3DRP pcl

IAS13 Padua, July 15-19, 2014

OpenPTrack

PEOPLE TRACKING FOR HETEROGENEOUS NETWORKS OF COLOR-DEPTH CAMERAS

Matteo Munaro[1,3], Alex Horn[2], Randy Illum[2], Jeff Burke[2], and Radu Bogdan Rusu[3]

[1] IAS-Lab at Department of Information Engineering, University of Padova
[2] Center for Research in Engineering, Media and Performance, University of California Los Angeles
[3] Open Perception, Inc., Menlo Park, CA

1st International Workshop on 3D Robot Perception with Point Cloud Library
Padova, July 15, 2014

DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE

INTELLIGENT AUTONOMOUS SYSTEMS LAB

- open source solution for **scalable**, **multi-imager** person tracking for education, arts, and culture applications

- **OpenPTrack** is led by UCLA REMAP and Open Perception. Key collaborators include the University of Padova and Electroland. Code is available under a BSD license. Portions of the work are supported by the National Science Foundation (IIS-1323767)

UCLA remap

ELECTROLAND

UNIVERSITÀ DEGLI STUDI DI PADOVA

http://openptrack.org

Single-camera tracking: state-of-the-art

Background subtraction in depth image [1]

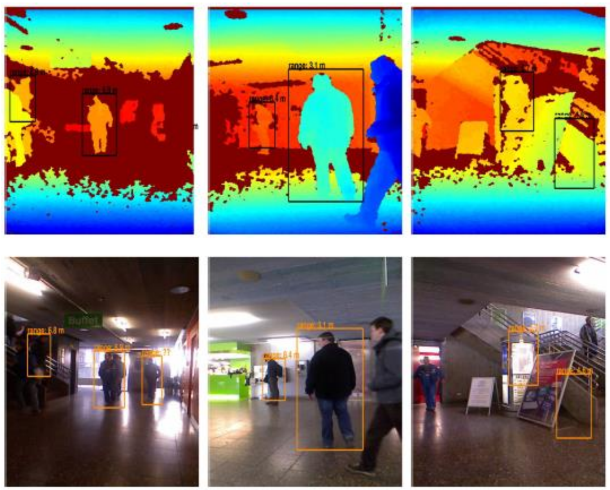
Real time on CPU

Static camera

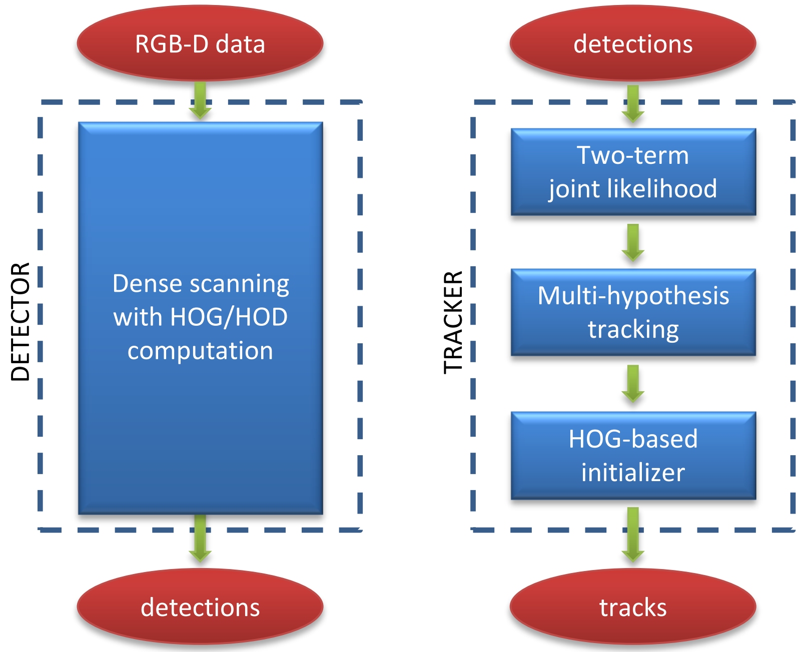
[1] http://www.primesense.com/solutions/nite-middleware/

Single-camera tracking: state-of-the-art

Dense scanning of RGB and depth images ([2],[3])



[2]



Moving camera

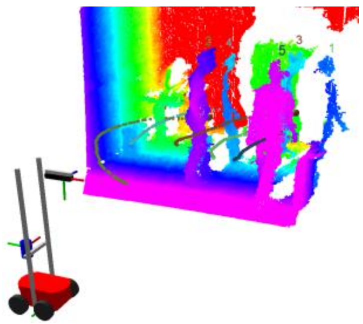
GPU for real time

[2] L. Spinello and K. O. Arras. *People detection in rgb-d data*. In IROS 2011.
[3] M. Luber, L. Spinello, and K. O. Arras, *People tracking in rgb-d data with on-line boosted target models*, in IROS 2011.

Single-camera tracking: state-of-the-art

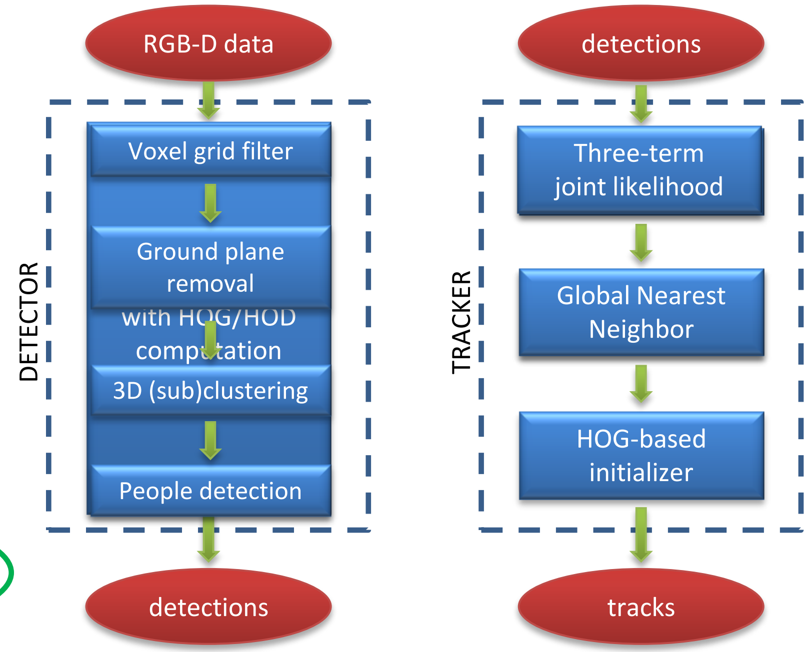
Based on point cloud sub-clustering

- Hypothesis: single ground plane



Moving camera

Real time on CPU

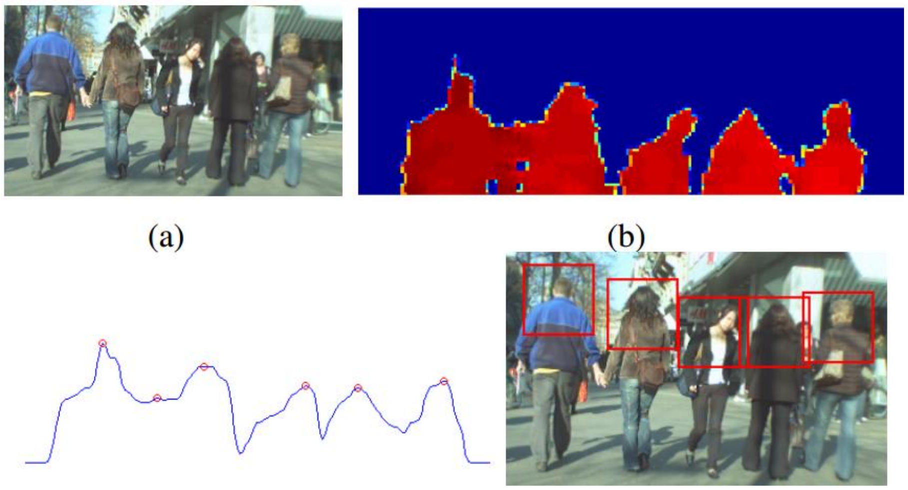


pcl

[4] M. Munaro, F. Basso, S. Michieletto and E. Menegatti. *Fast and robust multi-people tracking from RGB-D data for a mobile robot*. In Proceedings of the 12th Intelligent Autonomous Systems (IAS) Conference, Jeju Island (Korea), 2012.
[5] M. Munaro, F. Basso and E. Menegatti. *Tracking people within groups with RGB-D data*. In Proceedings of the International Conference on Intelligent Robots and Systems (IROS) 2012, Vilamoura (Portugal), 2012.
[6] M. Munaro, E. Menegatti. *Fast RGB-D People Tracking for Service Robots*. Autonomous Robots Journal, 2014.

