Visual-inertial tracking with sparse 3D information

DFKI Localization Workshop 2011

Markus Miezal (markus.miezal@dfki.de)
Research Department *Augmented Vision*
German Research Center for Artificial Intelligence

Contents

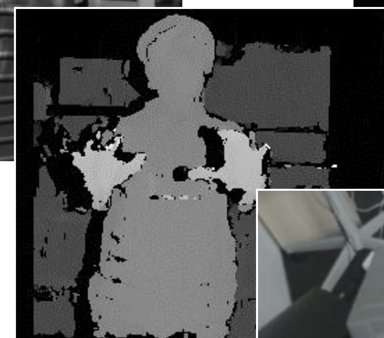
1. Augmented Vision Department
    - Core activities and competences
    - Project: Capture
    - Project: Cognito
2. Visual-inertial tracking with optical flow
3. Future Work

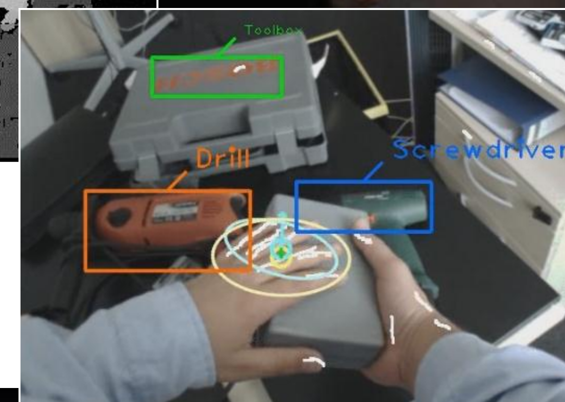
Augmented Vision Department

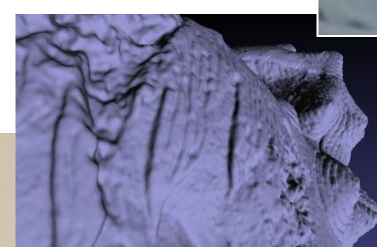
Core activities:

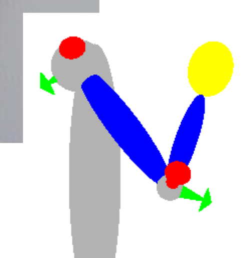
- Sensor fusion:
  - (body) motion tracking
  - activity recognition
- Computer vision:
  - 3D reconstruction
  - object recognition
- Visualization and rendering:
  - information visualization
  - realistic rendering
  - collaborative interaction

28 fulltime researchers









Augmented Vision Department

Application domains:

- Virtual engineering
- Ambient assisted living
- Safety and Security

Software platforms:

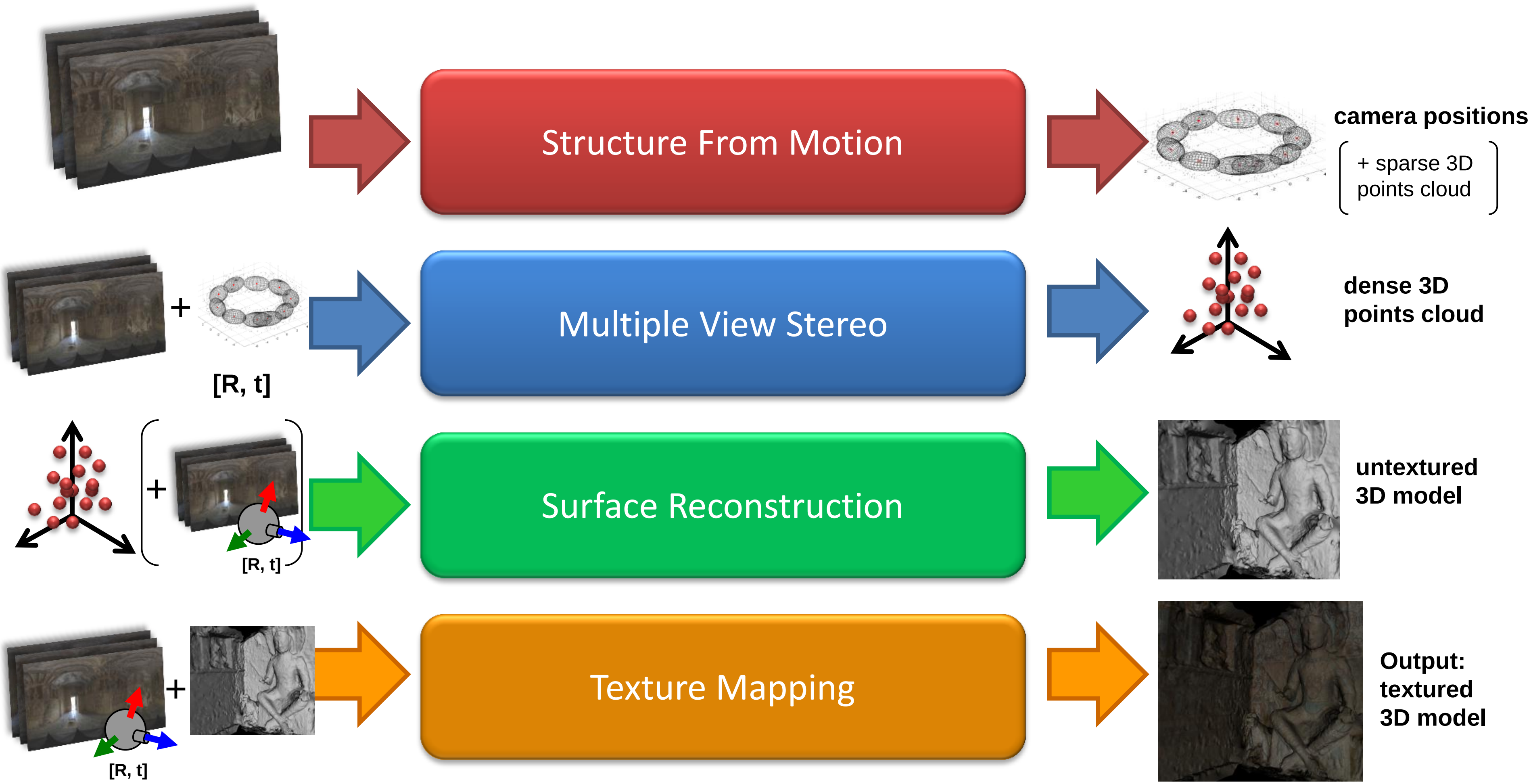
- **Argos:** data-driven parallel framework for scientific prototyping
- **Odysseys:** system for realistic rendering and collaborative work

Capture

3D-scene reconstruction with high resolution and high dynamic range spherical images

- Using special camera:
  - Civetta (360° x 180° HDR images, resolution 7393x14786 pixels)

Capture



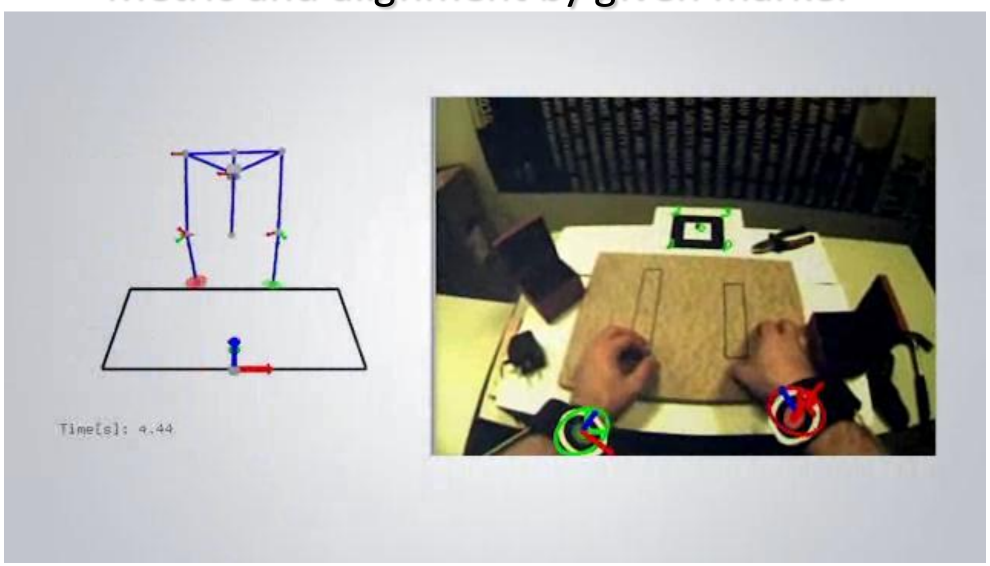
camera positions
+ sparse 3D points cloud
Structure From Motion
+
[R, t]
Multiple View Stereo
dense 3D points cloud
+
[R, t]
Surface Reconstruction
untextured 3D model
+
[R, t]
Texture Mapping
Output: textured 3D model

