IMAGE SPLICING DETECTION BASED ON GENERAL PERSPECTIVE CONSTRAINTS

M. Iuliani, G. Fabbri, **A. Piva**

Dept. of Information Engineering
University of Florence (Italy)
alessandro.piva@unifi.it

WIFS 2015, 16-19 November, Rome

Outline

Introduction

Proposed Method

Experimental Results

Conclusions & Questions

Introduction

Inherent traces are left behind in a digital image when image splicing is applied:

- **Signal Level Traces**
  - Splicing can leave demosaicking artefact, compression artefact, etc …
- **Physical Level Traces**

it is difficult to properly place the spliced object

- inconsistencies in lights
- inconsistencies in reflections
- inconsistencies in geometrical perspective

Introduction

Inherent traces are left behind in a digital image when image splicing is applied::

- **Signal Level Traces**
  - Splicing can leave demosaicking artefact, compression artefact, etc …
- **Physical Level Traces**

Pros :

- Robust to filtering, heavy compression & resizing

Cons:

- Require user interaction
- Work on very limited scenarios

State of the Art

- Yao et al. proposed a method to determine whether two subjects placed on the same plane have respective sizes satisfying perspective rules.
- by estimating the ratio of their height

Deviations from the expected value

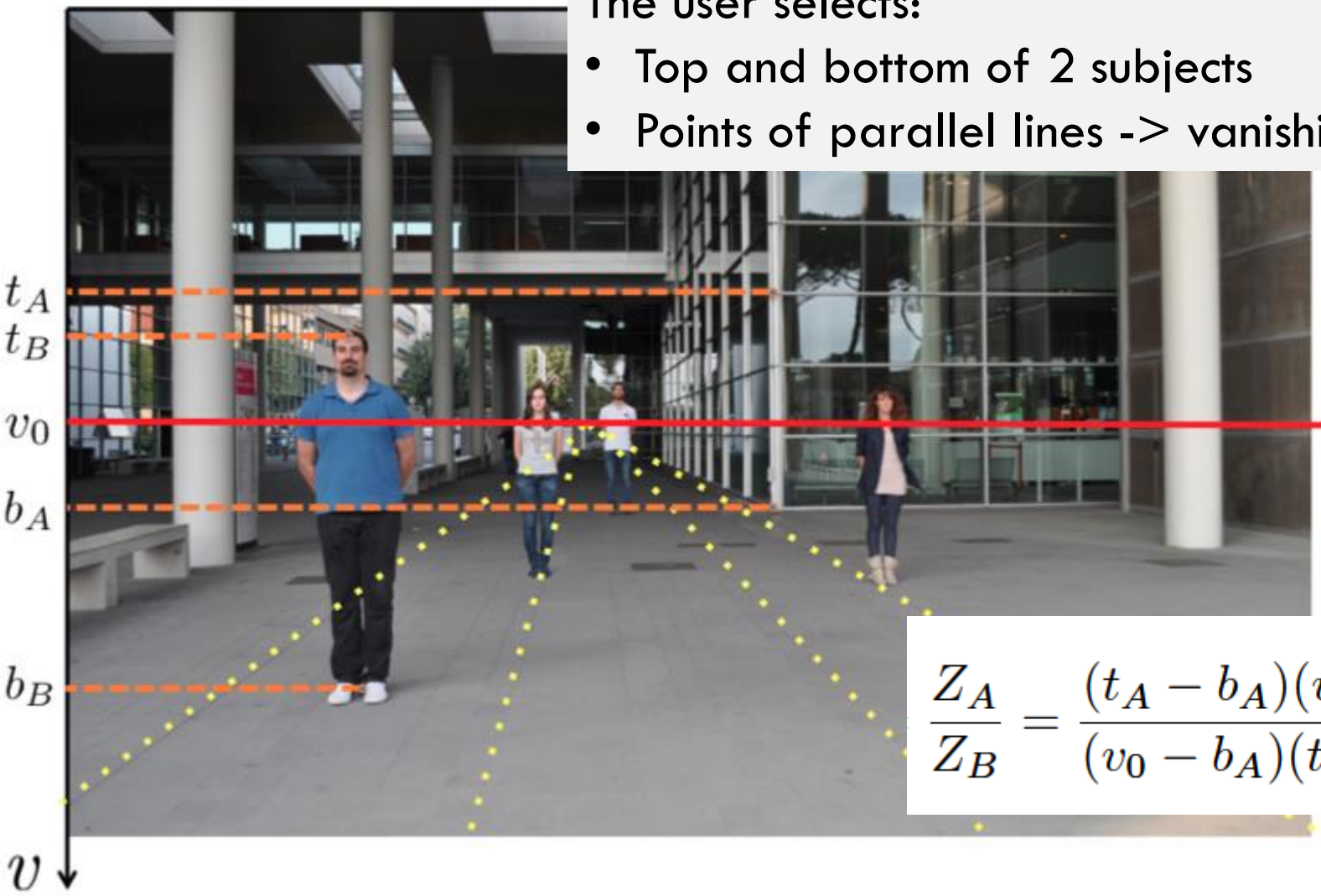
↓

Evidence of Tampering

Example

The user selects:

- Top and bottom of 2 subjects
- Points of parallel lines -> vanishing line



$$\frac{Z_A}{Z_B} = \frac{(t_A - b_A)(v_0 - b_B)}{(v_0 - b_A)(t_B - b_B)}$$

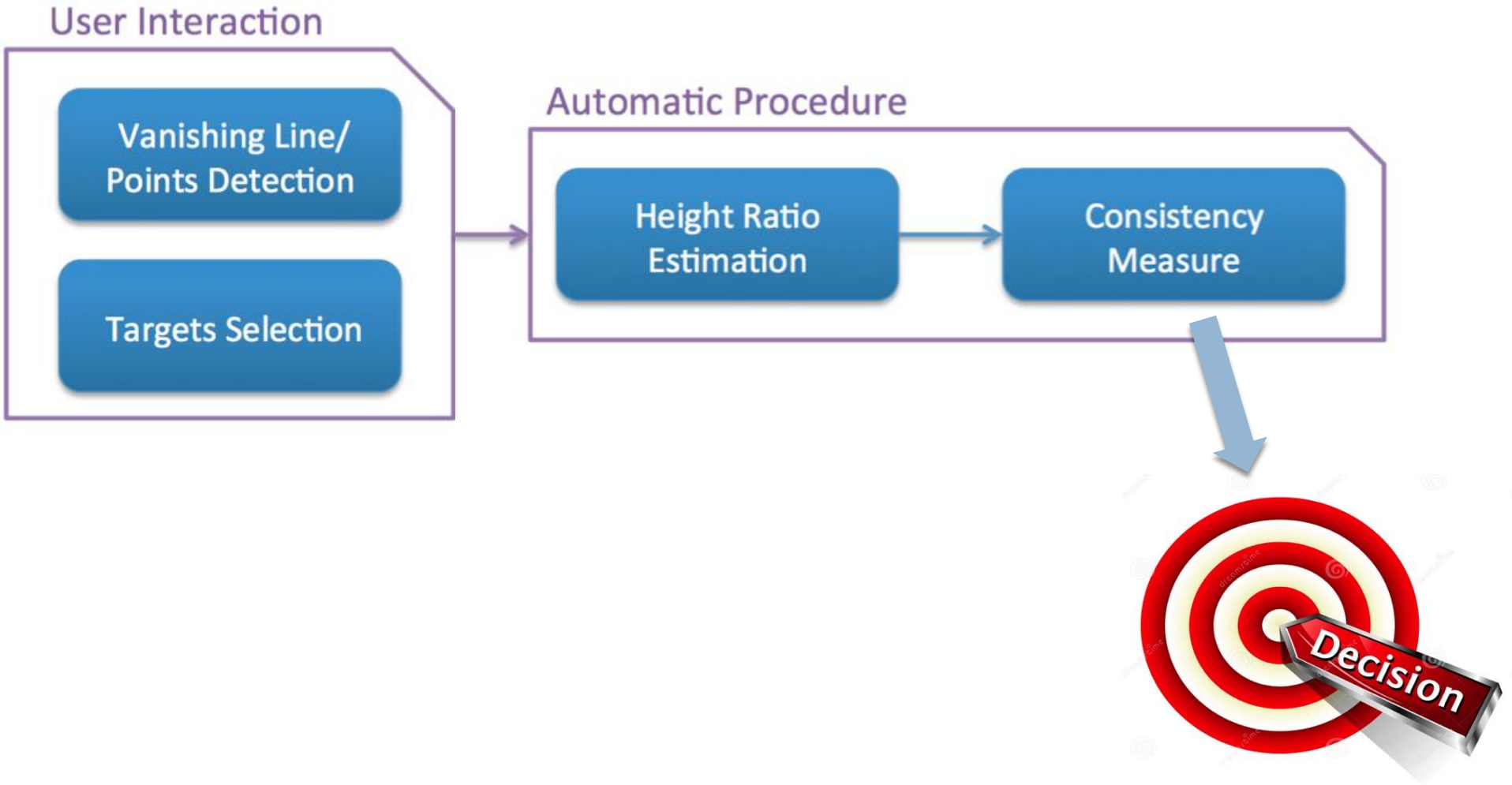
Contribution

- knowledge of camera parameters not required
- some a priori knowledge needed
- Needs buildings, streets, to obtain VL
- It works only if the picture is taken with no tilt & no roll, resulting almost useless on many images.

