UNIVERSITY OF Southampton
School of Electronics and Computer Science

Content-based Retrieval of Humans using Gait Biometrics

Sina Samangooei
Mark S. Nixon

UNIVERSITY OF Southampton
School of Electronics and Computer Science

Overview

- **Problem outline**
- **Semantic Biometrics**
- **Gait Biometrics**
- **LSA**
- **Results**
- **Conclusions and Future Work↑**

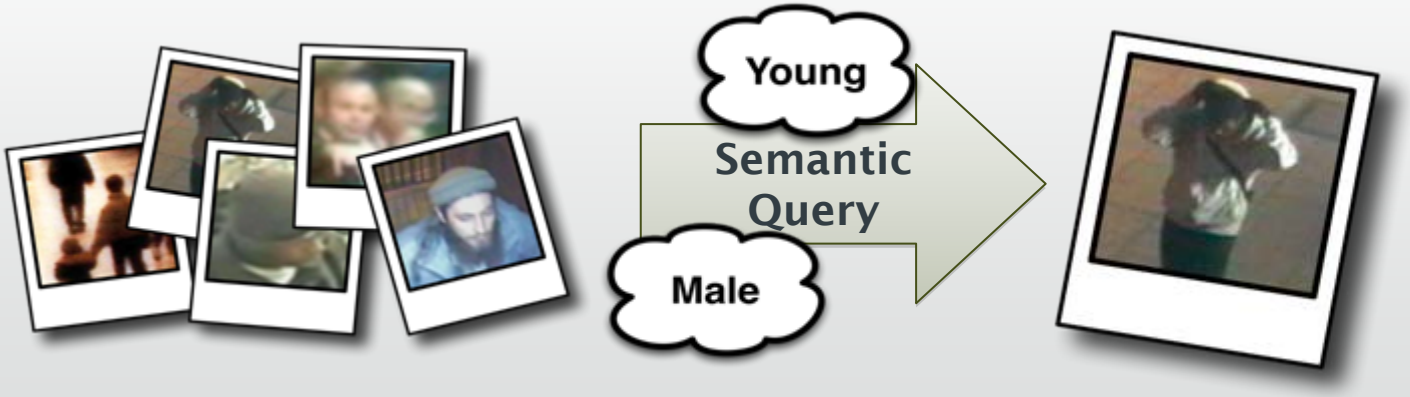
UNIVERSITY OF Southampton
School of Electronics and Computer Science

Problem Outline

UNIVERSITY OF Southampton
School of Electronics and Computer Science

Surveillance

- Increased interest in **surveillance** technologies
  - 3.2GBytes of CCTV footage capture **per hour per camera**
  - Estimated use of **6 million** CCTV cameras in the UK
- **Semantic queries** are an efficient method to facilitate **human exploration**



UNIVERSITY OF Southampton
School of Electronics and Computer Science

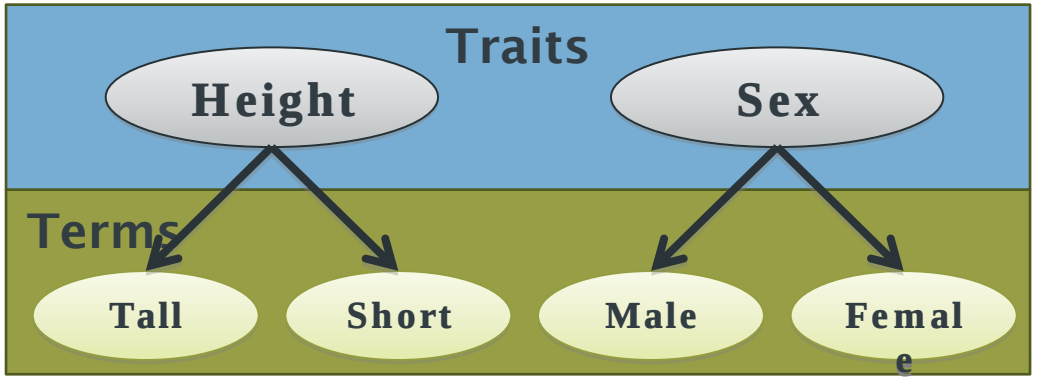
Automatic analysis required

- Videos require **semantic enrichment**
- Can be performed **manually**, however:
  - Too slow for live CCTV
  - Susceptible to various kinds of human error
- Analysis using **Biometrics** and **Automatic** Surveillance
  - Generate **low level** signatures automatically
- These signatures are not inherently **comparable** to human semantic queries
- This is the multimedia **semantic gap** problem

