People Reidentification Using Omnidirectional Cameras

J.Silla, A.Albiol, Al. Albiol, C.Sandoval
ICDP-2016

juliasilla@gmail.com

CONTENTS

INTRODUCTION

An important problem that appears in multi-camera systems is to re-identify people on a network of non overlapping cameras.

This is a key **challenge** for tracking applications

Some applications of people re-identification are:

- Locating the security personnel in a public space
- Determining the time people spends in a shop
- Finding suspicious people on a network of cameras
- Finding lost children in a shopping mall
- Locate people wearing a specific uniform
- ...

INTRODUCTION

There are different methods for re-identify people, some of them are based on the clothing appearance of the individuals.

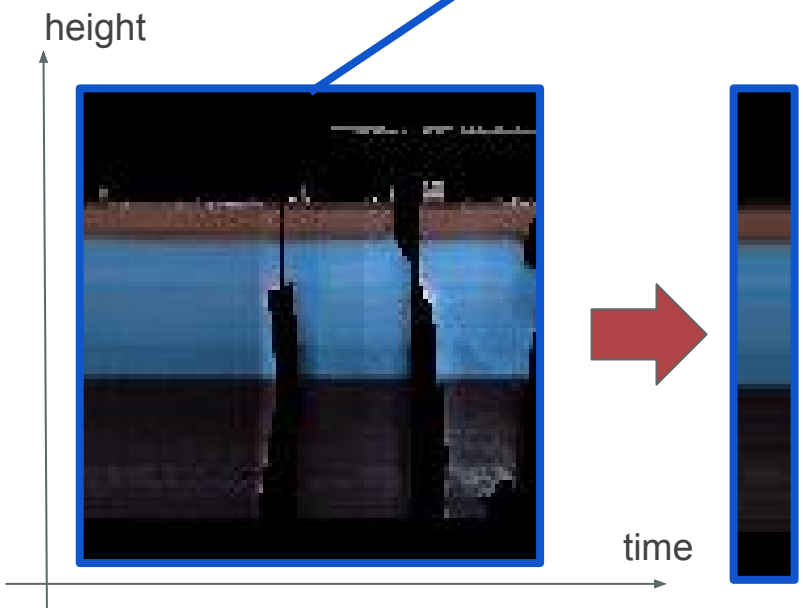
We based our study in a previous paper* that uses a 3D camera (Microsoft Kinect) to extract a descriptor of people based on the **clothing appearance.** This descriptor is known as **bodyprint**

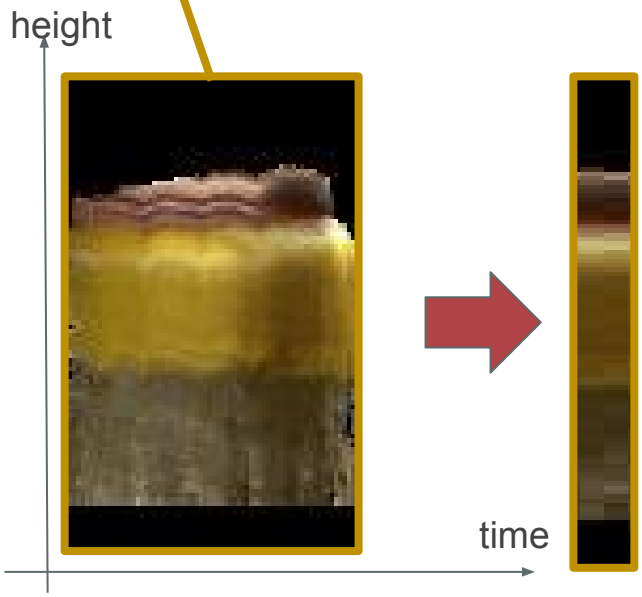
The bodyprint acts as a fingerprint in the sense that it is sufficiently distinctive to allow people discrimination.

They use of all the pixels associated to a person along time, known as **temporal signature**, is necessary because, due to occlusions, not all body parts are visible in all frames. Thus, the body print is the summary of the information of a temporal signature.

