

Content-based Retrieval of Humans using Gait Biometrics

Sina Samangooei
Mark S. Nixon

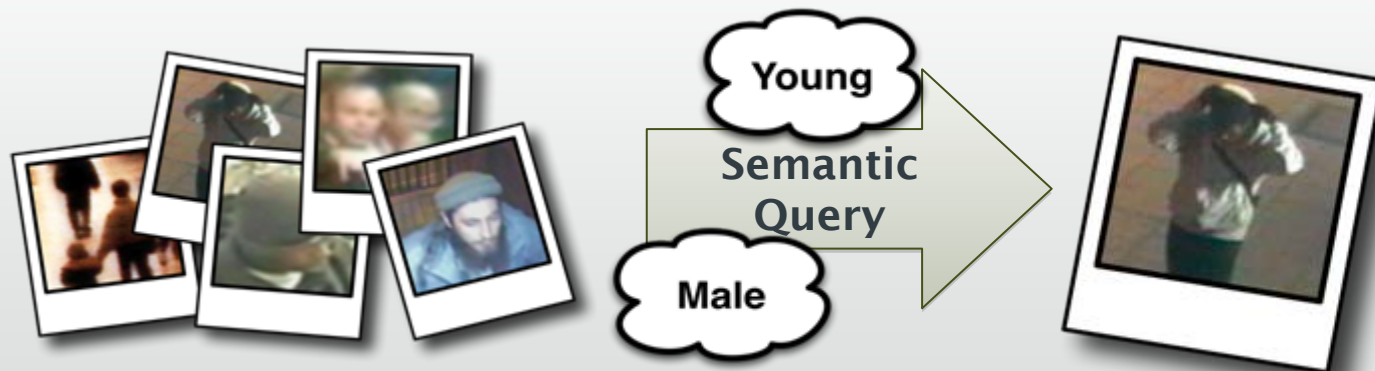
Overview

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

Problem Outline

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**

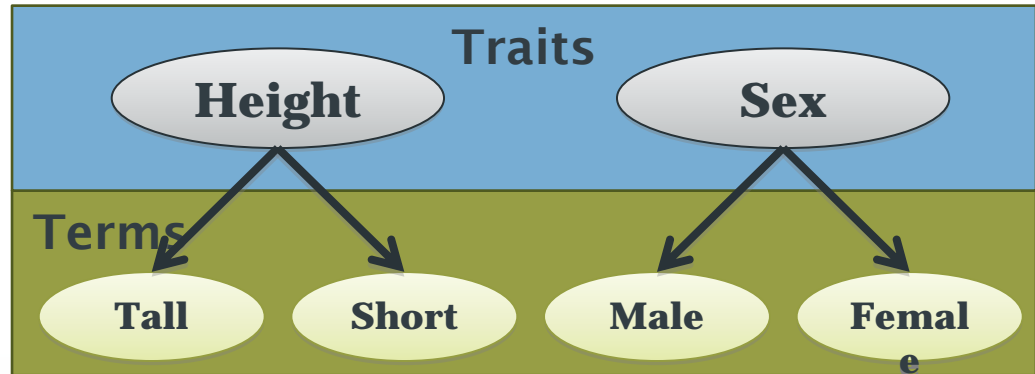


Automatic analysis required

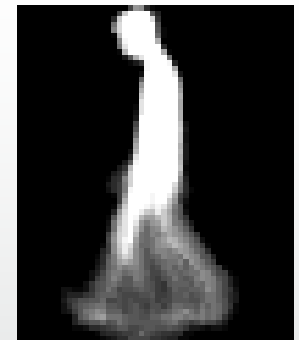
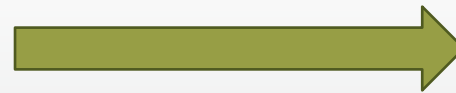
- Videos require **semantic enrichment**
- Can be performed **manually**, however:
 - Too slow for live CCTV
 - Susceptible to various kinds of human error
- Analysis using **Biometrics** and **Automatic Surveillance**
 - Generate **low level** signatures automatically
- These signatures are not inherently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

