Unsupervised Calibration of Camera Networks and Virtual PTZ Cameras

Horst Possegger, Matthias Rüther, Sabine Sternig,
Thomas Mauthner, Manfred Klopschitz,
Peter M. Roth, Horst Bischof

Motivation



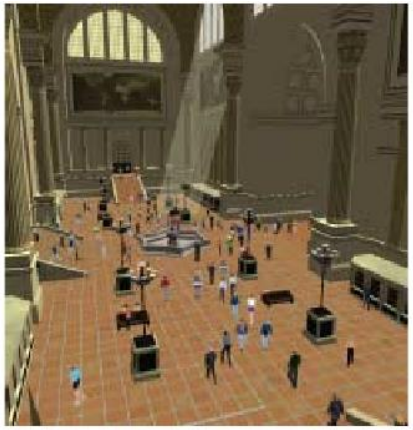

[Video]

Outline

- Virtual PTZ
- Unsupervised Calibration
- Experimental Results
- Conclusion

Related Work – Virtual Cameras

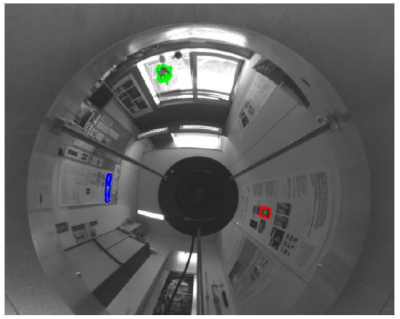
- Virtual vision approach [1]
  - Synthetic camera network
  - Virtual 3D environment
  - Modelling realistic human behaviour is complex



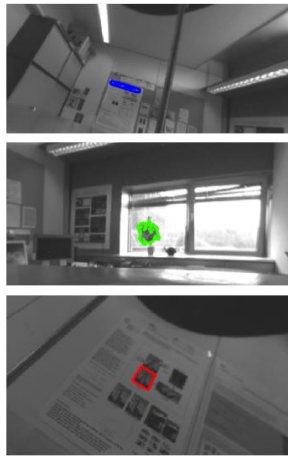
[1] F. Qureshi and D. Terzopoulos. Surveillance Camera Scheduling: A Virtual Vision Approach. *Multimedia Systems*, 12(3), 2006.

Related Work – Virtual Cameras

- Perspective views from omnidirectional images
  - Ishiguro *et al.* [1]
    - Approximate plenoptic representations
    - Visual modeling
  - Mauthner *et al.* [2]
    - Virtual camera planes for feature extraction
    - Region matching



[2]

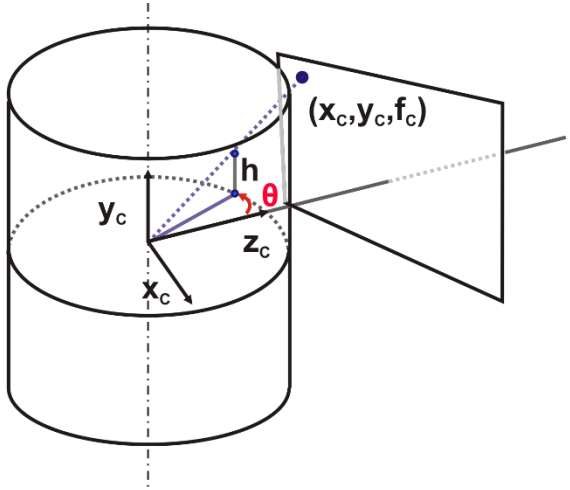


- Wu *et al.* [3]
  - Viewpoint invariant feature descriptors
  - 3D scene alignment

[1] H. Ishiguro, K. Ng, R. Capella, and M. Trivedi. Omnidirectional image-based modeling: three approaches to approximated plenoptic representations. *MVA*, 14(2), 2003.
[2] T. Mauthner, F. Fraundorfer, and H. Bischof. Region matching for omnidirectional images using virtual camera planes. *Proc. CVWW*, 2006.
[3] C. Wu, B. Clipp, X. Li, J.-M. Frahm, M. Pollefeys. 3D Model Matching with Viewpoint-Invariant Patches (VIP). *Proc.* CVPR, 2008.