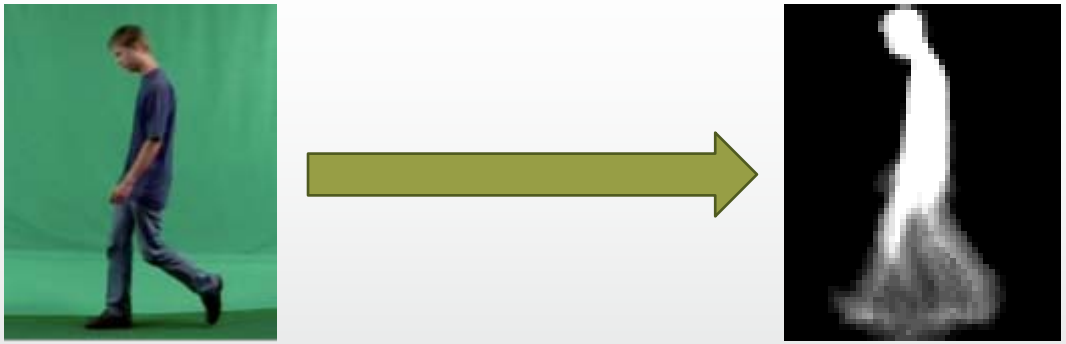
UNIVERSITY OF Southampton
School of Electronics and Computer Science

Exploring Human Descriptions

- We explore semantic descriptions of:
  - … physical **traits**
  - … semantic **terms**
  - … visible at a **distance**



- Find their correlation with **automatic gait signatures**



- In order to facilitate **semantic queries** against **automated gait signatures**

UNIVERSITY OF Southampton
School of Electronics and Computer Science

Semantic Biometrics

UNIVERSITY OF Southampton
School of Electronics and Computer Science

Terms and Traits

- What **traits are described** and what **terms used** depends on **situation**
  - **Mug shots** vs **CCTV**
- **Traits** chosen such that
  - They are visible at a **distance**
  - Mentioned **consistently**
- Complementary **qualitative terms** selected
  - To avoid issues with value judgments

UNIVERSITY OF Southampton
School of Electronics and Computer Science

Traits and Terms

Global Features

- Features mentioned most often in witness statements
- **Sex** and **age** quite simple
- **Ethnicity** undeniably important in terms of mention
  - Notoriously **unstable**
  - There could be anywhere between 3 and 100 ethnic groups
  - We've chosen 3 "main" **subgroups** and 2 extra to match UK Police force groupings

- **Global**
  - Sex
  - Ethnicity
  - Skin Colour
  - Age
- **Body Shape**
  - Figure
  - Weight
  - Muscle Build
  - Height
  - Proportions
  - Shoulder Shape
  - Chest Size
  - Hip size
  - Leg/Arm Length
  - Leg/Arm Thickness
- **Head**
  - Hair Colour
  - Hair Length
  - Facial Hair Colour/Length
  - Neck Length/Thickness

UNIVERSITY OF Southampton
School of Electronics and Computer Science

Traits and Terms

Body Features

- Based on **whole body** description **stability** analysis by MacLeod et al.
  - Features showing **consistency** by different **viewers** looking at the same **subjects**
- Mostly comprised of **5 point** qualitative measures
  - (Very Thin -> Very Fat, Very Short-> Very Long)
- Most likely candidate for **association** with gait

- **Global**
  - Sex
  - Ethnicity
  - Skin Colour
  - Age
- **Body Shape**
  - Figure
  - Weight
  - Muscle Build
  - Height
  - Proportions
  - Shoulder Shape
  - Chest Size
  - Hip size
  - Leg/Arm Length
  - Leg/Arm Thickness
- **Head**
  - Hair Colour
  - Hair Length
  - Facial Hair Colour/Length
  - Neck Length/Thickness