**A. Albiol, A. Albiol, J. Oliver, and J. M. Mossi. Who is who at different cameras: people re-identification using depth cameras. IET Computer Vision, 6(5):378–387, Sept 2012.*

INTRODUCTION





INTRODUCTION

Using the kinect sensor to extract bodyprints has some **advantages**:

- Kinect provides **3D** information with high accuracy
- It is possible to deduce the **height of the subject** and associate each height with color information

There are significant **drawbacks**:

- It does **not** work in **outdoors** environments
- It does **not** cover a **large area**

To address this problems of the Kinect sensor we have developed an alternative method for people re-identification using **omnidirectional cameras**

LOCATION OF PEOPLE

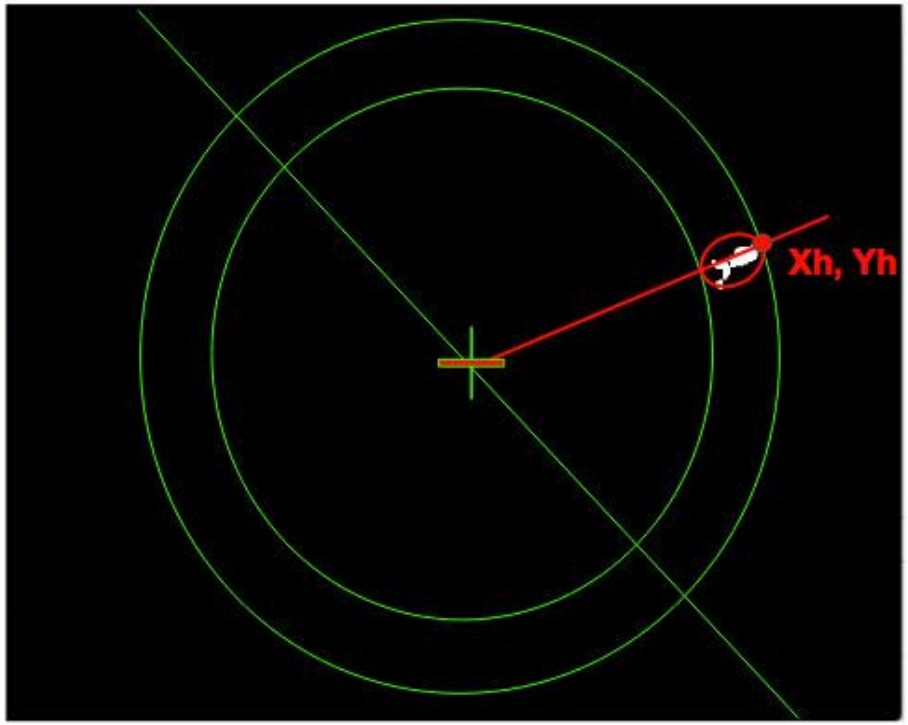
The first step of the method for bodyprint extraction and matching with omnidirectional cameras is to determine the **location of people**.

We use a basic **motion detection** and **tracking** algorithm

We locate **heads** --> the pixel of a blob that is further from the image center is the head of the individual

LOCATION OF PEOPLE

Head detector



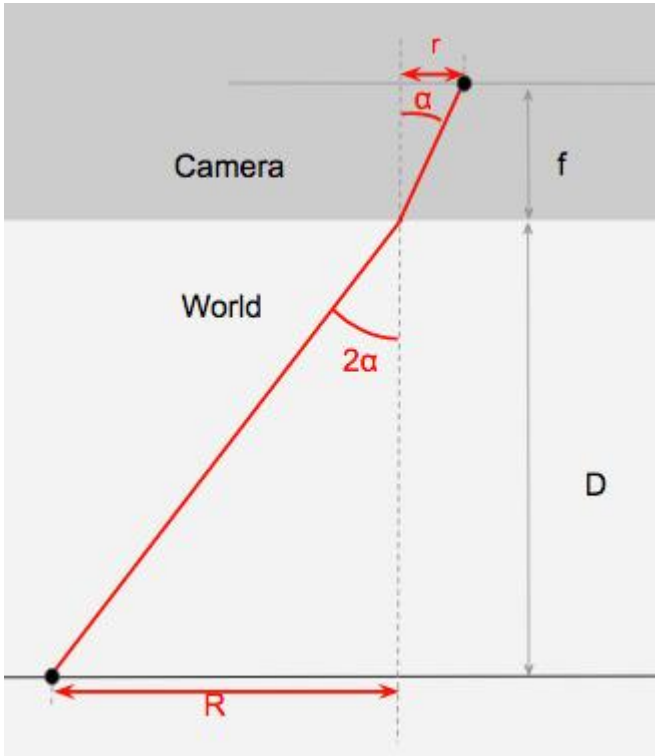
CAMERA MODEL

The camera model is the projection model of the fisheye lens that maps one point of the image with a ray in the space.

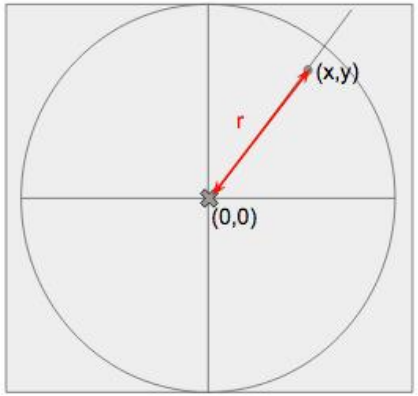
The distance of a point from the image center is dependent on the focal length of the optical system, and the angle from the optical axis

M3007 AXIS omnidirectional camera - **Stereographic** model.

