Tri-modal Person Re-identification with RGB, Depth, and Thermal Features

PBVS, CVPR 2013

June 24, 2013

Andreas Møgelmose[1]*, Chris Bahnsen[1], Thomas B. Moeslund[1], Albert Clapés[2], and Sergio Escalera[2]

am@create.aau.dk

[1]Visual Analysis of People Lab
Aalborg University
Denmark

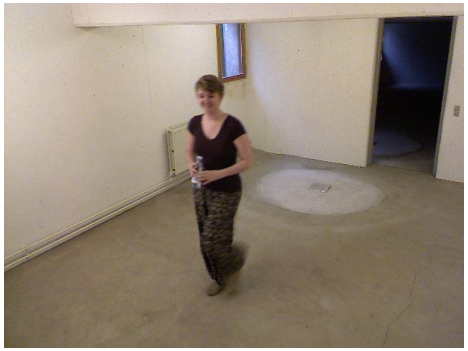
[2]Dept. Matemàtica Aplicada i Anàlisis
Computer Vision Center
Universitat de Barcelona

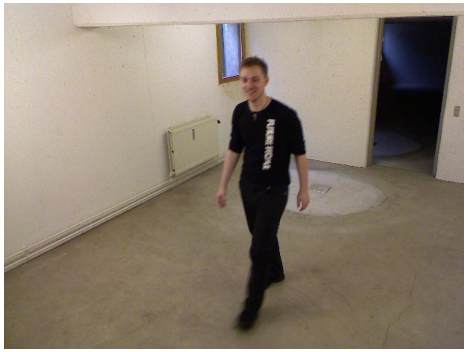
AALBORG UNIVERSITY
DENMARK

Re-identification

- The act of recognizing people who have passed by a camera earlier.
- Can be used in many situations:
  - Person flow analysis (think airport queues)
  - As a forensics tool (London's underground)
  - Anywhere with an established camera network, which benefits from tracking of people.
- Must update with new persons on the fly.
- Must work with soft biometrics, as hard biometrics are difficult to obtain.
- Previous work has been done in RGB, but the combination of modalities has not yet been tested.

Multi-modal features

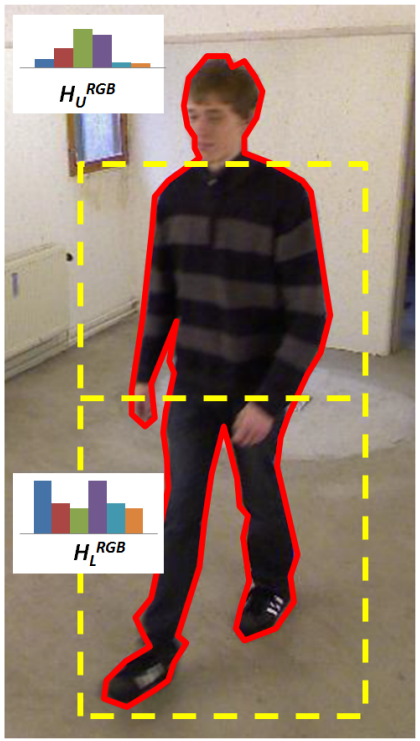


- RGB and depth captured with Microsoft Kinect for Xbox 360
- Thermal imagery captured with Axis Q1922

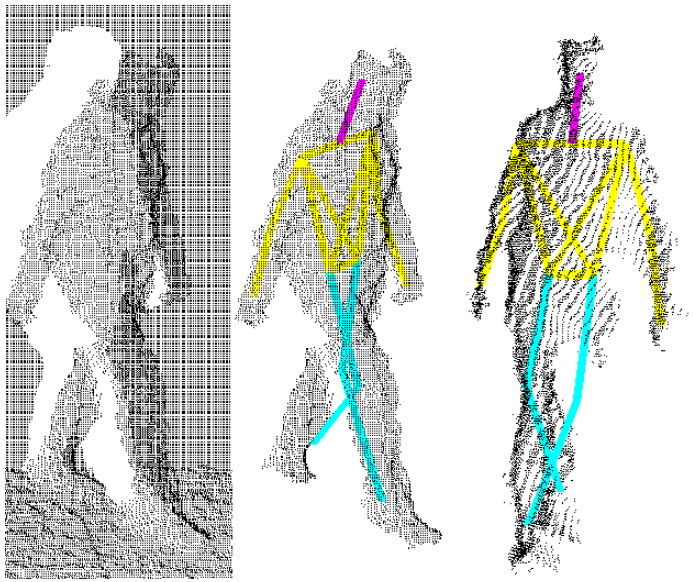
RGB features

- Color histograms.
- The body is split into two parts. One histogram is created for each
- In total a descriptor is a 120 bin histogram
- Averaged over all frames in a pass
- For re-identification, the descriptor is compared to the stored histograms using the Bhattacariyya-distance



