Fast One-Shot Learning for Identity Classification in Person Re-identification and Tracking

Andrew Tzer-Yeu Chen, Morteza Biglari-Abhari, Kevin I-Kai Wang

Embedded Systems Research Group
Department of Electrical and Computer Engineering

ICARCV 2018, Singapore

THE UNIVERSITY OF AUCKLAND
Te Whare Wānanga o Tāmaki Makaurau
NEW ZEALAND

Localising people over time

Targeting market research applications

- How many customers are there?
- What parts of the shop are they in?
- Are they going past high value items?
- Do we have enough staff?

Person Re-identification

- Matching people across multiple cameras
- Classification by appearance
- Then collect positions in each frame

Real-world Challenges

- Most approaches too slow for real-time
- Need to balance accuracy vs speed
- Lack of labelled training data

Why one-shot learning?

Supervised Learning
- Using labelled ground truth data
- Training can require a lot of data

Unsupervised Learning
- No training data needed, learns live
- Can diverge and fail to learn, hard to tell

Semi-supervised Learning
- Some labelled, some unlabelled data

One-shot Learning
- Only one labelled sample, rest unlabelled

Person Re-identification is difficult

- Variety of poses
- Partial obscuration
- Motion blur and focus issues
- Lighting sources cause colour changes

UoA-Indoor dataset captures 19 people from four camera views with 28,000 samples in total (750-2460 per identity)

Main Processing Steps

Image Input

↓

Person Detection (DPM)

↓

Feature Extraction (Colour and LBP Histograms)

↓

PCA and Metric Learning

↓

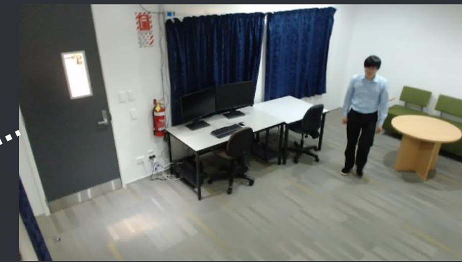
Compute Similarity to all Identity Classes

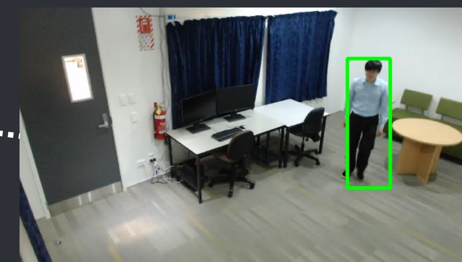
↓

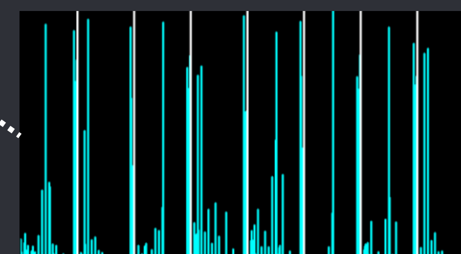
If match, Update Class Model

↓

Class/Identity Label Output







Evaluated Algorithms

Support Vector Machine (SVM) and Perceptron

- One-against-all Scheme
- Offline and Online Learning Schemes

Label Spreading

- Semi-supervised Learning
- Constructs a similarity graph based on labelled samples, uses unlabelled samples to shape the distribution

Gallery

- Similar to ViBe classification scheme
- One-shot Learning

- Build a model of N samples for each identity class
- Compare probe $\mathbf{x}$ to each gallery class sample s
- Calculate $||s\text{-}\mathbf{x}||$ and compare to threshold $\mathbf{DIS}$
- Declare match if $\mathbf{C(s)} >= \mathbf{numMin}$, where $\mathbf{C(s)}$ is the number of samples $\mathbf{s}$ that are close matches to $\mathbf{x}$
- Replace a random sample in the class with $\mathbf{x}$

- Reasonably robust against noise

Sequential K-means

- Each class has one cluster mean only
- One-shot Learning

- Compare probe $\mathbf{x}$ to each cluster mean $\mathbf{m}_c$
- Use Euclidean distance $||\mathbf{m}_c - \mathbf{x}||$
- Declare match for closest cluster mean (class $\mathbf{c}$)
- Update cluster mean with linear weighting

