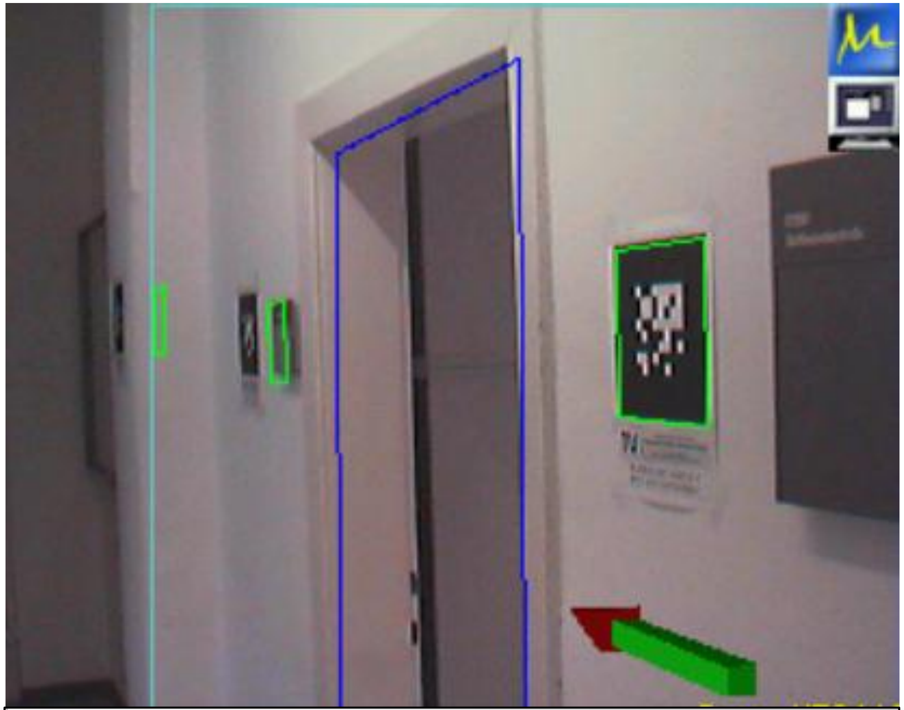
Photo-shoot localization for mobile AR based on registered images in virtualized reality models

Jun Nishida [1,2], Koji Makita [2], Tomoya Ishikawa [2],
Takashi Okuma [2], Masakatsu Kourogi [2], Thomas Vincent [3],
Laurence Nigay [3], Jun Yamashita [1],
Hideaki Kuzuoka [1], Takeshi Kurata [1,2]

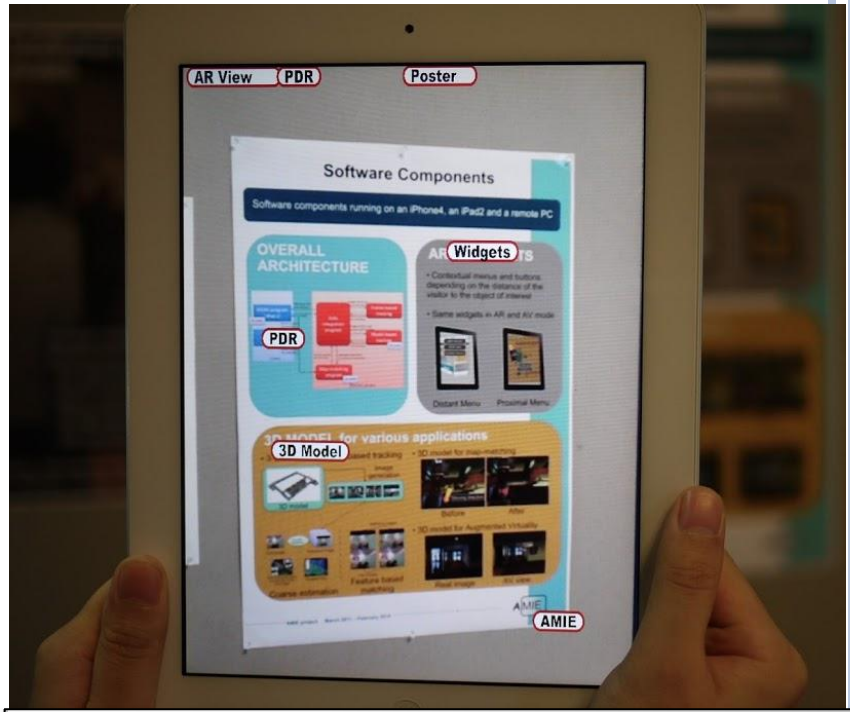
1. University of Tsukuba, Japan
2. Center for Service Research, AIST, Japan
3. Université Joseph Fourier, France

KJMR2013@OIST 2013.04.14

AR applications



D. Wagner, et al. “First steps towards handheld augmented reality” Proc. ISWC 2003.



Univ of Tsukuba, Demo in IDW/AD'12.

Camera position and orientation are needed for creating AR images

Research background ～localization methods～

Outdoor environment

- GPS（Global Positioning System）
  - Mainly used for navigation

Indoor environment

- Method using physical Infrastructure（WIFI, RFID, …）
  - Task : installation costs
- **Method using image processing (Marker, Feature points, …)**
  - Task : cost of creating the reference DB

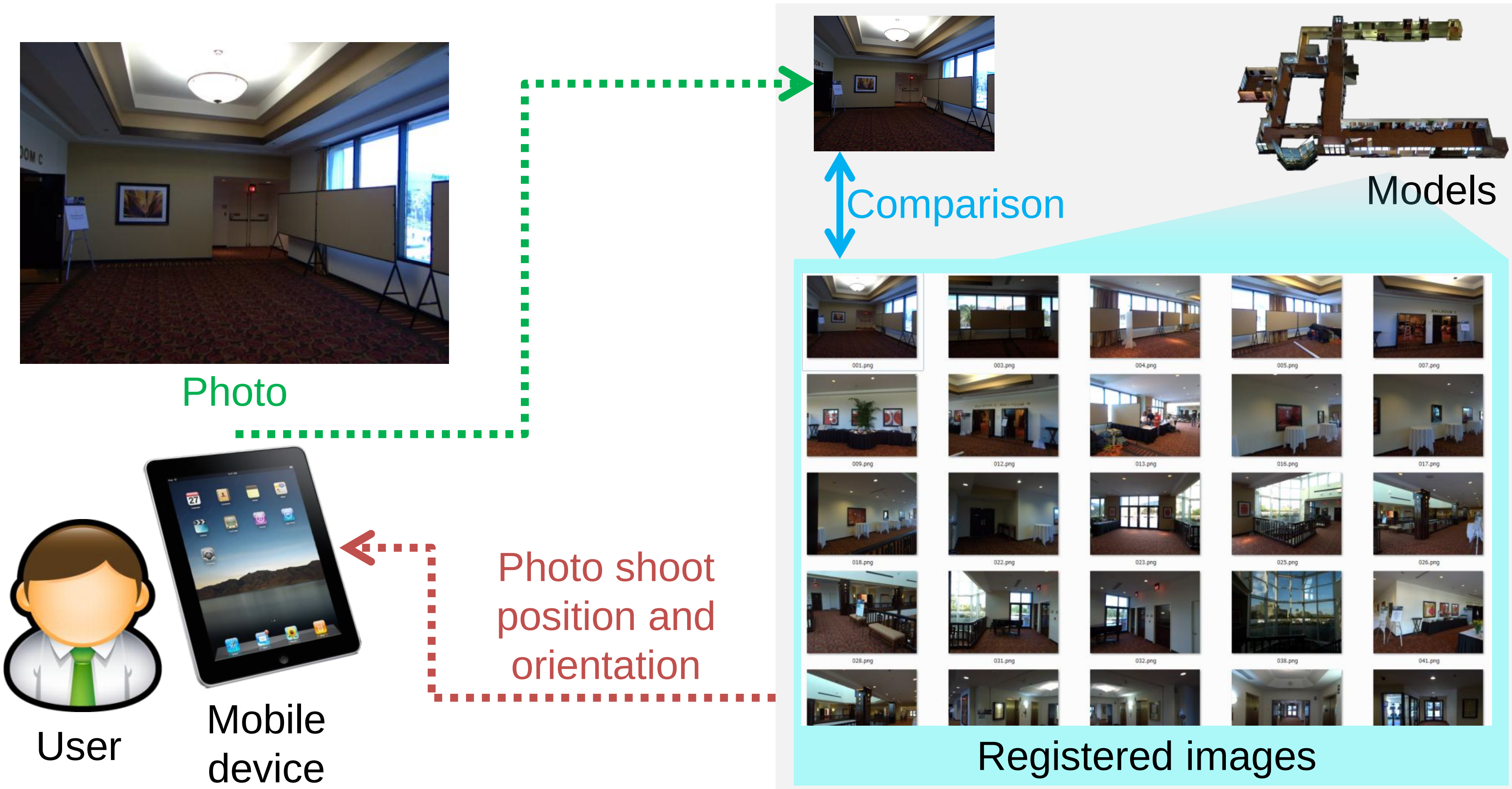
Effectiveness of indoor localization is heavily dependent on the condition, and there are still many tasks to ubiquitous localization.