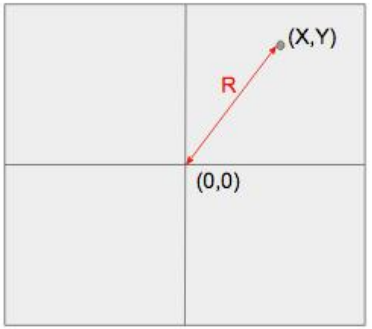
CAMERA MODEL



CAMERA COORDINATES



WORLD COORDINATES



$$r = \sqrt{x^2 + y^2}$$

$$\alpha = 2 \cdot \text{atan}(r/f)$$

$$R = (h_{cam} - h_{head}) \cdot \tan(\alpha)$$

$$X = x/r \cdot R$$

$$Y = y/r \cdot R$$

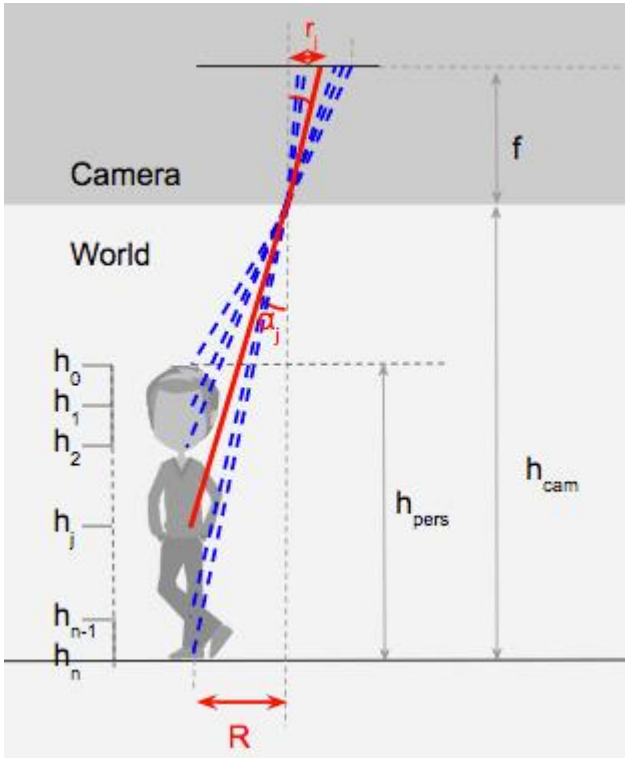
$$Z = 1.70 \quad \text{meters}$$

TEMPORAL SIGNATURES

Head coordinates (world): (X,Y,Z_h)