Single-camera tracking: state-of-the-art

- Height-based **point cloud labeling**
- **subclustering** in **depth** image



(a) (b)

(c) (d)

Real time on CPU

Moving camera

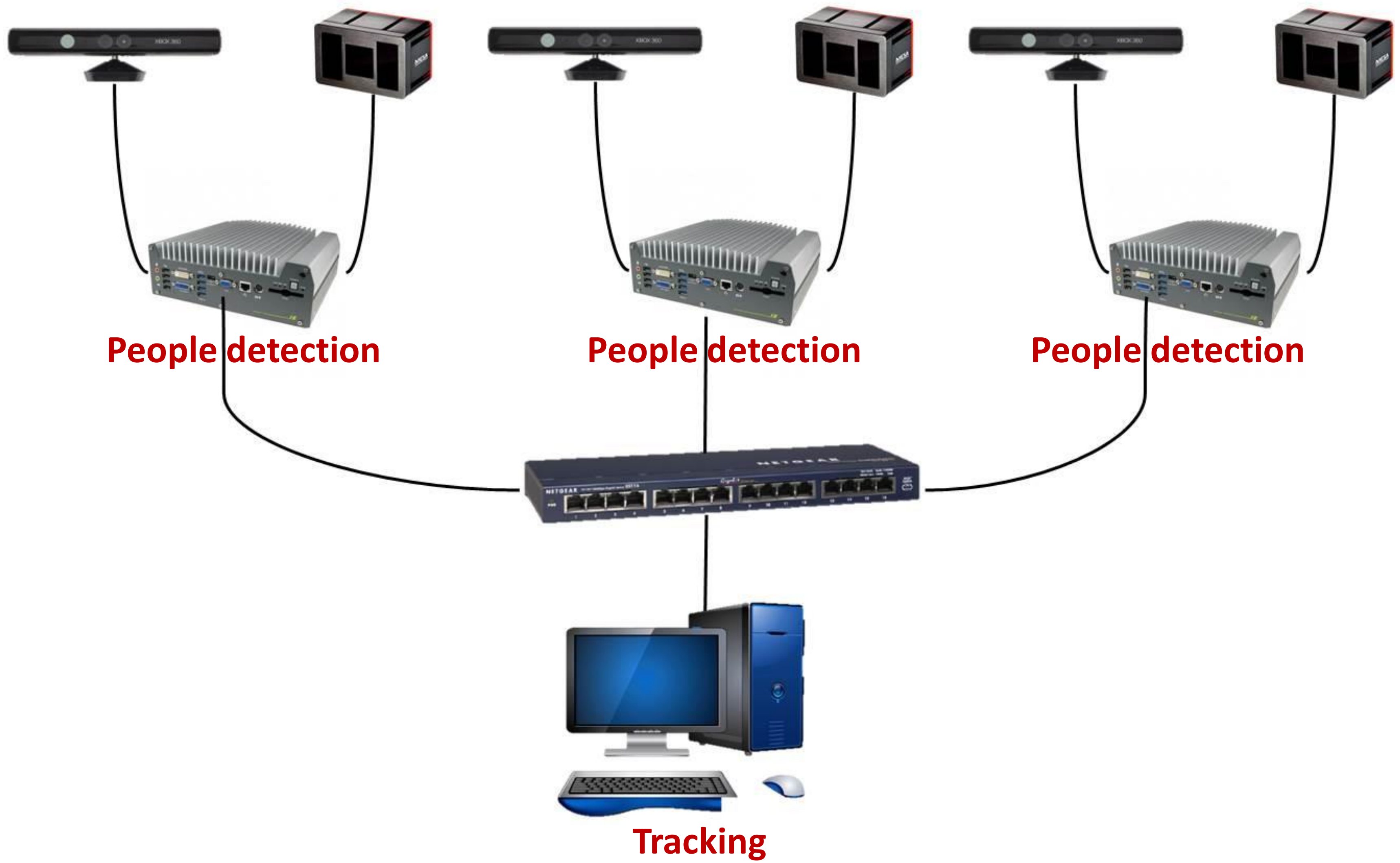
[7] Omid Hosseini Jafari, Dennis Mitzel, Bastian Leibe. *Real-Time RGB-D based People Detection and Tracking for Mobile Robots and Head-Worn Cameras*. ICRA 2014.

Multi-camera tracking

- **Multiple** cameras
  - Create **redundancy**
  - **Reduce** the problem of **occlusion**
- Problem **poorly** studied in literature [8]
- Main problems in network of RGB-D sensors:
  - Network **calibration**
  - **Scalability**

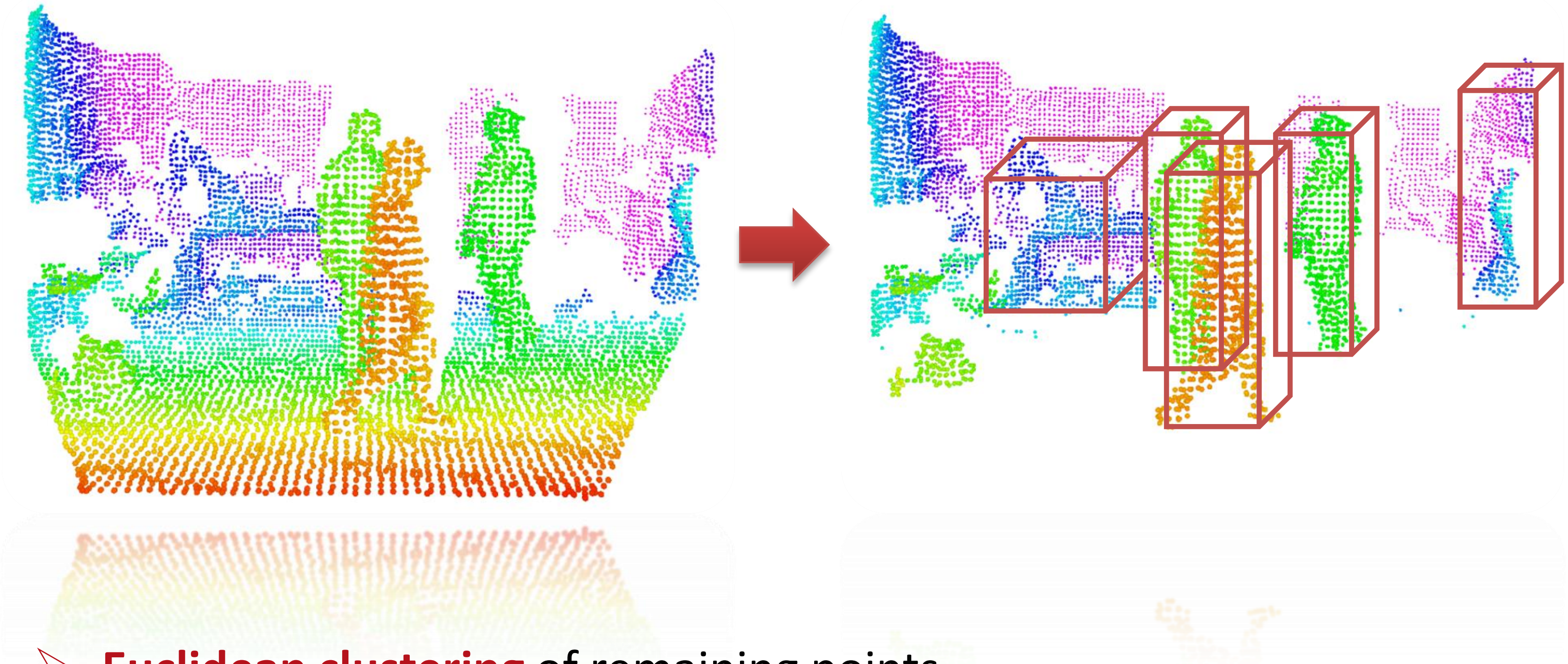
[8] M. Munaro, F. Basso, S. Michieletto, E. Pagello and E. Menegatti. A Software Architecture for RGB-D People Tracking Based on ROS Framework for a Mobile Robot. Frontiers of Intelligent Autonomous Systems, Volume 466, pp 53-68, Springer 2013.