From Panoramic Imagery ...

- Image stitching [1]
  - Compute 3D ray
  - Intersect with cylinder
  - Map onto flat surface



- Panoramic video
  - Spherical camera
  - *E.g.* Point Grey Ladybug3 with 6 sensors, 2 MP each

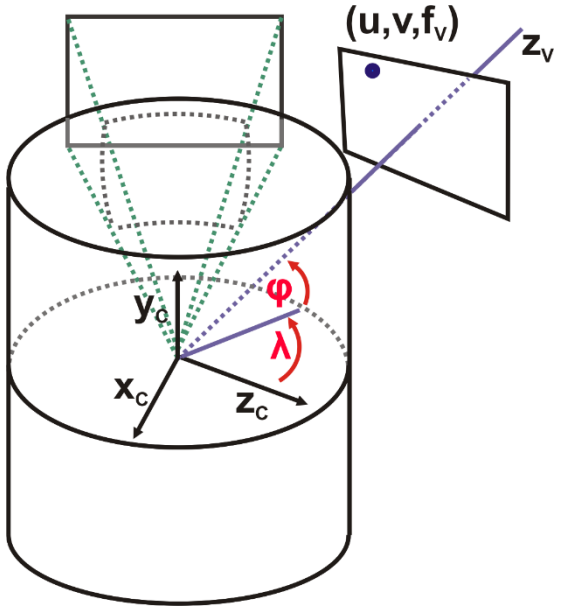
[www.ptgrey.com]

[1] R. Szeliski. Image Alignment and Stitching: A Tutorial. *FTCGV*, 2(1), 2006.

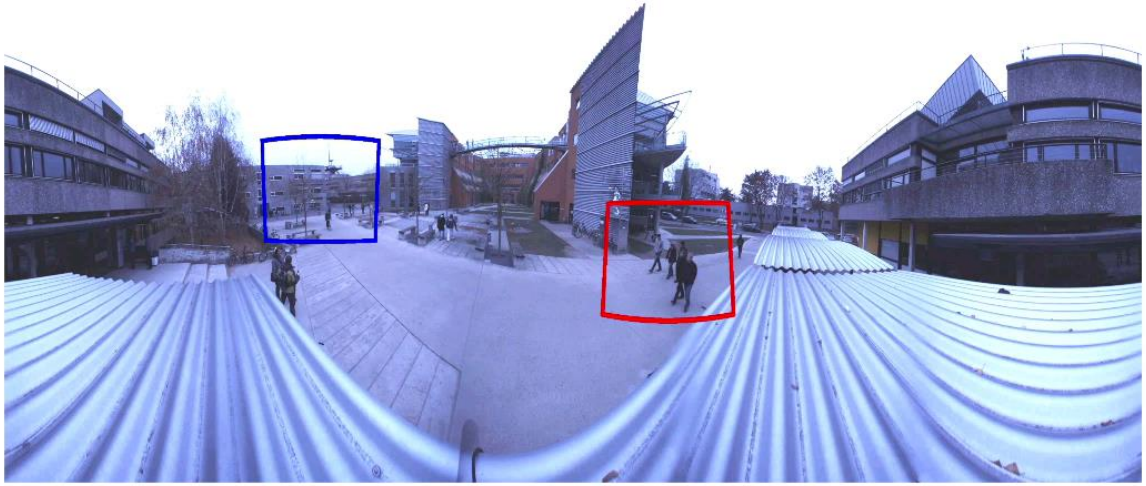
... to Virtual PTZ

- Virtual pinhole camera meets panorama projection
  - Optical center coincides with the center of the cylinder
  - Resample virtual image plane using panorama

- Simulate a PTZ camera off-line
  - Pan/tilt almost arbitrarily
    - Pan – Rotation around y-axis
    - Tilt – Rotation around x-axis
  - Zoom by changing the focal length