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**

Semantic Biometrics

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

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



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



Traits and Terms

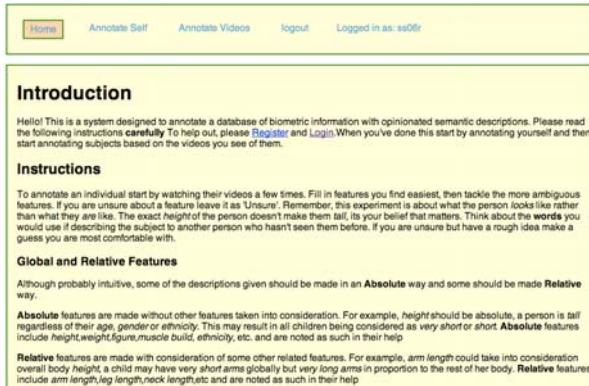
Head Features

- 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**



Annotation Interface



The screenshot shows the top navigation bar with links for Home, Annotate Self, Annotate Videos, Logout, and Logged in as: ss00r. Below this is the 'Introduction' section, which explains the system's purpose and provides instructions on how to use the interface. The 'Instructions' section follows, detailing the process of annotating videos and the types of features to be recorded. A large downward arrow is positioned below the instructions section.

Home Annotate Self Annotate Videos Logout Logged in as: ss00r

Introduction

Hello! This is a system designed to annotate a database of biometric information with opinionated semantic descriptions. Please read the following instructions **carefully** To help out, please **Register** and **Login**. When you've done this start by annotating yourself and then start annotating subjects based on the videos you see of them.

Instructions

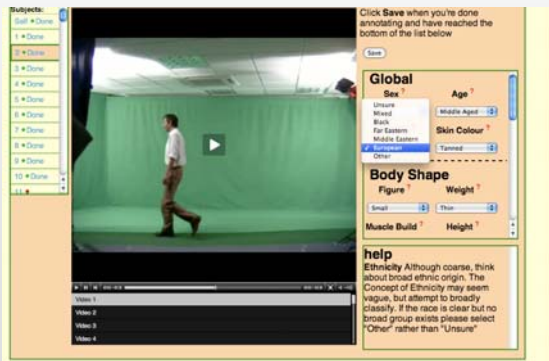
To annotate an individual start by watching their videos a few times. Fill in features you find easiest, then tackle the more ambiguous features. If you are unsure about a feature leave it as 'Unsure'. Remember, this experiment is about what the person looks like rather than what they are like. The exact height of the person doesn't make them tall, to your belief that matters. Think about the **words** you would use if describing the subject to another person who hasn't seen them before. If you are unsure but have a rough idea make a guess you are most comfortable with.

Global and Relative Features

Although probably intuitive, some of the descriptions given should be made in an **Absolute** way and some should be made **Relative** way.

Absolute features are made without other features taken into consideration. For example, height should be absolute, a person is tall regardless of their age, gender or ethnicity. This may result in all children being considered as very short or short. **Absolute** features include height, weight, figure, muscle build, ethnicity, etc. and are noted as such in their help.

Relative features are made with consideration of some other related features. For example, arm length could take into consideration overall body height, a child may have very short arms globally but very long arms in proportion to the rest of her body. **Relative** features include arm length, leg length, neck length, etc and are noted as such in their help.



The screenshot shows the annotation interface. On the left, there is a list of subjects with checkboxes. The main area is a video player showing a person walking in a green screen environment. On the right, there is a form for entering biometric data. The form has sections for 'Global' (Sex, Age, Skin Colour, Ethnicity), 'Body Shape' (Figure, Weight, Muscle Build, Height), and a 'help' section. A 'Save' button is located above the form.

Subjects: 10 #Done

Click Save when you're done annotating and have reached the bottom of the list below

Save

Global

Sex ? Age ?

Unsure Middle Aged -2

Black Skin Colour ?

Far Eastern Tanned -5

Middle Eastern

Other

Body Shape

Figure ? Weight ?

Small -3 Thin -2

Muscle Build ? Height ?

help

!Ethnicity Although coarse, think about broad ethnic origin. The Concept of Ethnicity may seem vague, but attempt to broadly classify. If the race is clear but no broad group exists please select "Other" rather than "Unsure"