$$\mathbf{m}_c = \beta\mathbf{x} + (1-\beta)\mathbf{m}_c$$

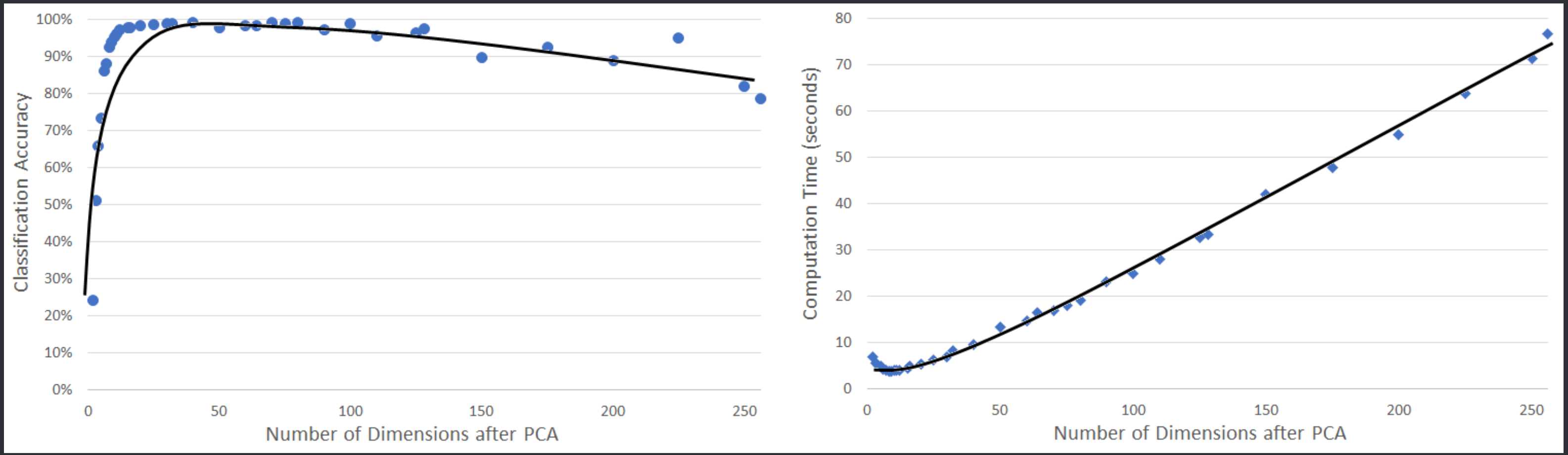
- Light on computation and memory
- Dependent on metric learning

Experimental Results

Evaluated on UoA-Indoor

	Method	Accuracy (%)	Classification Time (s)
Supervised	Offline SVM	98.82	8.377
Supervised	Offline Perceptron	91.35	0.391
Semi-Supervised	Label Spreading	72.92	1.772
One-shot	**Sequential K-means**	**67.22**	**0.282**
One-shot	Gallery	48.41	4.720
One-shot	Online Perceptron	34.44	29.797
One-shot	Online SVM	29.61	28.891