- Extension to general perspective conditions
  - Verify it on real case

    images exchanged through social networks

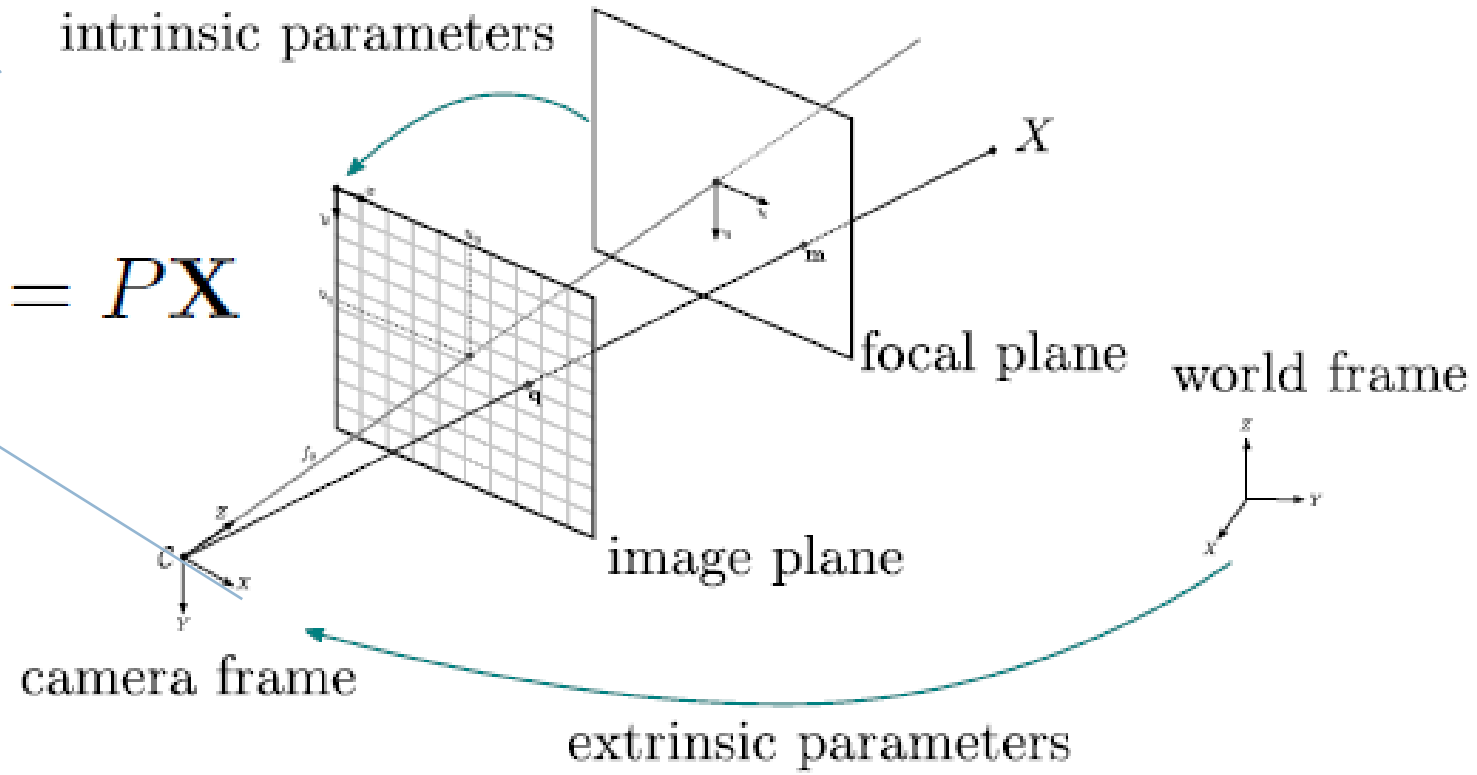
Detection Scheme



Pinhole camera model

X = (X,Y,Z,1) and **x** = (x, y,1) are the homogeneous coordinates of 3D world points and 2D image points