- **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**

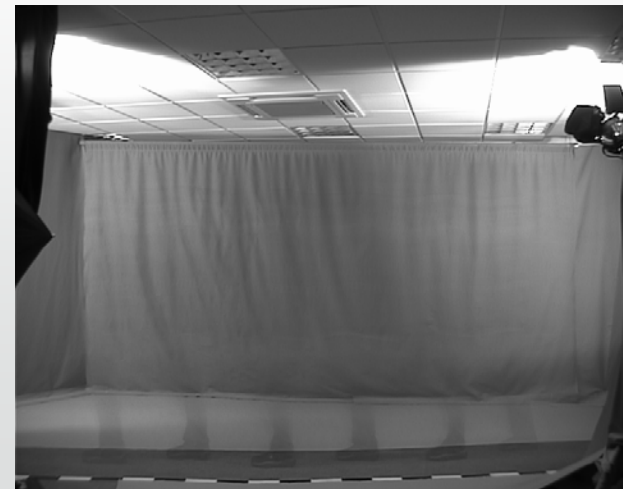
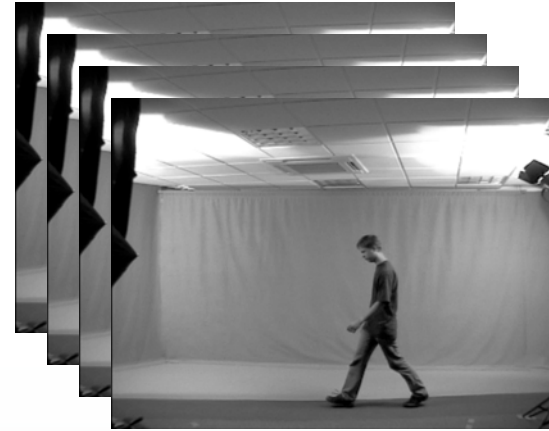
Gait Biometrics