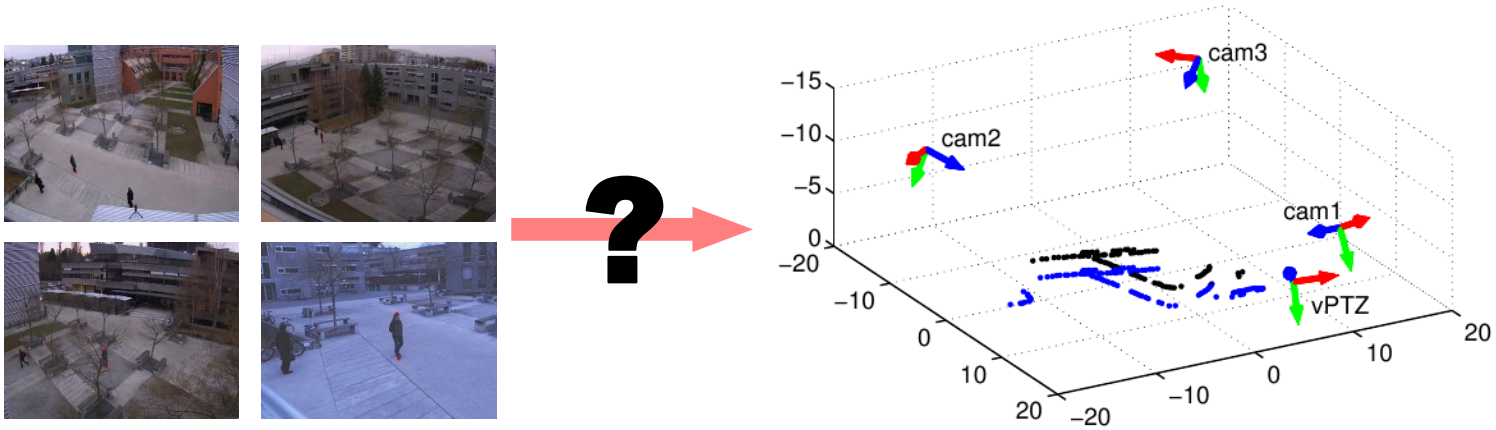
Virtual PTZ



[Video]

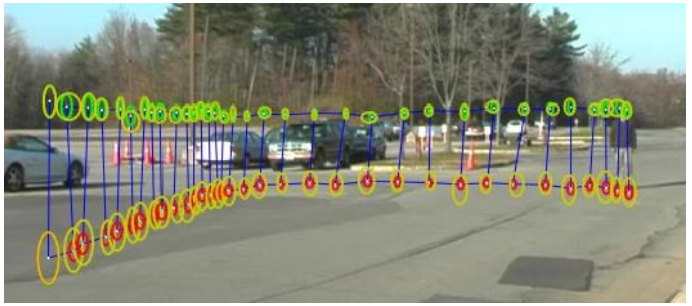
Unsupervised Calibration

- Motivation
  - Camera network of static and PTZ cameras
  - Extrinsic calibration necessary
  - Simple to use
  - Simplify development using vPTZ



Related Work – Self-Calibration

- Point correspondences from moving objects
  - Lee *et al.* [1]
    - Centroids of moving objects
    - Additional refinement steps
  - Lv *et al.* [2]
    - Vanishing points
    - Estimated from pedestrians
    - Sensitive to noise
  - Krahnstoever and Mendonça [3]
    - Statistical model of human motion
    - Reduces noise sensitivity

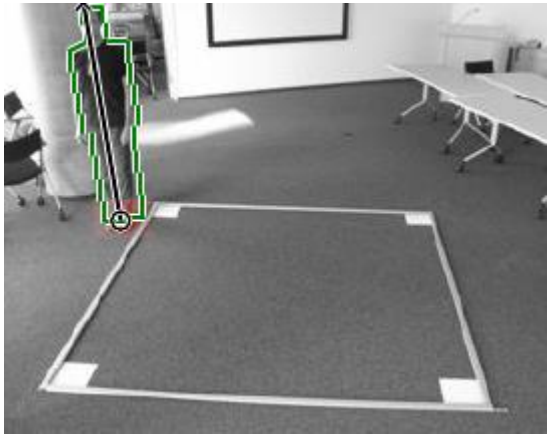


[KM06]

[1] L. Lee, R. Romano, G. Stein. Monitoring Activities from Multiple Video Streams: Establishing a Common Coordinate Frame. *PAMI*, 22(8), 2000.
[2] F. Lv, T. Zhao, and R. Nevatia. Self-Calibration of a camera from video of a walking human. *Proc. ICPR*, 2002.
[3] N. Krahnstoever and P. Mendonça. Autocalibration from Tracks of Walking People. *Proc. BMVC*, 2006.