Objective

Construction of the absolute localization system using virtualized reality models

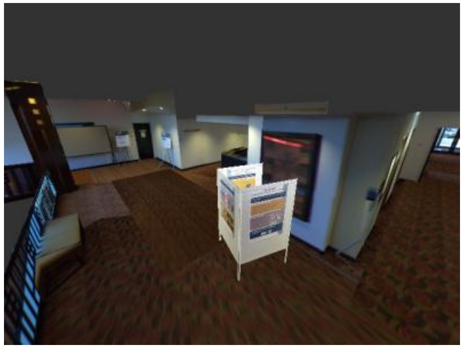
-Approach-

<u>Comparison of the photo to the registered images of the models</u>

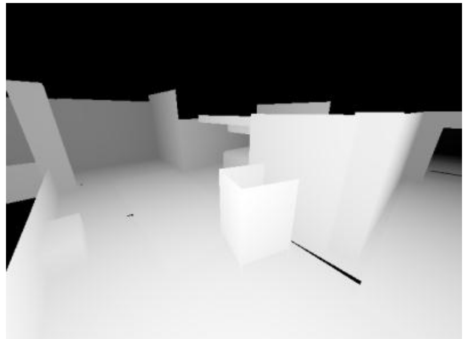


How to compare the photo to registered images ?

- Definition of "Registered image"
  → contains shooting position, orientation, and depth information.



Registered image



Depth image

Point cloud
(created by Univ. of Tsukuba)

- We compare the photo to the registered images using feature point, edge, and surface information.

Photo

Registered images
with shooting position, orientation, and depth information

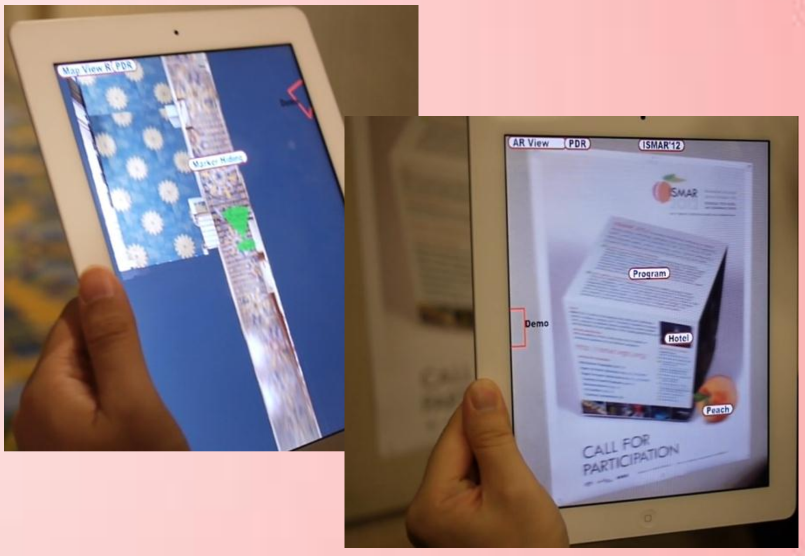
→

Estimation using
◆Feature point
◆Edge
◆Surface information

→

Output
Photo shoot position and orientation

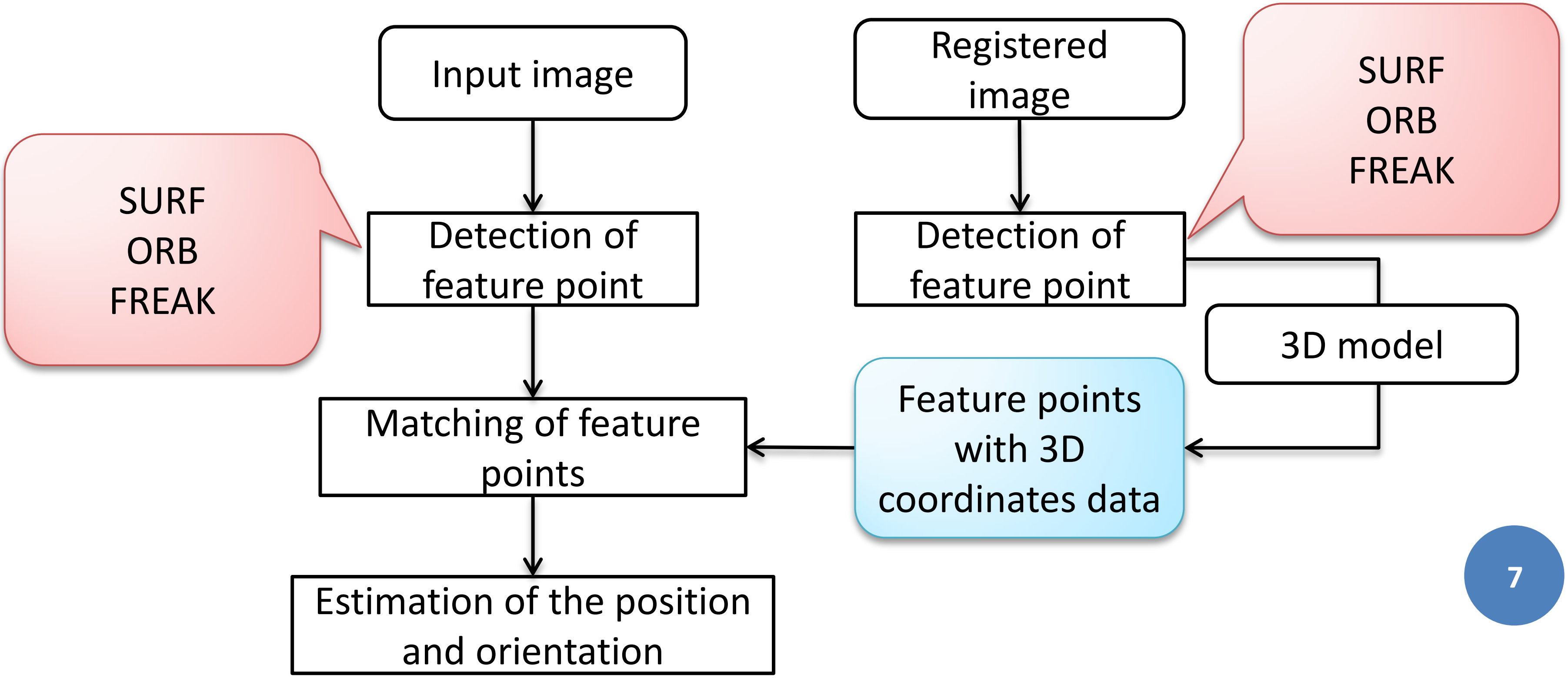
Road map for localization

Phase 2	Absolute localization > In real time > In global coordinate system Our goal	
Phase 1	Relative motion estimation > In real time > In local coordinate system We are going to apply some previous methods.	Absolute localization > Not in real time > In global coordinate system Currently we are researching this part.