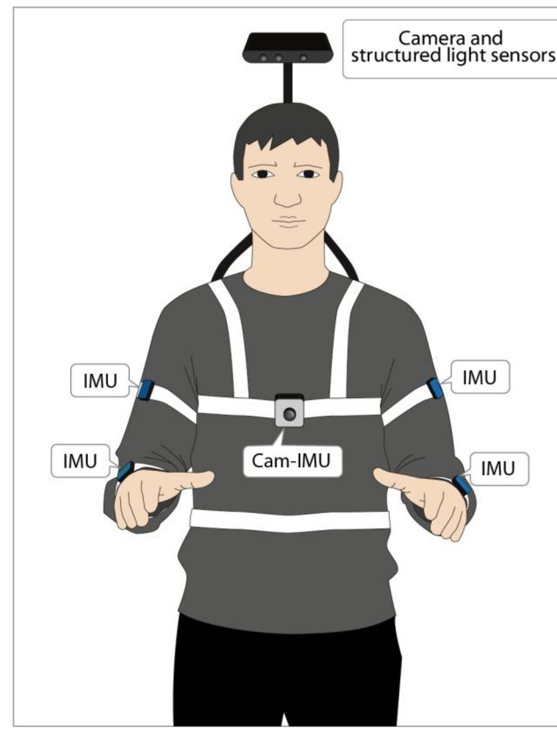
Future work: Localization

Cognito

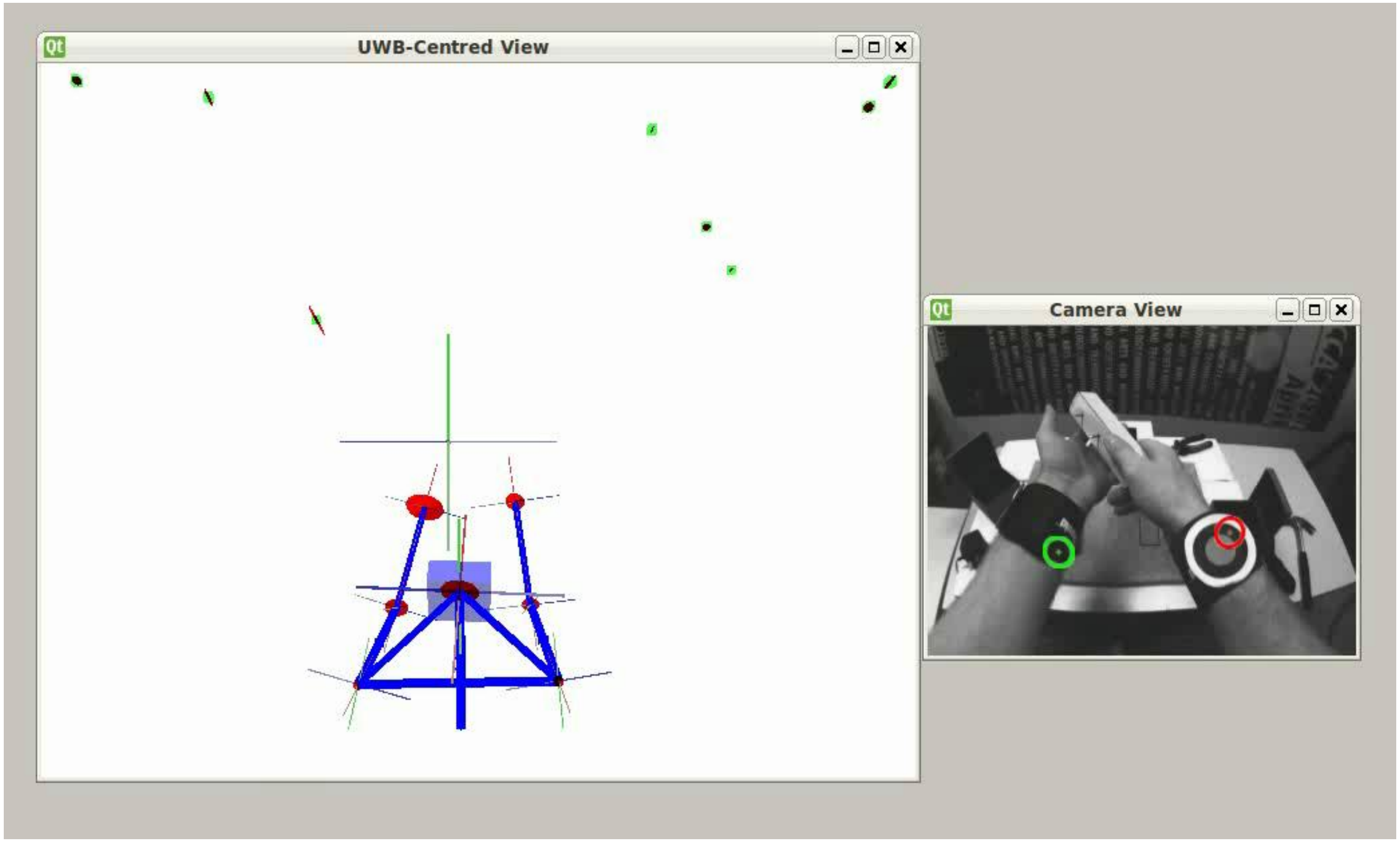
Cognitive Workflow Capturing and Rendering with On-Body Sensor Networks