$$\mathbf{x} = K[R|t]\mathbf{X} = P\mathbf{X}$$

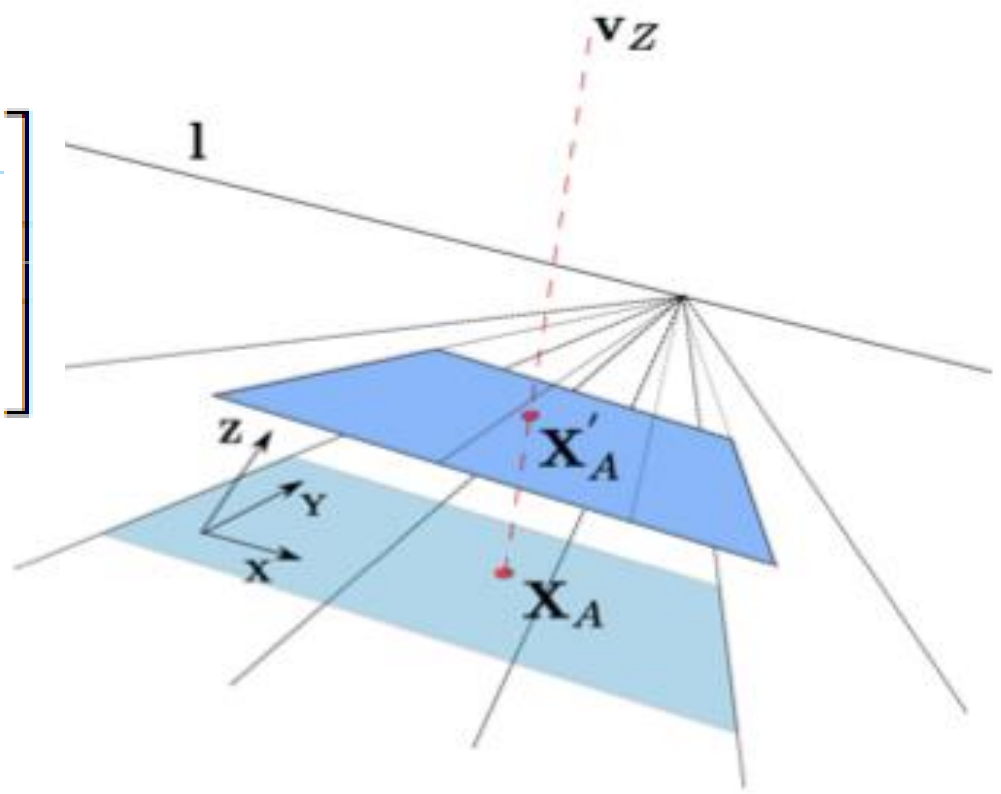


P: 3x4 Projection Matrix

Tampering Detection

Bottom and top of object A

$$\mathbf{x}_A = P\begin{bmatrix} X_A \\ Y_A \\ 0 \\ 1 \end{bmatrix} \quad \mathbf{x}'_A = P\begin{bmatrix} X_A \\ Y_A \\ Z_A \\ 1 \end{bmatrix}$$



Height can be determined up to a scale factor

$$\alpha Z_A = \frac{||\mathbf{x}_A \times \mathbf{x}'_A||}{(\bar{\mathbf{l}} \cdot \mathbf{x}_A)||\mathbf{v}_Z \times \mathbf{x}'_A||}$$

l: Vanishing line of the reference plane

$\mathbf{v}_Z$: vanishing point of the vertical direction

Tampering Detection

Height Ratio between two objects A and B can be determined

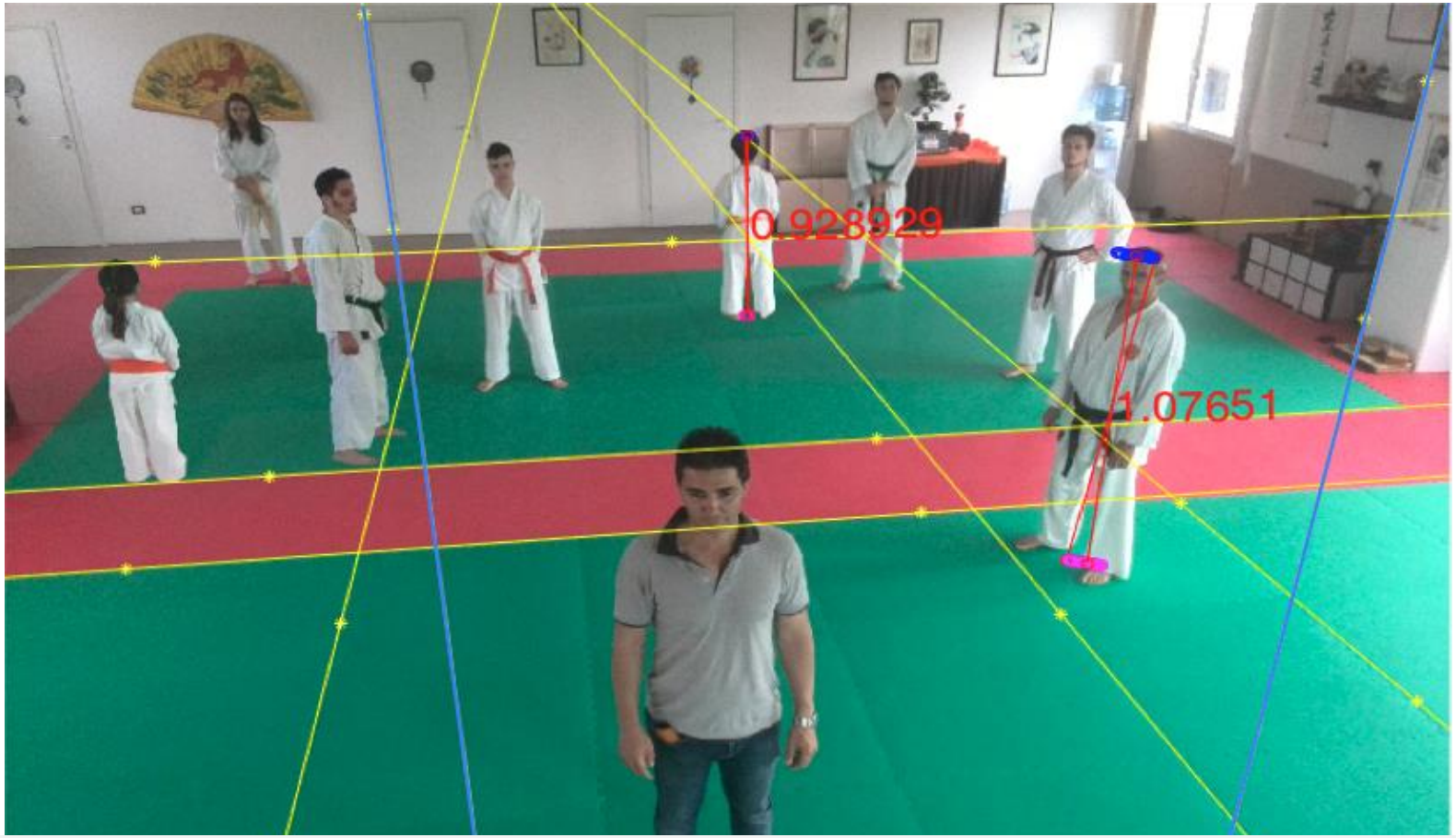
$$\kappa = \frac{Z_A}{Z_B} = \frac{||\mathbf{x}_A \times \mathbf{x}'_A||}{||\mathbf{x}_B \times \mathbf{x}'_B||} \frac{(\bar{\mathbf{l}} \cdot \mathbf{x}_B)||\mathbf{v}_Z \times \mathbf{x}'_B||}{(\bar{\mathbf{l}} \cdot \mathbf{x}_A)||\mathbf{v}_Z \times \mathbf{x}'_A||}$$