Localization method using feature points

Processing steps

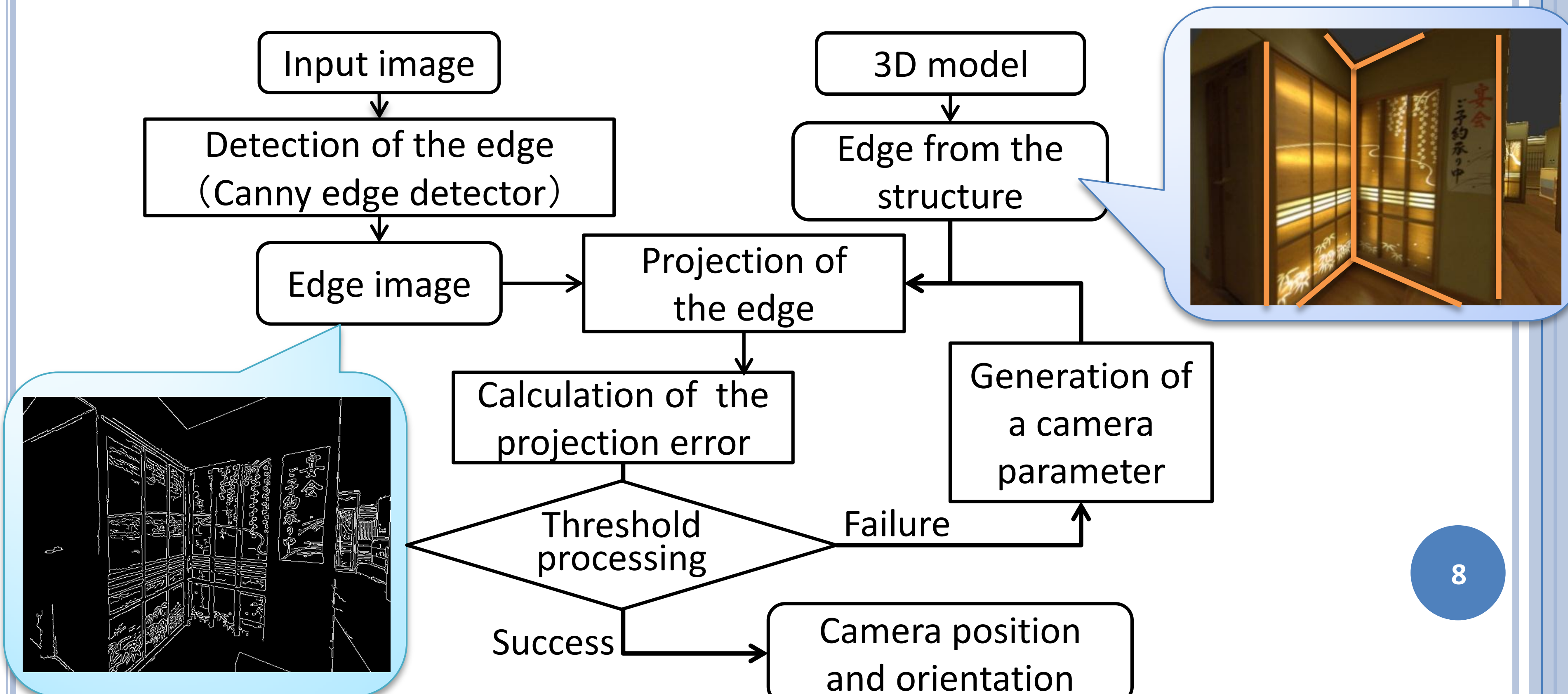
1. Detection of feature points from the input image and the registered image
2. Matching of feature points
3. Estimation of camera parameters from a set of feature points by RANSAC



Localization method using edge information

<u>Processing steps</u>

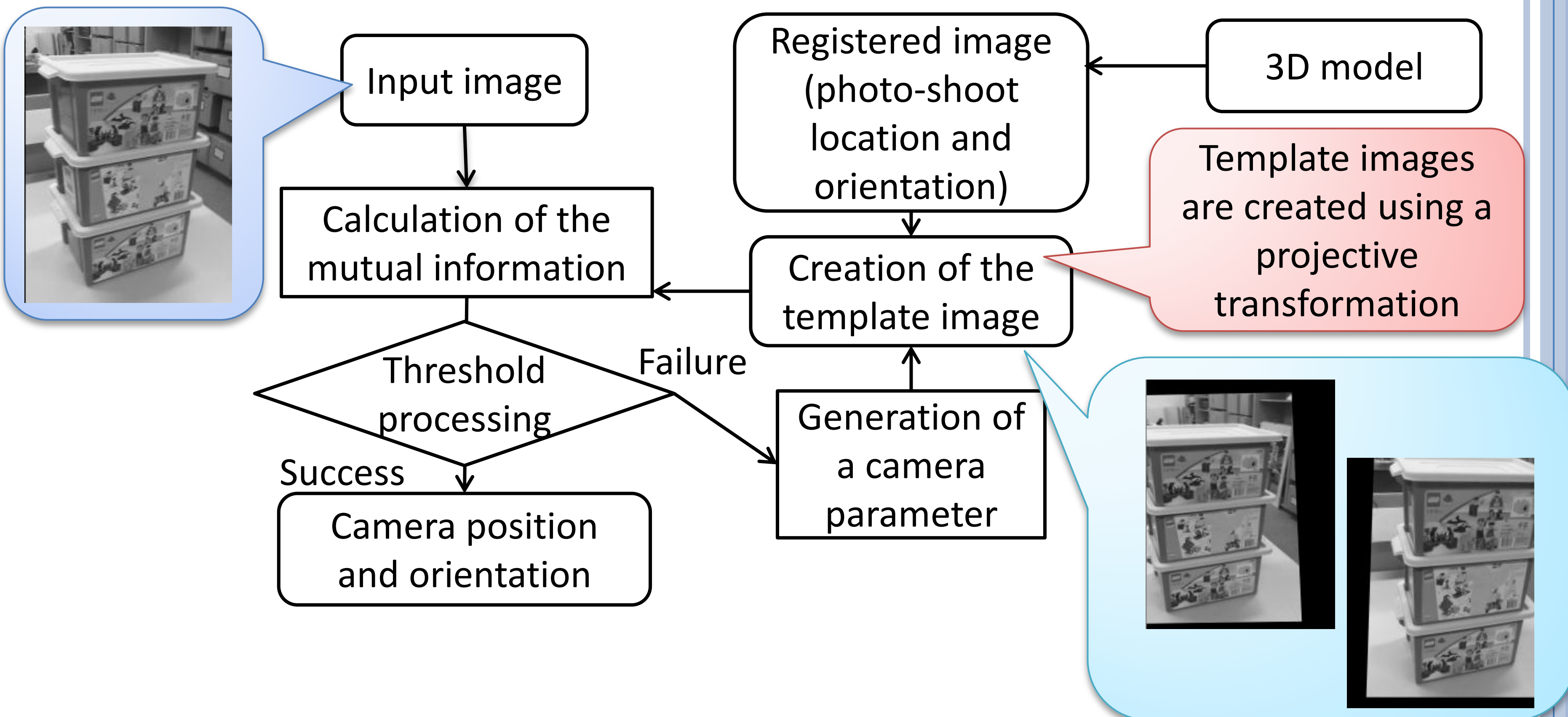
1. Detection of the edge from the input image (Canny)
2. Projection of the edge from the structure of the model using generated camera parameters
3. Calculation of the projection error of the edge
4. Estimation of camera parameters using threshold processing



Localization method using mutual information(*)

Processing steps

1. Creation of the template image
2. Calculation of the mutual information
3. Estimation of camera parameters using threshold processing



(*) A. Dame, et al : "Accurate real-time tracking using mutual information, "Proc. IEEE International Symposium on Mixed and Augmented Reality (ISMAR2010), pp. 47-56, 2010.