Related Work – Self-Calibration

- Puwein *et al.* [1]
  - Network of PTZ cameras
  - Geometric constraints from foot trajectories of soccer players
  - Refinement via detected field lines
- Micusik and Pajdla [2]
  - Simultaneous calibration and foot-head homology estimation
  - Silhouette imagery of a synthetic 3D model
  - Human needs to stand still



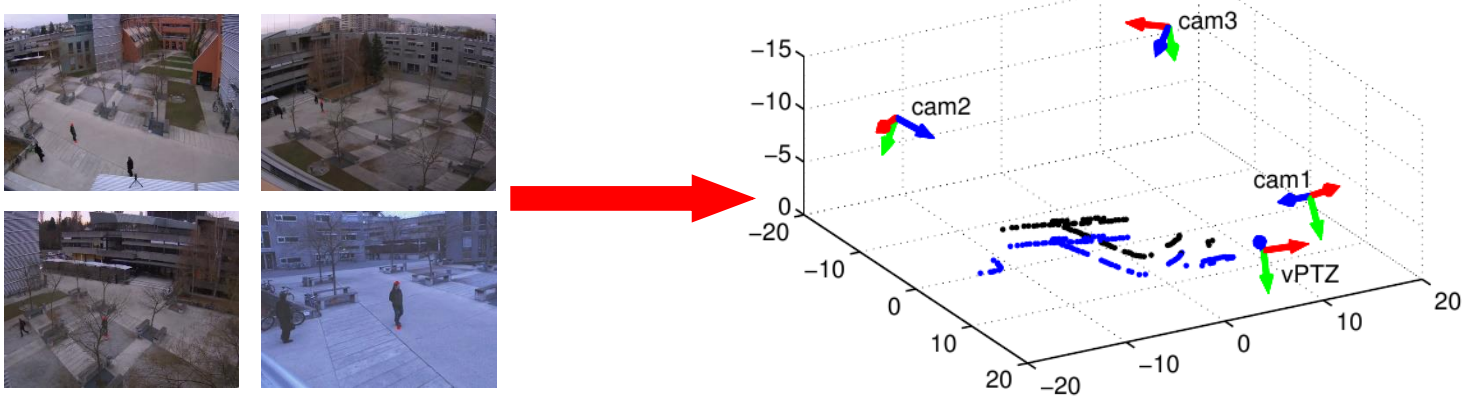
[MP10]

[1] B. Micusik and T. Pajdla. Simultaneous surveillance camera calibration and foot-head homology estimation from human detections. *Proc. CVPR*, 2010.
[2] J. Puwein, R. Ziegler, L. Ballan, and M. Pollefeys. PTZ Camera Network Calibration from Moving People in Sports Broadcasts. *Proc. WACV*, 2012.

Self-Calibration Method

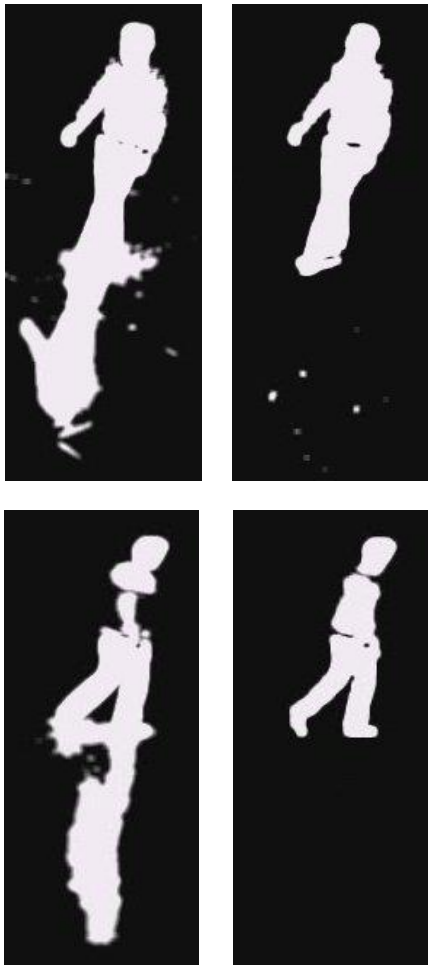
- Track pedestrian
  - Extract point correspondences
  - Remove outliers