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

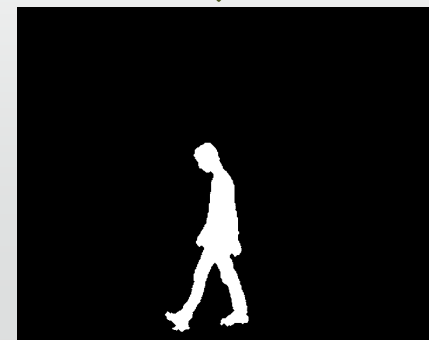
Average Silhouettes Signature

- 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



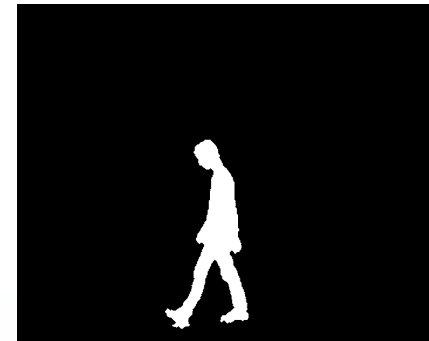
Average Silhouettes Signature

- 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



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



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**



Latent Semantic Analysis

Latent Semantic Analysis

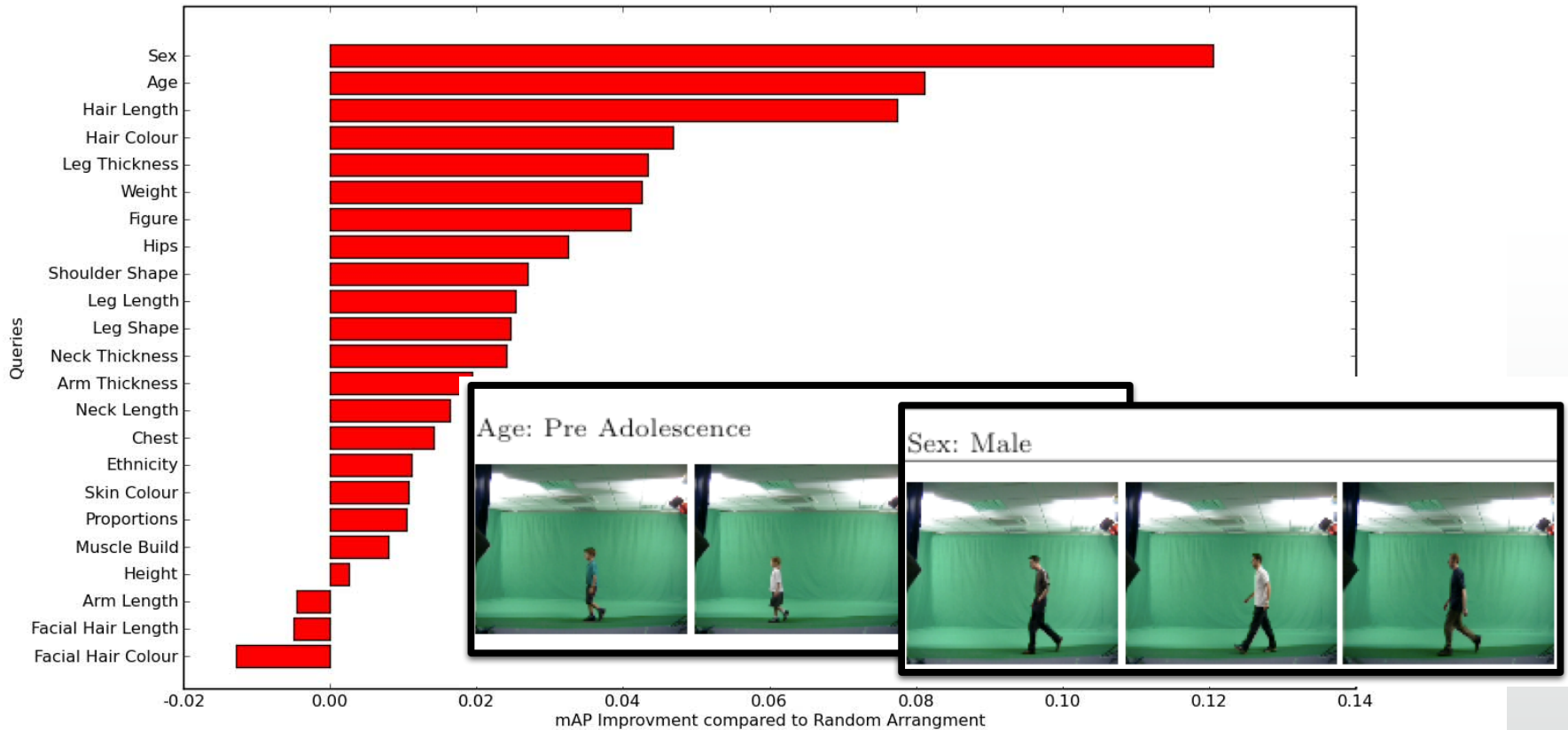