Experiments

Setup of experiments

- Scene 1 : Virtual scene
  - Input images :
    Generated images from the model
  - Registered images :
    Generated images from the model
  - Resolution of images :
    640 * 480 [pix]
  - # of feature points : about 800
  - # of edges : 7

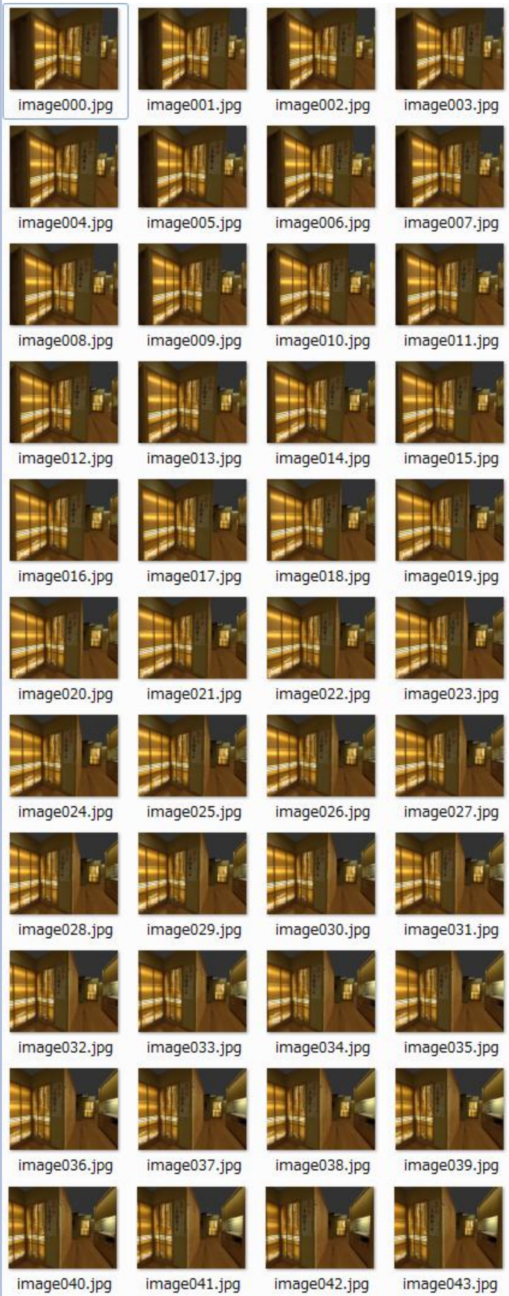
- Scene 2, 3 : Real scene
  - Input images :
    Photo taken by “iPad 2 (Apple)”
  - Registered images :
    Photo taken by “iPad 2 (Apple)”
  - Resolution of images :
    180 * 240 [pix]
  - # of feature points : about 200
  - # of edges : 7

Scene 1

Scene 2

Scene 3

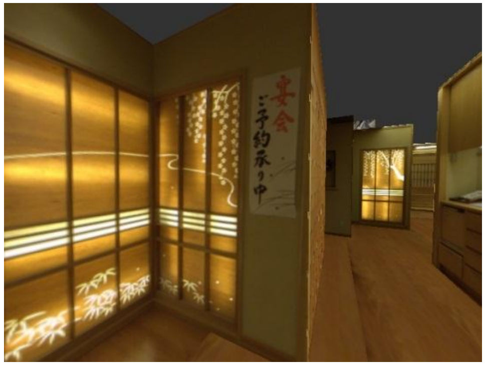
Scene 1 : Virtual scene



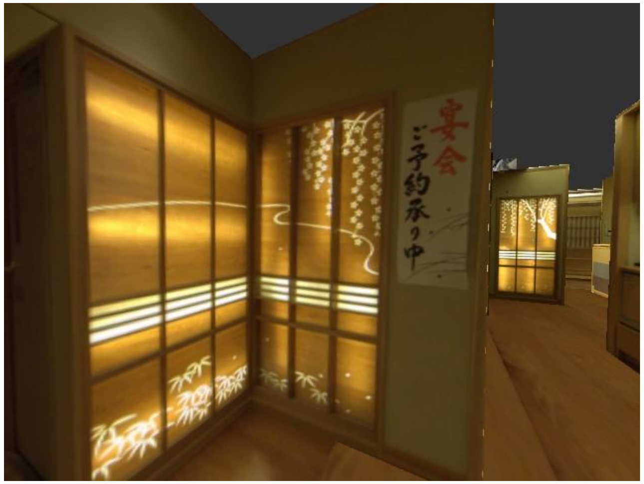
Input (44 images)