- Modified Bundle Adjustment



Detecting Head and Foot Locations

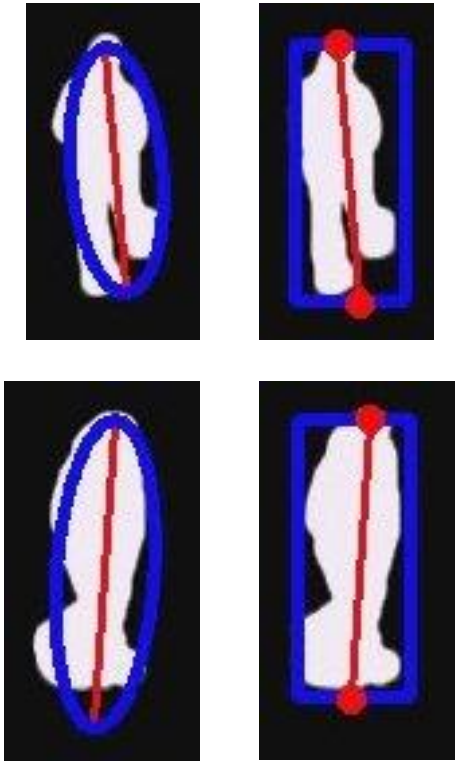
- Adaptive background models
  - Intensities and HSV in parallel
  - Combination reduces
    - Weak reflections
    - Penumbras (soft shadows)
- Blob tracking
  - Size similarity
  - Spatial proximity



Intensity Saturation

Detecting Head and Foot Locations

- Ellipse fitting
  - Corresponding foreground blob
  - Eigenanalysis of covariance matrix
  - First eigenvector axis upright
  - Intersect eigenvector axis with minimum bounding box
- Outlier Removal
  - Pairwise homographies
  - RANSAC
  - Intersection of consensus sets



Reprojection Errors

- Static cameras [1]
  - Normalized camera
    $$\hat{P} = [R|t], \quad x = \hat{P}X$$
  - Normalize image measurements using known intrinsics
- Reprojection error for static camera *j*
  $$E^s = \|x_j - \hat{P}_j X\|^2$$

- (v)PTZ
  - Pan/tilt/zoom parameters
  - Spherical camera coordinate system
  - Image measurements
    $$(\omega, \psi)^\top$$
  - 3D world points
    $$(\Omega, \Psi)^\top$$
- Reprojection error for (v)PTZ camera
  $$E^d = \|(\omega, \psi)^\top - (\Omega, \Psi)^\top\|^2$$

[1] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2nd ed., 2004.

Camera Calibration

- Non-linear least squares optimization problem
  - Static cameras N_s
  - (v)PTZ cameras N_d

$$\arg\min_{R_j,t_j,a_i,b_i} \sum_{j\in N_s,i} E^s_{i,j} + \sum_{j\in N_d,i} E^d_{i,j}$$

$$E^s_{i,j} = \|x_{i,j} - \hat{P}_j X_i\|^2$$

$$E^d_{i,j} = \|(\omega_{i,j}, \psi_{i,j})^\top - (\Omega_{i,j}, \Psi_{i,j})^\top\|^2$$

- Solution using iterative Levenberg-Marquardt
  - Initialized without knowledge of real positions
  - Sufficiently accurate estimates