Foot coordinates (world): $(X,Y,0)$

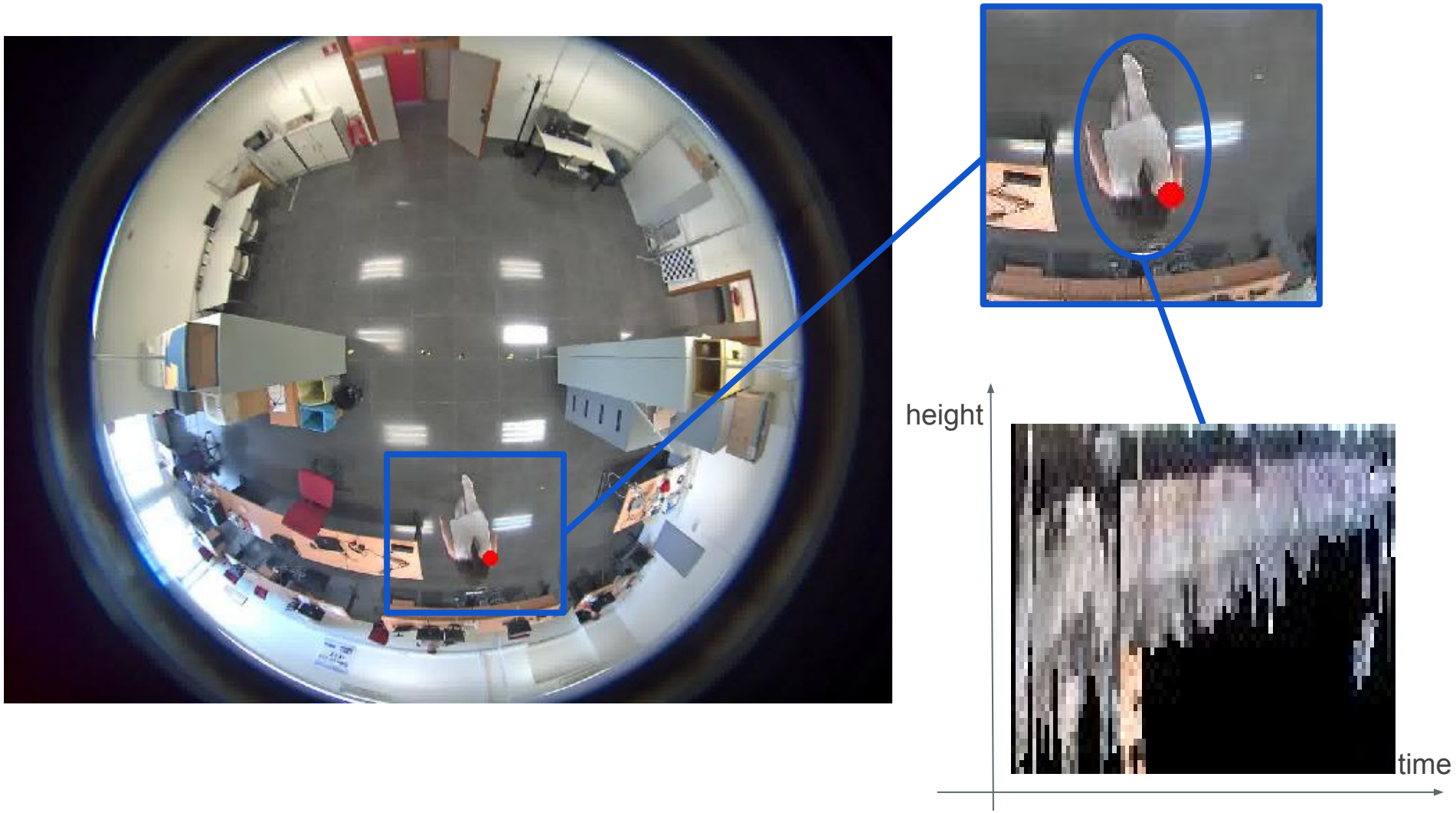
For a point at hj (X,Y,h_j), the value of r_j can be obtained by:



$$\alpha_j = \text{atan}\left(\frac{h_{cam} - h_j}{R}\right)$$

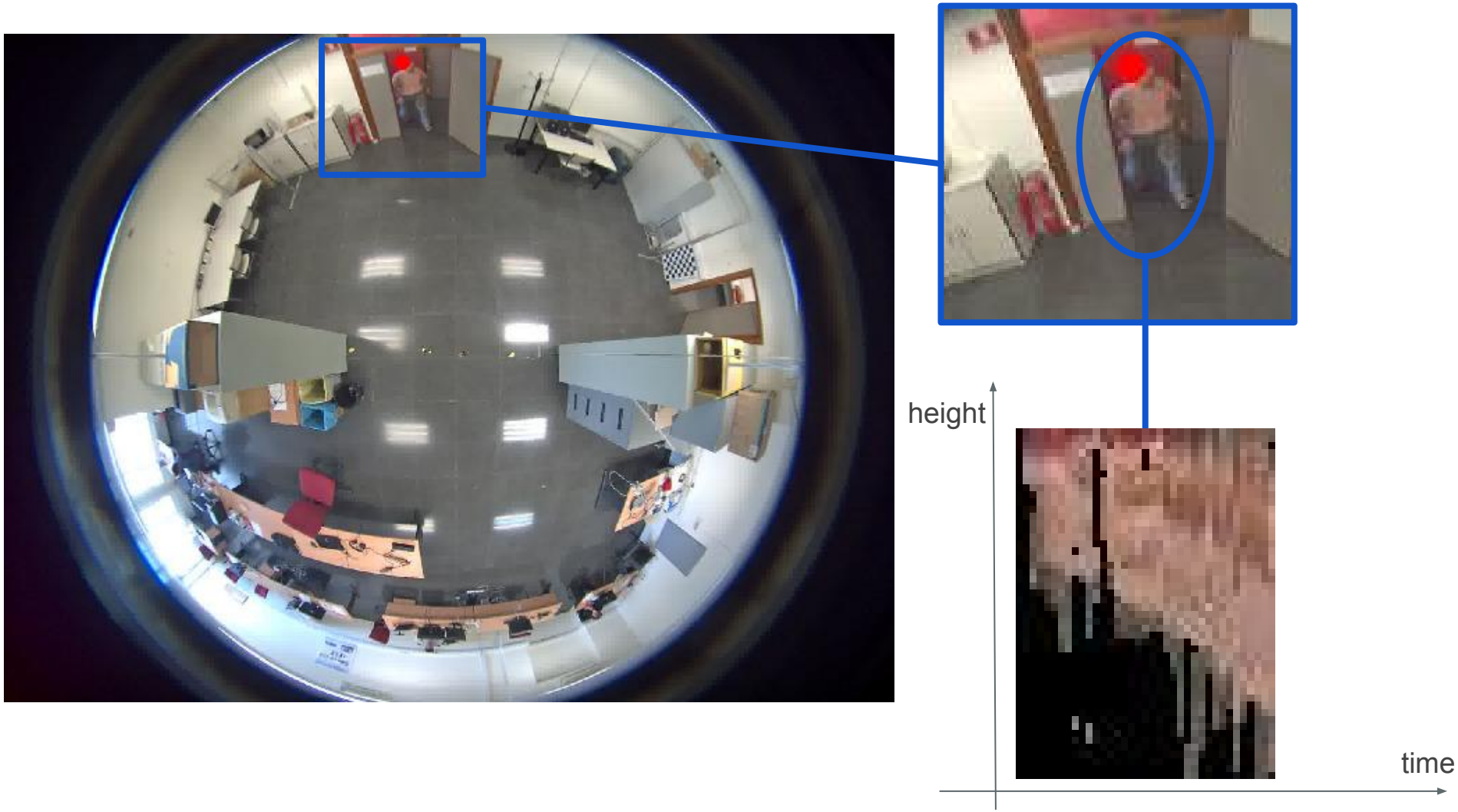
$$r_j = f \cdot \tan(\alpha_j/2)$$

TEMPORAL SIGNATURES

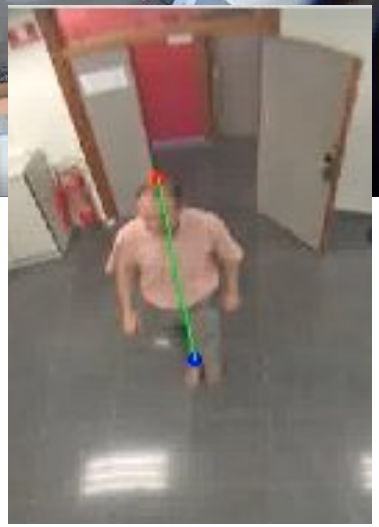


TEMPORAL SIGNATURES

TEMPORAL SIGNATURES



TEMPORAL SIGNATURES



TEMPORAL SIGNATURES

The line of the waist is **not horizontal** along time, but it is descending or ascending

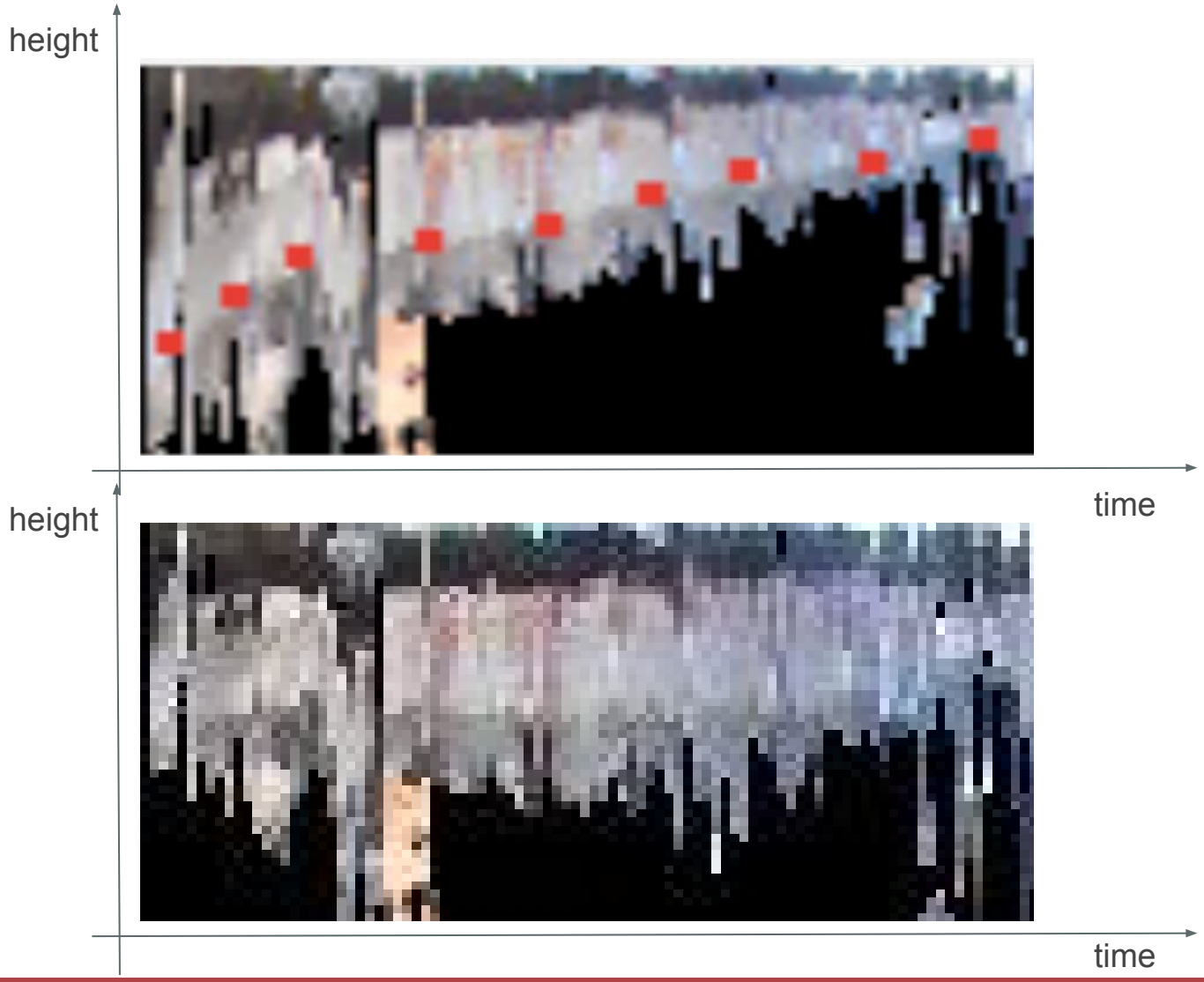
This is an effect of the camera perspective