We need:

- Top and bottom of objects A and B
- Vanishing line
- Vertical vanishing point

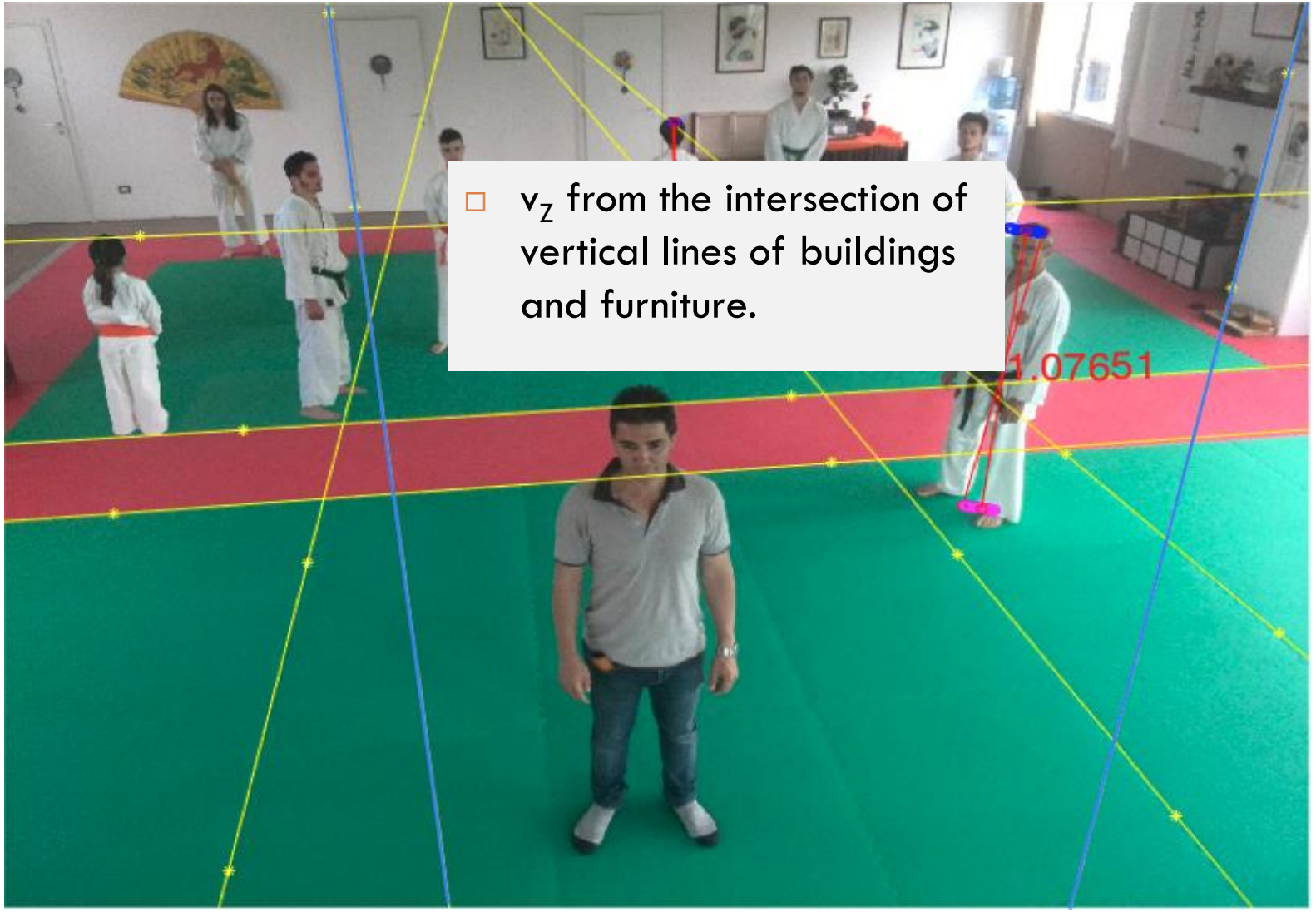
Example

Example



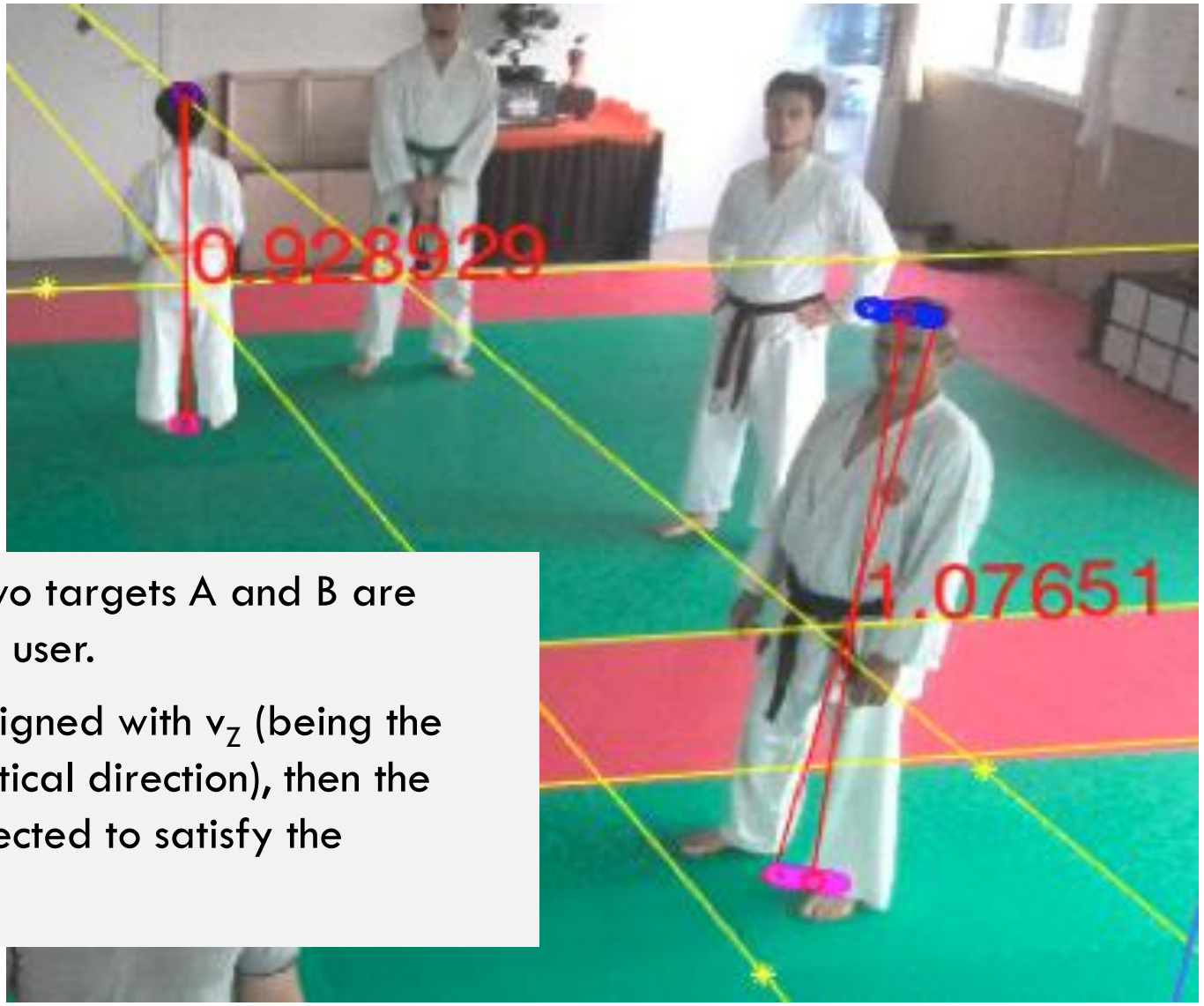
- The vanishing line can be identified by the cross product of the