Distributed Multi-camera tracking



People Detection on Kinect data

- **Ground plane** estimation and **removal**

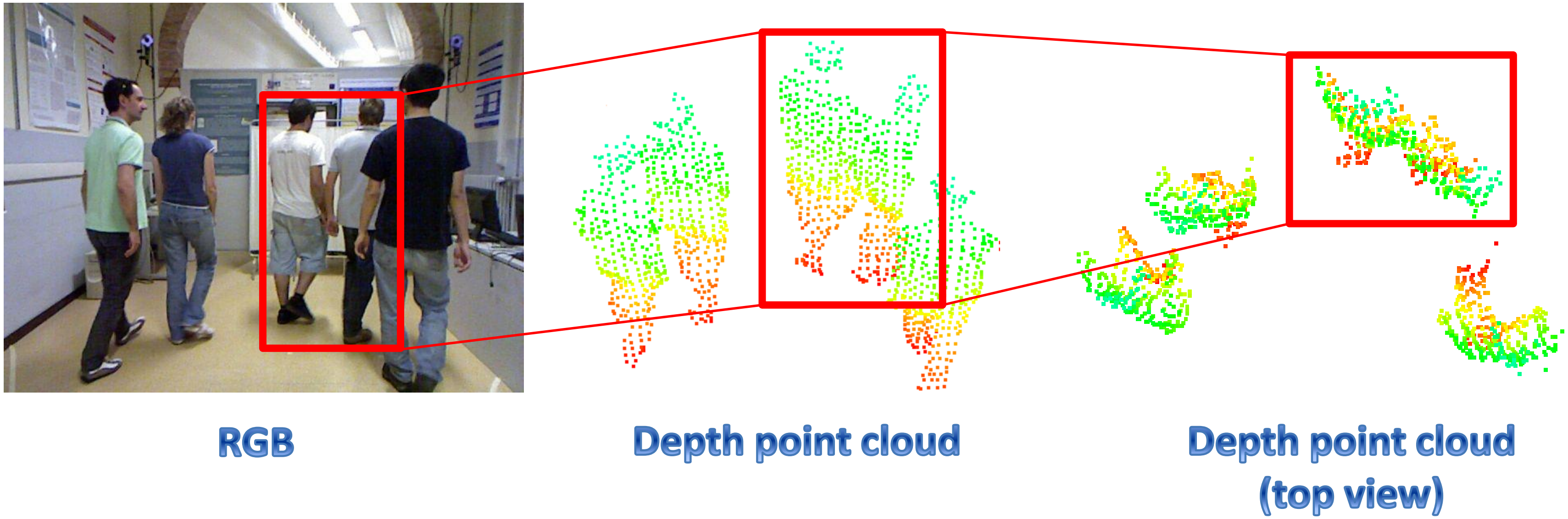


- **Euclidean clustering** of remaining points

Detection on Point Cloud: main issues

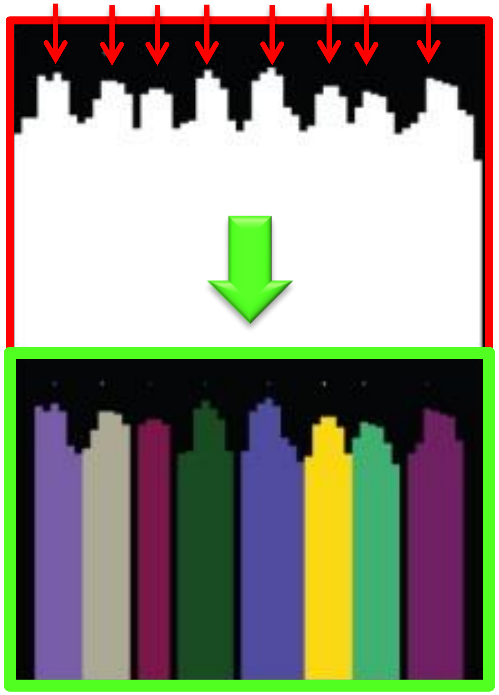
- People vertically **splitted** into more clusters

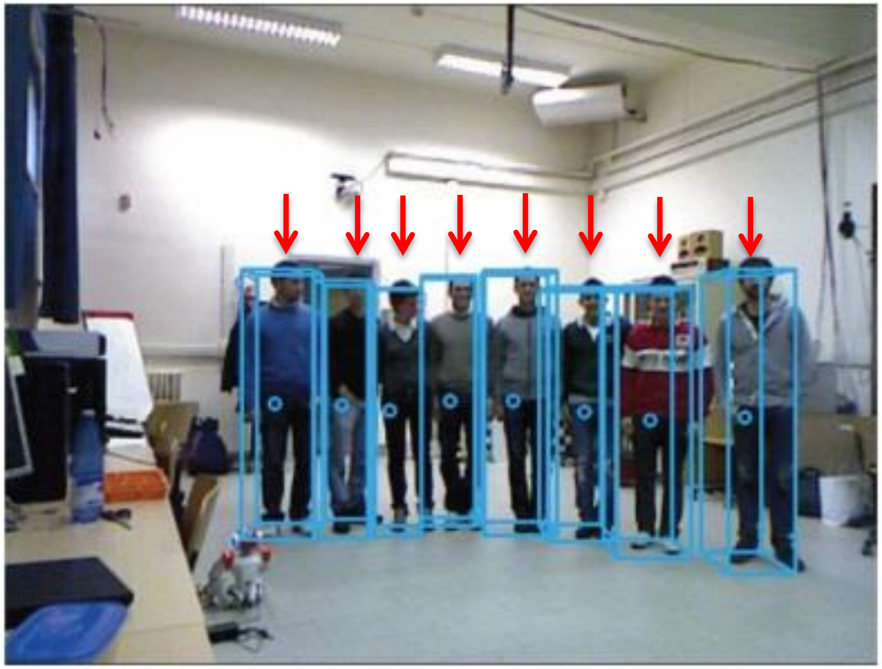
- More people **merged** into the same cluster



Sub-clustering procedure

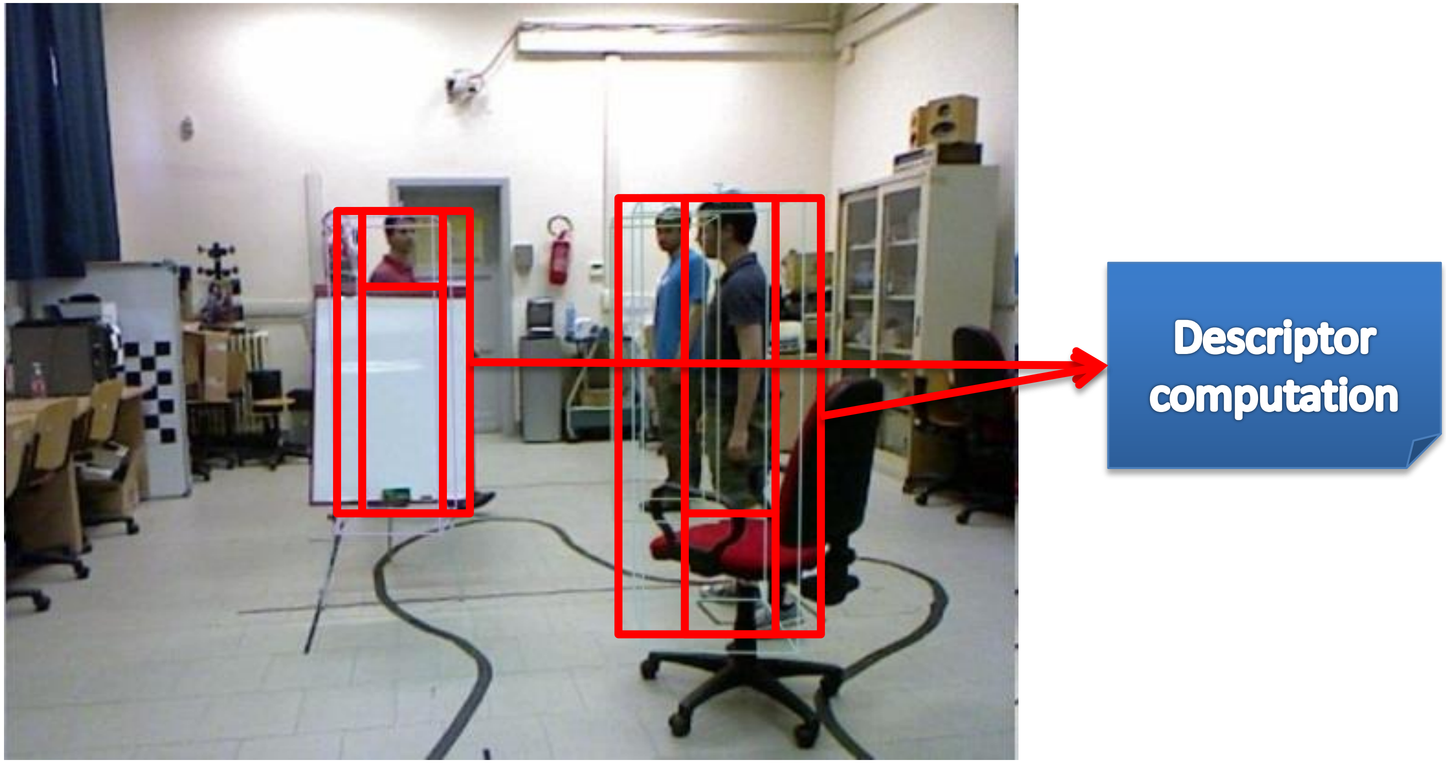
Subdivision of big clusters by means of **head detection** (peaks in height from the ground plane)





