An Integrated Feature Selection Strategy For Monocular SLAM

Presenter: Kuan-Ting Yu
Advisor: Li-Chen Fu

NTU Intelligent Robot Lab

Industrial Technology Research Institute

September, 2011

Outline

- Introduction
- Feature Extraction
- Bottom-up Feature Selection
- Top-down Feature Selection
- Experimental Results
- Conclusions

Outline

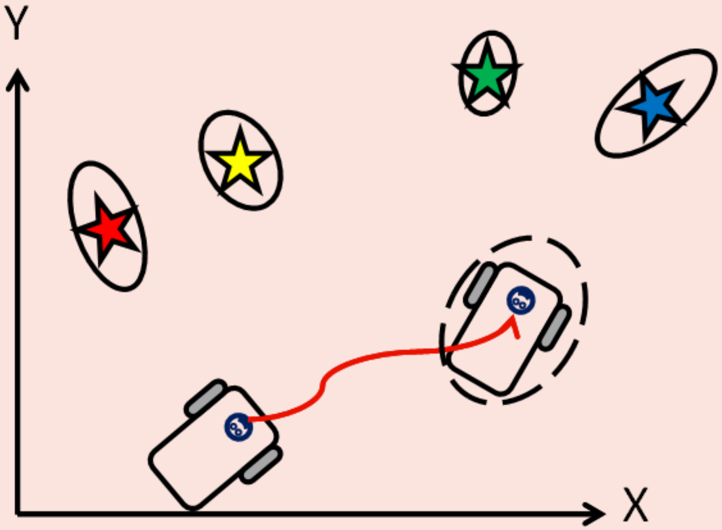
- Introduction
- Feature Extraction
- Bottom-up Feature Selection
- Top-down Feature Selection
- Experimental Results
- Conclusions

Introduction

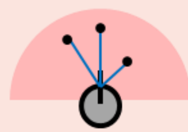
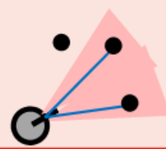
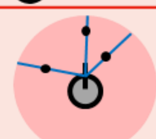
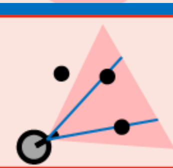
- To perform desired tasks in indoor environments, how to estimate robot's position and map about surroundings is a critical problem.

SLAM

- **S**imultaneous **l**ocalization **a**nd **m**apping (SLAM)
  - Incrementally build a map of this environment while simultaneously determining its location
  - Given:
    1. Robot's odometry
    2. Observations of nearby features
  - Estimate:
    1. Location of the robot
    2. Map of features



Sensors

Sensor	Diagram	Cost	Weight
Sonar array		medium	lightweight
Laser range finder		high	moderate weight
Stereo camera		high	lightweight
Omnidirectional camera		high	moderate weight
Monocular camera		low	lightweight

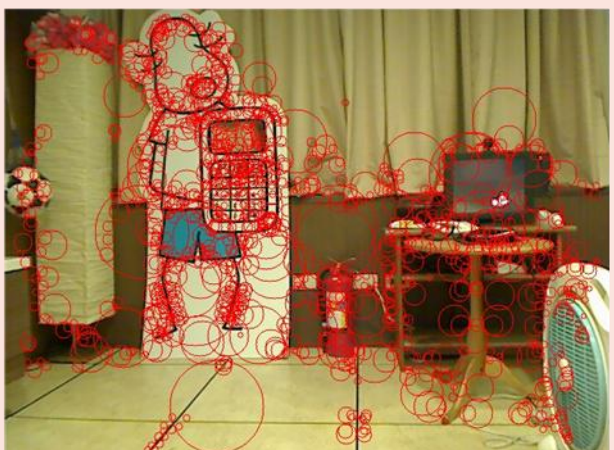
vSLAM

Generic SLAM	Vision-based SLAM
• Landmark extraction →	• Interest points (**features**)
• Data association →	• Interest points matching
• State predict • State update & landmark update	Probabilistic framework