UNIVERSITY OF Southampton
School of Electronics and Computer Science

Traits and Terms

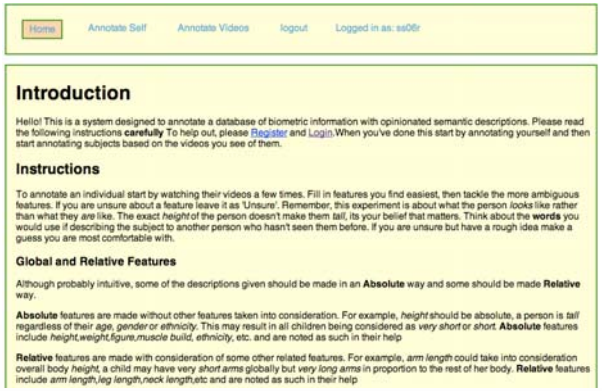
Head Features

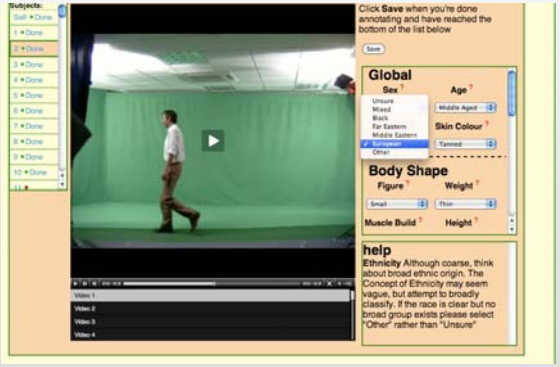
- Mentioned **consistently** by people even at **long distances**
- Prominent area of **gaze**
- **Hair Length** and **colour** inherently connected with **style**
  - **Hundred** of different hair styles
  - **Avoided** due to unfamiliarity of annotators

- **Global**
  - Sex
  - Ethnicity
  - Skin Colour
  - Age
- **Body Shape**
  - Figure
  - Weight
  - Muscle Build
  - Height
  - Proportions
  - Shoulder Shape
  - Chest Size
  - Hip size
  - Leg/Arm Length
  - Leg/Arm Thickness
- **Head**
  - Hair Colour
  - Hair Length
  - Facial Hair Colour/Length
  - Neck Length/Thickness

UNIVERSITY OF Southampton
School of Electronics and Computer Science

Annotation Interface





- **Web interface** constructed to gather annotations against any source
- Designed to deal with **issues of human description**
  - *Memory issues:* View a **subject** as many times as **required**
  - *Defaulting:* Explicitly asked to fill out **every feature**
  - *Value Judgments:* Categorical qualitative values.
  - *Subjective variables:* Collect **description of annotators**