Experiments

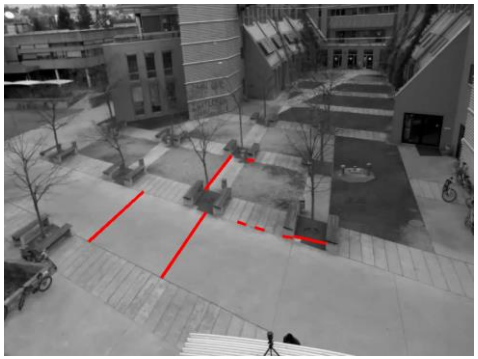





- Outdoor
  - 3 static Axis P1347
  - 1 Point Grey Ladybug3
- Indoor
  - 4 static Axis P1347
  - 1 Point Grey Ladybug3
- Data sets publicly available
  http://lrs.icg.tugraz.at/download.php#vptz

Results

- Evaluation
  - Project known points onto ground plane
  - Compare distance measurements
    - Lengths between 0.5 and 6.0 meters (outdoor)
    - Lengths between 3.6 and 5.1 meters (indoor)

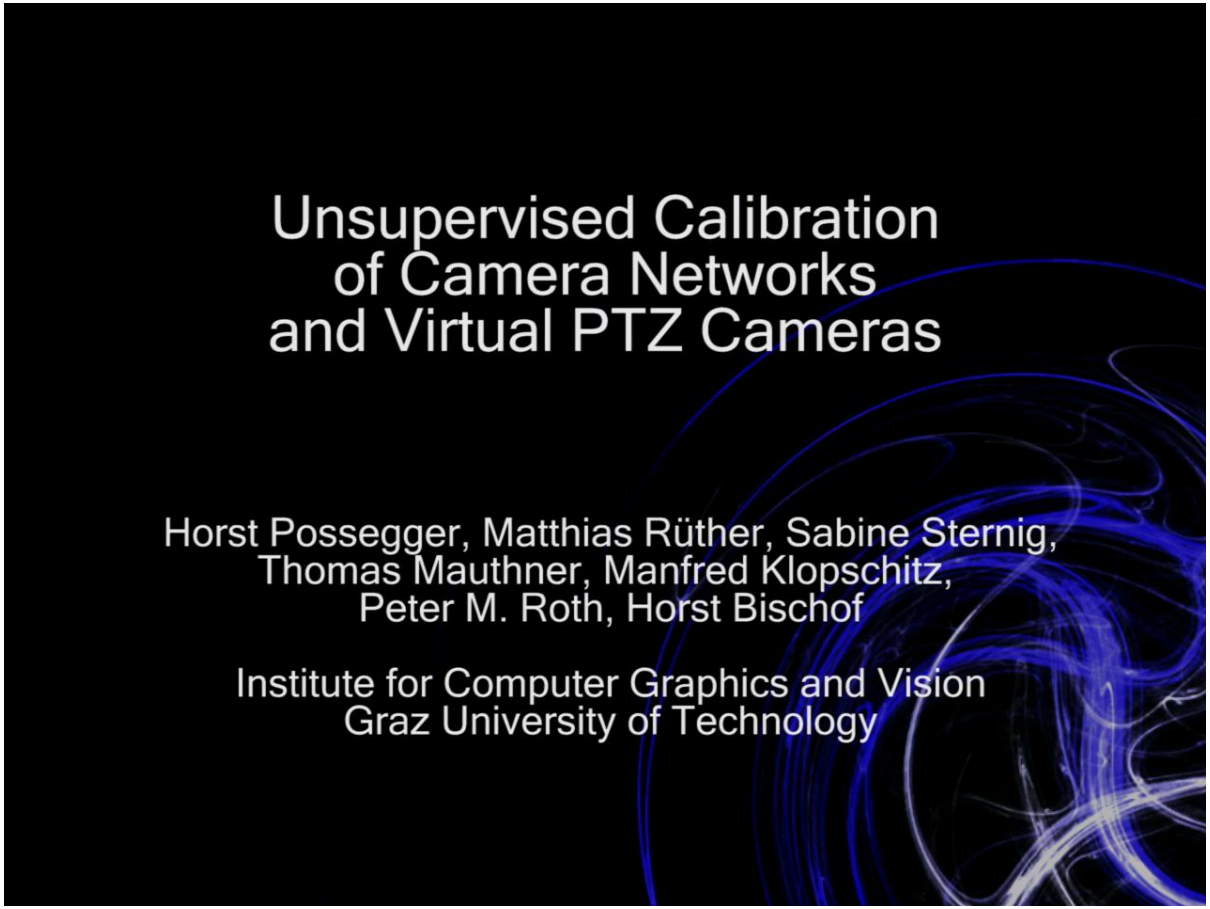


Mean reprojection errors [%]					
Data set	**Camera**				
	1	**2**	**3**	**4**	**vPTZ**
Outdoor	2.8	3.7	12.0	n.a.	4.5
Indoor	3.8	2.9	4.4	2.4	5.8

Conclusion

- Virtual PTZ
  - Panoramic imagery
  - Simulate a real PTZ camera off-line
  - Simplifies development
- Camera network calibration
  - Extrinsic parameters
  - Correspondences from a walking human
  - Sufficiently accurate estimates
- Future work