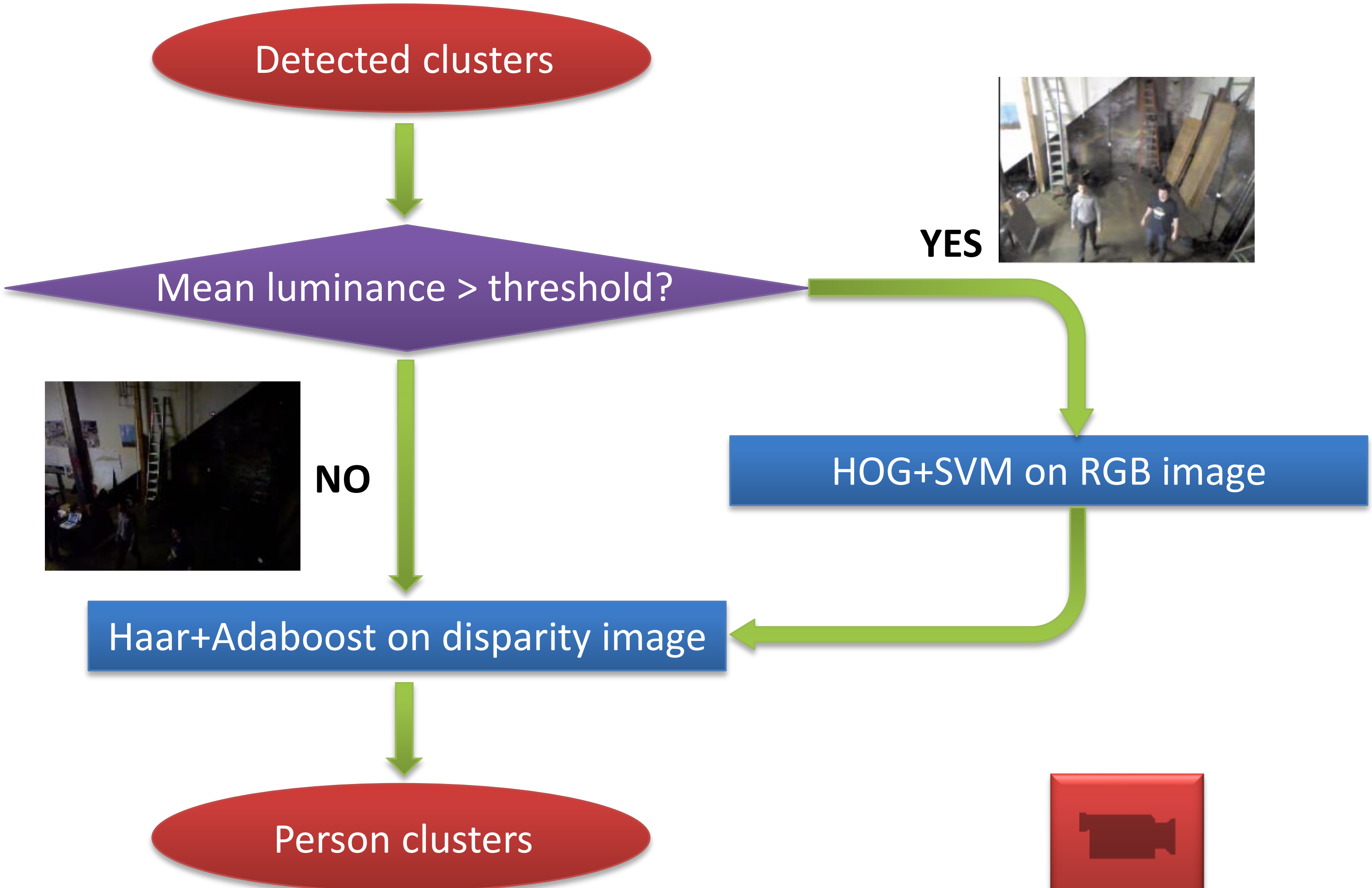
HOG evaluation

- Candidate **clusters** are **extended** until the ground plane for achieving robustness to lower limbs occlusion
- **HOG detector** [10] applied to image patches that are projection of 3D clusters onto the RGB image



[10] N. Dalal and B. Triggs. Histogram of Oriented Gradients for Human Detection, CVPR 2005.

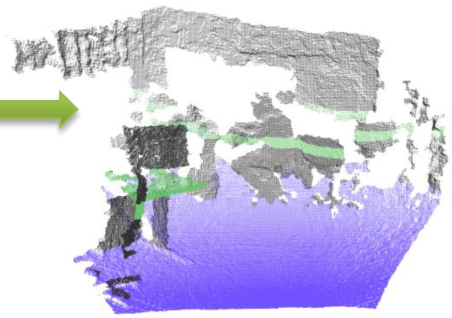
Classification scheme adaptation to brightness changes



Automatic ground plane estimation

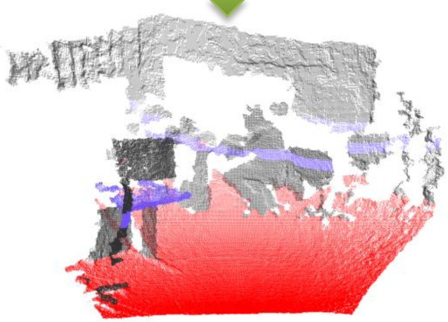
Input point cloud

OrganizedMultiPlaneSegmentation

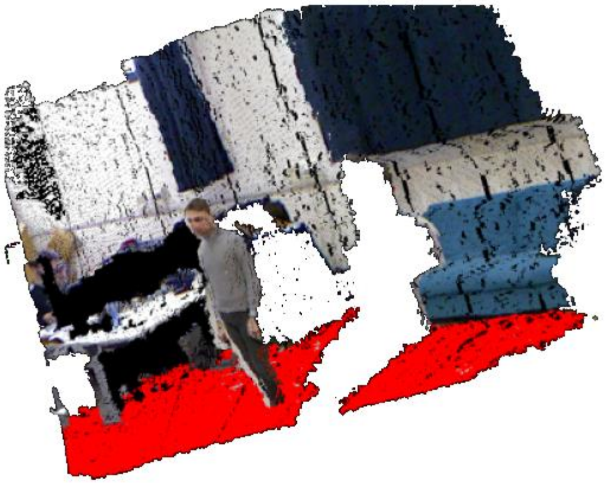


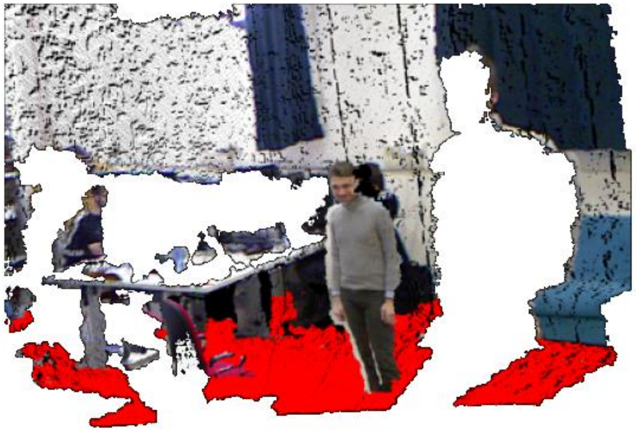
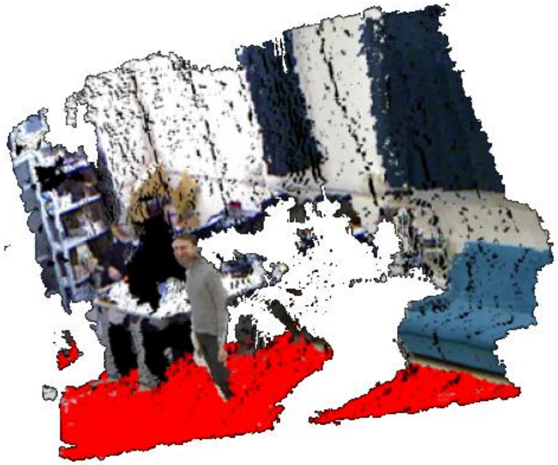
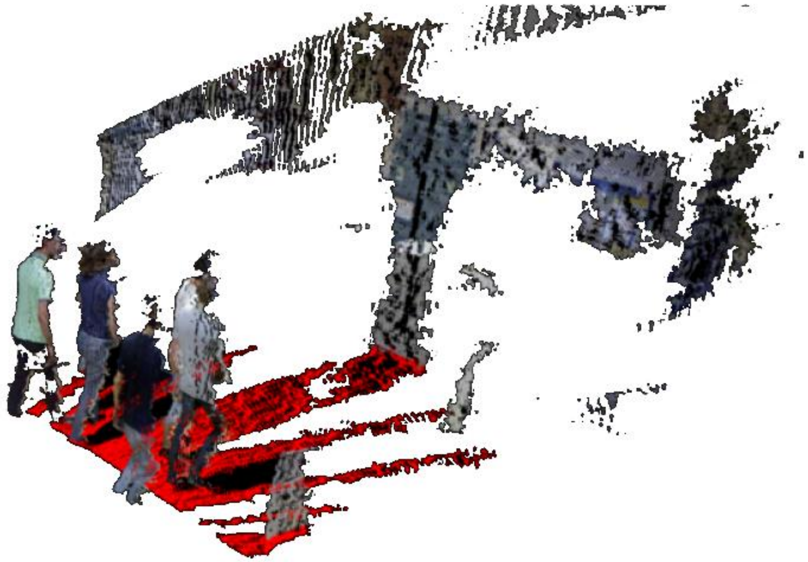
Check on horizontality