It is necessary to **compensate this effect** before extracting the bodyprint. Otherwise the average of the temporal signature along time will be done using pixels belonging to different heights.

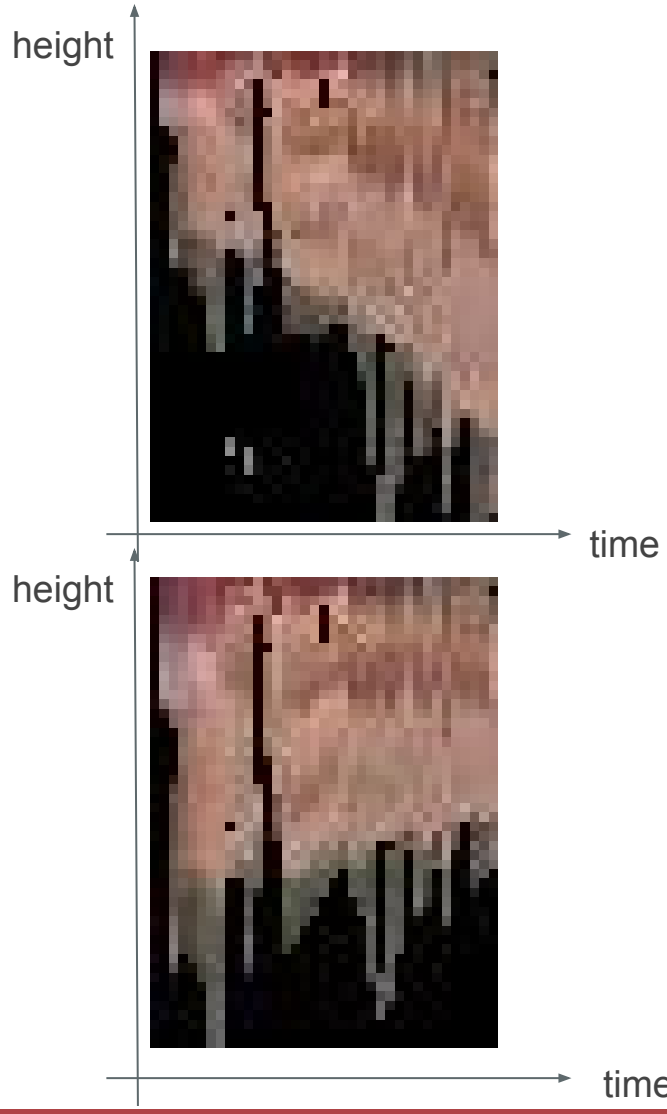
The deviation is a function of the distance to the image center

To compensate it, we estimate the polinomial that better fits the slope of the bodyprint