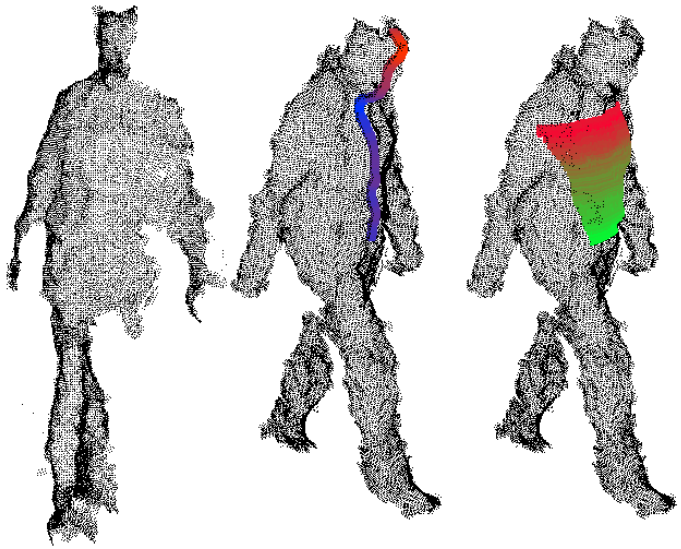
Depth features
Pre-processsing

- Figure-ground segmentation
- Skeleton extraction from raw depth data
- Coordinate transformation
- Smoothing



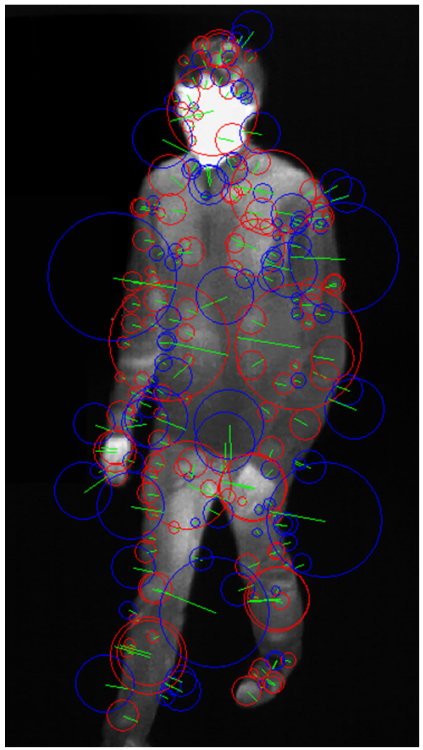
Depth features
Feature extraction

- Frontal curve model
- Thoracic geodesic distances
- Anthropometric distances
- Averaged over all frames in a pass