  - Multiple object tracking
  - Controlling (v)PTZ cameras

Thank you for your attention



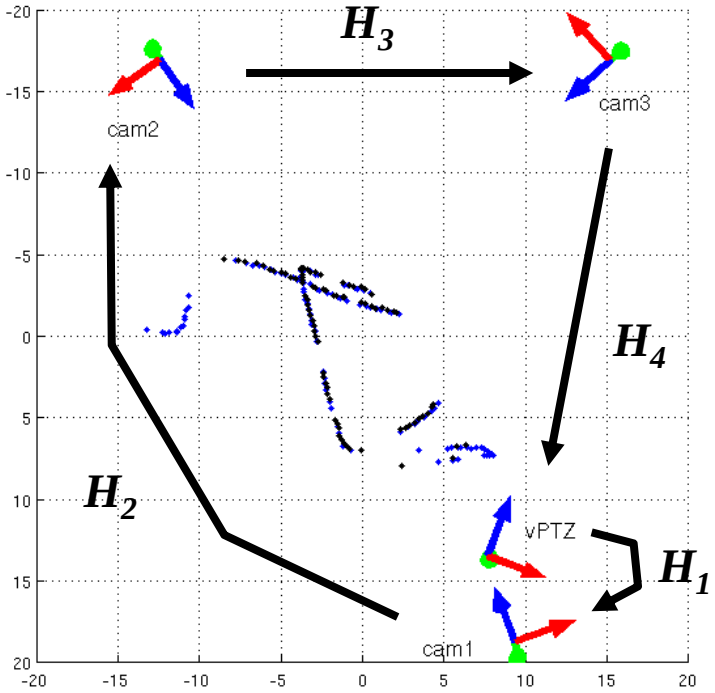
Data sets & sample implementation publicly available
http://lrs.icg.tugraz.at/download.php#vptz

[Video]

Sample Views

Outlier Removal

- Pairwise homographies
  - N camera pairs
  - RANSAC
  - Intersection of consensus sets
- Maximum set
  - Spatially neighboring cameras
  - Slightly improves results



RE – Static Cameras

- 3D foot and head points

$$X = (a, b, 0, 1)^\top, \quad X = (a, b, h, 1)^\top$$

- Pixel-based reprojection error [HZ04]
  - Normalized camera j
    - Intrinsics K_j known

$$\hat{P} = [R|t], \quad x = \hat{P}X$$

  - Normalized correspondences
    - From image measurements m_j

$$x_j = K_j^{-1} m_j$$

- Reprojection error

$$E^s = \|x_j - \hat{P}_j X\|^2$$

[HZ04] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2nd ed., 2004.