#0

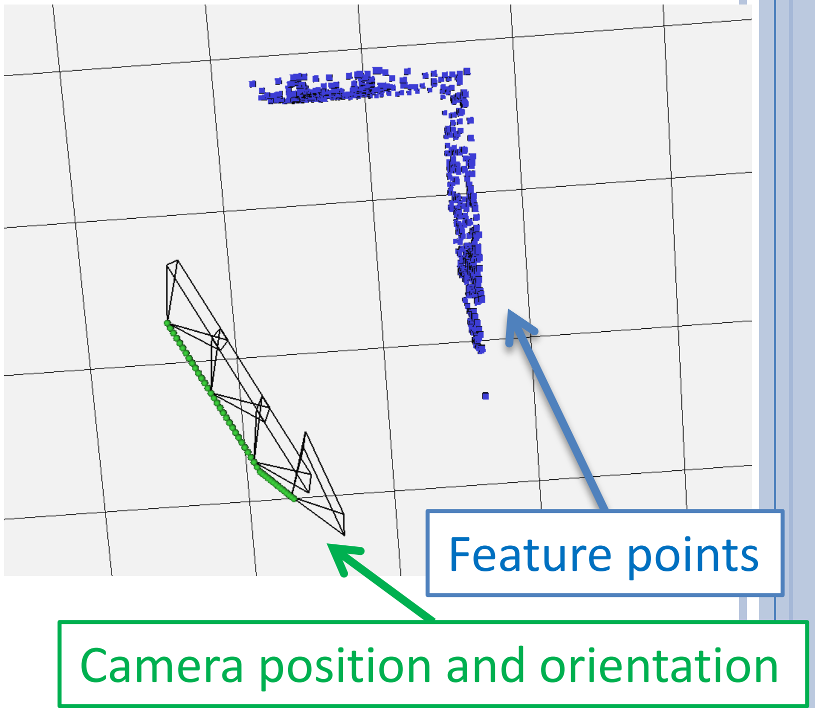
#22



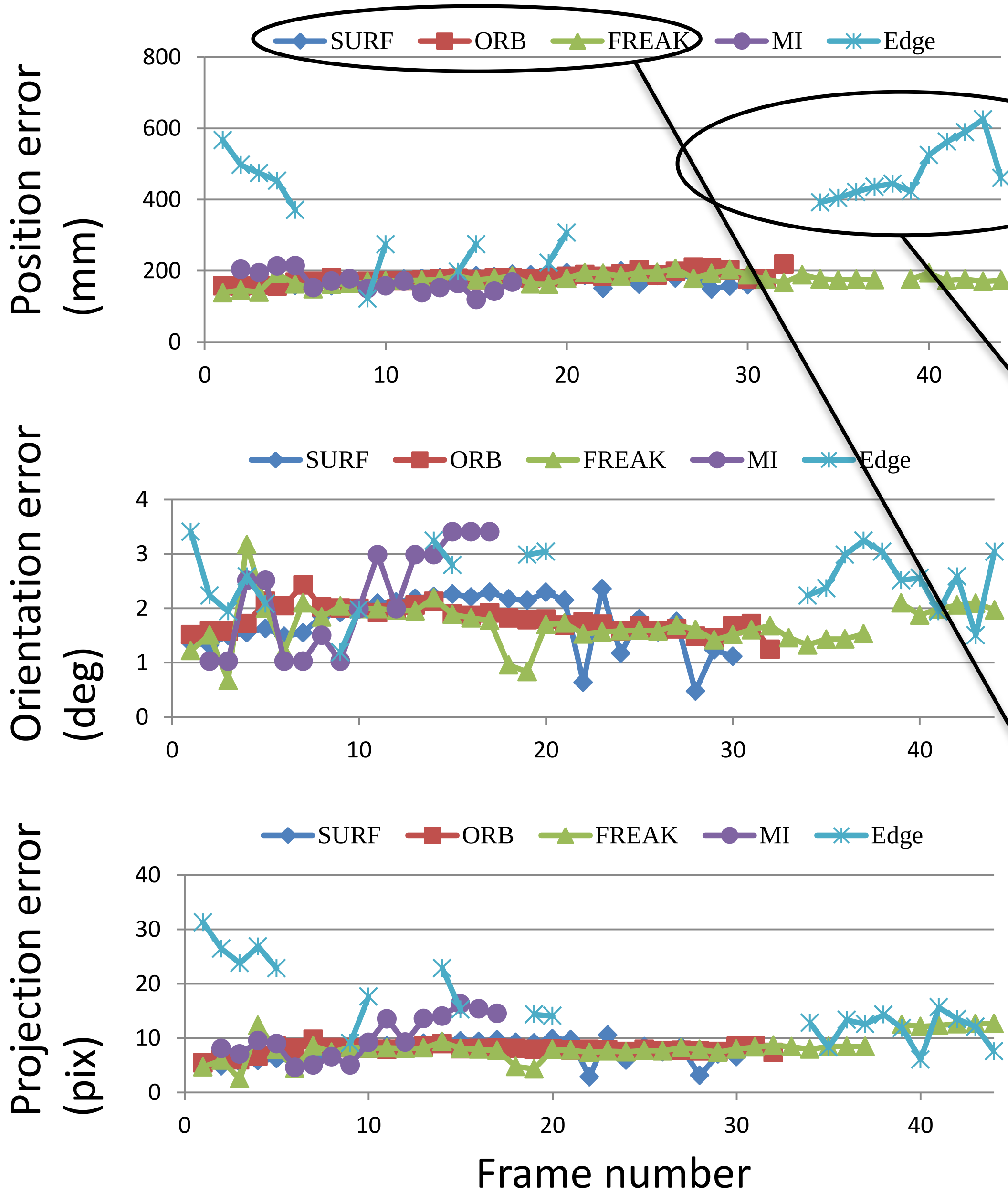
#43



Registered image



Result　～Scene1～

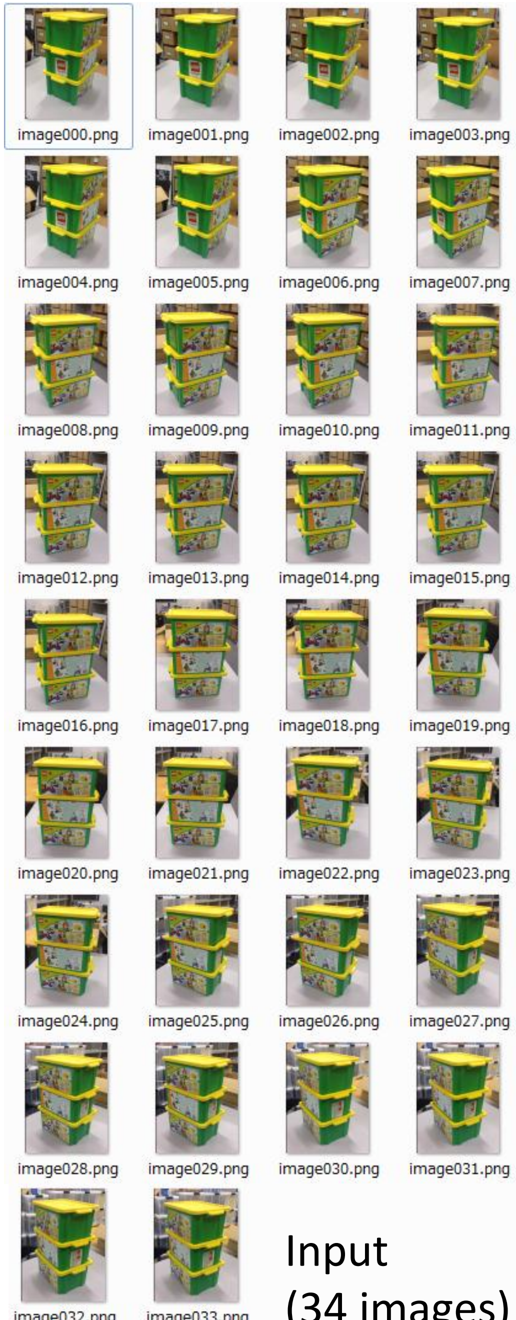


～Virtual environment～
Best scene for calculating the correspondence of features

Despite position error is large, the method using edge information judged the result as success.

Results using feature points are good in many frames.
（SURF, ORB, FREAK）

Scene 2 : Real scene



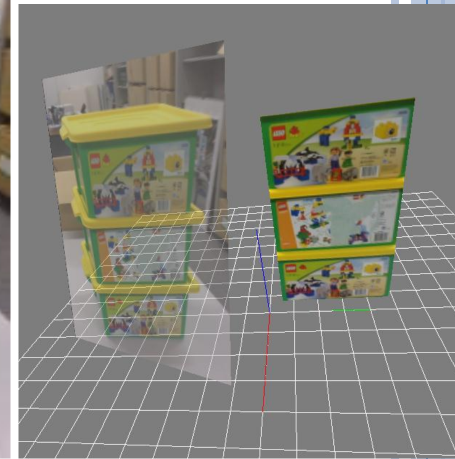
Input
(34 images)

