CONIX RESEARCH CENTER

Computing on Network Infrastructure for Pervasive Perception, Cognition, and Action

Capricorn: Towards Real-time Rich Scene Analysis Using RF-Vision Sensor Fusion

Ziqi Wang, Ankur Sarker, Jason Wu, Derek Hua, Gaofeng Dong, Akash Deep Singh, Mani Srivastava (UCLA)

Motivation

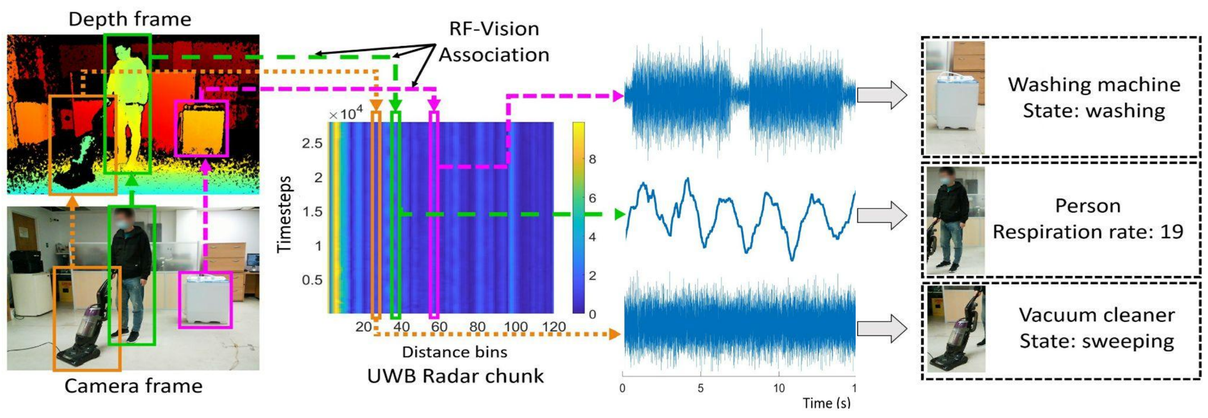
Rich scene analysis:

- Simultaneous estimation of both extrinsic and intrinsic object states
- Extrinsic states: visually observable properties (type, shape, location, etc); acquired by vision sensors like cameras
- Intrinsic states: internal physical/biological activities (machine operating status, human vital signals, etc); however, they are often *overlooked* by vision-based scene analysis

Capricorn:

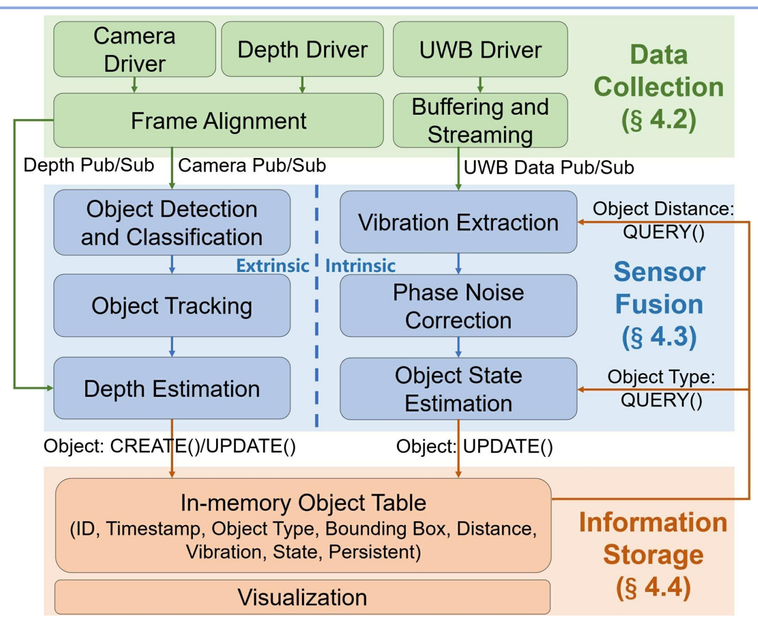
- A real-time RF-vision sensor fusion system efficiently builds a cross-modal correspondence between visual pixels and RF time series to better understand the complex natures of a scene
- RF sensors used in wireless vibrometry only provide time series; it's challenging to associate these series with multiple objects

Overview



- The extrinsic sensing pipeline uses object detection, classification, and tracking algorithms to infer the object types and their locations (i.e., bounding boxes and distance) in the scene
- The intrinsic sensing pipeline uses the distance information estimated above to extract vibration signals for each object from a three-dimensional RF data stream (i.e., time, distance, intensity)
- Combine vibration signals and object type information to estimate the objects' intrinsic states (i.e., machine operating states, human respiration rate)