TEMPORAL SIGNATURES



TEMPORAL SIGNATURES



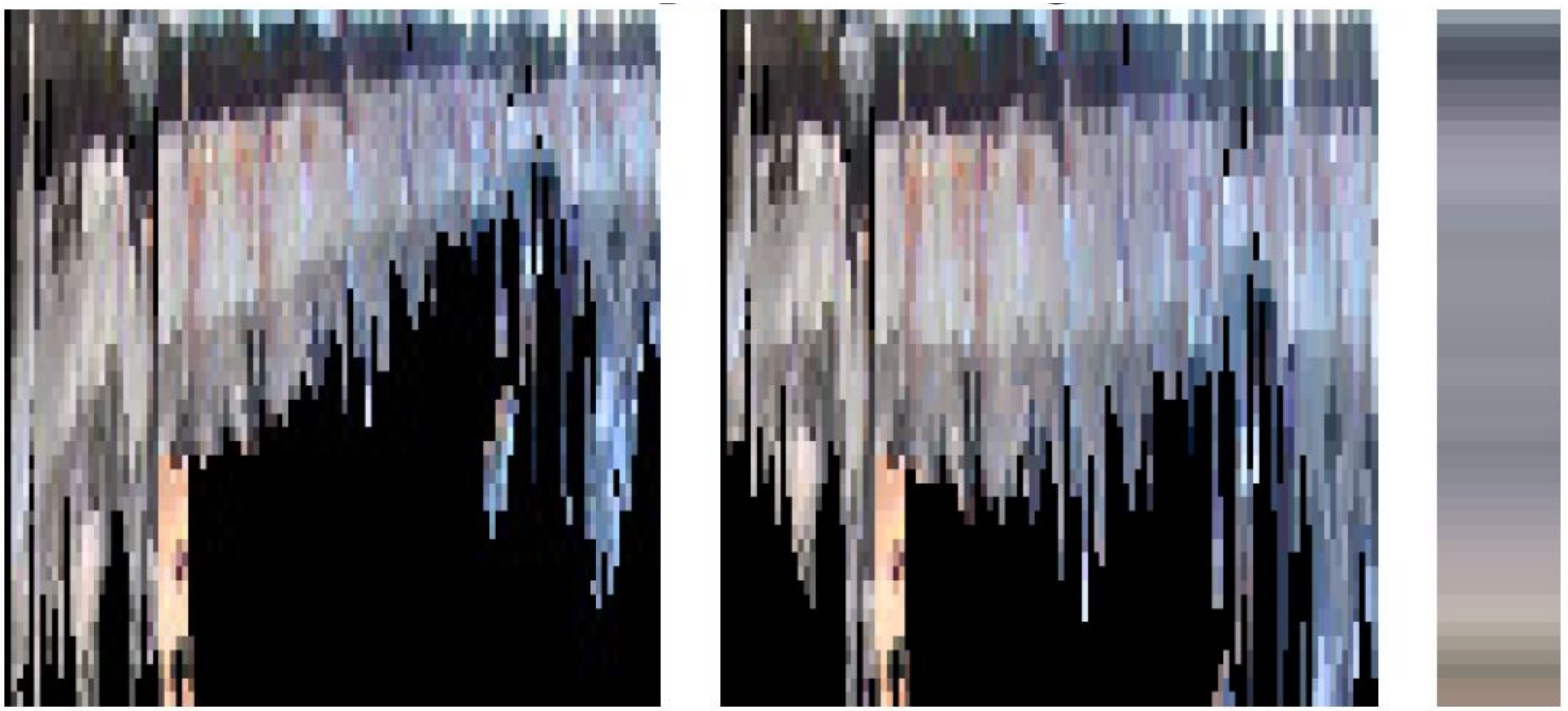
BODYPRINT EXTRACTION

The bodyprint is the resume of the temporal signature into a vector that is the average of the pixels along time

We discard black pixels since they belong to those areas where the value of the signature could not be obtained (due to occlussions)