#0

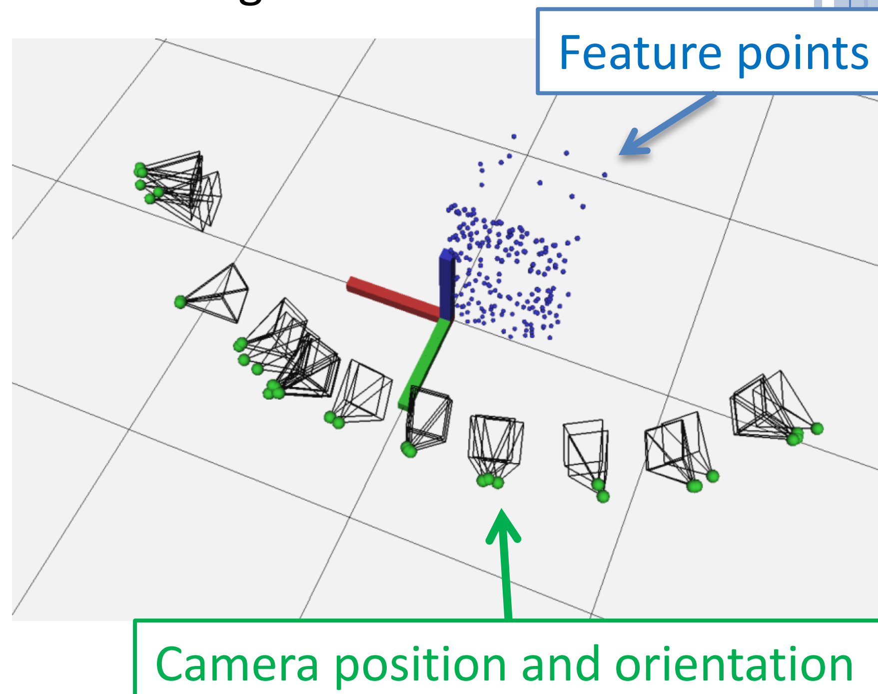
#17

#33

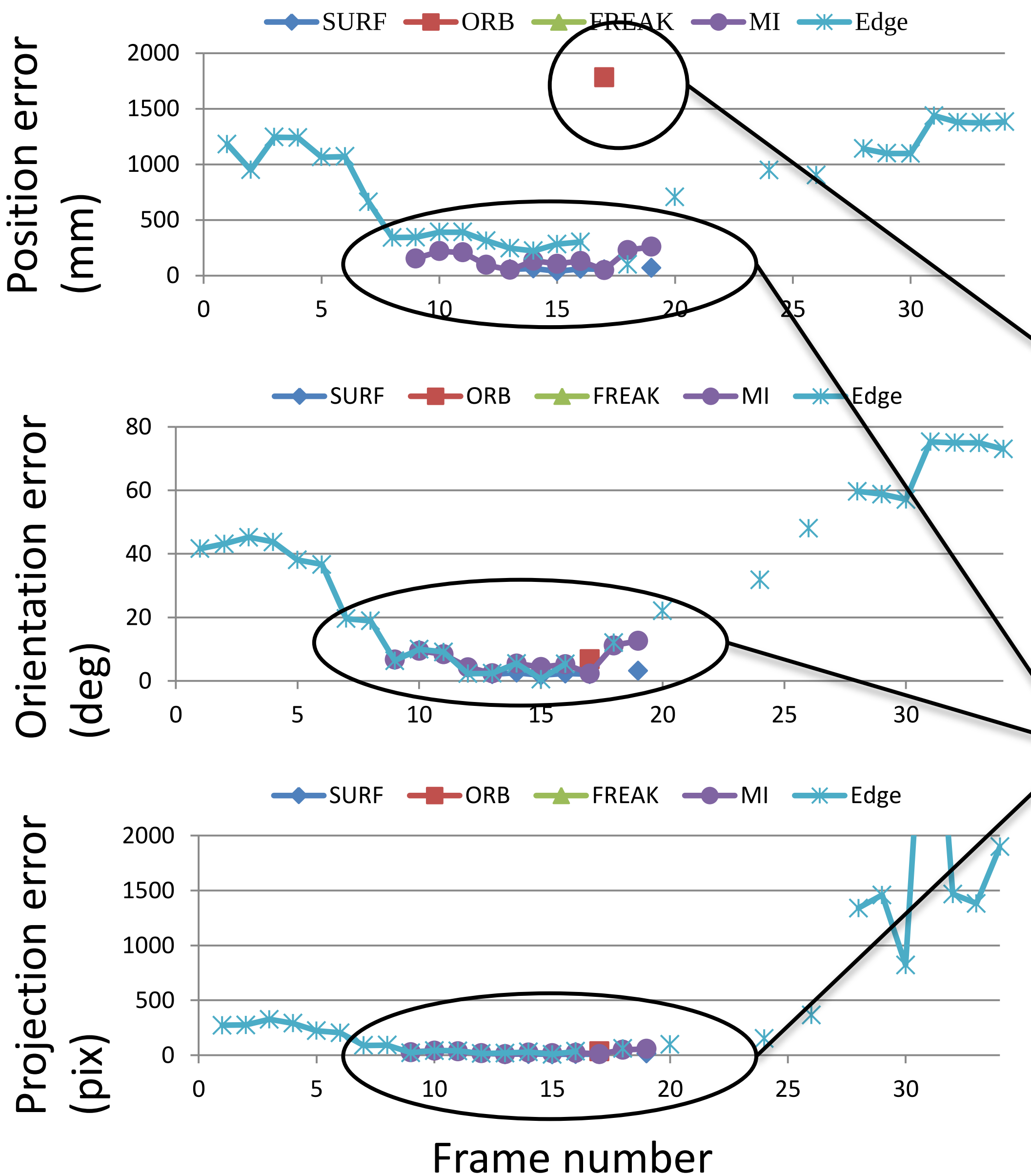
Registered image



View of the modeler



Result ～Scene2～

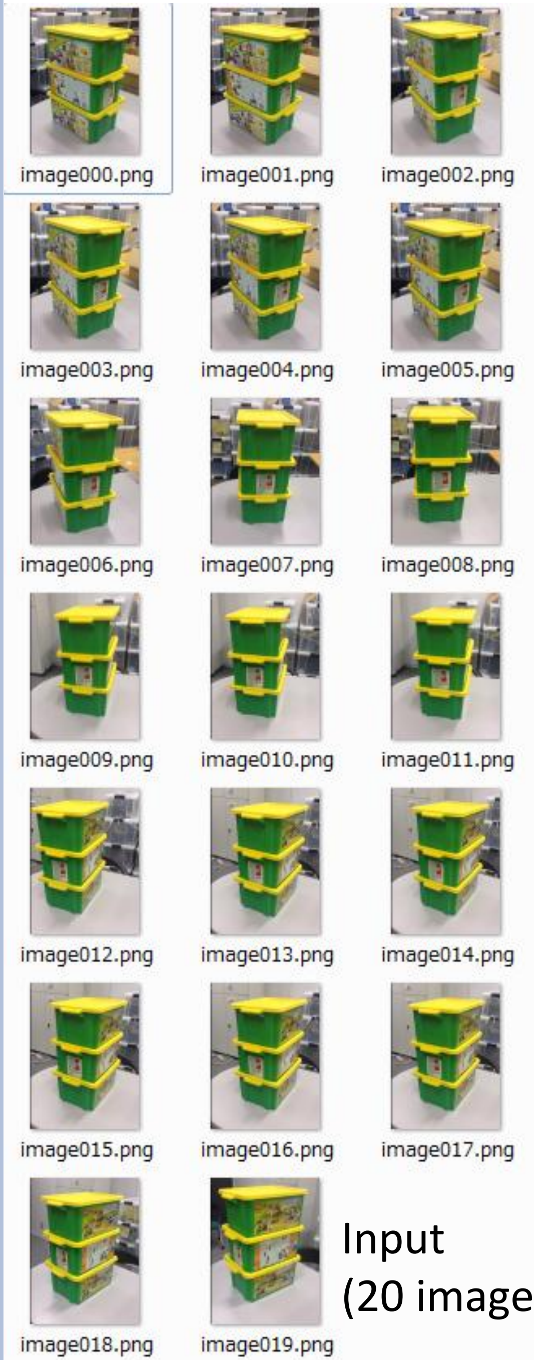


Real environment with complex textures

Results using ORB included large position error in the front-back direction.

Results using Mutual information (MI) and SURF were good, and judged as success.

