Tracking competitions for evaluating visual SLAM techniques

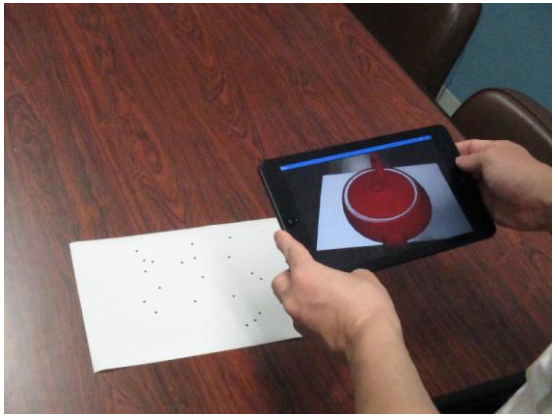
Hideaki Uchiyama
Kyushu University

2016/10/08

Laboratory for Image and Media Understanding

九州大学
KYUSHU UNIVERSITY

Introduction



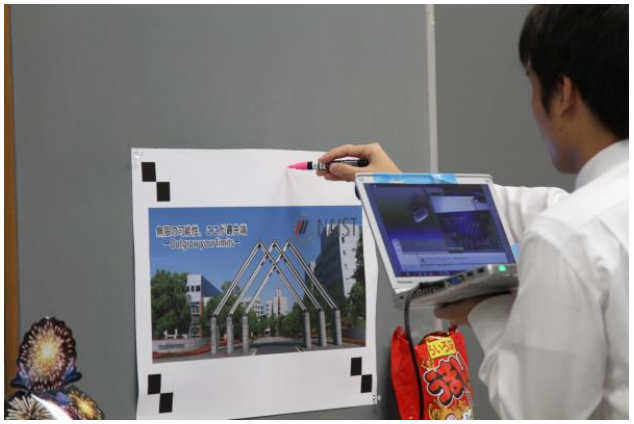
Overlay of teapot on a marker

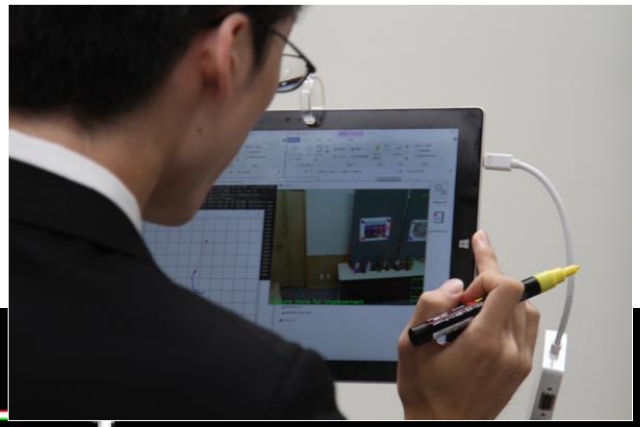
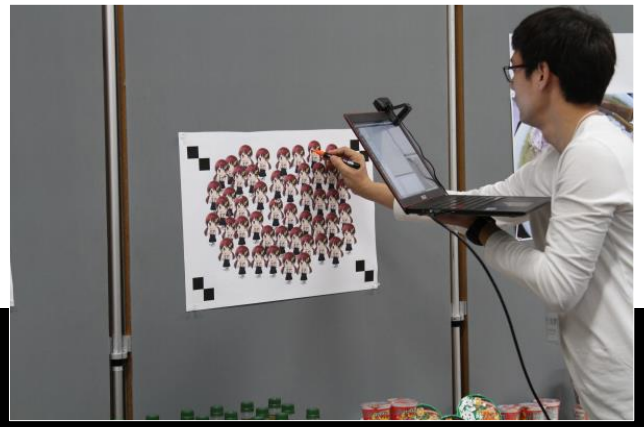
- About me
  - Assistant professor at Kyushu University
  - Working on augmented reality (AR)
  - Tracking competition committee in ISMAR2015

- About ISMAR
  - International Symposium on Mixed and Augmented Reality
  - Visual tracking papers presented like PTAM and KinectFusion

What is tracking competition?

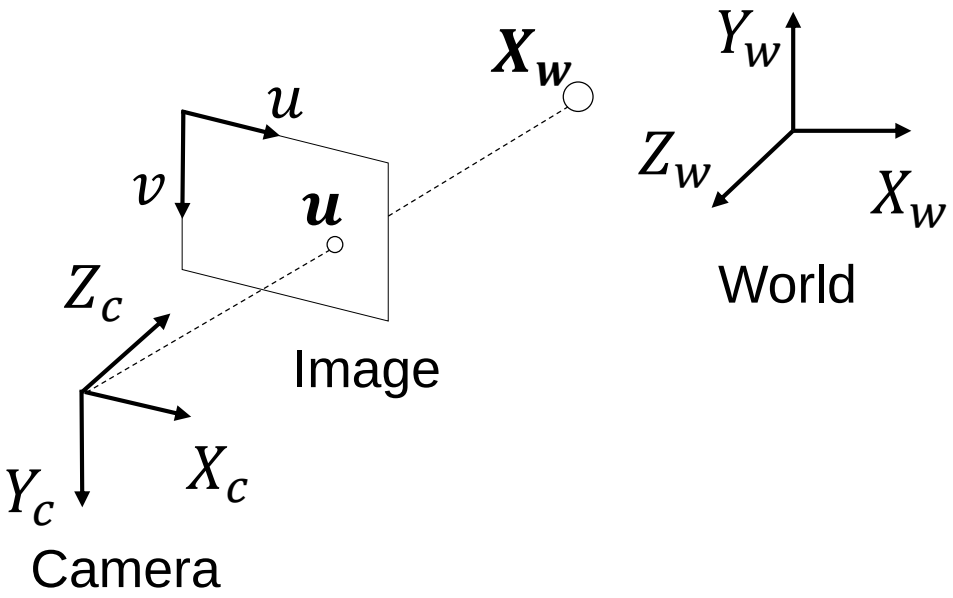
- Participants develop a vision based mobile guidance system for some tasks and compete at a conference site
  - Started since 2008 in ISMAR



What is visual tracking in AR?