- We use **Latent Semantic Analysis** to facilitate **content based retrieval**
- This process involves formulating **observations of subjects** as a matrix of
 - Documents (**subjects**) with
 - Terms (**annotations** and **average silhouette signatures**)
- Based on **observations**, we construct a 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**

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

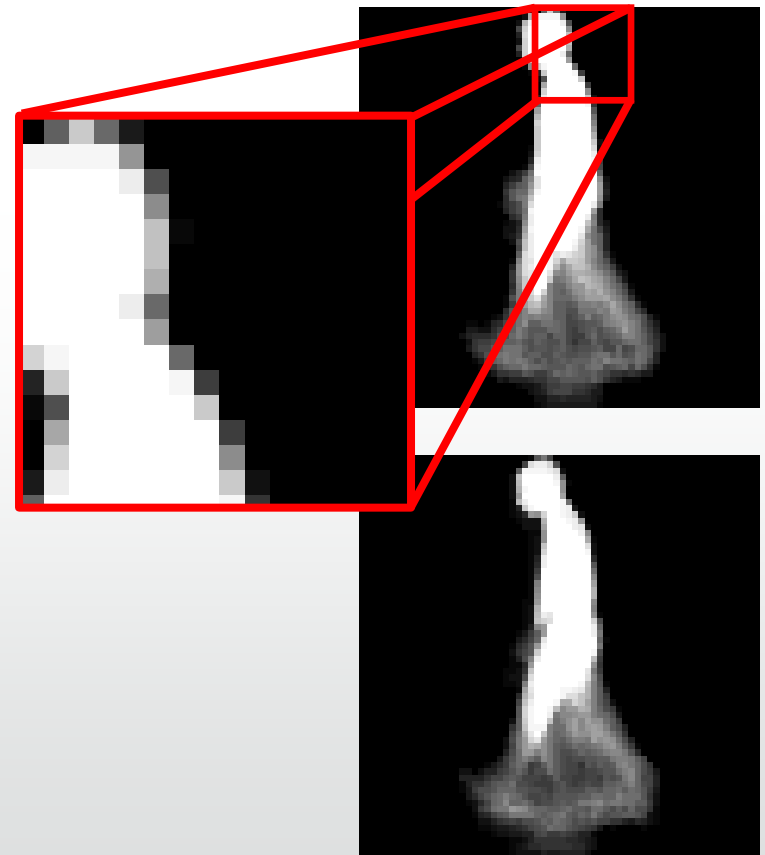
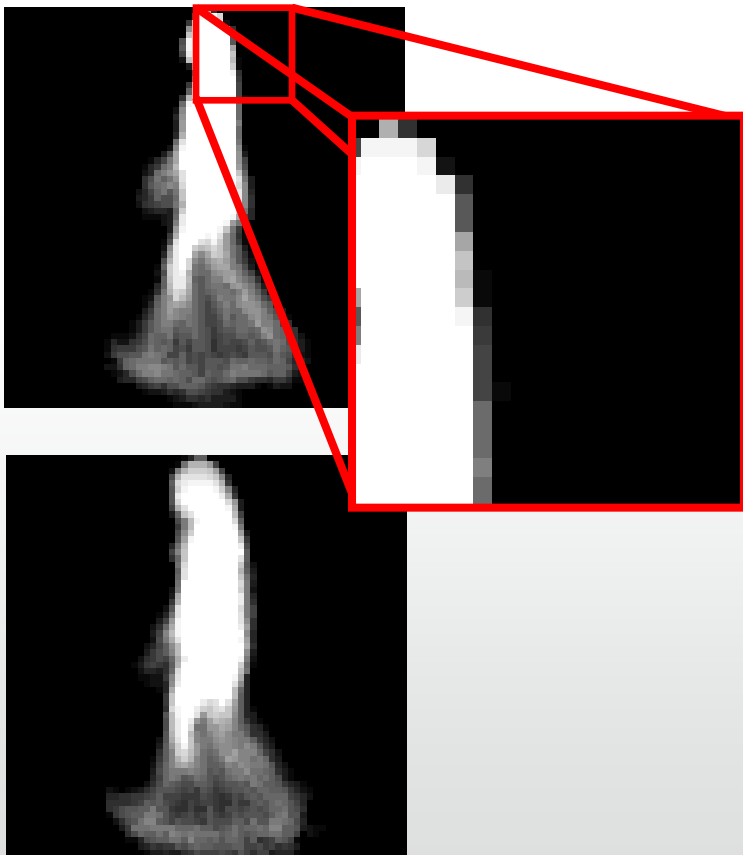
Results