vanishing points of two non parallel directions of the reference plane.
- At least two lines are needed to estimate a vanishing point .

Example



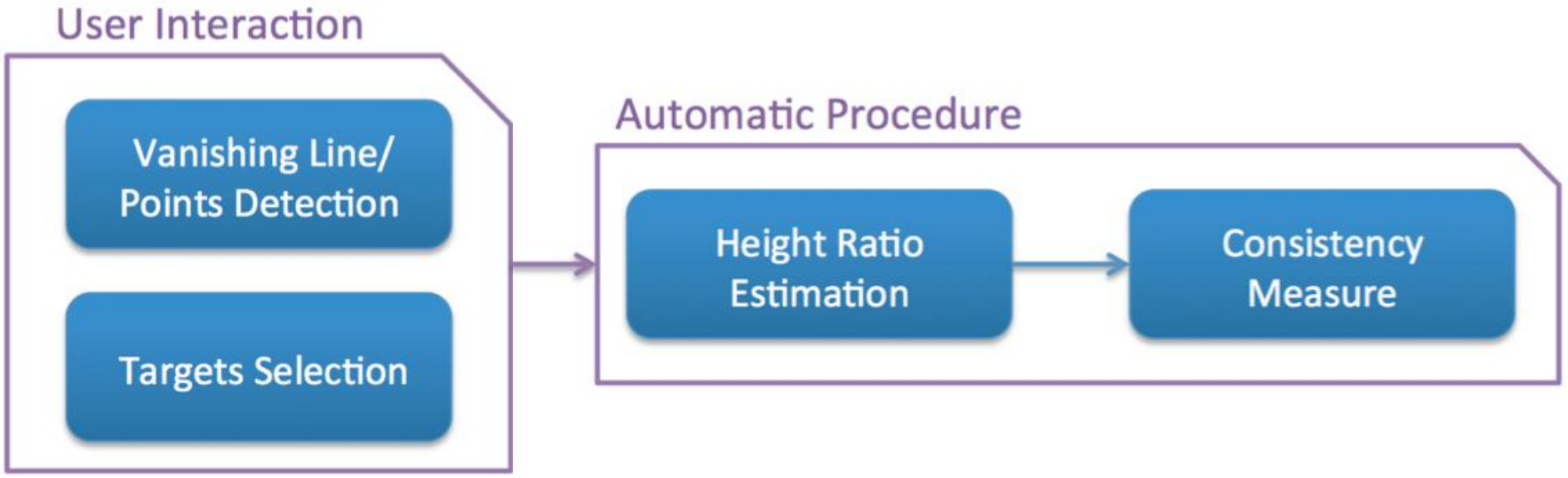
- v_Z from the intersection of vertical lines of buildings and furniture.