vSLAM Challenge

Huge amount of features

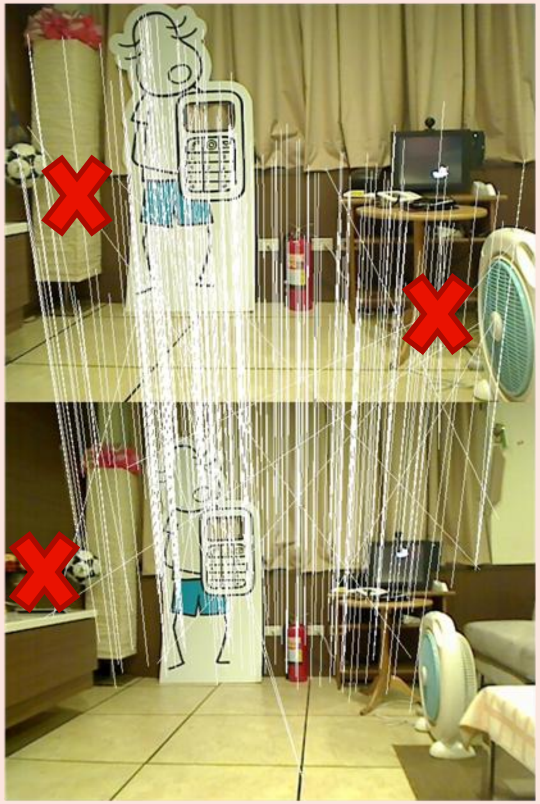
- Degrade performance
- Cause mismatch



Solution Idea

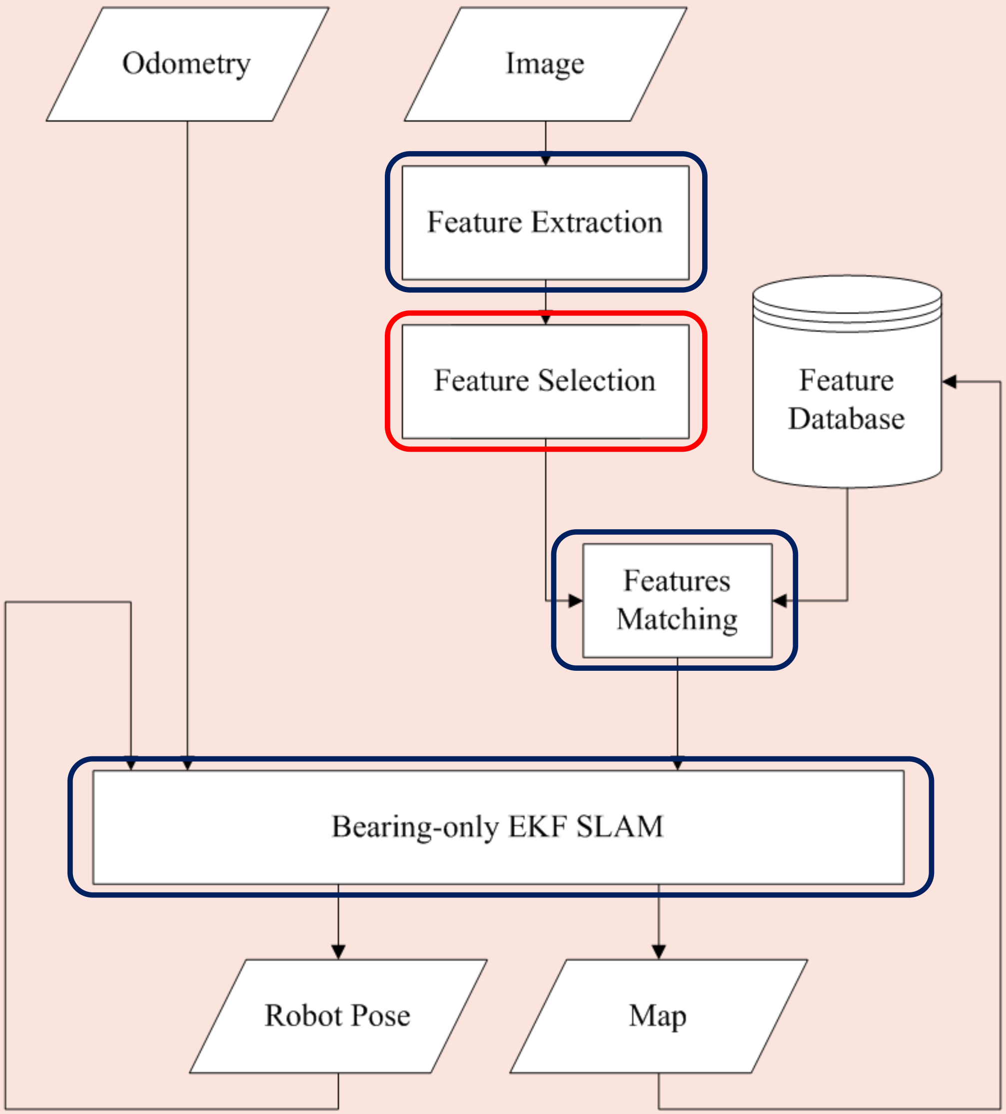
Feature selection strategy

- Remove useless features
- Keep good features



System Overview

1. Landmark extraction
2. Data association
3. State predict
4. State update &
   landmark update



Outline

- Introduction
- Feature Extraction
- Bottom-up Feature Selection
- Top-down Feature Selection
- Experimental Results
- Conclusions

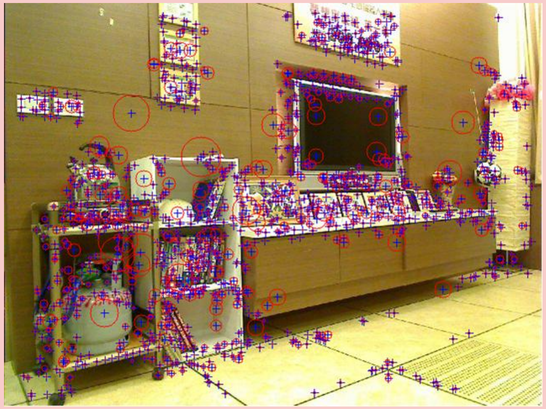
What is Feature?

- Features (interest points):
  - Easy to find their correspondences

- Desired features:
  - *Distinctive*: outstanding, easily matched
  - *Invariant*: invariant to scale, viewpoint change