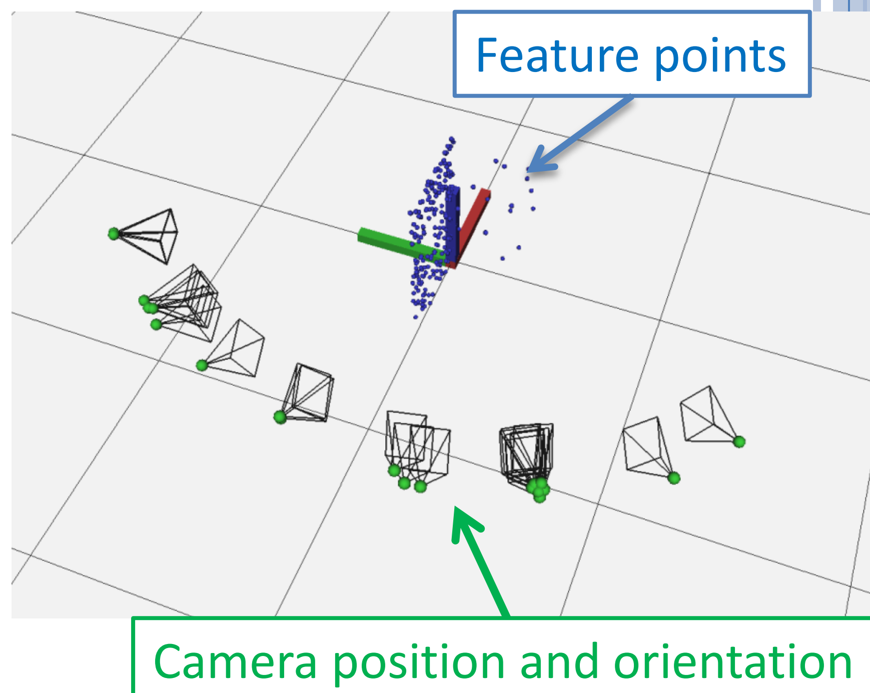
Scene 3 : Real scene



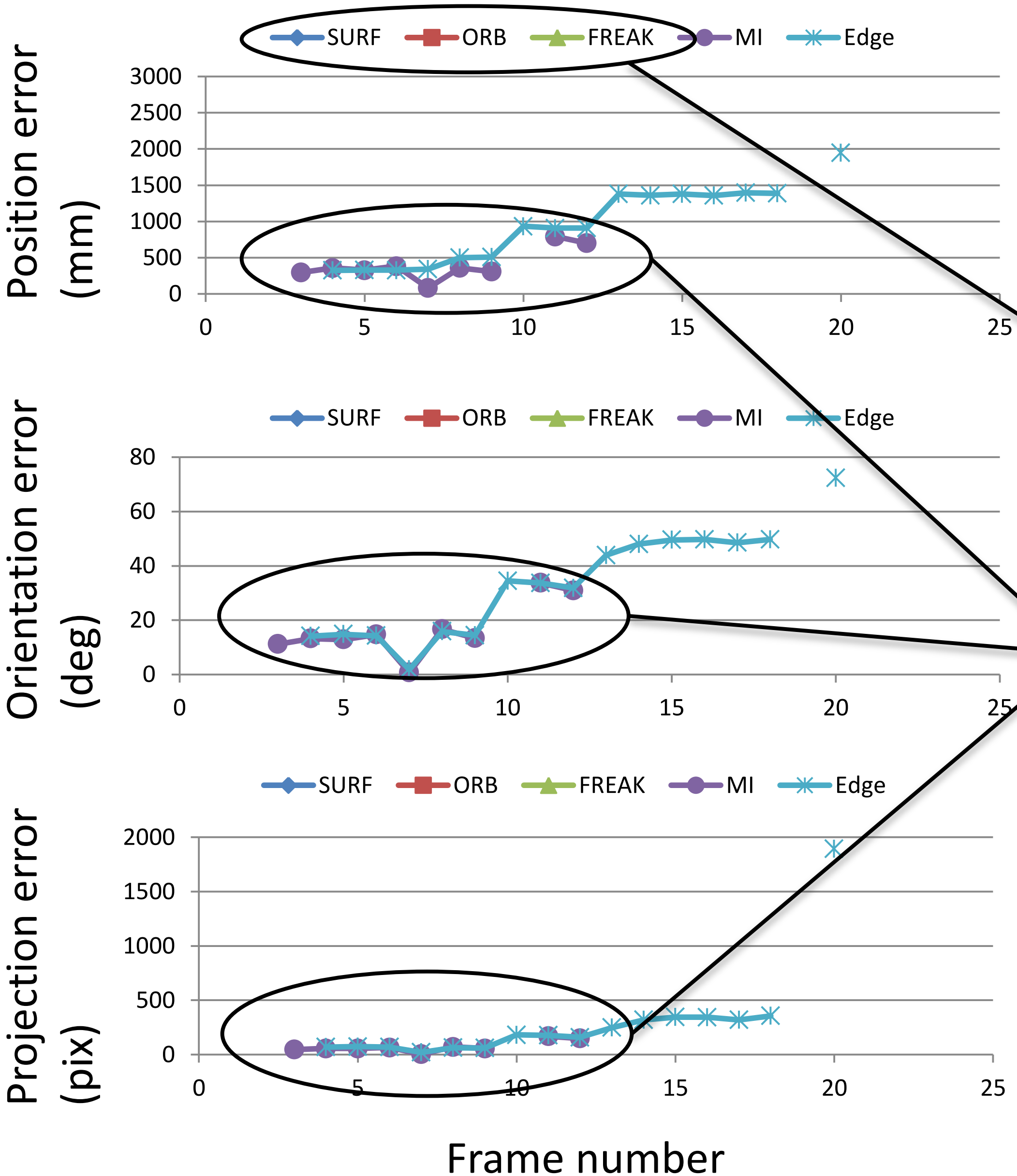
Input
(20 images)

Registered
image

View of the
modeler



Result ～Scene3～



Real environment with less complex textures

Results using feature points were failed in all frames.

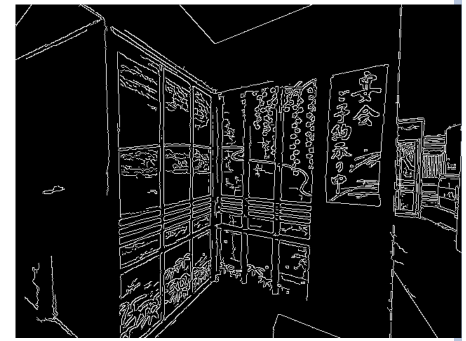
Results using mutual information was comparatively good.

Summary

- We proposed a method to estimate the position and orientation of the camera with real image and registered image in virtualized reality models
- Experimental results in 3 scenes
  - Among the methods using feature points, SURF method was good.
  - Localization with Mutual Information (MI) was highly accurate both in Scene 2 and Scene 3, but it was based on lots of templates. Therefore, speeding up of the calculation is needed for actual use in future.