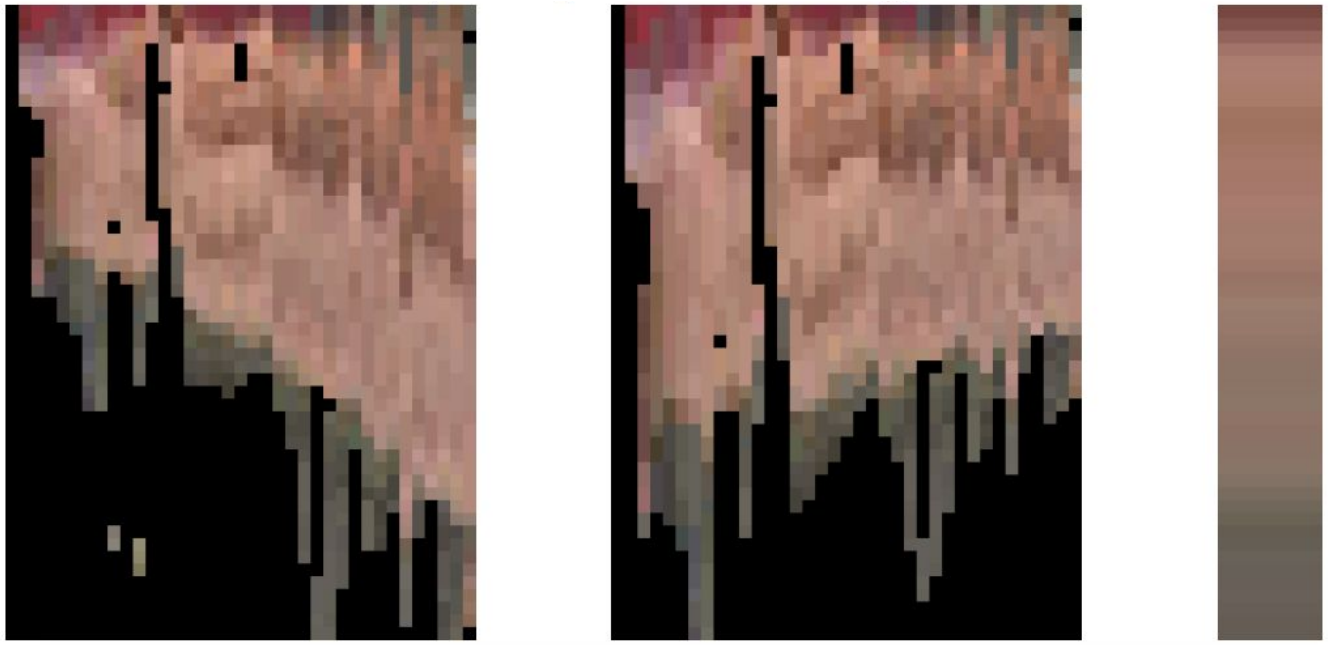
BODYPRINT EXTRACTION

Example



BODYPRINT EXTRACTION

Example



MATCHING METRIC

We have chosen as matching metric the **maximum euclidean distance of the RGB** components of the bodyprint vectors

$$d_R = \sqrt{\sum_l (\overline{\mathbf{BP}}_{\mathbf{1R}}(l) - \overline{\mathbf{BP}}_{\mathbf{2R}}(l))^2}$$

$$d_G = \sqrt{\sum_l (\overline{\mathbf{BP}}_{\mathbf{1G}}(l) - \overline{\mathbf{BP}}_{\mathbf{2G}}(l))^2}$$

$$d_B = \sqrt{\sum_l (\overline{\mathbf{BP}}_{\mathbf{1B}}(l) - \overline{\mathbf{BP}}_{\mathbf{2B}}(l))^2}$$

$$d = \max\{d_R, d_G, d_B\}$$

- l: length of the bodyprint,
- d_R, d_G and d_B: distances of each R, G and B color components
- BP_{iR}, BP_{iG} and BP_{iB} are the R, G and B components of the bodyprint vector i.

RESULTS

The scenario has the following characteristics:

- Presence of areas with different illumination: natural and artificial light
- People at any position: frontal, lateral, rear views
- Occlussions due to furniture (tables and chairs)
- The camera height is 3.20 meters
- The bodyprint height 50 pixels
- The area we cover is $5m^2$

RESULTS

We compared each bodyprint vector with the rest of bodyprints in our database .

We consider a correct match when the bodyprint that has the smallest distance to the one we are comparing with belongs to the same person.

The performance obtained is **87%**

Some samples of our database that are not matched correctly belong to tracks that are in the periphery of the image and have very short duration (less than 1 second).

Using the track duration to discard samples: **>90%** performance

CONCLUSIONS AND FUTURE WORK

- This is a preliminary study of a method for people re-identification, based on bodyprint extraction and comparison with omnidirectional cameras
- The results (87% performance) show that it is possible to re-identify people using omnidirectional cameras, leaving room for future improvements of the topic
- The conditions of the scenario include areas with different illumination, frontal, lateral and rear views of people and occlusions.
- We are covering a large area simulating a real scenario for people reidentification (retail, public spaces).

CONCLUSIONS AND FUTURE WORK

- These results should be evaluated on a larger database of people and on different scenarios.
- We made some assumptions, such as a fixed person height, camera pointing vertically to the ground. It would be interesting to anlayze how changing those values affects the performance.
- The matching method can also be improved by comparing bodyprints that have similar positions/trajectories