Feature Detectors

- Scale Invariant Feature Transform (SIFT) [Lowe, 2004] is one of the best feature detectors, but slow
- Speeded Up Robust Features (SURF) [Bay, *et al.*, 2006]
  - SURF has good performance as SIFT and faster

	SIFT	SURF
Images		
Extraction time	863.998 ms	267.634 ms

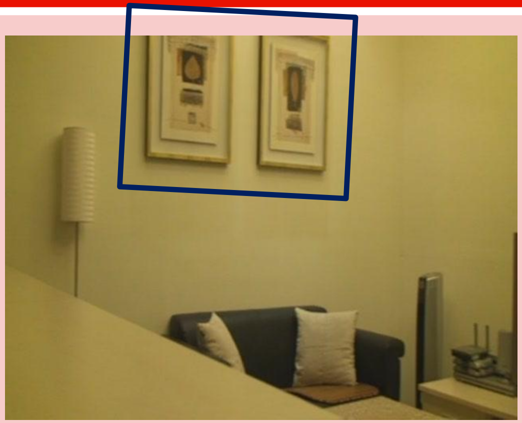
Feature Extraction Challenge

- Hundreds of SURF points are extracted
  - Increase computational time
  - Need higher data storage
- But, fewer features are desired in practice
  - Remove useless features
  - Keep repeatable features

- A feature selection strategy is necessary

Feature Selection

- Human never process the whole image at once
  - Focus on some regions of interest (ROIs)
- A natural solution — visual attention system