Example



- Top and bottom of the two targets A and B are manually selected by the user.
- Each couple should be aligned with v_Z (being the target aligned to the vertical direction), then the selected have to be corrected to satisfy the geometric constraint.

Detection Scheme



Consistency Feature

α Ground Truth Height Ratio

κ Estimated Height Ratio

We assume that in authentic images $(\kappa - \alpha) \sim N(0, \sigma^2)$

Consistency Feature

$$C = 2F(-|\alpha - \hat{\kappa}|, 0, \sigma^2)$$

F : CDF

Anomaly $0 < C < 1$ Good Estimation

$$C < \tau \Rightarrow$$ EVIDENCE OF TAMPERING

Experimental Setup

- **Dataset**

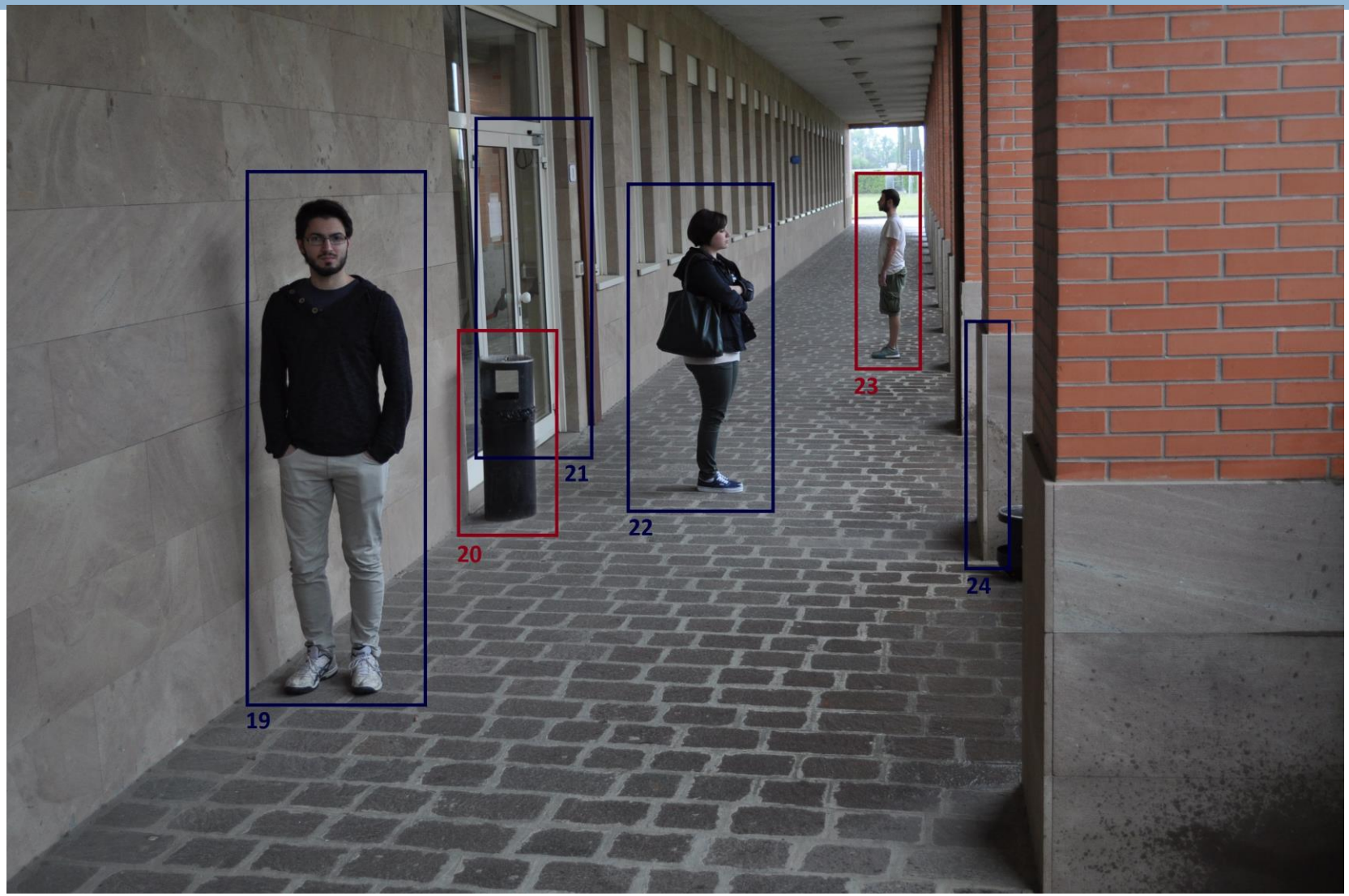
  7 high res images (6-8 megapixel) containing both authentic and tampered elements

- **Collected Data**

Authentic Targets	Tampered Targets	Number of Pictures	Autentic Couples	Tampered Couples
4	2	5	30 = 6x5	40 = 8x5
6	3	1	15	18
6	0	1	15	0
		TOT	**60**	**58**

The height of each analyzed subject is known !