The lowest plane is selected



Automatic ground plane estimation

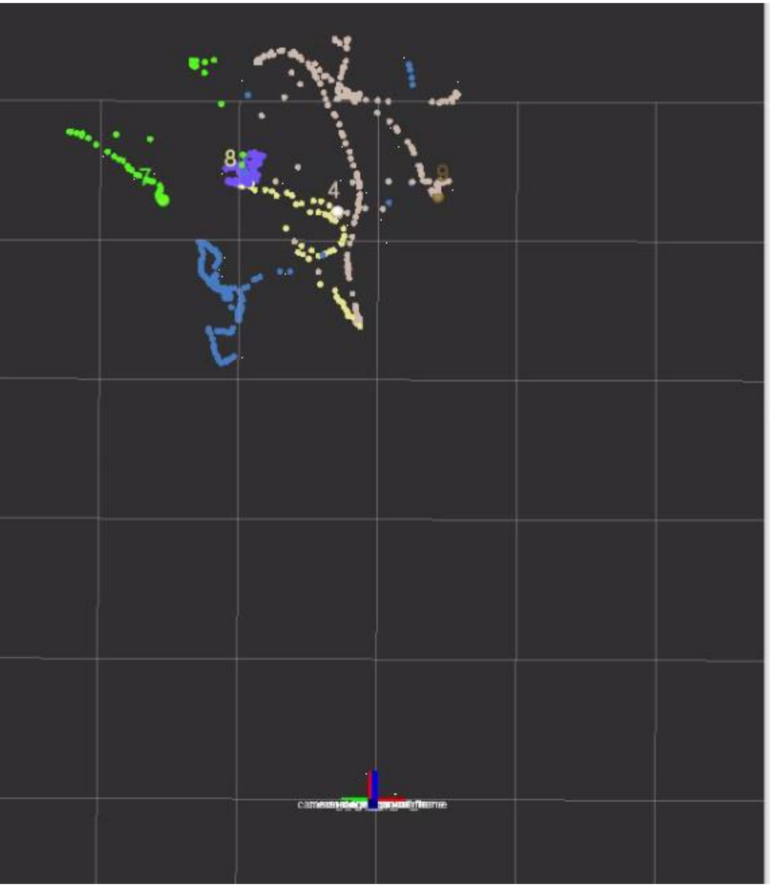


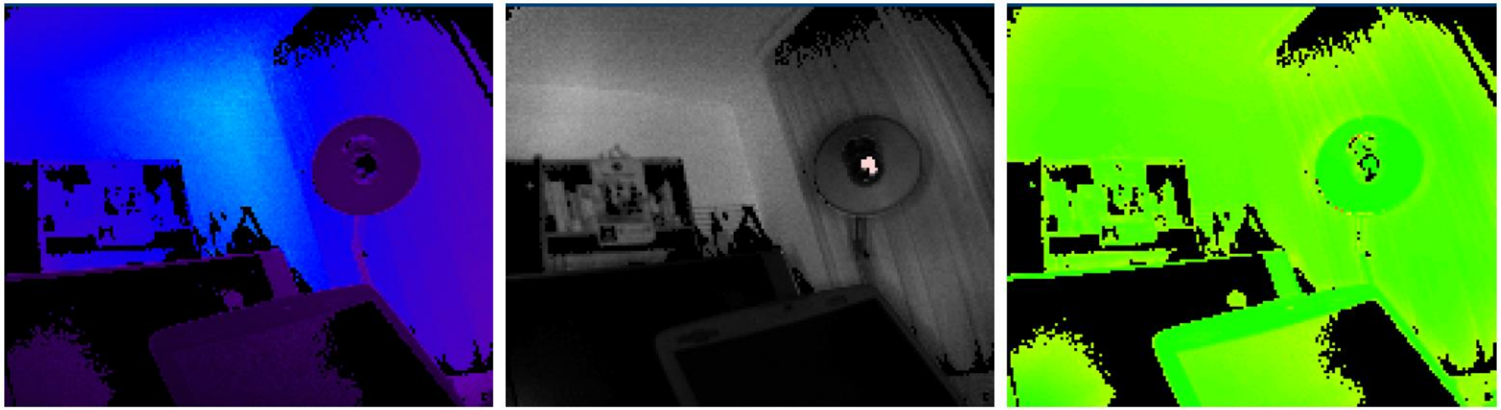
People detection for very tilted sensors

Only for clustering operations (e.g. computing a cluster's bounding box):

- the point cloud is **rotated** so that the **ground** plane is aligned with camera **xOz** plane



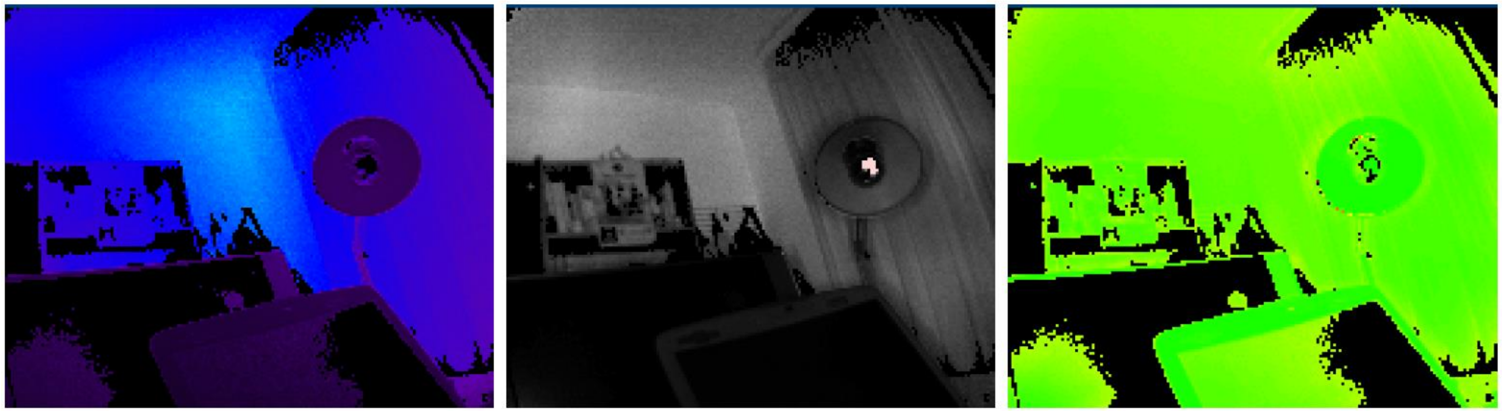
People Detection on SwissRanger (ToF) data



(a) Depth (b) Intensity (c) Confidence

- Point cloud **filtering** based on **confidence** image
- Depth-based **clustering** (as for Kinect)
- **Infrared** intensity image **equalization**
- **HOG+SVM** classification on equalized intensity image

People Detection on SwissRanger (ToF) data



(a) Depth (b) Intensity (c) Confidence

(a) (b) (c) (d)

Fig. 7. People detection results on SwissRanger infrared intensity images.

Centralized tracking-by-detection node

- An **Unscented Kalman Filter** (UKF) is used to predict people positions along the two **world** axes (x, y)
- Human motion is modeled with a **constant velocity** model
- The tracking node receives **detections** from all distributed detection nodes
- A **cost matrix** is created based on the **Mahanolobis distance** between the new detections and the predicted position of all active tracks

$$D_M^{i,j} = \tilde{\mathbf{z}}_k^{\mathrm{T}}(i,j) \cdot \mathbf{S}_k^{-1}(i) \cdot \tilde{\mathbf{z}}_k(i,j)$$

$$\tilde{\mathbf{z}}_k(i,j) = \mathbf{z}_k(i,j) - \hat{\mathbf{z}}_{k|k-1}(i).$$

- The track-detection association is done with the **Global Nearest Neighbor** method

Multi-camera calibration

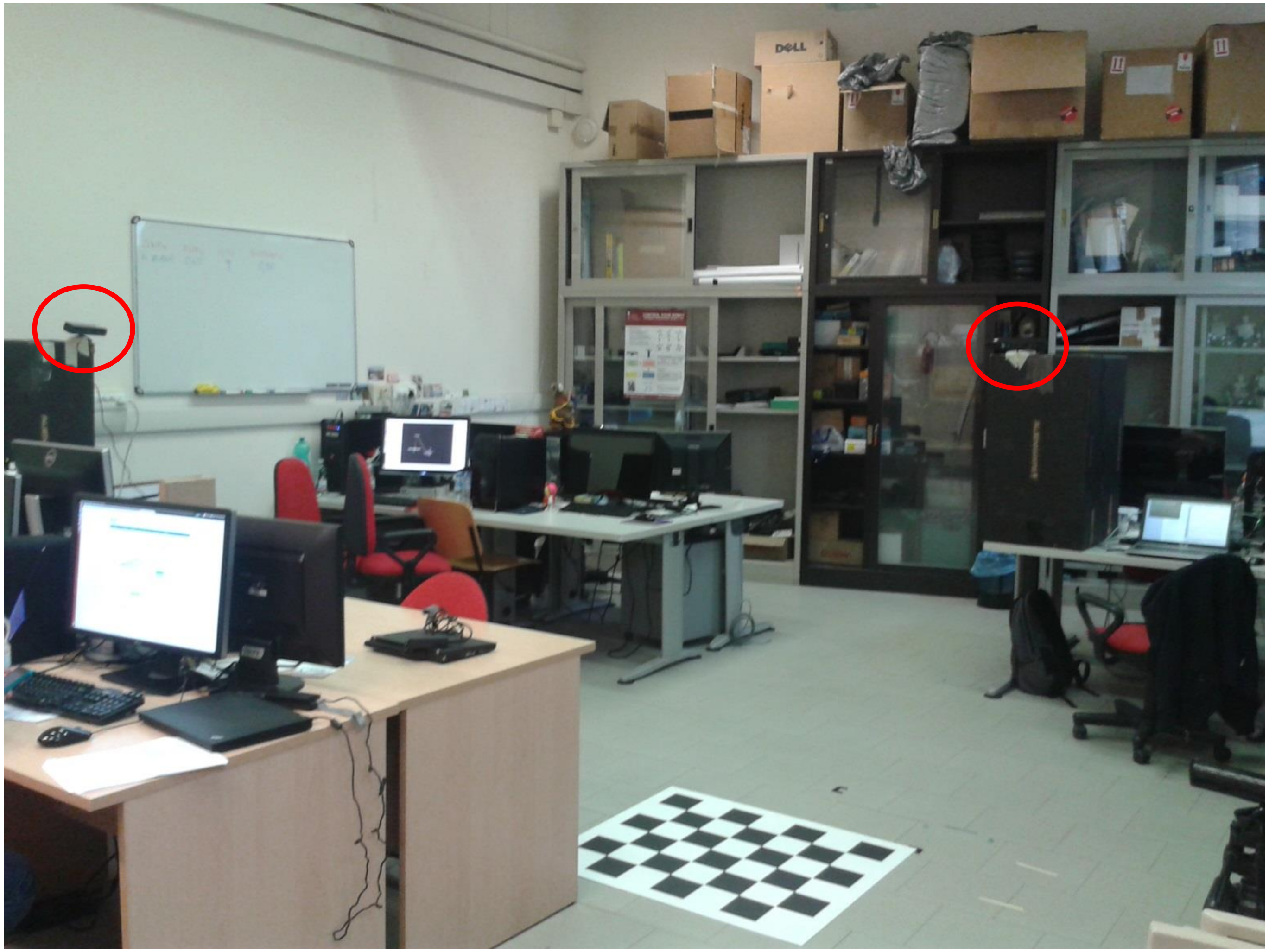
- A **checkerboard** is moved in front of the cameras
- If two cameras see the checkerboard at the same time, the **transformation** between them can be estimated
- A **transformation tree** is built starting from the first camera that sees the checkerboard
- All images are then used in an **optimization** problem [9] which optimizes:
  - The pose of every **camera**
  - The pose of every **checkerboard**