UNIVERSITY OF Southampton
School of Electronics and Computer Science

Gait Biometrics

UNIVERSITY OF Southampton
School of Electronics and Computer Science

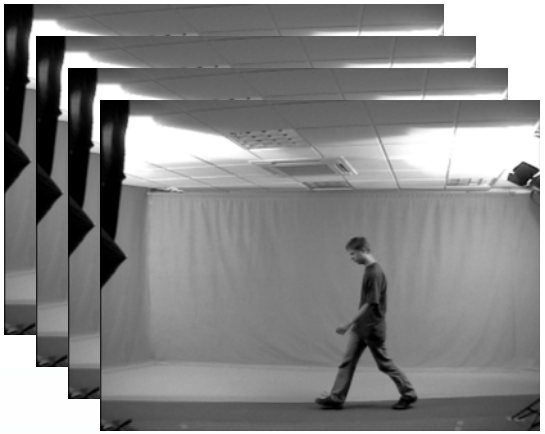
Automatic Gait Signatures

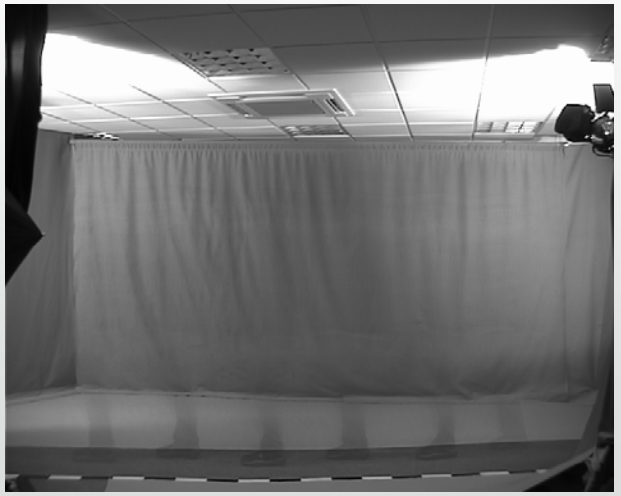
- Gait biometrics **work** under the **constraints** of CCTV
  - Long **distance** to camera, noisy data etc.
- Several **gait signatures** can be generated from video
  - **Statistical** vs **model based**
- Average Silhouette **baseline** algorithm chosen for these preliminary tests
- We use the Southampton dataset
  - **115** subjects
  - At least 6 **fronto-parallel** videos of natural gait

UNIVERSITY OF Southampton
School of Electronics and Computer Science

Average Silhouettes Signature

- Video background calculated using median of each frame
- This background is taken from each frame and pixels thresholded resulting in a binary image
- Normalise silhouettes by height to account for distance
- Add all silhouettes together and divide by the number of frames
- Resulting image is the signature