- Estimation of 6 DoF camera pose (PnP problem)
  - not only for AR, but for visual odometry/SLAM



Coordinate systems

Camera pose

$$s\tilde{\boldsymbol{u}} = \boldsymbol{K}[\boldsymbol{R}|\boldsymbol{t}]\tilde{\boldsymbol{X}}_{\boldsymbol{w}}$$

Image point | Camera parameters | World point

Projection model

Standard benchmarking

- Dataset
  - KITTI Vision Benchmark Suite (stereo, laser, large scale outdoor)
    http://www.cvlibs.net/datasets/kitti/
  - TUM RGB-D SLAM Dataset (RGBD, indoor)
    http://vision.in.tum.de/data/datasets/rgbd-dataset

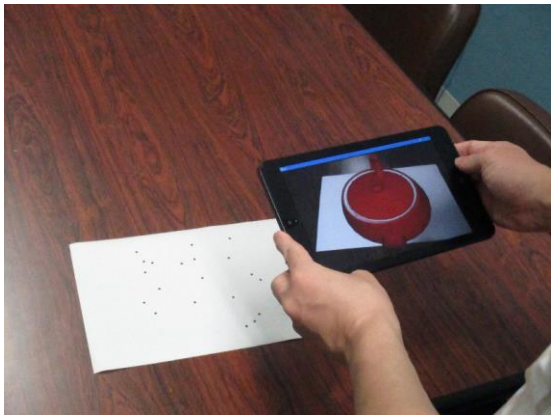
- Evaluation criteria
  - Rotation error
  - Translation error

$$s\tilde{\boldsymbol{u}} = \boldsymbol{K}[\boldsymbol{R}|\boldsymbol{t}]\tilde{\boldsymbol{X}}_{\boldsymbol{w}}$$

Rotation error

Translation error

Another benchmarking in AR



- Dataset
  - Metaio (closed)
  - TrakMark (synthetic and real images)
    http://ypcex.naist.jp/trakmark/

- Evaluation criteria in TrakMark
  - Projection error of virtual objects

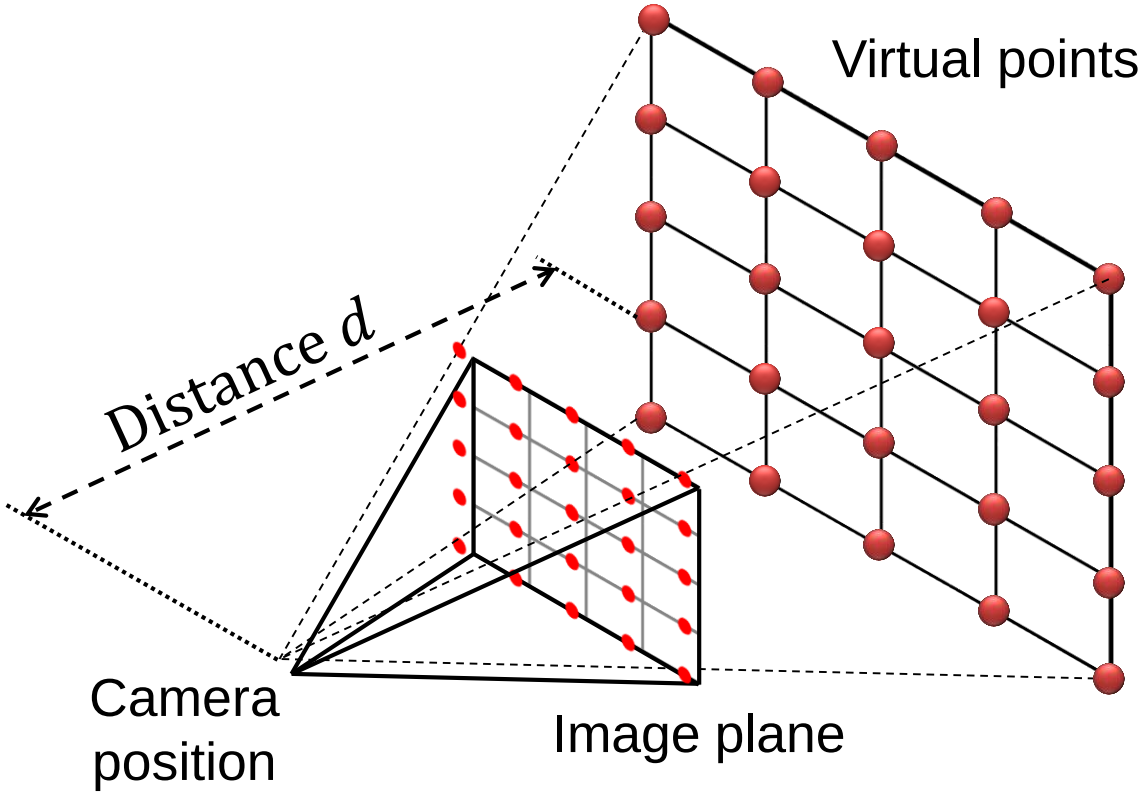
$$s\tilde{\boldsymbol{u}} = \boldsymbol{K}[\boldsymbol{R}|\boldsymbol{t}]\tilde{\boldsymbol{X}}_{\boldsymbol{w}}$$ ← Virtual objects

$$E = \sum_{i} \left\| \boldsymbol{u}_i(\boldsymbol{R}_e, \boldsymbol{t}_e) - \boldsymbol{u}_i(\boldsymbol{R}_g, \boldsymbol{t}_g) \right\|$$

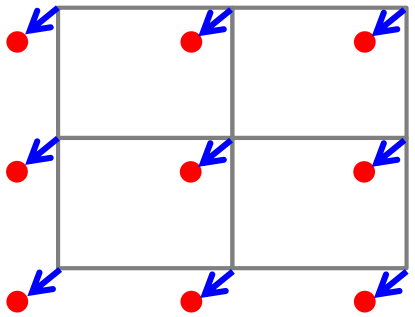
Estimated　　Ground truth

Projection error of virtual points

- Error in image plane is important in AR



Projection of virtual points



Error in image

Tracking competition as onsite benchmarking

- Compared to dataset based benchmarking
  - Parameter optimization is less influential
  - Actual performance is demonstrated
  - Usability of systems is also demonstrated
  - BUT, hard to measure camera pose

- Issues on onsite benchmarking
  - Design of evaluation criteria, rules, environment for fair comparison
  - Management at a competition site
  - Attractiveness for audience