Results:

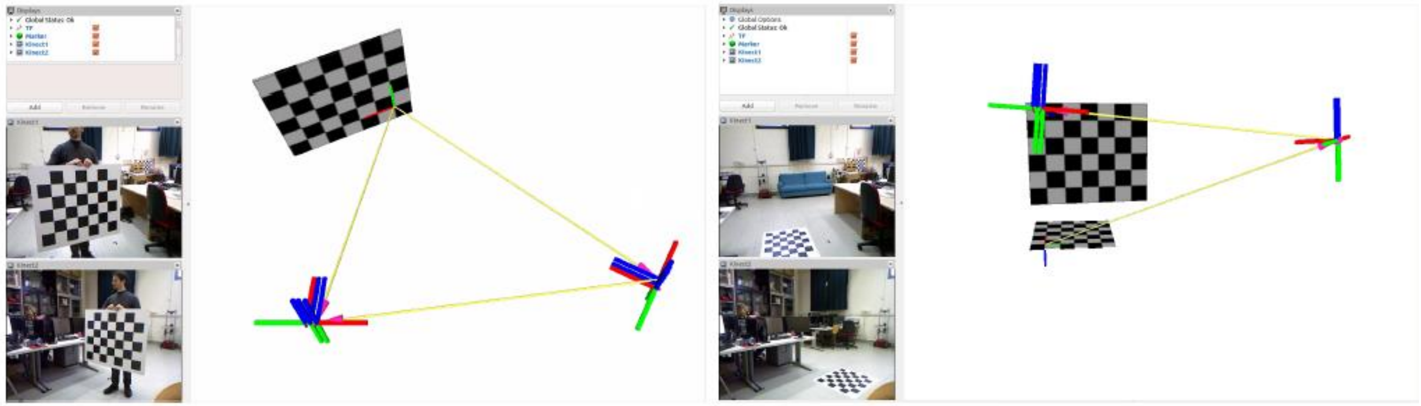
- Extrinsic parameters are estimated in **real time**
- The **ground** plane equation can be calibrated too.

[9] Filippo Basso, Riccardo Levorato and Emanuele Menegatti. *Online Calibration for Networks of Cameras and Depth Sensors*. In ICRA 2014.

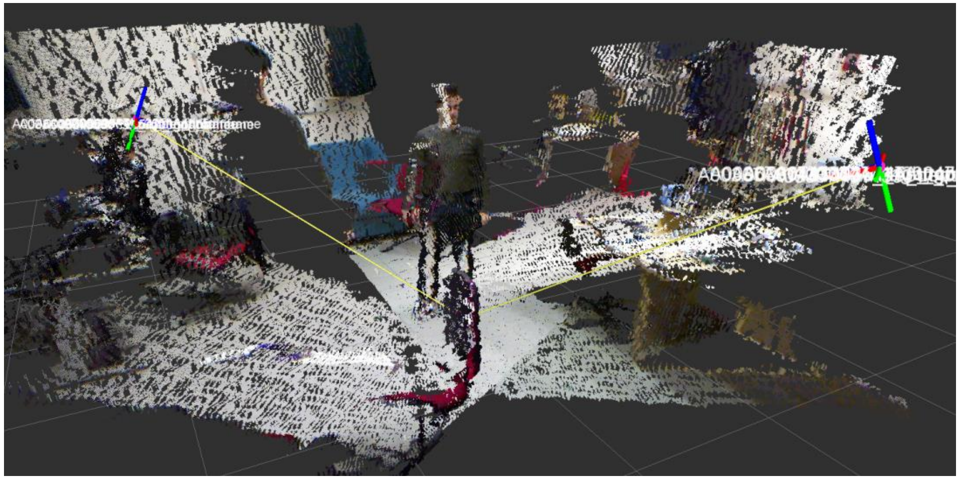
External view of camera network with two Kinects

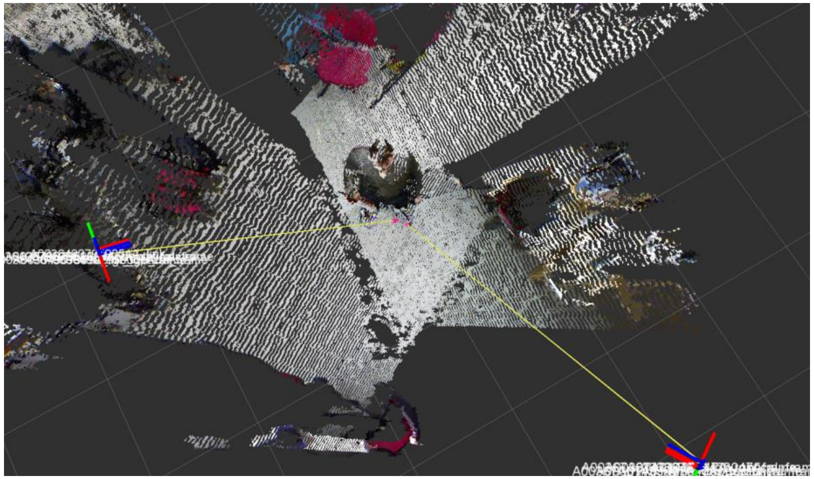


Multi-imager calibration: 2 Kinects

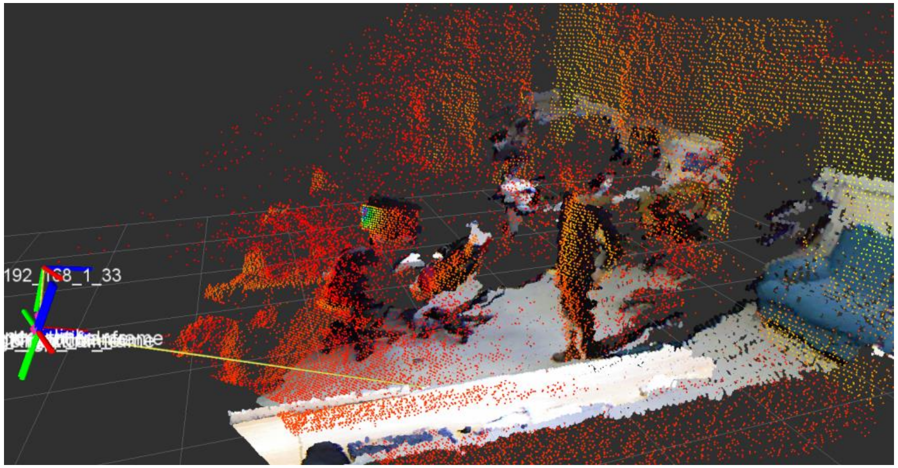


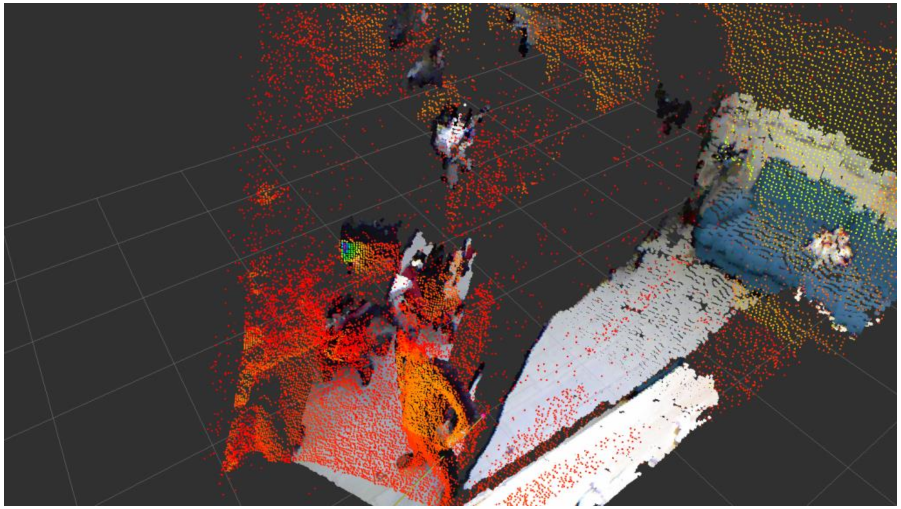
(a) Pairwise calibration (b) Ground plane calibration