Experiment Results



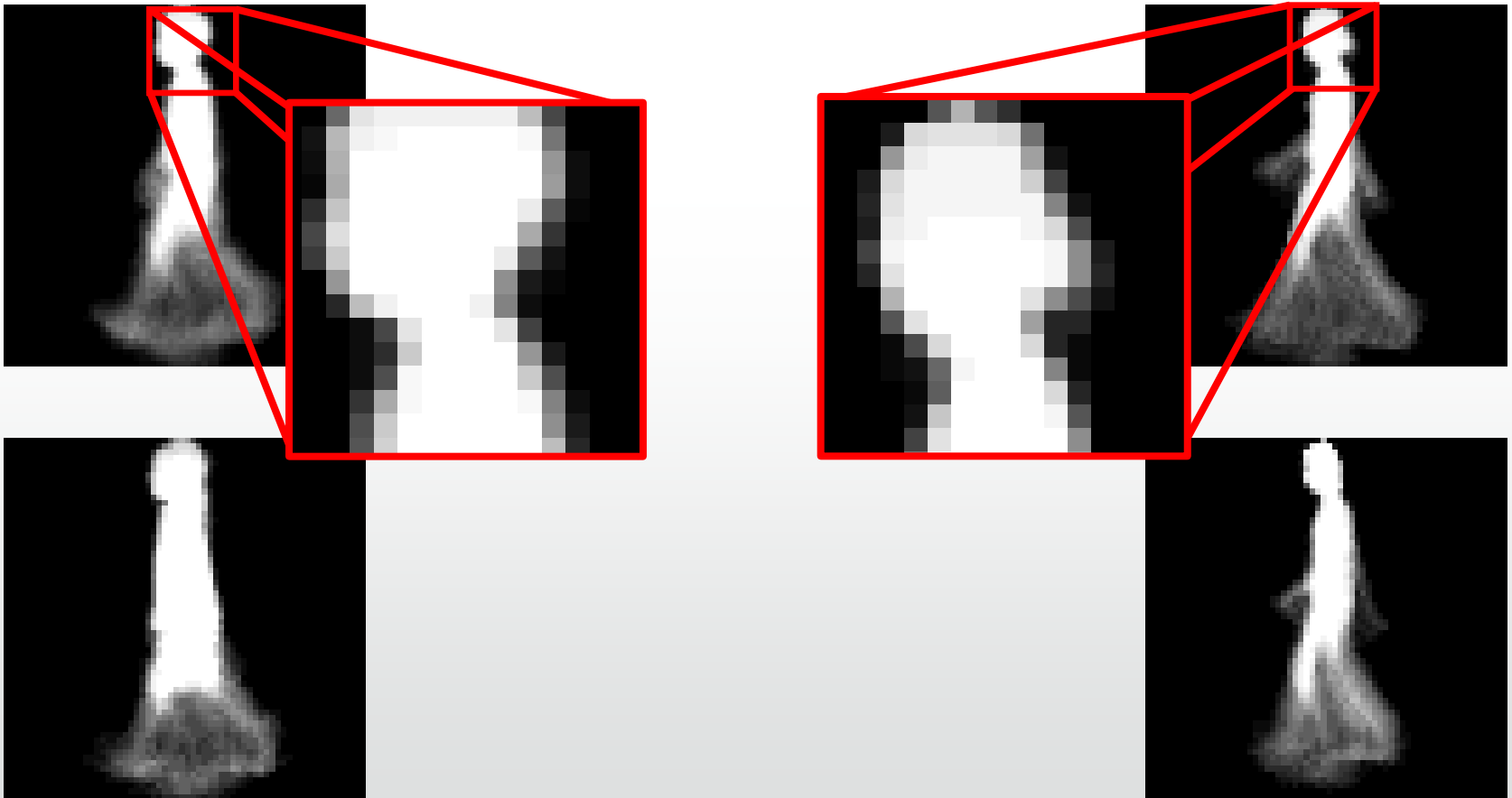
Successful Results

- Hair Length (Long vs Short)



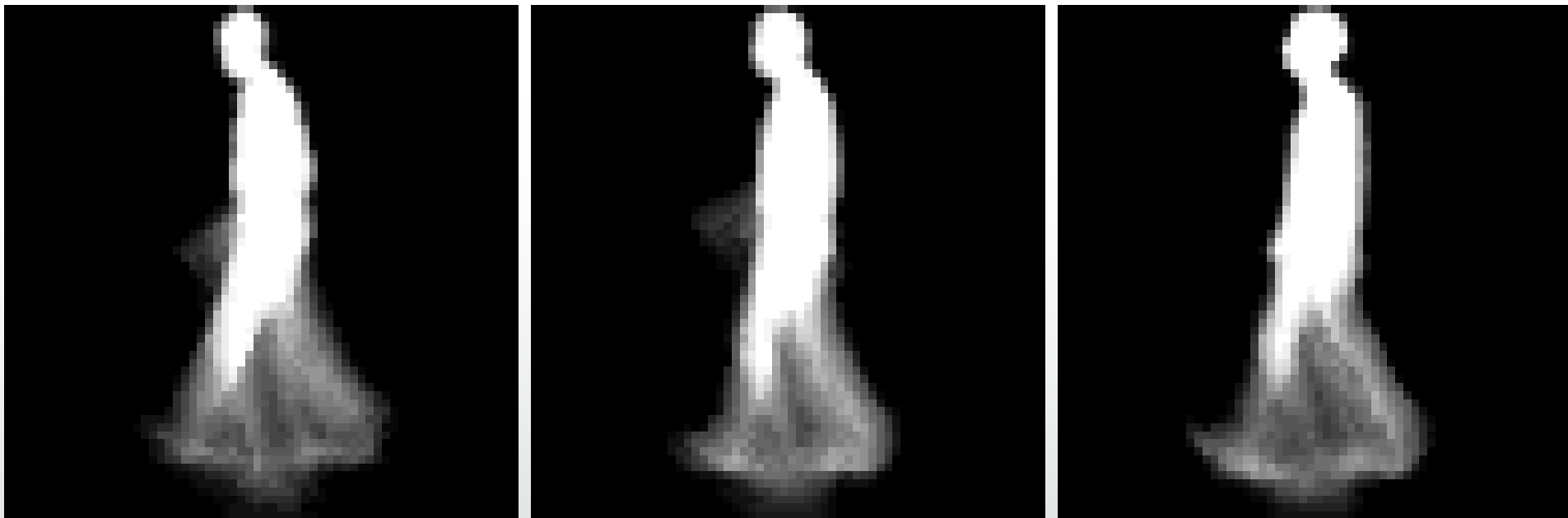
Successful Results

- Age (Pre-Adolescent vs Young Adult)



Failed Results

- Skin Colour (Black vs White vs Tanned)



Conclusions

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**

Future Work

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



Questions?