UNIVERSITÀ di VERONA
Dipartimento di INFORMATICA

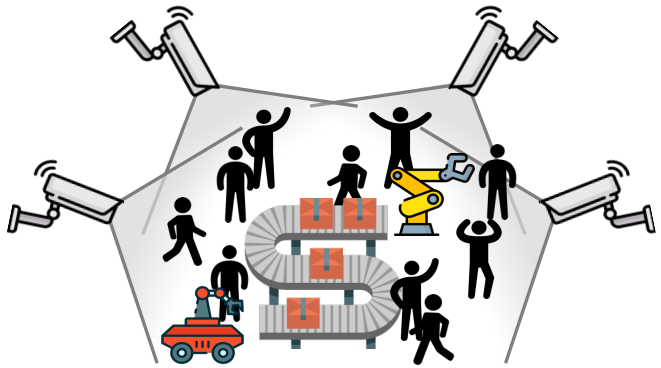
ICE Lab
INDUSTRIAL COMPUTER ENGINEERING

Offloading Real-time Multi-camera 3D Human Pose Estimation on Edge Computing Devices

Mirco De Marchi
Michele Boldo
Enrico Martini
Nicola Bombieri
Franco Fummi

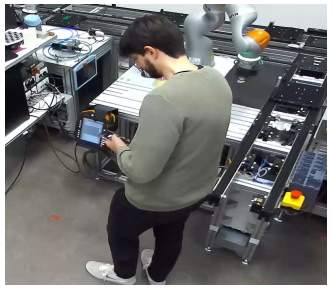
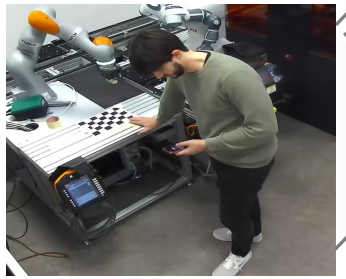
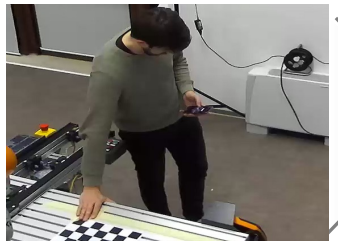
Roadmap

- **Goal:** 3D Human Pose Estimation in ICE LAB, useful for a variety of other applications for the industrial environment, such as:
  - Human Robot Interaction
  - Safety on Industrial Environment
  - Gesture Recognition
- **Issues to be addressed:**
  - **Privacy:** In working environment, it is not possible to send raw intelligible data, such as video stream, to a centralized server to compute it
  - **Occlusions:** working environments are full of machinery that may occludes the body parts to the camera, generating incomplete information
- **Proposed solution:**
  - Offloading the 3D HPE on Edge Computing Devices
  - Distribute edge devices on multiple views



BeFine Architecture

BeFine: distributed 3D HPE at the edge

RGB+D

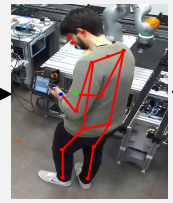
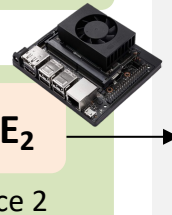
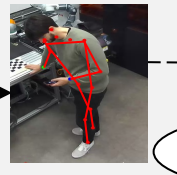
RGB+D

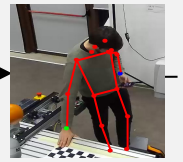
RGB+D

3D HPE$_1$
Edge device 1

3D HPE$_2$
Edge device 2

3D HPE$_n$
Edge device n



MQTT

3D Skelet.

***Shared* communication Network (e.g., Ethernet, WiFi)**

Variable bandwidth

Variable delay

Protocol-dependent communication

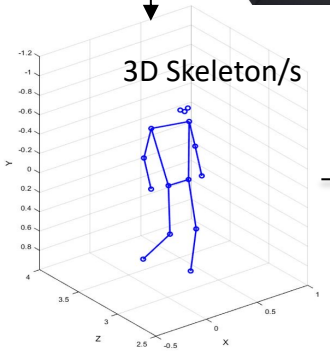
Potential packet loss

3D Skelet.

Centralized compute unit

docker

Data synchronization + aggregation



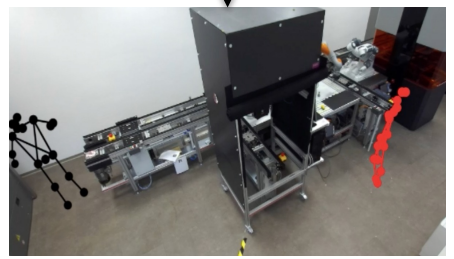
Various ICE Lab Applications:

Digital shadow

HRI Collision Prediction

Gesture Recognition

kafka



Main challenges

NVIDIA JETSON XAVIER NX:
Low Power: ~10 W
Low Cost: ~300 $

Hardware constraints → High Performance SW for embedded platforms (C++/CUDA)
Optimized Deep Neural Network (Nvidia TensorRT)

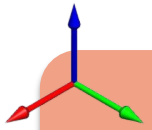
MQTT kafka

Standard industrial network protocol on shared communication channel

Due to the variability of the transmission network it is mandatory to have a global ordering of the messages

↓

Two-level synchronization → Network Time Protocol
→ After the network



Multi-view multi-person Spatial Fusion

Common 3D coordinates sistem → Clustering and optimization algorithm used for merging the various view and associate the «skeletons» related to the same person

For more details let’s meet at live demo…

Thanks for the attention

PARCO Parallel Computing Lab
Michele Boldo, PhD Student
Parallel **Co**mputing Lab
michele.boldo@univr.it