Tracking Competition in ISMAR2008

- Task
  - Pick up objects in a room using own guidance system
- Procedure
  - Setup the system using known 3D points within 5 hours
  - Receive 3D coordinates of objects from organizers
  - Overlay object locations onto camera images and pick up objects
- Evaluation criteria
  - Number of correctly picked objects
  - Time needed to pick the objects
  - Time needed for setup

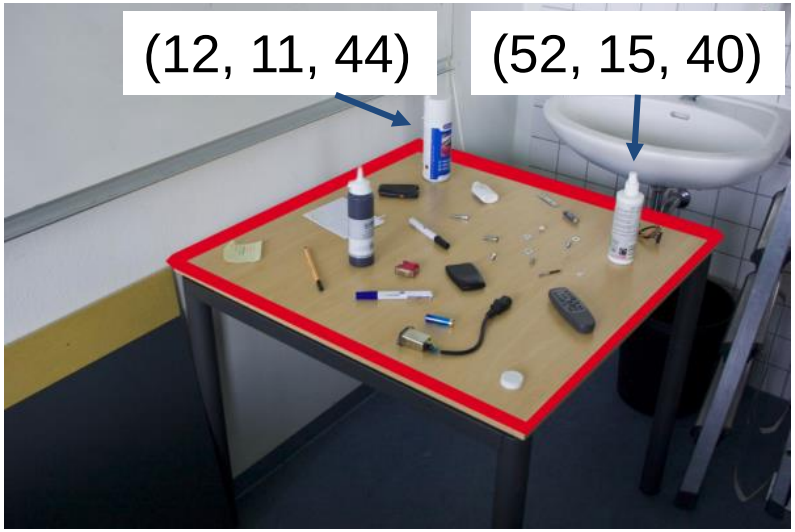
Model based tracking or relocalization

http://ismar2008.vgtc.org/doku2867.html?id=program-competition

Set-up by organizers

- Measure 3D coordinates of reference markers using geodetic equipment
- Prepare questions and their answers

Reference point



Picking area

Tracking Competition in ISMAR 2014

- Four competitions organized by Volkswagen
  - Tracking of a rotation vehicle
  - Tracking and learning on different vehicles
  - Tracking with high accuracy
  - Tracking inside an unknown area

Model based tracking or relocalization

Odometry/SLAM

No winner

http://ismar2014.vgtc.org/ismar/2014/info/overview/vw-tracking-challenge.html

Tracking Competition in ISMAR 2015

- Two competitions organized by SIGMR in VRSJ
  - Onsite for tracking inside an unknown environment
  - Offsite for dataset based benchmarking
    http://ypcex.naist.jp/trakmark/tracking-competition/
- Sample code provided for beginners
  - A simple implementation of monocular visual SLAM
    “Abecedary tracking and mapping”
    https://github.com/CVfAR/ATAM

Flow of onsite competition

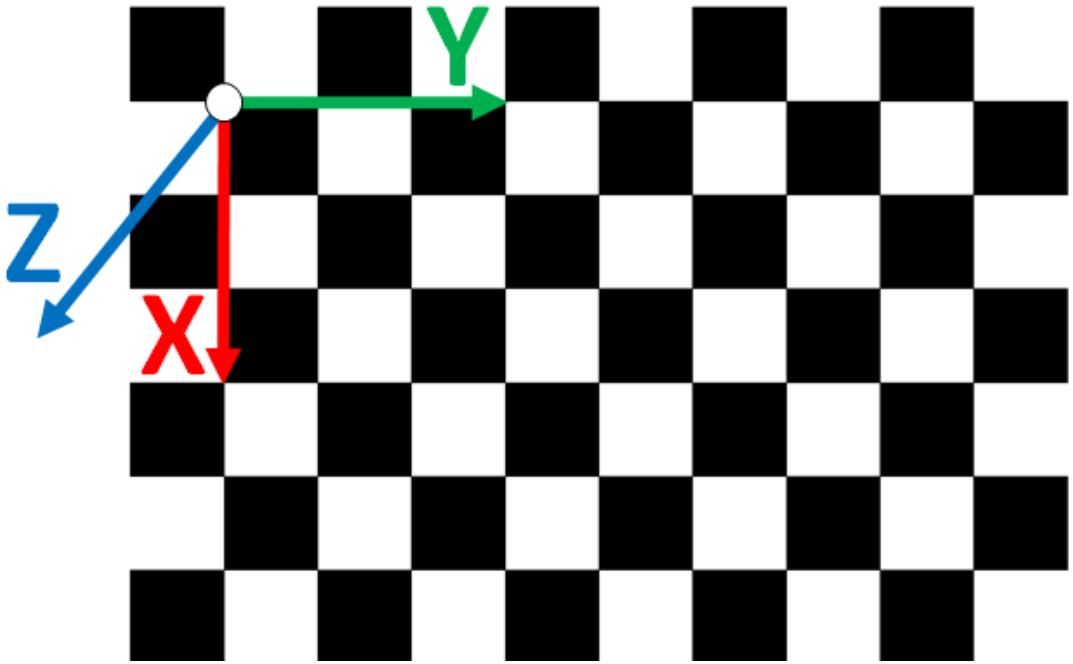
- Receive 3D coordinates of points for marking
- Acquire world coordinate system at starting area
- Track camera poses in an unknown area
- Visualize and mark up challenge points at challenge areas