- Sensor fusion mainly for upper body tracking
- SLAM used for global localization (provided by the University of Bristol)
  - Pure monocular visual SLAM
  - Metric and alignment by given marker





Cognito



Future work: Add Visual-inertial data to SLAM system

Visual-inertial tracking with optical flow

Sensor qualities

+ Absolute measurements

- Occlusion
- Usually slow
- Motion blur
- No alignment
- Usually unknown scale

CamIMU

+ Fast
+ Metric and alignment

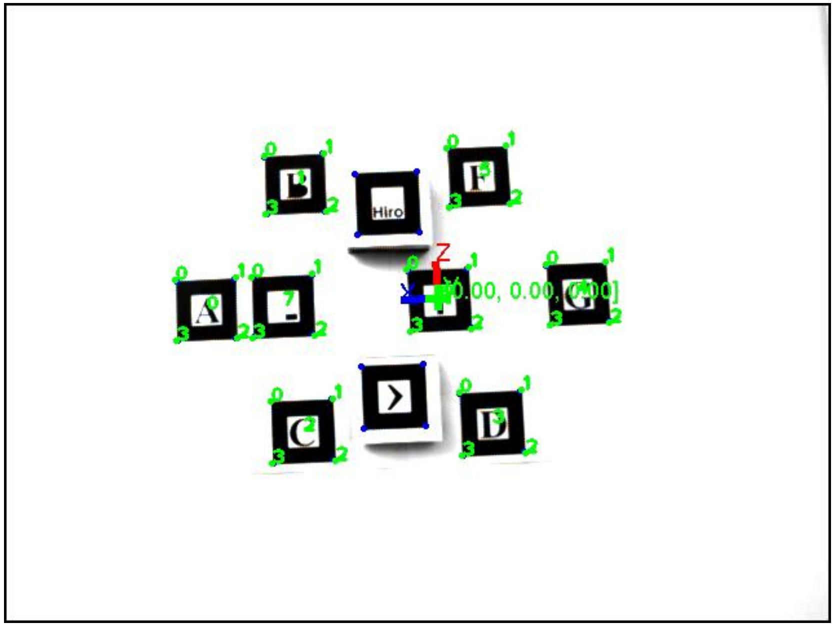
- Relative Measurements
- Eventually drifts

Sensor fusion

- EKF with state $\left(s_w, \dot{s}_w, \ddot{s}_w, q_{sw}, \omega_s, b_\omega\right)^T$
- Time update (100 Hz):
  - Constant acceleration
  - Constant rotation speed
- Measurement:
  - IMU: y_ω and y_a (100 Hz)
  - Vision: 2D/3D correspondences (reprojection error) (25 Hz)

Problems

- IMU induces a drift when only a few 2D/3D correspondences are available



Use 2D/2D correspondences instead

- Cheap and fast acquisition
- No 3D information needed

Measureing optical flow

- Derive the point projection from world to camera by t
- 3D information for OF is not available: Eliminate z.
- Inner product gives 1D innovation

