for Predictive HRC Collision Detection", ASME IMECE, 2021.

[5] J. Flowers, and G. Wiens, "Comparison of Human Skeleton Trackers Paired with a Novel Skeleton Fusion Algorithm", ASME MSEC, 2022.