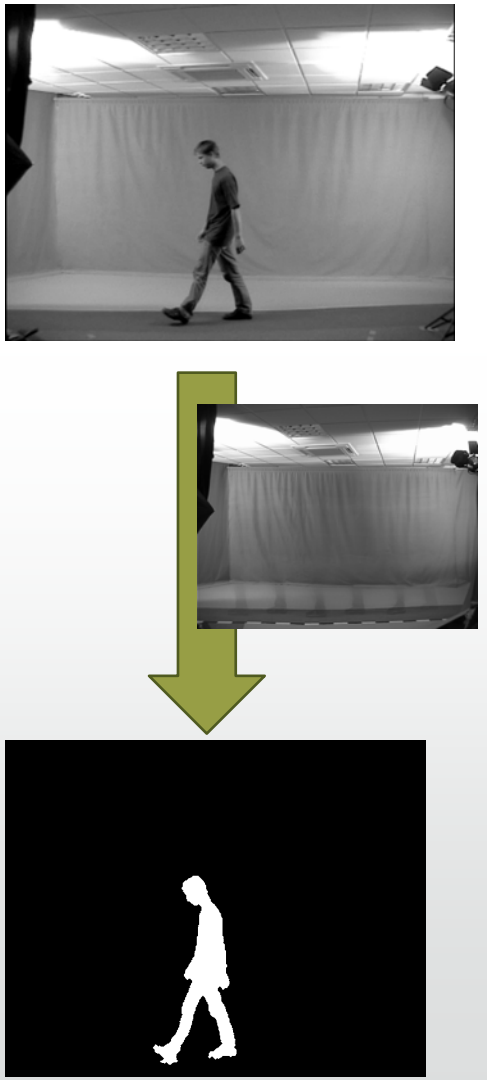




UNIVERSITY OF Southampton
School of Electronics and Computer Science

Average Silhouettes Signature

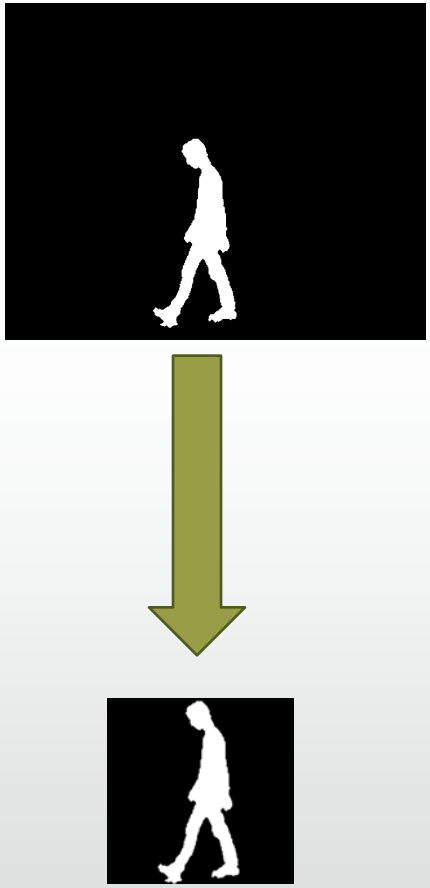
- Video background calculated using median of each frame
- This background is taken from each frame and pixels thresholded resulting in a binary image
- Normalise silhouettes by height to account for distance
- Add all silhouettes together and divide by the number of frames
- Resulting image is the signature



UNIVERSITY OF Southampton
School of Electronics and Computer Science

Average Silhouettes Signature

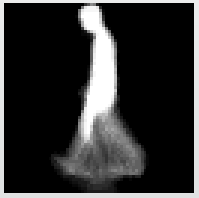
- Video background calculated using median of each frame
- This background is taken from each frame and pixels thresholded resulting in a binary image
- Normalise silhouettes by height to account for distance
- Add all silhouettes together and divide by the number of frames
- Resulting image is the signature



UNIVERSITY OF Southampton
School of Electronics and Computer Science

Average Silhouettes Signature

- Video background calculated using median of each frame
- This background is taken from each frame and pixels thresholded resulting in a binary image
- Normalise silhouettes by height to account for distance
- Add all silhouettes together and divide by the number of frames
- Resulting image is the signature



UNIVERSITY OF Southampton
School of Electronics and Computer Science

Latent Semantic Analysis

UNIVERSITY OF Southampton
School of Electronics and Computer Science

Latent Semantic Analysis

- We use **Latent Semantic Analysis** to facilitate **content based retrieval**
- This process involves formulating **observations of subjects** as a matrix of
  - Documents (**subjects**) with
  - Terms (**annotations** and **average silhouette signatures**)
- Based on **observations**, we construct a linear algebraic **semantic space**
  - The axis of which are the **eigenvectors** of the co-occurrence matrices
  - We ignore axis with lower **eigenvalues**, those likely to represent noise
  - (related to **PCA**)
- In this space similar **geometric position** implies similar **meaning**