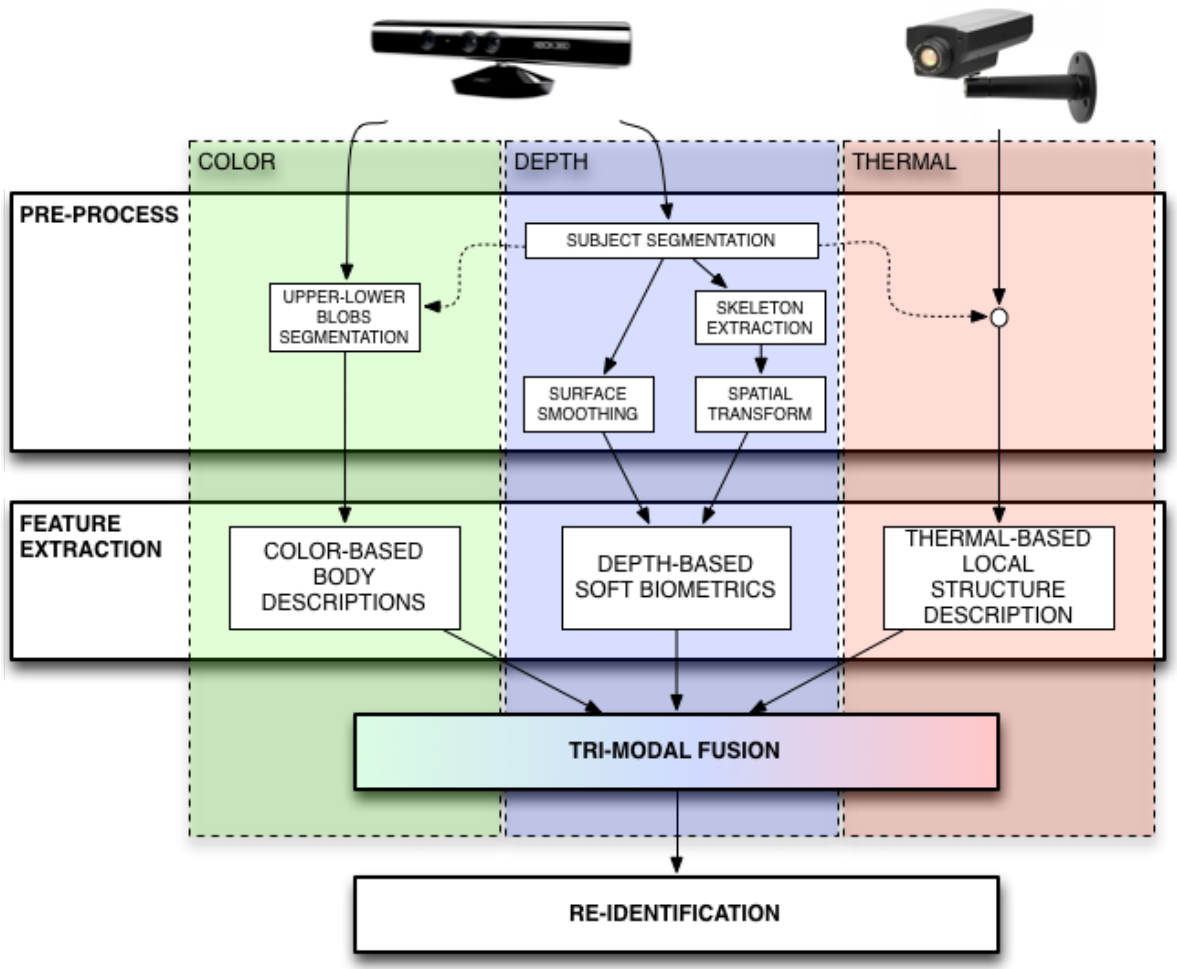


Thermal features

- SURF descriptors extracted within the contour
- Location not taken into account
- Around 150 descriptors per person per frame
- No convenient way to average, so the model consists of all descriptors from the pass



System flow


COLOR
DEPTH
THERMAL
PRE-PROCESS
SUBJECT SEGMENTATION
UPPER-LOWER BLOBS SEGMENTATION
SKELETON EXTRACTION
SURFACE SMOOTHING
SPATIAL TRANSFORM
FEATURE EXTRACTION
COLOR-BASED BODY DESCRIPTIONS
DEPTH-BASED SOFT BIOMETRICS
THERMAL-BASED LOCAL STRUCTURE DESCRIPTION
TRI-MODAL FUSION
RE-IDENTIFICATION

Fusion

- **Id voting**

  Combines the distance measures from each modality:

  $$C(U_1, U_2) = \alpha \cdot \frac{1}{d_{\text{RGB}}(H_1, H_2)} + \beta \cdot \frac{1}{d_{\text{depth}}(D_1, D_2)} + \gamma \cdot d_{\text{thermal}}(S_1, S_2)$$

- **Determine if new**

  If $C \leq T_N \rightarrow$ Subject new
  If $C > T_R \rightarrow$ Re-identify subject
  If $C > T_N \wedge C \leq T_R \rightarrow$ Ignore subject (too close to existing)

- **Re-identify or enroll**

  Assign ID with the highest score, or add model to database

Overall results

- Tested on tri-modal dataset with 35 subjects doing 2 passes each.

	Correct new	Wrong new	Correct ID	Wrong ID	Ignored
Run 1	32	4	21	1	12
Run 2	33	3	21	1	12
Run 3	33	4	24	0	9
Run 4	32	3	23	0	12
Run 5	32	3	19	1	15
Average	**32.4**	**3.4**	**21.6**	**0.6**	**12**
Percentage[1]			**97.3%**	**2.7%**	

Re-identification results.

[1] Percentage of attempted re-identifications.

Results for individual modalities

	Correct new	Wrong new	Correct ID	Wrong ID	Ignored
All modalities	32.4	3.4	21.6	0.6	12
RGB only	33.4	7.8	14.6	0	14.2
Depth only[1]	30.2	20.6	3.2	16	0
Thermal only	33.8	11.6	14	1	9.8

Average re-identification results for individual modalities

[1] With other T_N and T_R thresholds.

- The largest number of correct IDs are obtained with the combined system
- Depth features do more harm than good
- While RGB-only has the least wrong IDs it is a the expense of more ignored subjects and a much larger number of wrong new classifications

Results for individual modalities
Hallway dataset

- Tested on bi-modal dataset with 9 subjects doing 2 passes each.



	Correct new	Wrong new	Correct ID	Wrong ID	Ignored
All modalities	8	0	6	0.4	3.6
RGB only	8	3	2	0	5
Thermal only	9	2	5	0.4	1.6

Average re-identification results for individual modalities, hallway dataset

Thank you