At starting area

At challenge area

At starting area

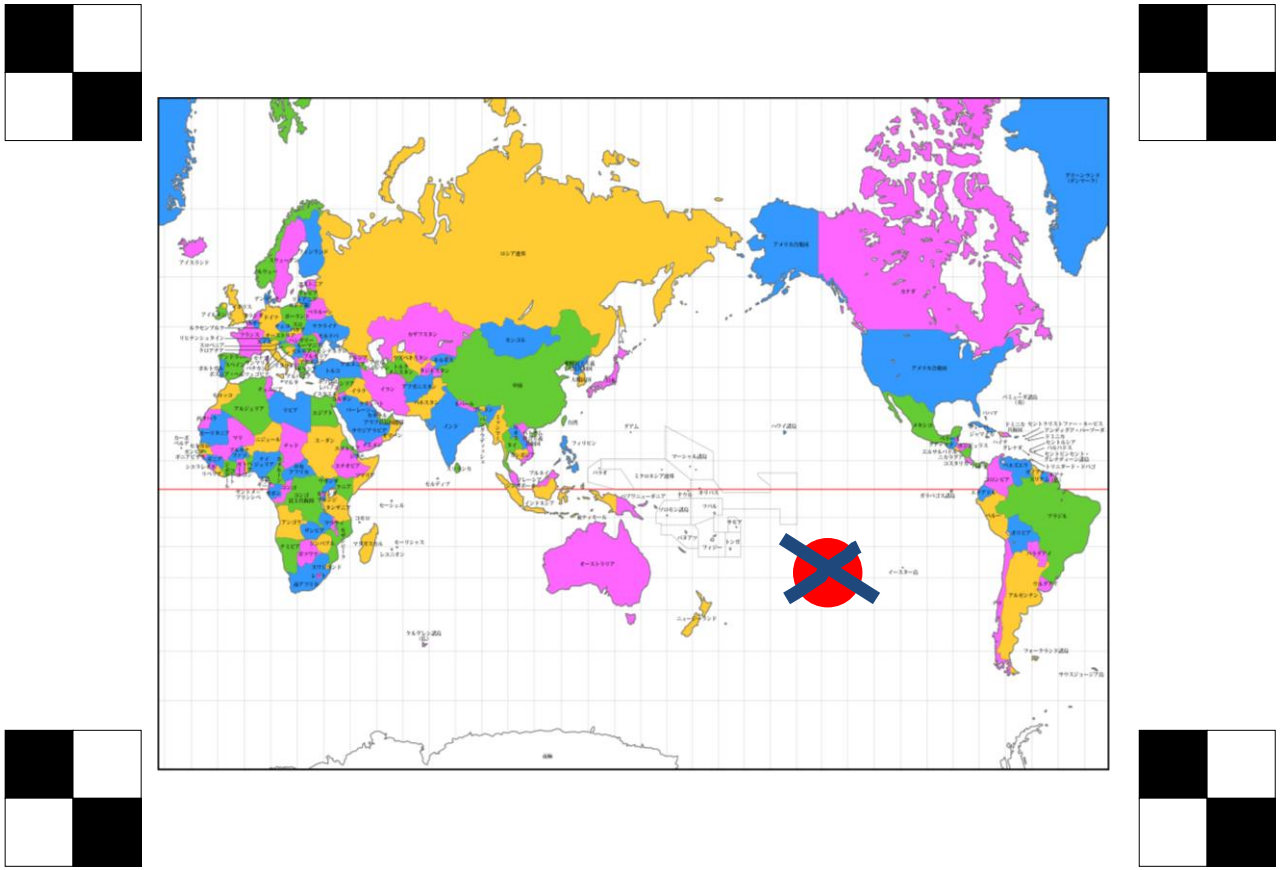
- Acquire world coordinate system



World coordinate system on a chessboard

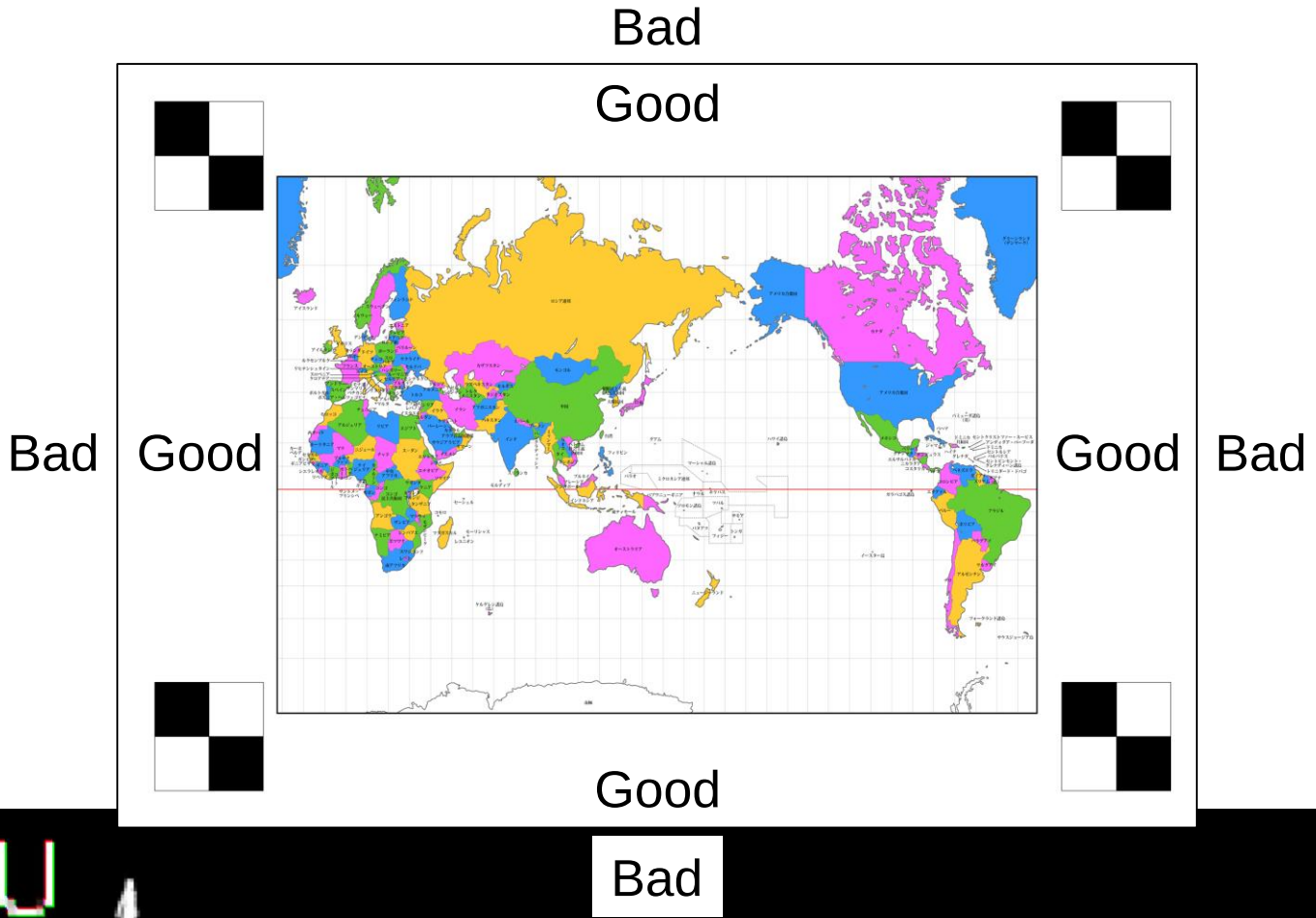
At challenge area

- Visualize a challenge point and mark it up on a poster



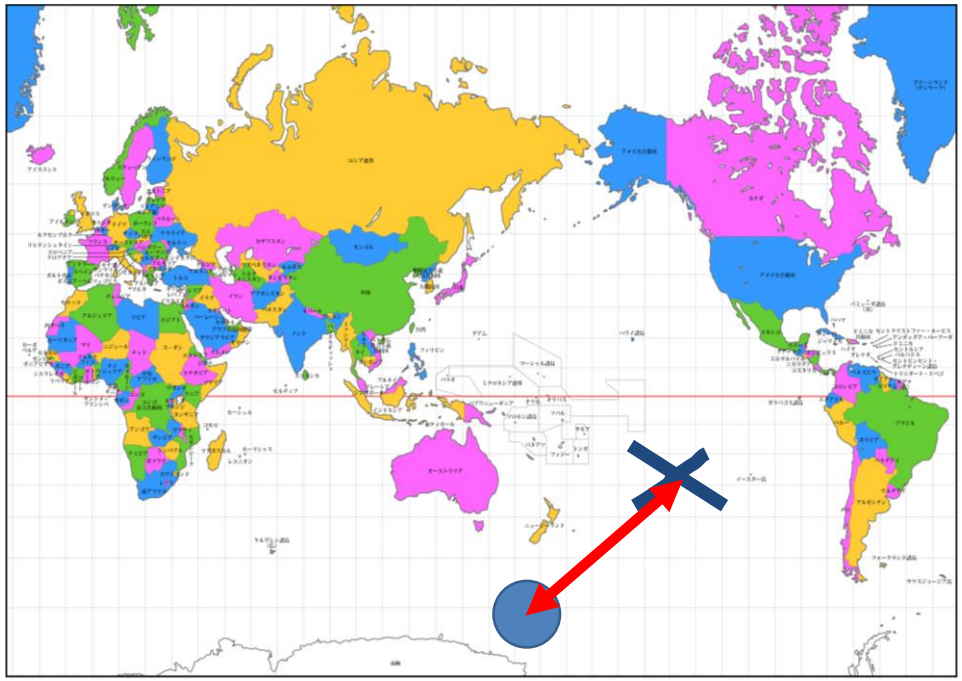
One important rule

- Cannot mark up a point if not visualized inside a paper
  - Limit maximum error



Difference with previous competitions