UNIVERSITY OF Southampton
School of Electronics and Computer Science

Experiment Design

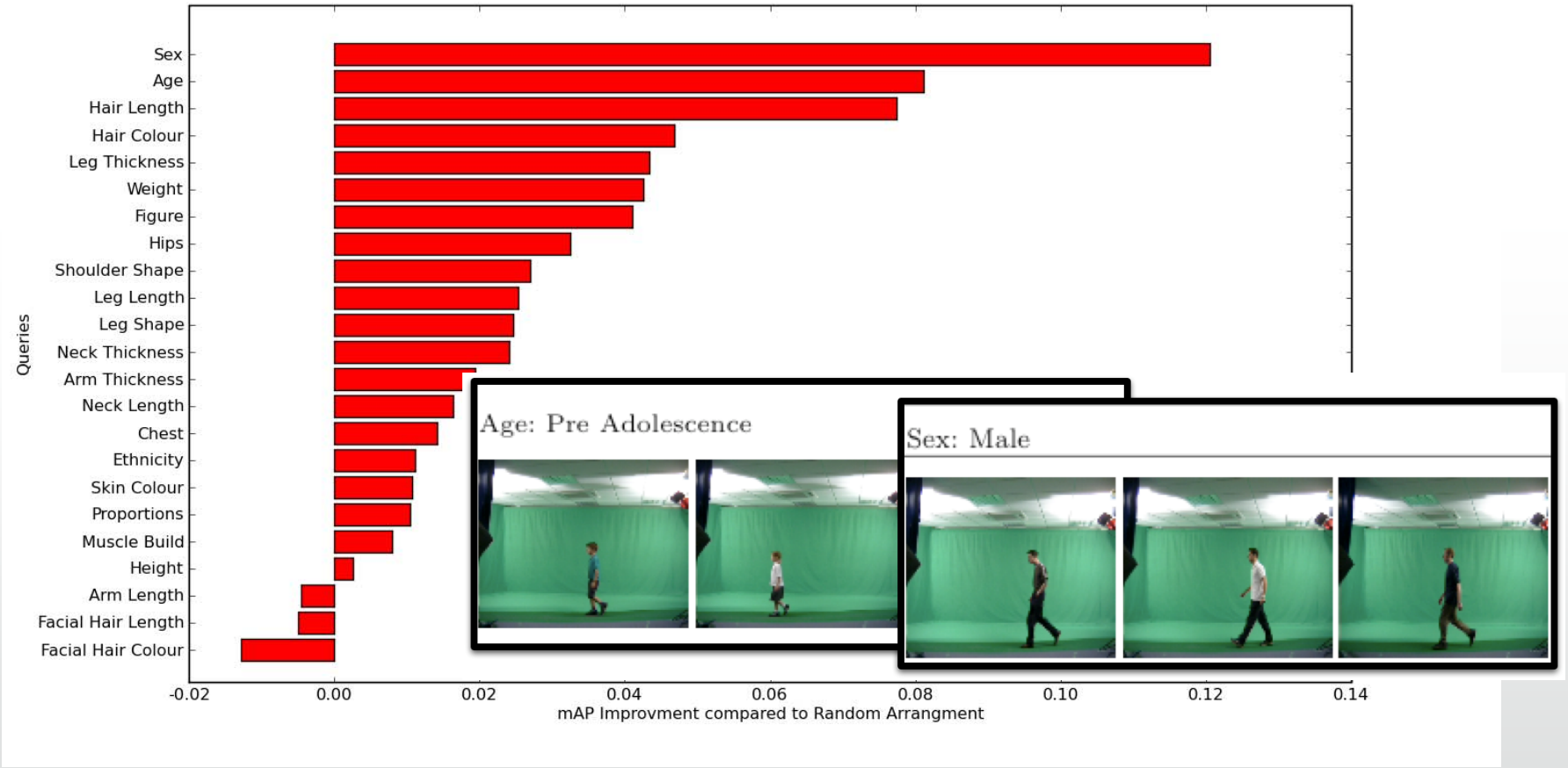
- For each subject
  - Each **video sample** is concatenated with average **semantic annotation**
- Randomly split the set of **subjects** in half
  - **training** and **test**
- Construct **semantic space** from the **training** set
- Project **test** set into the space with **semantic features set to zero**
- Construct a **document** per **semantic annotation** and **project** into the space
- Order subjects by **cosine** distance to query

