

Cross-modality Learning for Ubiquitous Human Identification

CHRIS XIAOXUAN LU, University of Oxford

ANDREW MARKHAM, University of Oxford

NIKI TRIGONI, University of Oxford

Biometric information, such as facial and vocal features are critical for identifying and authenticating individuals. This is an enabling service for smart-spaces, allowing building management agents to easily monitor ‘who is where’, anticipating user needs and tailoring the local environment and user experiences.

Although biometric recognition, especially through the use of deep neural networks, has achieved stellar performance with large datasets, the majority of approaches require supervised learning, that is, to be trained with tens or hundreds of images of users in different conditions. In this talk, we motivate that this enrolment effort is unnecessary if the smart-space has access to a wireless identifier e.g., through a smart-phone’s MAC address. By learning and refining the noisy and weak association between a user’s smart-phone and her biometric observations, AutoTune can fine-tune a deep neural network and tailor it to the environment, users and conditions of a particular space, e.g., lighting variability and viewing angles.

In particular, we introduce a novel soft-association technique which limits the impact of erroneous decisions taken early on in the training process from corrupting the clusters. We also show how this zero-effort enrolment system becomes better over time at recognizing and identifying users by increasing the inter-cluster distance and minimizing the intra-cluster distance. Through extensive experiments of different biometrics with multiple users on different sites, we demonstrate the ability of AutoTune to design an environment-specific, continually evolving biometric recognition system with entirely no user effort. We believe that this framework is a step towards Mark Weiser’s vision of the third wave of computing which recedes into the background of our physical spaces.

This talk is based on the previous work in [1–3].

CCS Concepts: •**Human-centered computing** →**Ubiquitous and mobile computing**; *Ambient intelligence*;

Additional Key Words and Phrases: Human Identification; Biometric Learning; Smart Space; IoT

ACM Reference format:

Chris Xiaoxuan Lu, Andrew Markham, and Niki Trigoni. 2018. Cross-modality Learning for Ubiquitous Human Identification. 1, 1, Article 1 (January 2018), 1 pages.

DOI: 10.1145/3131672.3136991

REFERENCES

- [1] Chris Xiaoxuan Lu, Xuan Kan, Bowen Du, Changhao Chen, Hongkai Wen, Andrew Markham, Niki Trigoni, and Jack Stankovic. 2019. AutoTune: Autonomous Learning for Face Recognition in the Wild. *In Submission* (2019).
- [2] Chris Xiaoxuan Lu, Xuan Kan, Stefano Rosa, Bowen Du, Hongkai Wen, Andrew Markham, and Niki Trigoni. 2017. Towards Self-supervised Face Labeling via Cross-modality Association. In *ACM Conference on Embedded Network Sensor Systems (SenSys)*. ACM.
- [3] Chris Xiaoxuan Lu, Hongkai Wen, Sen Wang, Andrew Markham, and Niki Trigoni. 2017. SCAN: learning speaker identity from noisy sensor data. In *ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*. IEEE.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2018 ACM. XXXX-XXXX/2018/1-ART1 \$15.00

DOI: 10.1145/3131672.3136991