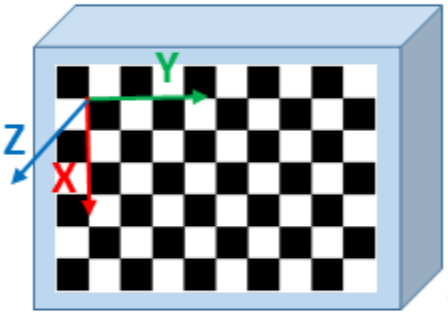
- Quantitative evaluation in metric 3D space



1. Receive the 3D coordinates of challenge points from the jury.

challenge.txt
10
1 2 3 5
2 10 11 9
3 23 22 40
:

2. Run your SLAM system and register it with the world coordinate system.



Starting area

Challenge area

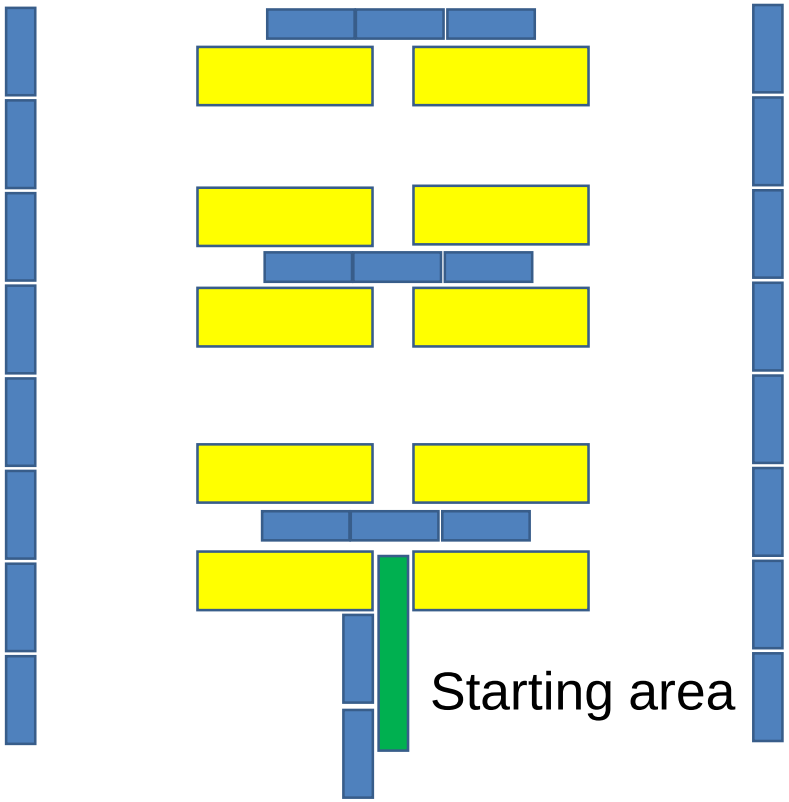
3. Move your device to find the challenge points and mark them up (e.g. put star stickers).