UNIVERSITY OF Southampton
School of Electronics and Computer Science

Results

UNIVERSITY OF Southampton
School of Electronics and Computer Science

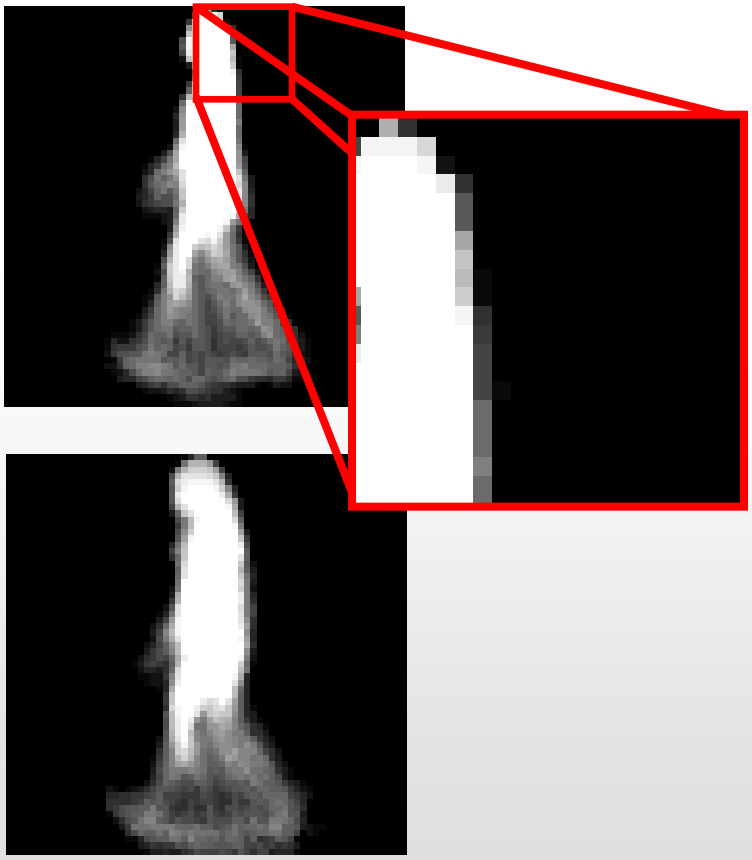
Experiment Results

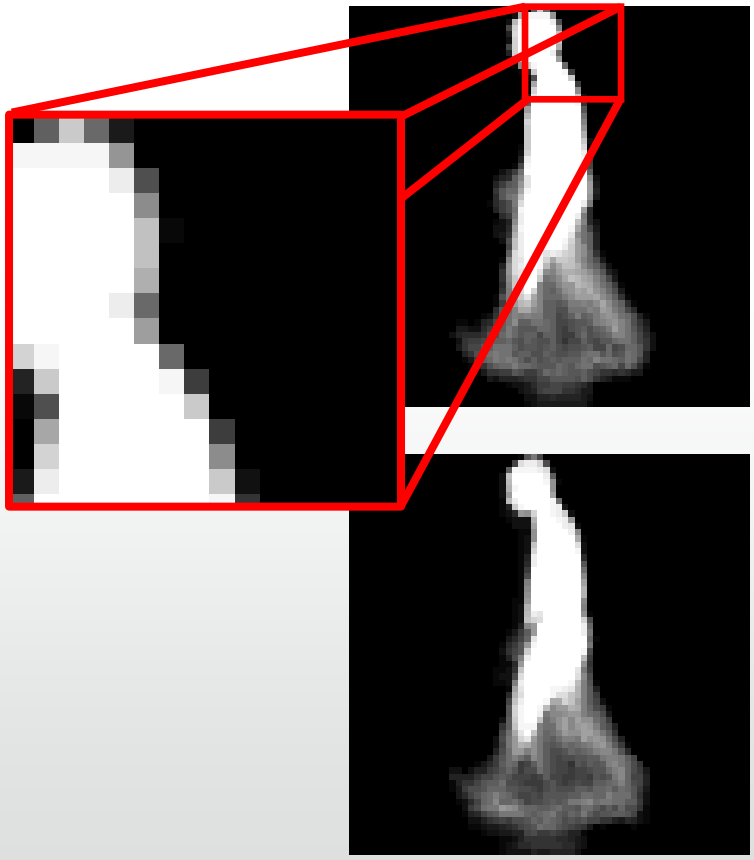


UNIVERSITY OF Southampton
School of Electronics and Computer Science

Successful Results

- Hair Length (Long vs Short)