Multi-imager calibration: 1 Kinect and 1 Mesa SwissRanger SR4500





Calibration of Kinects and SwissRangers

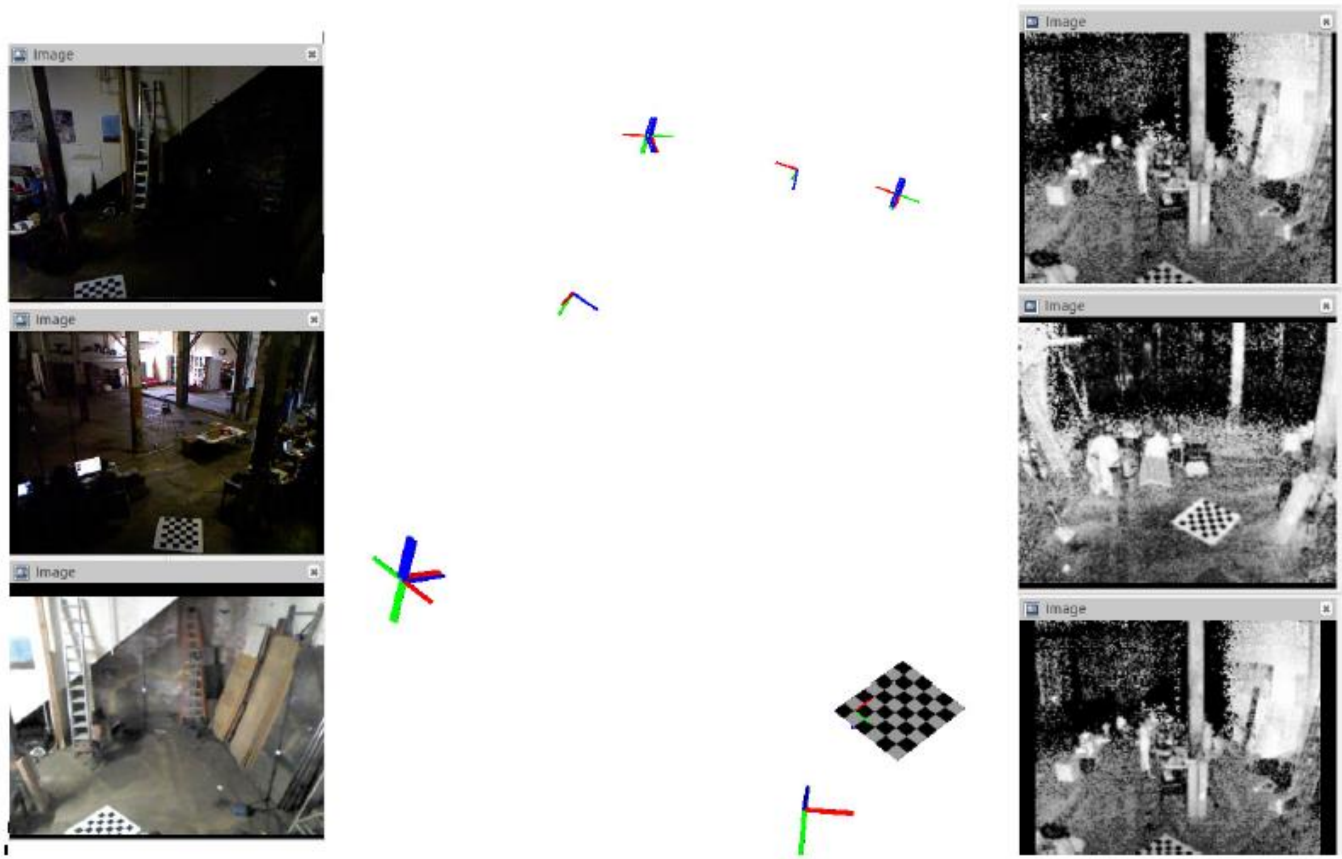
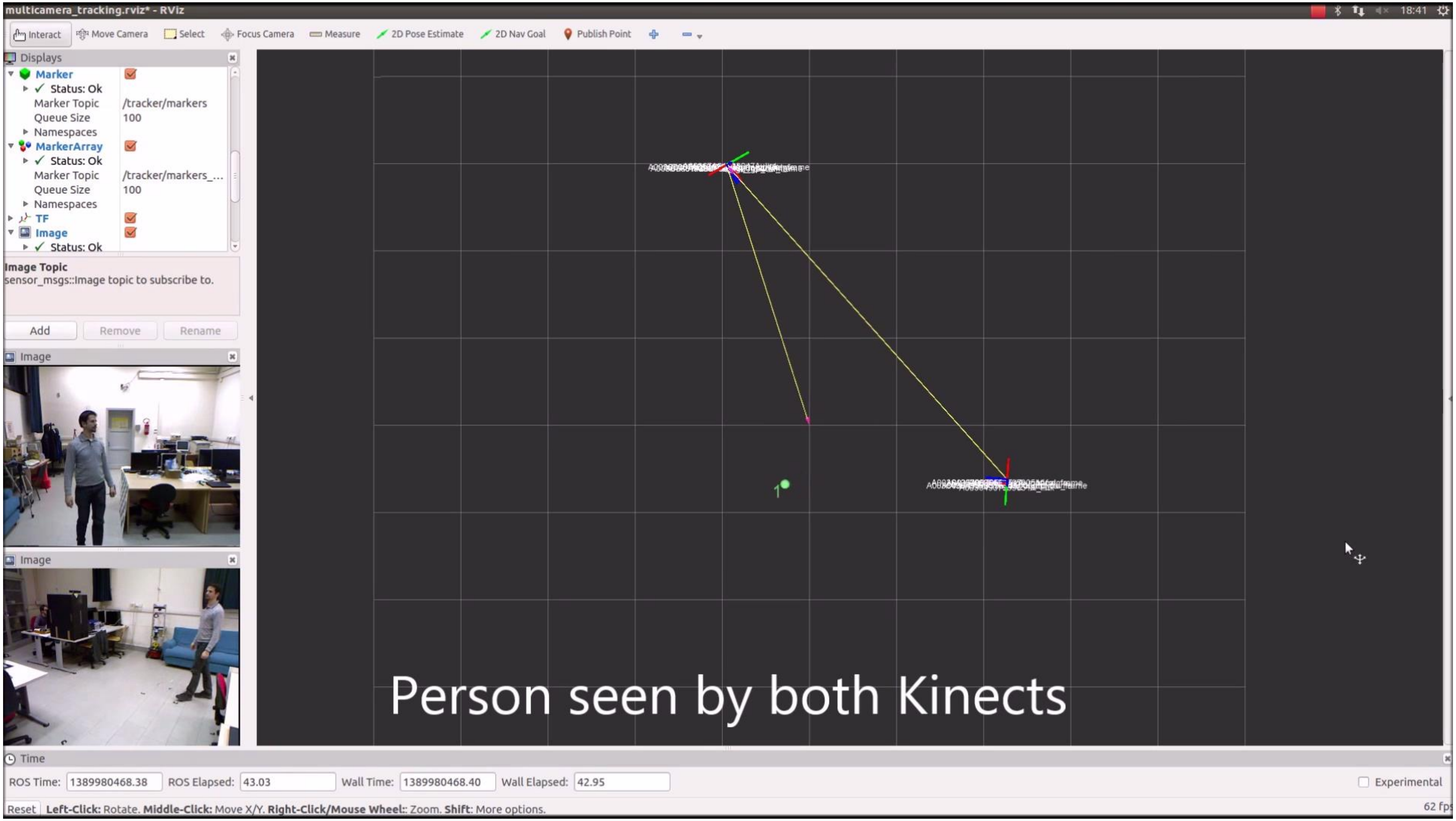
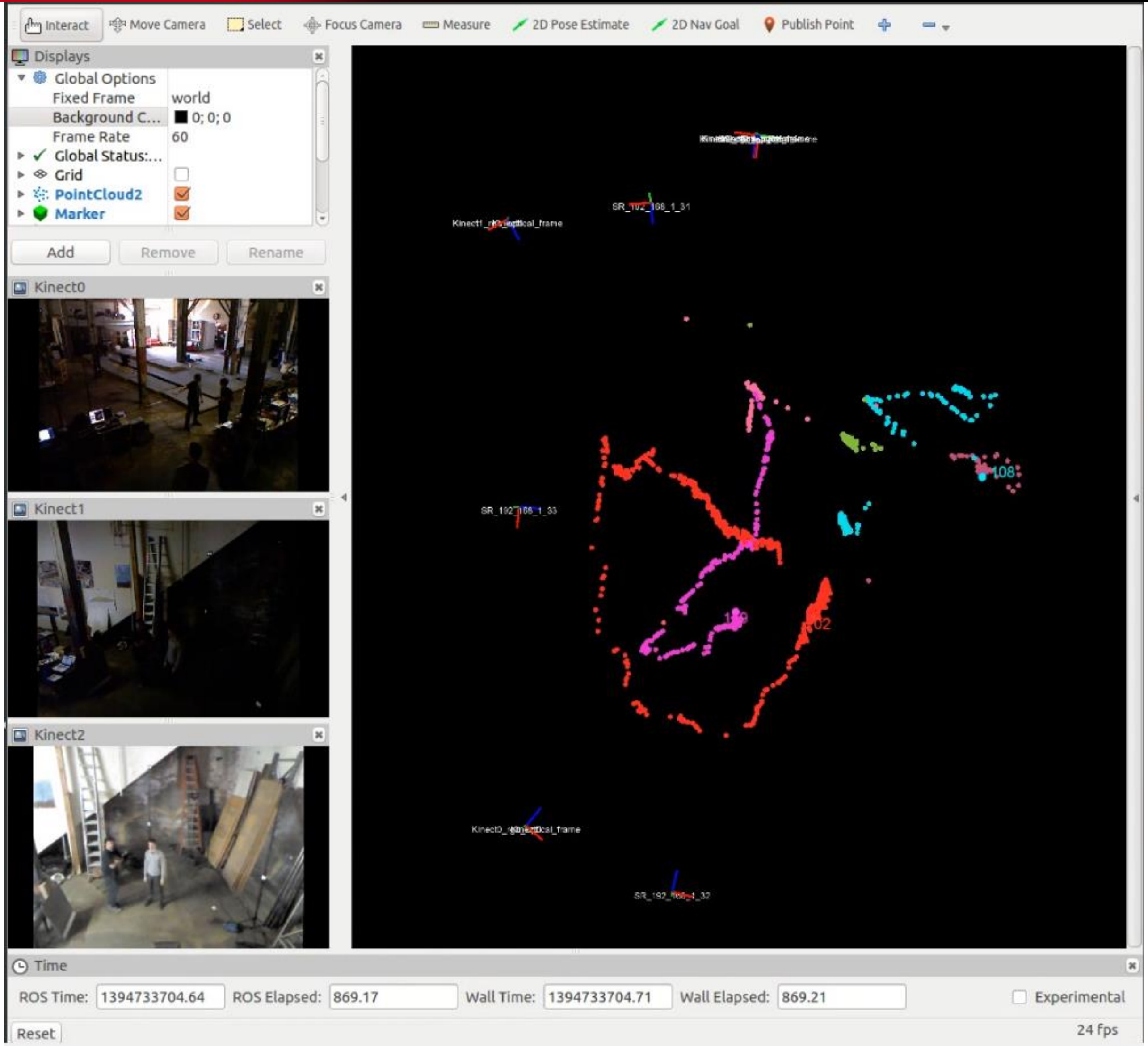