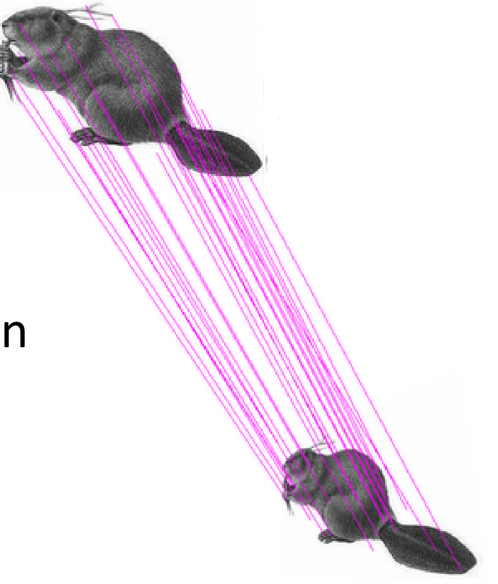
$$0 = h(x) = (\dot{\tilde{m}}_n + e_{\dot{m}})^T (v_c^{cw} \times (\tilde{m}_n + e_m)) + (\tilde{m}_n + e_m)^T (\omega_c^{cw} \times (v_c^{cw} \times (\tilde{m}_n + e_m)))$$

$$\omega_c^{cw} = -R_{cs}\omega_s$$

$$v_c^{cw} = \omega_c^{cw} \times (R_{cs}T_s^{sc}) - R_{cs}Rot(q_{sw})\dot{s_w}$$

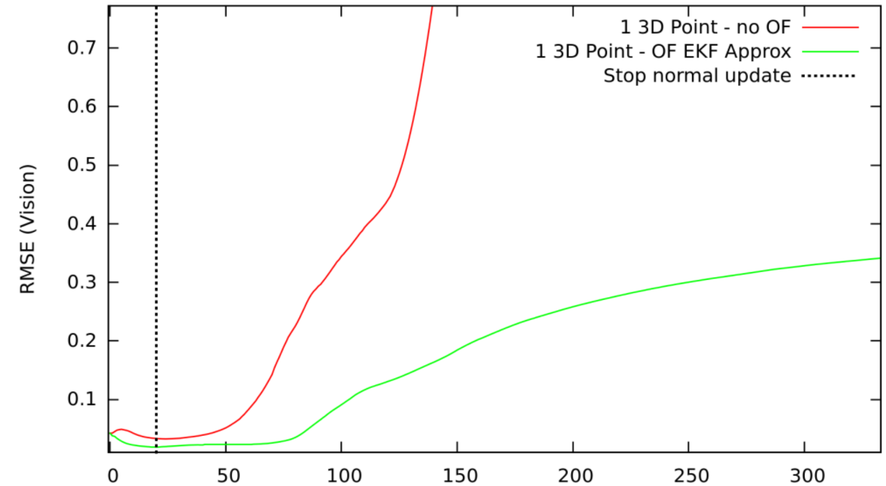
Optical flow acquisition

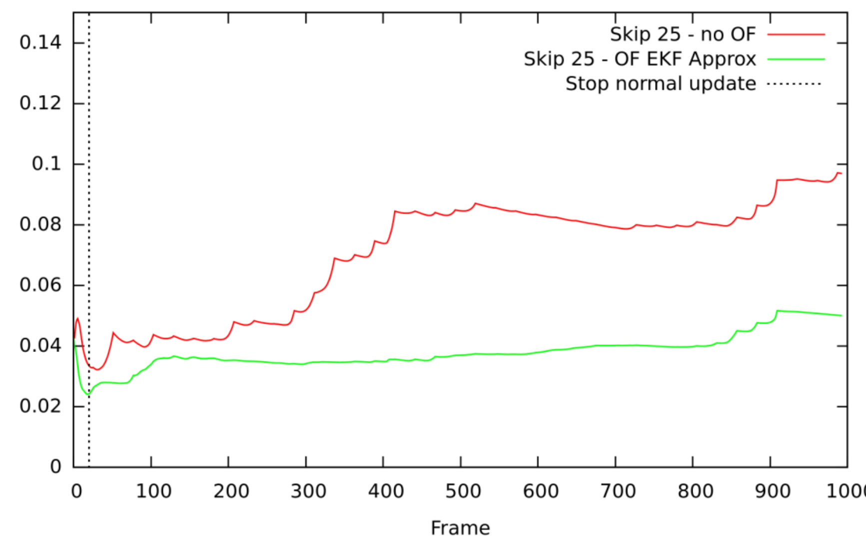
1. Get point matches
   - KLT, SURF, SIFT, Marker points…
2. Derive point speed from consecutive points using:
   - Euler approximation: $\dot{m} = \frac{m_2 - m_1}{\Delta t}$
   - Approximate using KF
     - Time update: Constant Velocity/Acceleration
     - Measure: Point position