Dataset

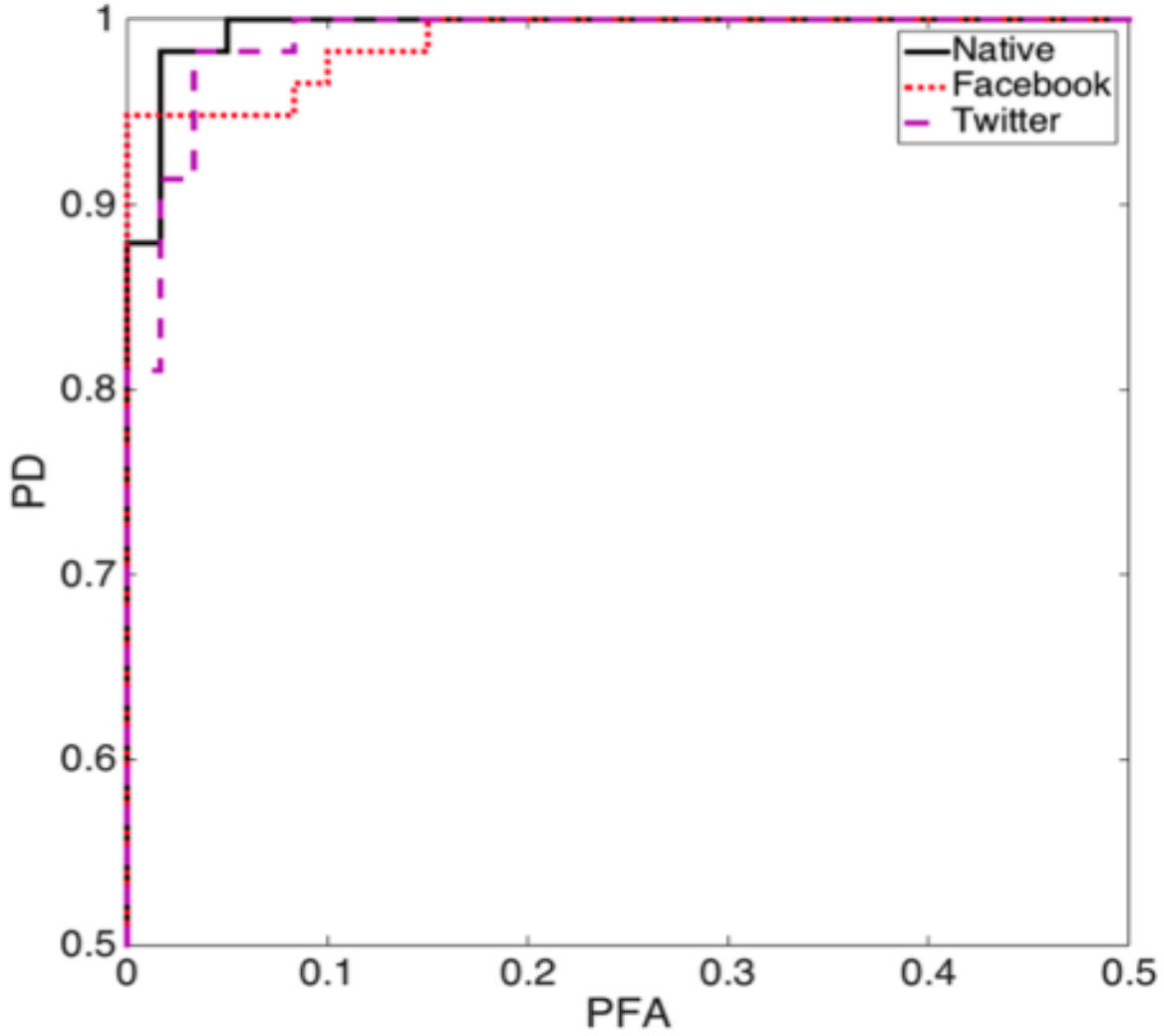


Dataset

High Resolution detail

The Pain of Being Social

Performance



Dataset	AUC
Native	0.980
Facebook	0.978
Twitter	0.978

Conclusions

- **Improvement of state of the art technique**

  Applicable under general perspective condition

  Applicable on images exchanged through social network (Facebook, Twitter)

What Now?

- **Results confidence analysis**

  different perspective condition

  Different user behavior

- **Further extensions**

  Compare subjects on parallel planes

  Lower user interaction

Thank you!

alessandro.piva@unifi.it