Tuning Sequential K-means



Balance accuracy and computation time

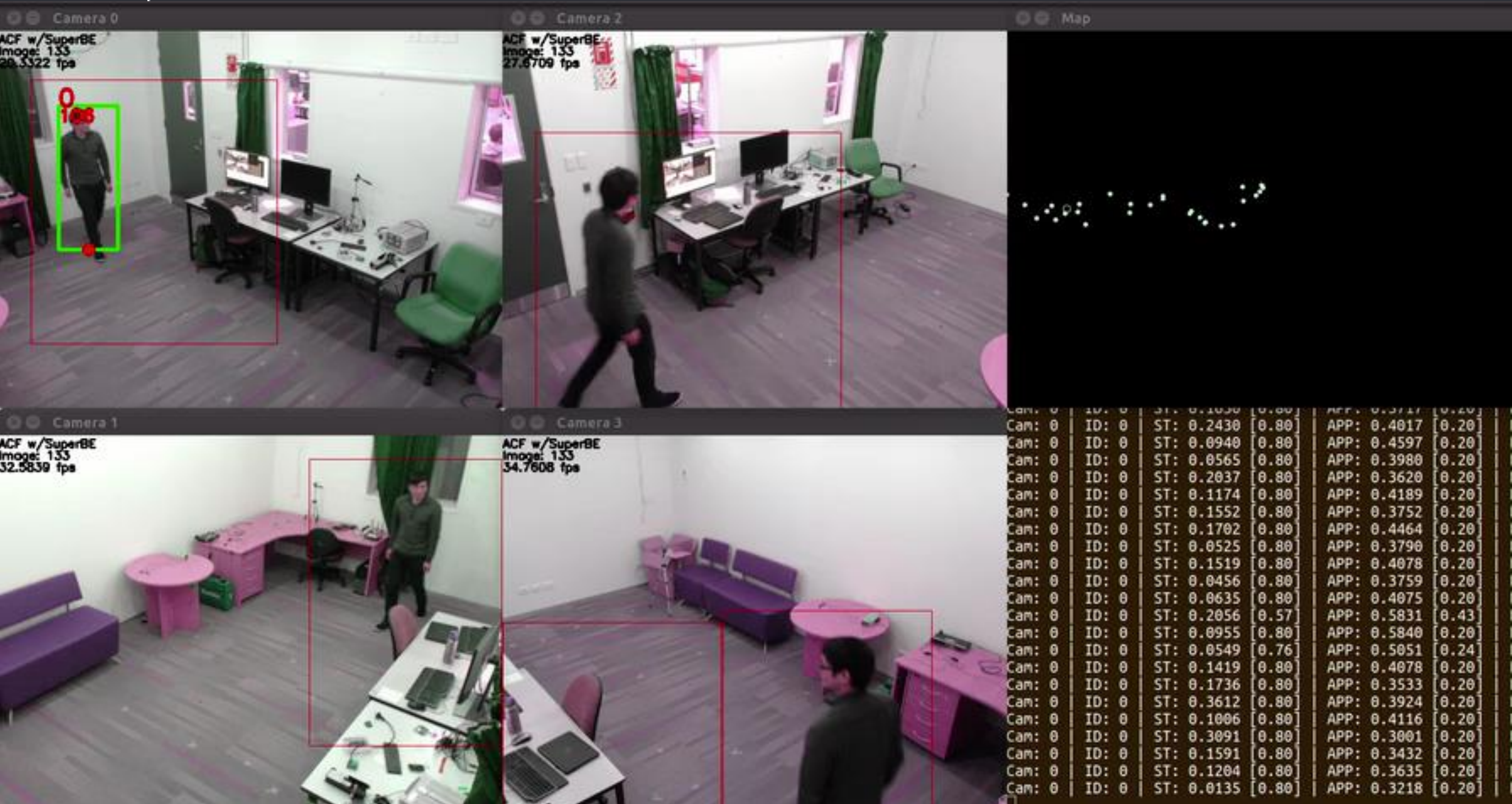
PCA set to use 50 dimensions

67.22% -> **72.65%**

Competitive r=1 result for re-identification

Conclusions

- Fast person re-identification across multiple camera views
- One-shot learning methods are needed in the real-world
- Sequential k-means outperforms several other choices on UoA-Indoor


Camera 0
ACF w/SuperBE
Image: 133
20.3322 fps
0
Camera 2
ACF w/SuperBE
Image: 133
27.6709 fps
Map
Camera 1
ACF w/SuperBE
Image: 133
32.5839 fps
Camera 3
ACF w/SuperBE
Image: 133
34.7608 fps
Cam: 0 | ID: 0 | ST: 0.2430 [0.80] | APP: 0.4017 [0.20]
Cam: 0 | ID: 0 | ST: 0.0940 [0.80] | APP: 0.4597 [0.20]
Cam: 0 | ID: 0 | ST: 0.0565 [0.80] | APP: 0.3980 [0.20]
Cam: 0 | ID: 0 | ST: 0.2037 [0.80] | APP: 0.3620 [0.20]
Cam: 0 | ID: 0 | ST: 0.1174 [0.80] | APP: 0.4189 [0.20]
Cam: 0 | ID: 0 | ST: 0.1552 [0.80] | APP: 0.3752 [0.20]
Cam: 0 | ID: 0 | ST: 0.1702 [0.80] | APP: 0.4464 [0.20]
Cam: 0 | ID: 0 | ST: 0.0525 [0.80] | APP: 0.3790 [0.20]
Cam: 0 | ID: 0 | ST: 0.1519 [0.80] | APP: 0.4078 [0.20]
Cam: 0 | ID: 0 | ST: 0.0456 [0.80] | APP: 0.3759 [0.20]
Cam: 0 | ID: 0 | ST: 0.0635 [0.80] | APP: 0.4075 [0.20]
Cam: 0 | ID: 0 | ST: 0.2056 [0.57] | APP: 0.5831 [0.43]
Cam: 0 | ID: 0 | ST: 0.0955 [0.80] | APP: 0.5840 [0.20]
Cam: 0 | ID: 0 | ST: 0.0549 [0.76] | APP: 0.5051 [0.24]
Cam: 0 | ID: 0 | ST: 0.1419 [0.80] | APP: 0.4078 [0.20]
Cam: 0 | ID: 0 | ST: 0.1736 [0.80] | APP: 0.3533 [0.20]
Cam: 0 | ID: 0 | ST: 0.3612 [0.80] | APP: 0.3924 [0.20]
Cam: 0 | ID: 0 | ST: 0.1006 [0.80] | APP: 0.4116 [0.20]
Cam: 0 | ID: 0 | ST: 0.3091 [0.80] | APP: 0.3001 [0.20]
Cam: 0 | ID: 0 | ST: 0.1591 [0.80] | APP: 0.3432 [0.20]
Cam: 0 | ID: 0 | ST: 0.1204 [0.80] | APP: 0.3635 [0.20]
Cam: 0 | ID: 0 | ST: 0.0135 [0.80] | APP: 0.3218 [0.20]

Thank You!

ANY QUESTIONS?

andrew.chen@auckland.ac.nz

THE UNIVERSITY OF AUCKLAND
Te Whare Wānanga o Tāmaki Makaurau
NEW ZEALAND