RE – Active Cameras

- (v)PTZ
  - Pan/tilt/zoom parameters at frame n

$$(\lambda_n, \varphi_n, f_n)$$

  - Spherical camera coordinate system
  - Image measurements

$$\omega = \lambda_n + \operatorname{atan} \frac{u_n}{f_n}, \quad \psi = \varphi_n + \operatorname{atan} \frac{v_n}{f_n}$$

  - 3D world points

$$(r_1, r_2, r_3) = [R|t] \, (a_n, b_n, z_n, 1)^\top$$

$$\Omega = \operatorname{atan} \frac{r_1}{r_3}, \qquad \Psi = \operatorname{atan} \frac{r_2}{r_3}$$

- Reprojection error

$$E^d = \| (\omega, \psi)^\top - (\Omega, \Psi)^\top \|^2$$

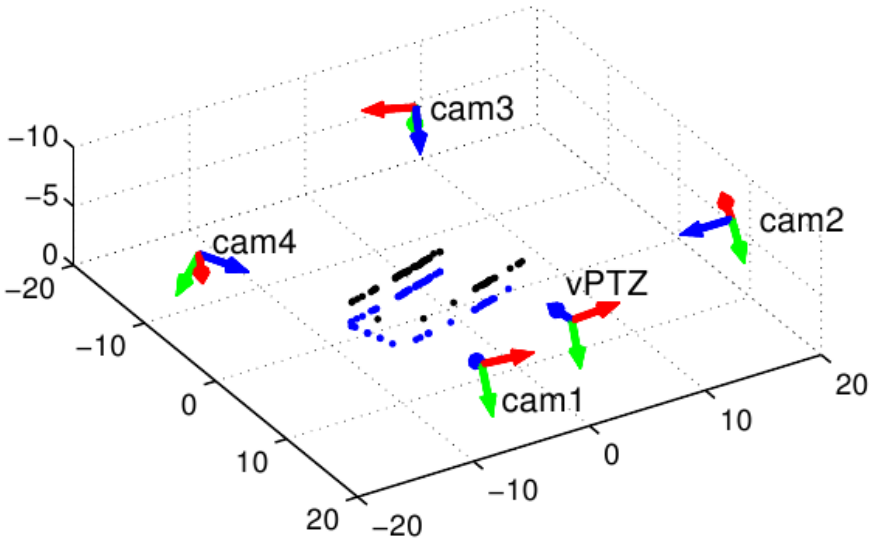
Calibration – From vPTZ to PTZ

- Common geometric model
  - Imaging optics aligned with rotation axes
  - Not true in general, but vPTZ follows ideal model
  - Nevertheless, sufficiently accurate

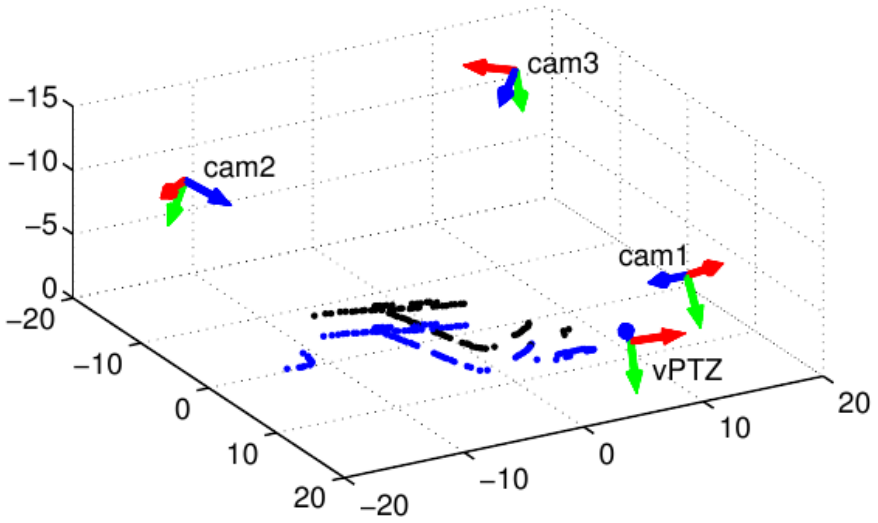
- Varying lens distortion
  - Estimate for several zoom levels
  - Interpolation
  - Undistorted image measurements

Results – Setup

Indoor



Outdoor



Axes units are meters. z-axis is negative, because cameras face in positive direction along their optical axis.

[Video]