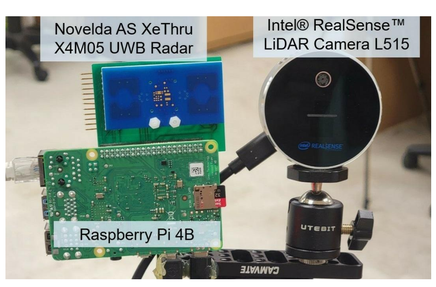
System Design



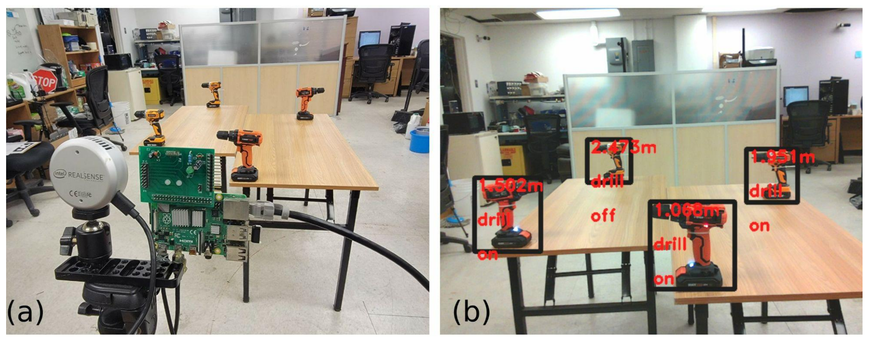
- Data collection unit collects raw data from different sensors simultaneously and feeds them to the sensor fusion unit in a publish-subscribe (Pub-Sub) pattern
- Sensor fusion unit processes sensor data and estimates extrinsic and intrinsic object states
- Information storage unit generates an in-memory table to store the inference results and facilitate the fusion between different modalities

Results

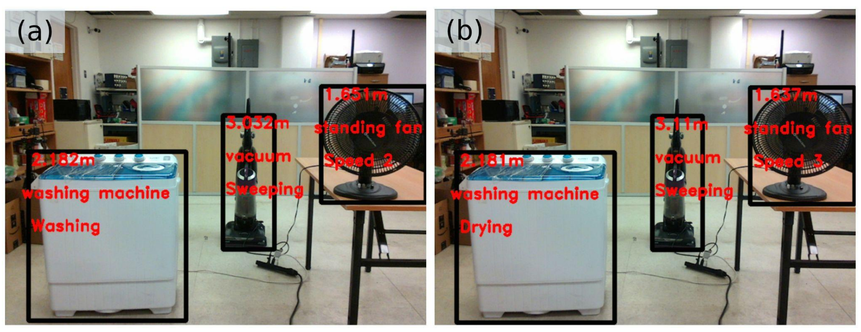
Demo



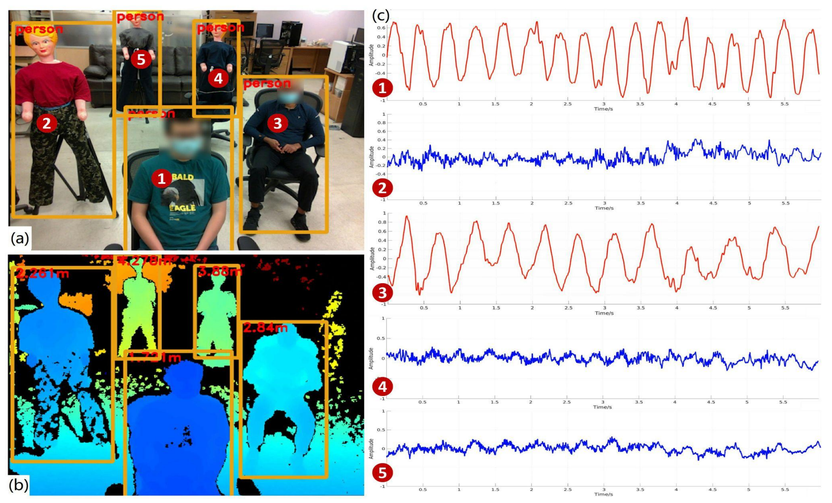
Hardware platform



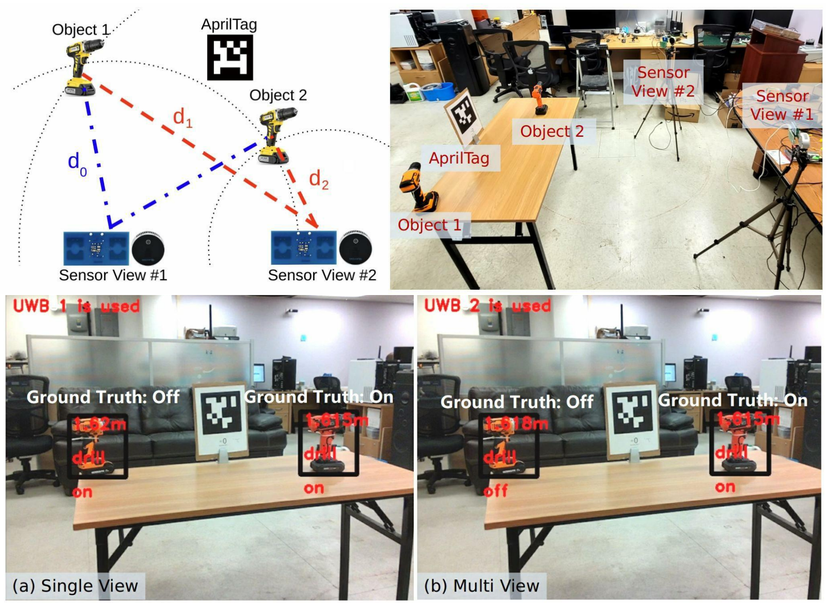
Workshop scene: drill state detection



SmartHome scene: appliance usage tracking



Multi-person respiration rate estimation



Multi-view version to distinguish objects at the same distance

Complex event detection: "a person doing laundry"

Contact: wangzq312@ucla.edu

This work was supported in part by the CONIX Research Center, one of six centers in JUMP, a Semiconductor Research Corporation (SRC) program sponsored by DARPA.

www.conix.io