Fig. 3. Extrinsic calibration results for a network composed of three Kinects (images on the left) and three SwissRangers SR4500 (images on the right).

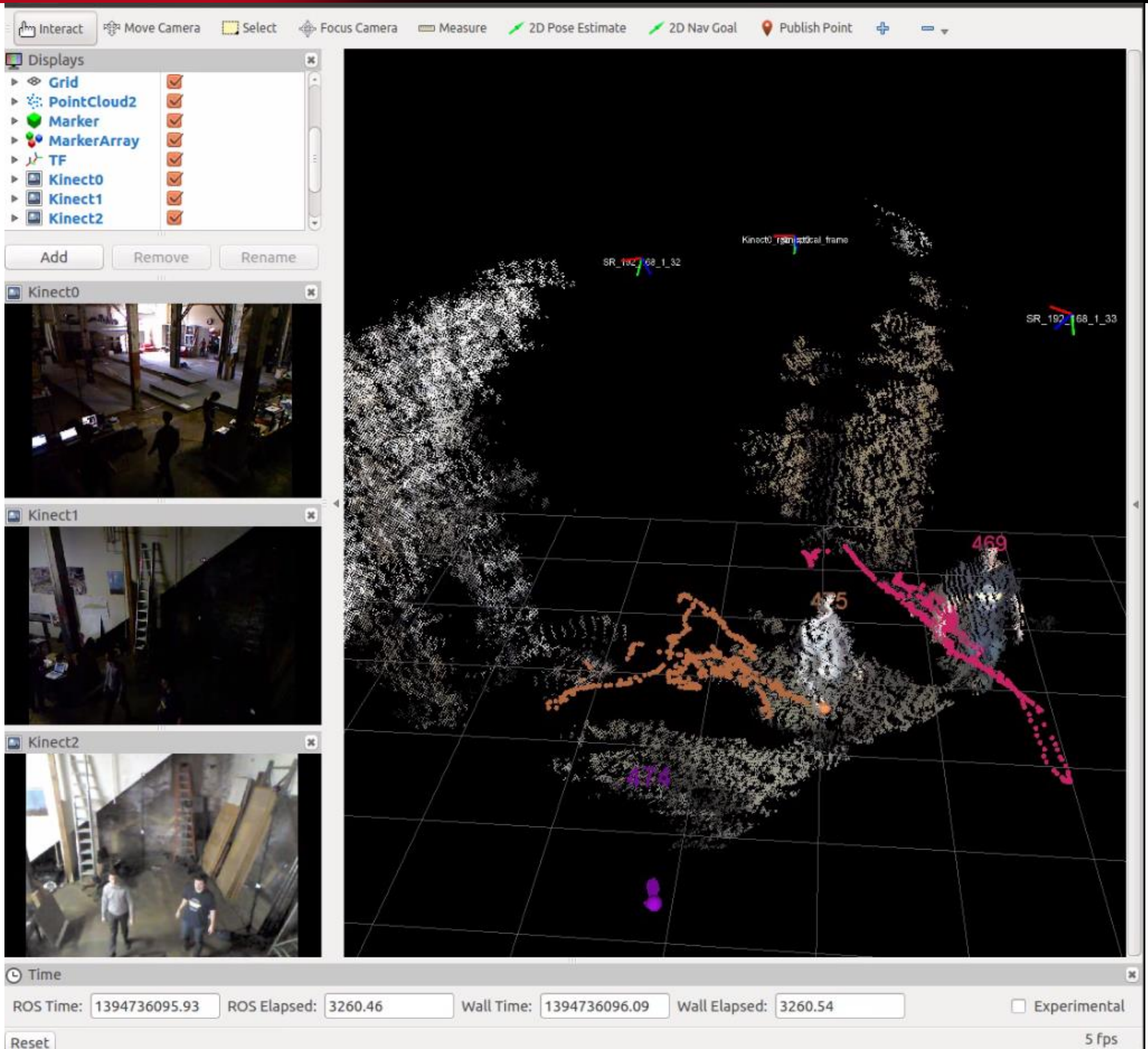
Results: Tracking from two Kinect cameras



Results: Tracking from three Kinect cameras



Results: Tracking from three Kinect cameras

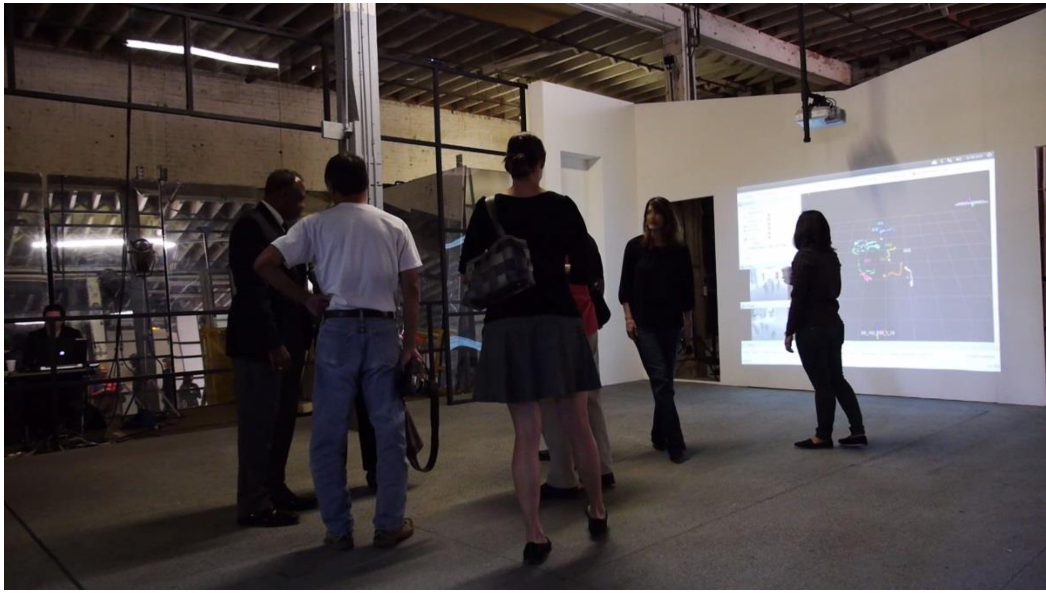


Università degli Studi di Padova

OpenPTrack installations

IAS-Lab

- **Chiparaki**, Los Angeles (CA), March 2014
  - Six cameras (three Kinects, three SwissRangers)
  - 8x6x4m Pavilion
  - Tracking output drove an interactive mural



- **UCLA Lab School**, Los Angeles (CA)
  - Applied to children
  - Tracking output drove a science simulation (on-screen avatars)

Conclusions

- **Open source** multi-camera people tracking
- **Distributed** infrastructure for people detection
- **Real-time** and user-friendly **calibration** procedure
- Works with both Microsoft **Kinect** and Mesa **SwissRanger** SR4500
- People detection **adaptation** to brightness conditions

Future works:

- **Improve** extrinsic **calibration**
  - RGB+depth
  - Depth map correction
  - Handle network delays
- **Stereo** camera support
- Increase **ID persistence** (height feature, re-identification signatures?)
- Increase detection accuracy for cameras placed very high and for detecting **children**
- Real-time parameter configuration (**GUI**)

OpenPTrack

Any questions?

matteo.munaro@dei.unipd.it
http://www.dei.unipd.it/~munaro
http://openptrack.org

DIPARTIMENTO
DI INGEGNERIA
DELL'INFORMAZIONE

IAS-Lab