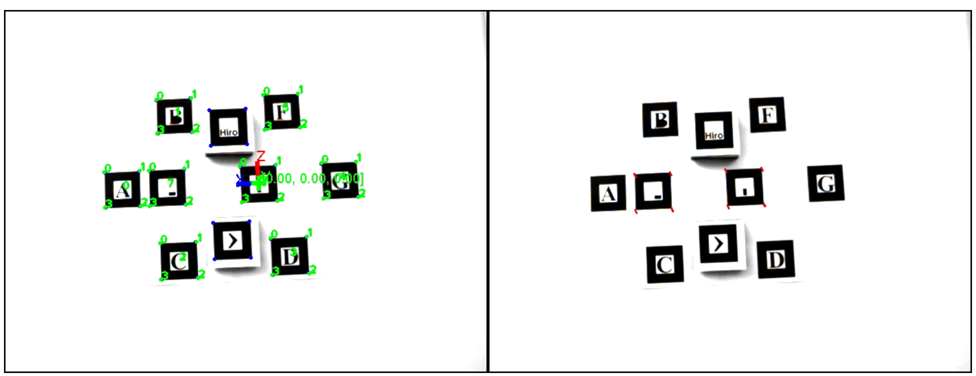


Results (so far)

- Improvements are visible (only when less 3D information is available)





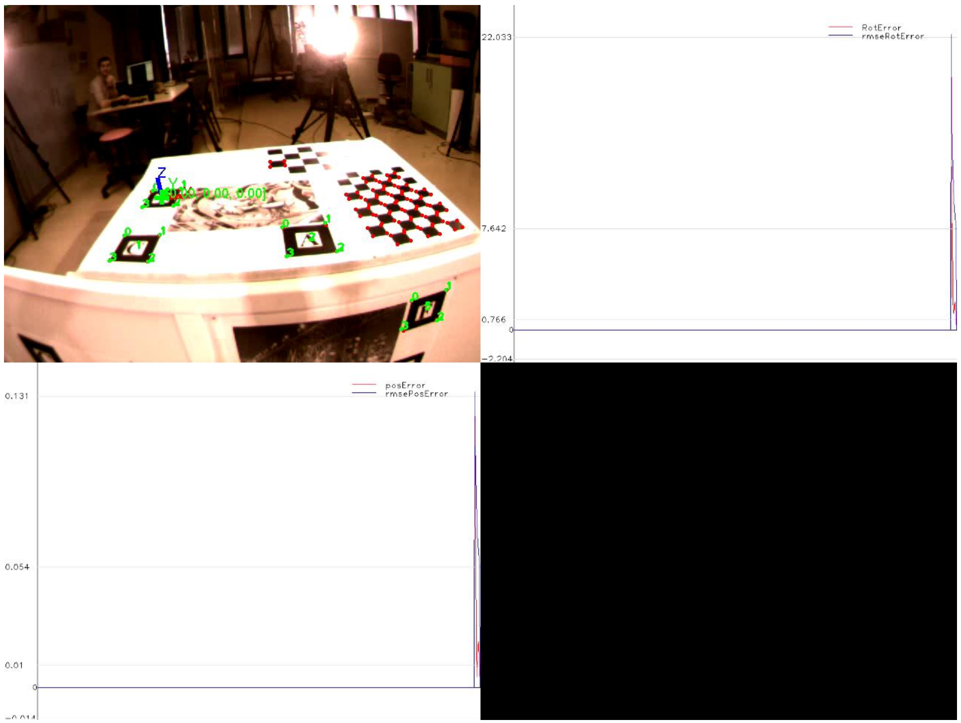


Known problems (Calibration)

- A Hand-Eye calibration (R_{cs}, t_{cs}) is crucial for accuracy
- State of the art:
  1. Relative rotation
     - Align down vector of vision (chessboard detection) with down vector from IMU
  2. Relative translation
     - Optimize t_{cs} over normalized residuals of a filter
- Calibration is highly dependent on parameters and Hardware setup

Future work

- Evaluation on Vicon data



Questions?

Thank you!