	Scene2		**Scene3**	
	SURF	MI	SURF	MI
Accuracy	✔	✔	×	✔
Calculation Time	✔	△	✔	△

Future works



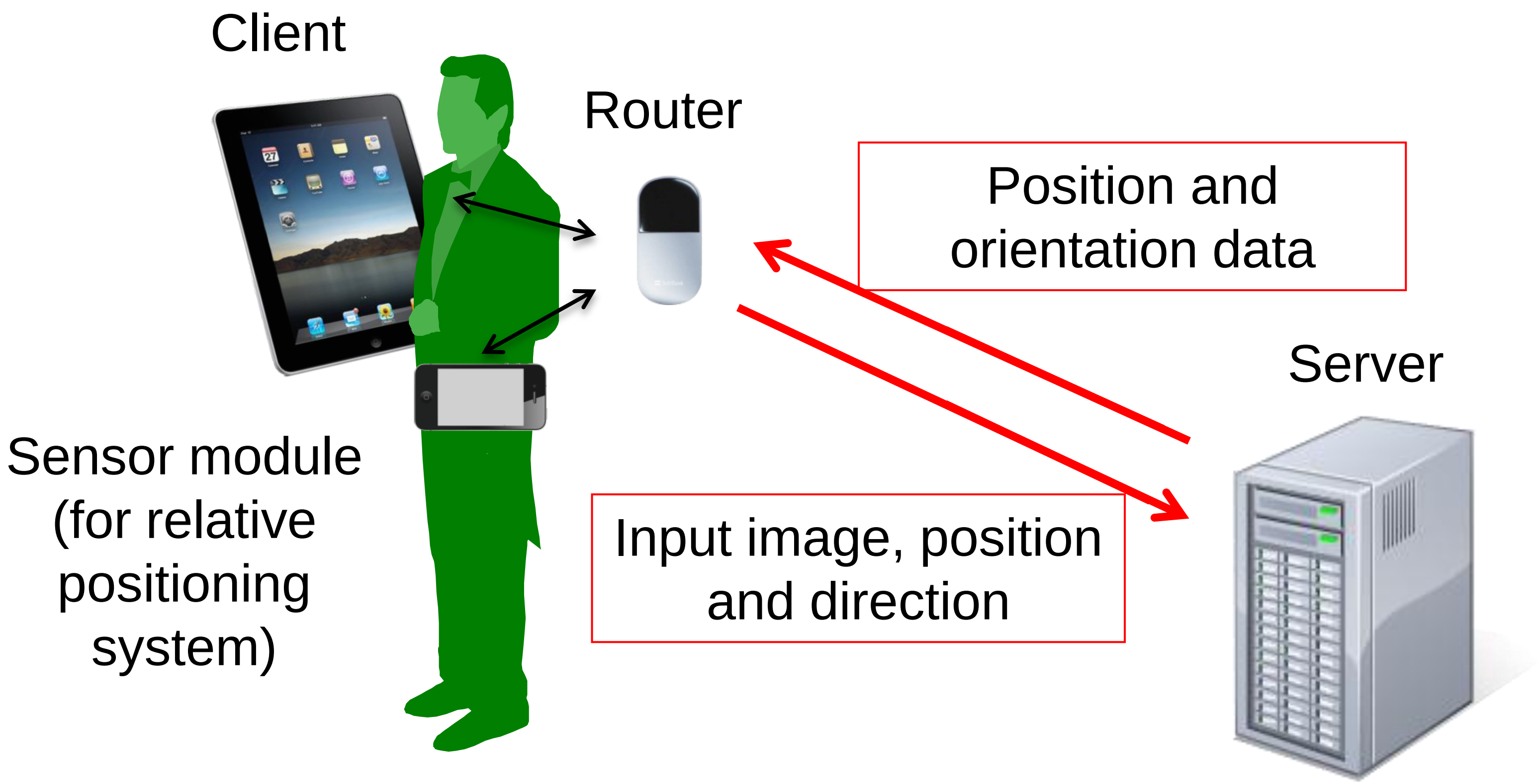
- Edge detection from registered image
  - Currently, only the edge on structure is available
- Parallel processing in mutual information method for speeding up
- Combination use of the mutual information method and the SURF method for covering wide area
- Utilization of fast access method for the database
- Introduction of human-computer interactions
  - Displaying position and orientation on the map that are appropriate for the localization.
- Benchmarking of the localization method with “TrakMark”

We plan to standardize benchmarking methods in “ISO / JTC1 / SC24 / WG9 （Augmented reality continuum concepts and reference model）”

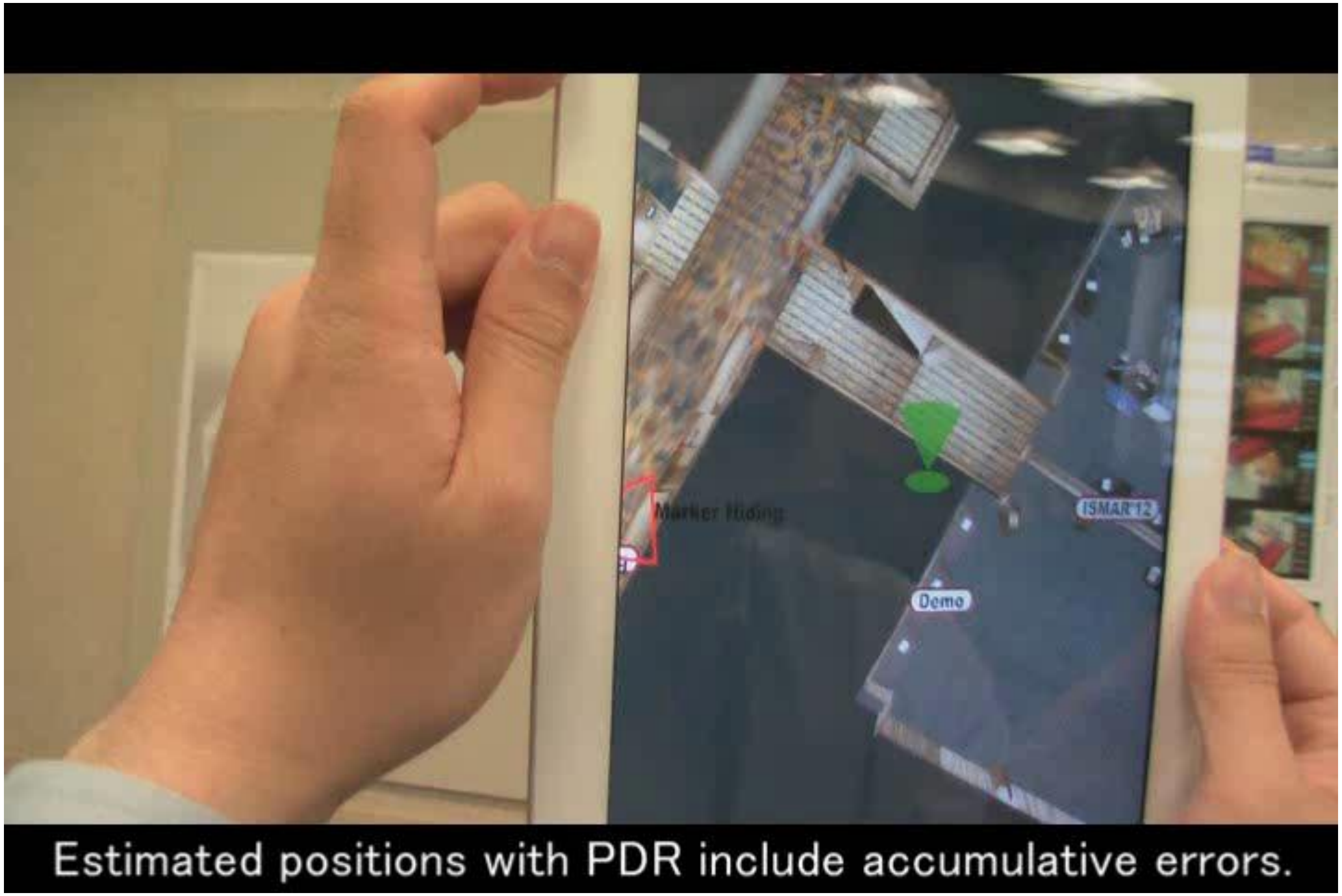
Application

System architecture for combining the relative positioning system

- Server-client system
- Always estimate the relative position and orientation using the relative positioning system (PDR)
- Absolute position and orientation are estimated in the server when the camera image is sent to the server



Appearance of the demonstration (in ISMAR2012)



Start without the position information

↓

Estimation of position and orientation (with SURF)

↓

Correction of the position and orientation

↓

The relative positioning using PDR system

Thank you !

- TrakMark
  http://trakmark.net/
- Center for Service Research, AIST
  http://unit.aist.go.jp/cfsr/en/
- Koji Makita (Post doctoral research scientist)
  k.makita@aist.go.jp

Acknowledgements
This work was supported by Strategic Japanese-French Cooperative Program on Information and Communications Technology Including Computer Sciences (ANR and JST).