Evaluation criteria

- Number of marked points
- Average error
- Time

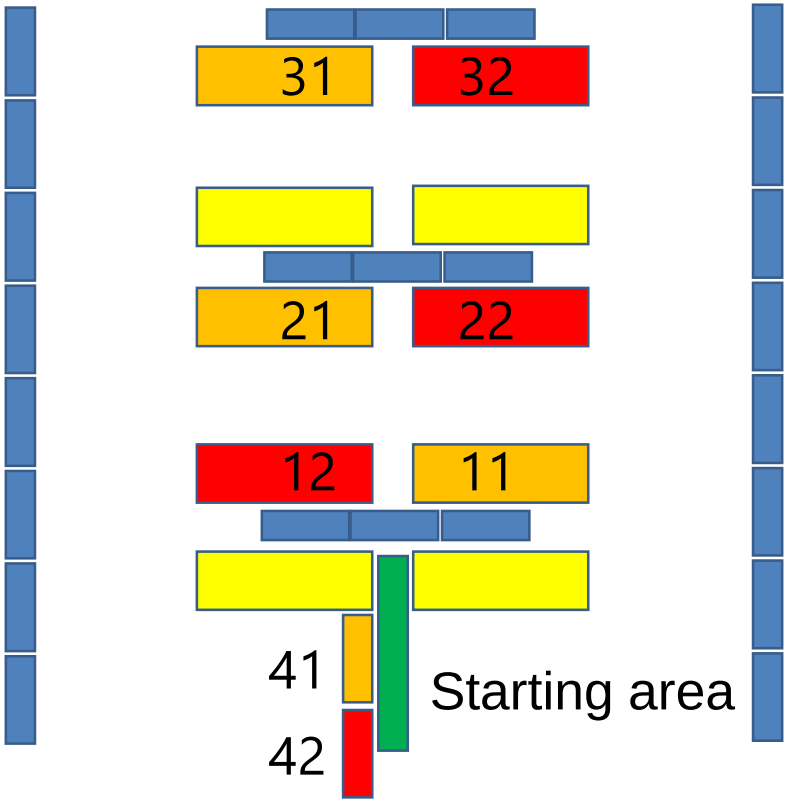
Design of environment

- 8m x 8m space with walls, tables and objects like an office



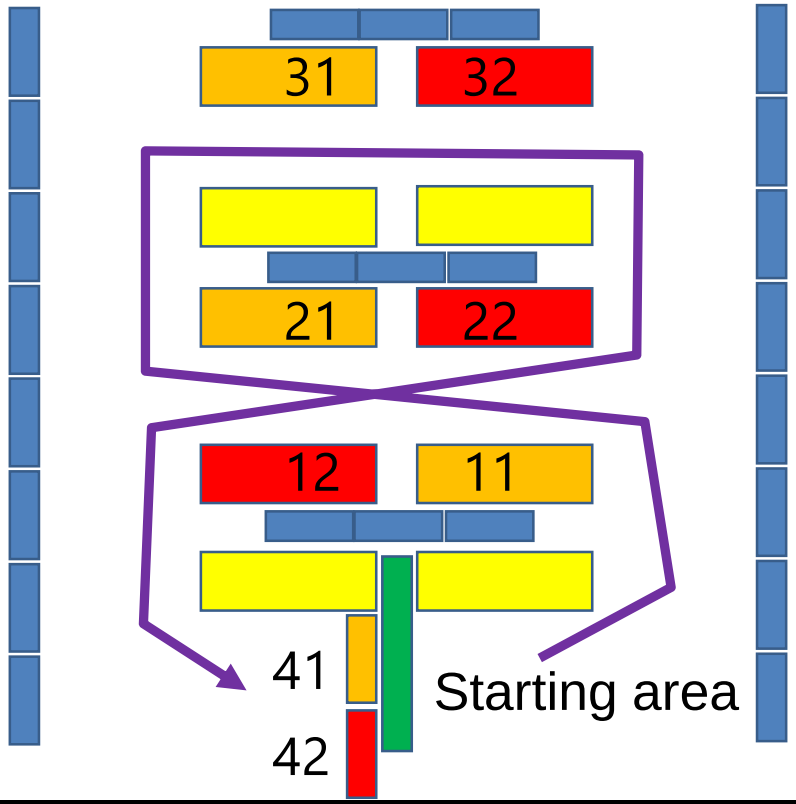
Locations of challenge areas

- 11 -> 21 -> 31 -> 41 for 1st trial
- 12 -> 32 -> 22 -> 42 for 2nd trial



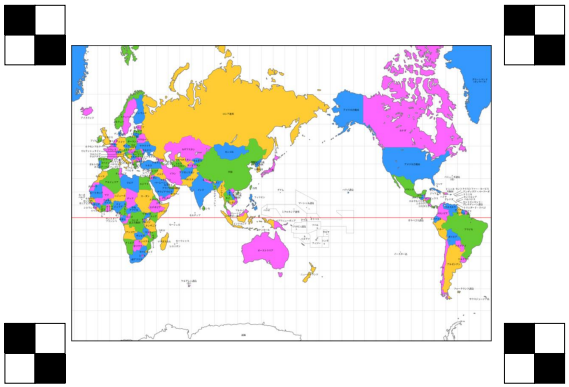
Route to be followed by all participants

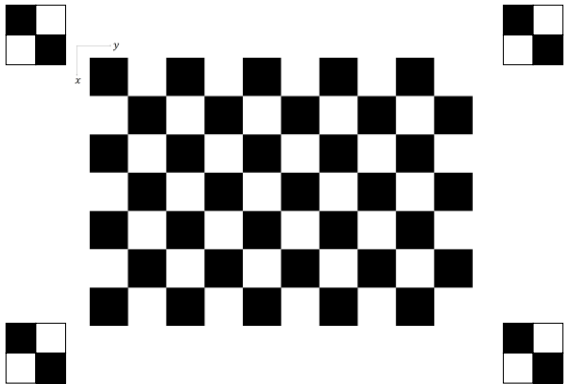
- Measure error accumulation fairly
- Check loop closure included or not