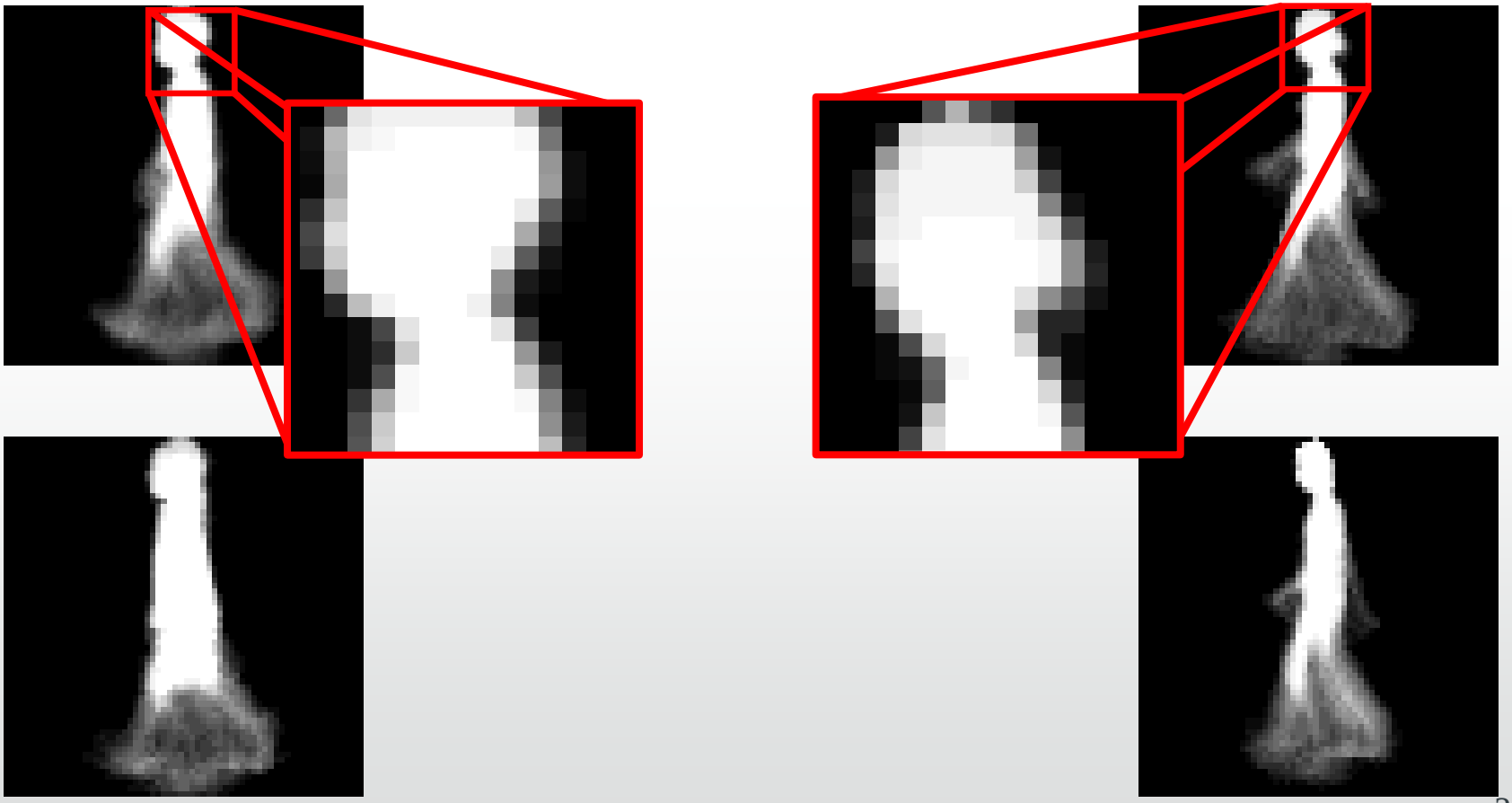




UNIVERSITY OF Southampton
School of Electronics and Computer Science

Successful Results

- Age (Pre-Adolescent vs Young Adult)



UNIVERSITY OF Southampton
School of Electronics and Computer Science

Failed Results

- Skin Colour (Black vs White vs Tanned)

UNIVERSITY OF Southampton
School of Electronics and Computer Science

Conclusions

UNIVERSITY OF Southampton
School of Electronics and Computer Science

Conclusions

- Outlined a set of **features** used by people to describe **each other** at a **distance**
- **Examined correlation** with an automatic gait analysis technique
- Exploited these correlations to allow **content-based retrieval**

UNIVERSITY OF Southampton
School of Electronics and Computer Science

Future Work

- More **Data**!
  - Explore other **automatic features**
  - More **semantic data**
- **Trait** and **Term** expansion
  - Better annotations for **humans**

UNIVERSITY OF Southampton
School of Electronics and Computer Science

Questions?