Design of challenge areas

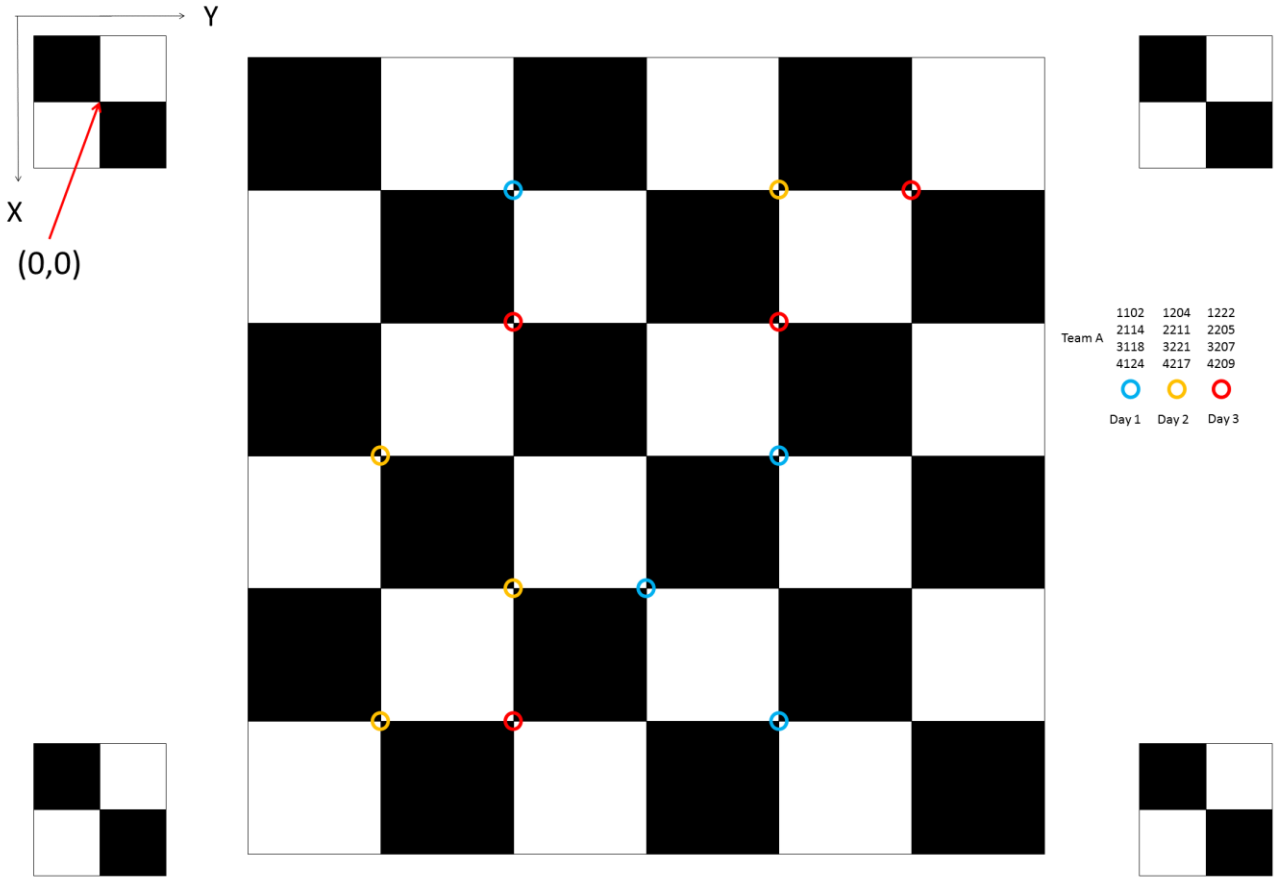
- Use textured posters



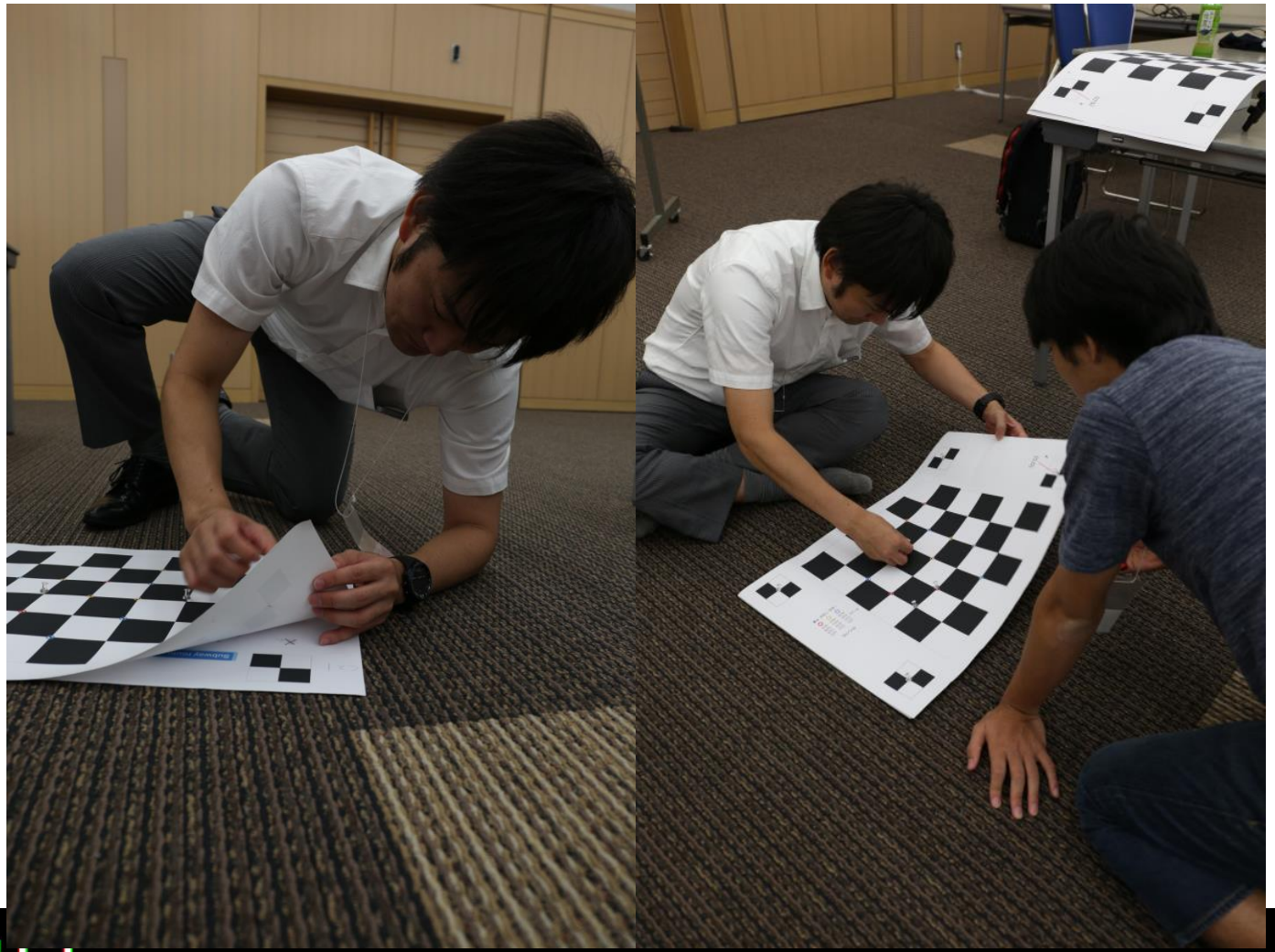


Design of challenge points

- Put points on a grid



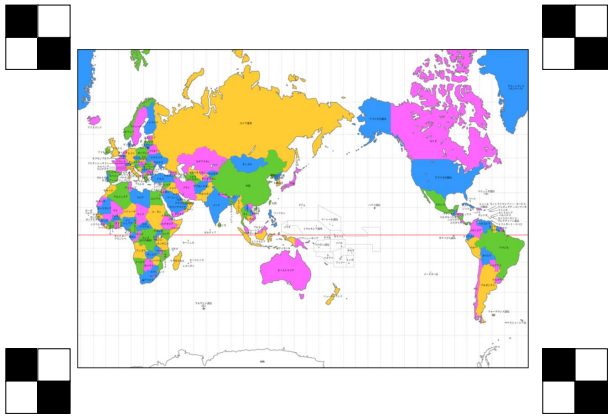
How to check answers

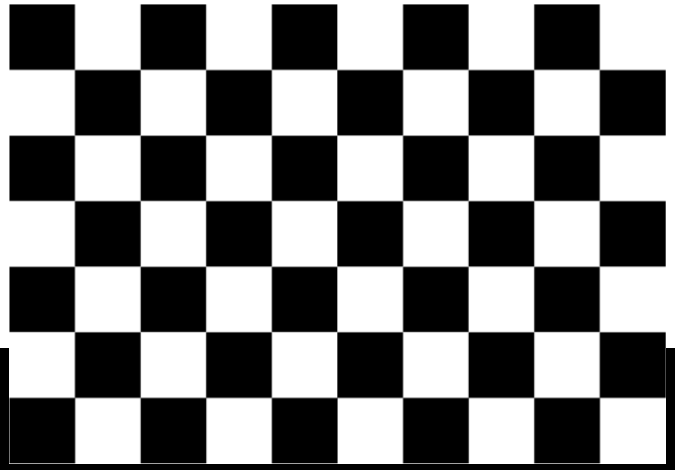


Calibration

- Use total station to measure four corners at challenge areas and at starting area, and merge them

Total station

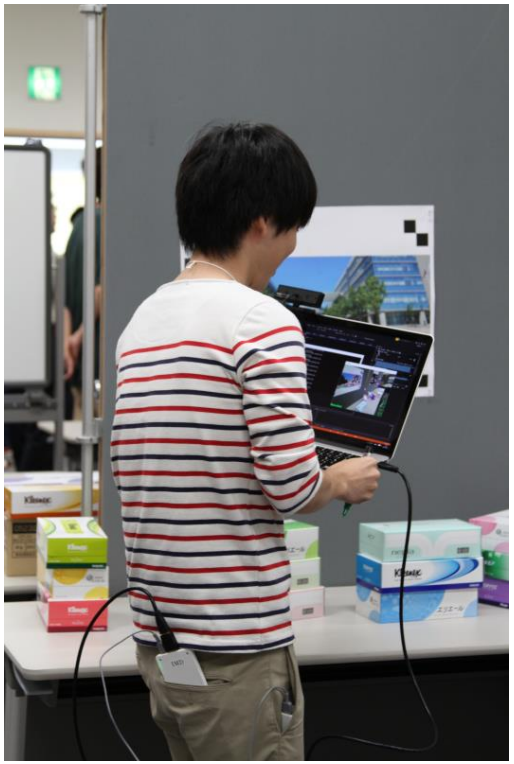




Visualization for audience

- Show participant's display using wireless HDMI system

Screens



Wireless HDMI system

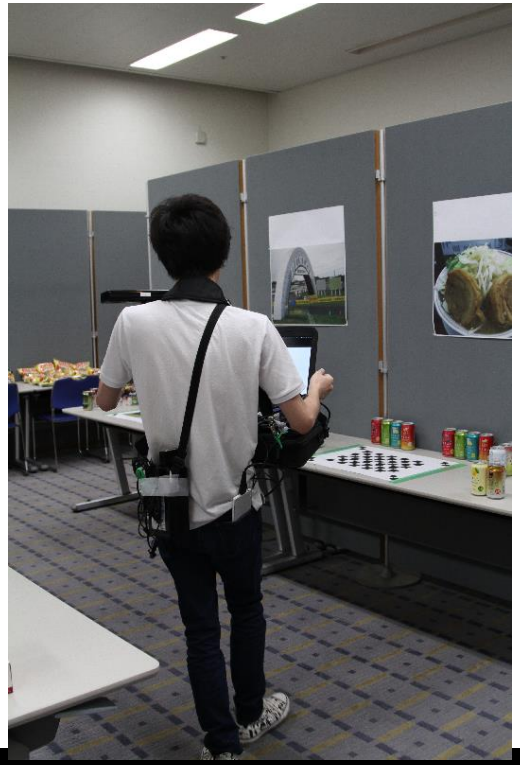
Results

- One team completed

	1st challenge				2nd challenge			
	1	2	3	4	1	2	3	4
Fraunhofer	55	328	239		197	291		
Voxar	42				231			
CEST	54	291	53		227	271	245	262

Tracking Competition in VRSJ2016

- Follow onsite tracking competition in ISMAR2015
  - Chessboard on a desk
  - Challenge areas not only on walls, but on desks



Winner's mobile device

Results

- Globally OK, but localy not so good?
- Error variance is big?

	1st challenge					2nd challenge				
	time	1	2	3	4	time	1	2	3	4
CEST	5:09	28	82	311	22	10:00	333	202	60	
Kumamoto U.	10:00					10:00				
Keio U. A	4:35	62	278	80	78	3:36	122	93	221	120
Keio U. B	10:00	114	363			10:00	160			
U. Tsukuba A	10:00					10:00				
U. Tsukuba B	10:00	97	41	105		10:00	149	21	134	
U. Tsukuba C	10:00	38	99			10:00	70	129		
U. Tsukuba D	10:00					10:00				
OPUCT	10:00	2	76	78		10:00	48	211	121	
Ritsumei U.	10:00	108				10:00	58			

Next issues

- Design of tasks
  - Any realistic scenario for SLAM?
  - More complicated tasks?
- More challenging environment
  - Less textured scene?
  - Moving objects or illumination changes?
- Fair conditions for all participants
- Onsite competitions on other research topics