Bottom-up approach	Top-down approach
	
Image-driven	Knowledge-driven

Outline

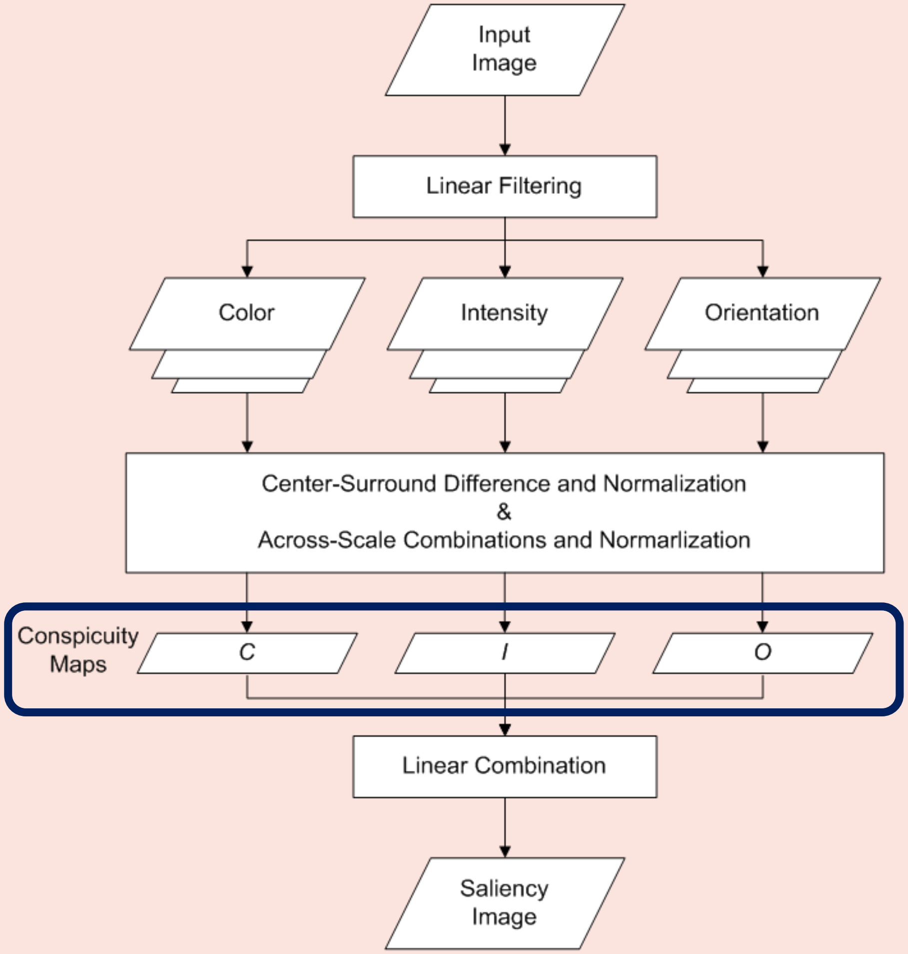
- Introduction
- Feature Extraction
- Bottom-up Feature Selection
- Top-down Feature Selection
- Experimental Results
- Conclusions

Feature Selection – Bottom-up Approach

- Visual attention system
  - Saliency model

    [Itti *et al.*, 1998][1]

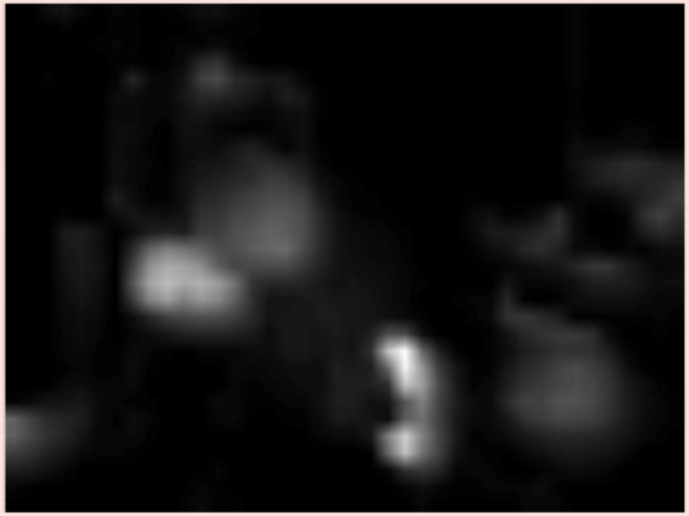
  - Replace the original RGB color space with *CIE L*a*b*
    - Mimic color opponencies of human vision



[1] L. Itti, C. Koch, and E. Niebur, "A Model of Saliency-Based Visual Attention for Rapid Scene Analysis," *PAMI, 1998*.

Feature Selection – Bottom-up Approach

- Results of modified saliency model



C

I

O

$$Sal = N(\tfrac{1}{3}(C + I + O))$$

Outline

- Introduction
- Feature Extraction
- Bottom-up Feature Selection
- Top-down Feature Selection
- Experimental Results
- Conclusions

Feature Selection - Top down approach

- Sometimes, bottom-up ROIs are not enough
  - For example:

- Integrates top-down approach to achieve flexible feature selection

Feature Selection - Top down approach

- Human robot interaction (HRI) can be applied to object learning

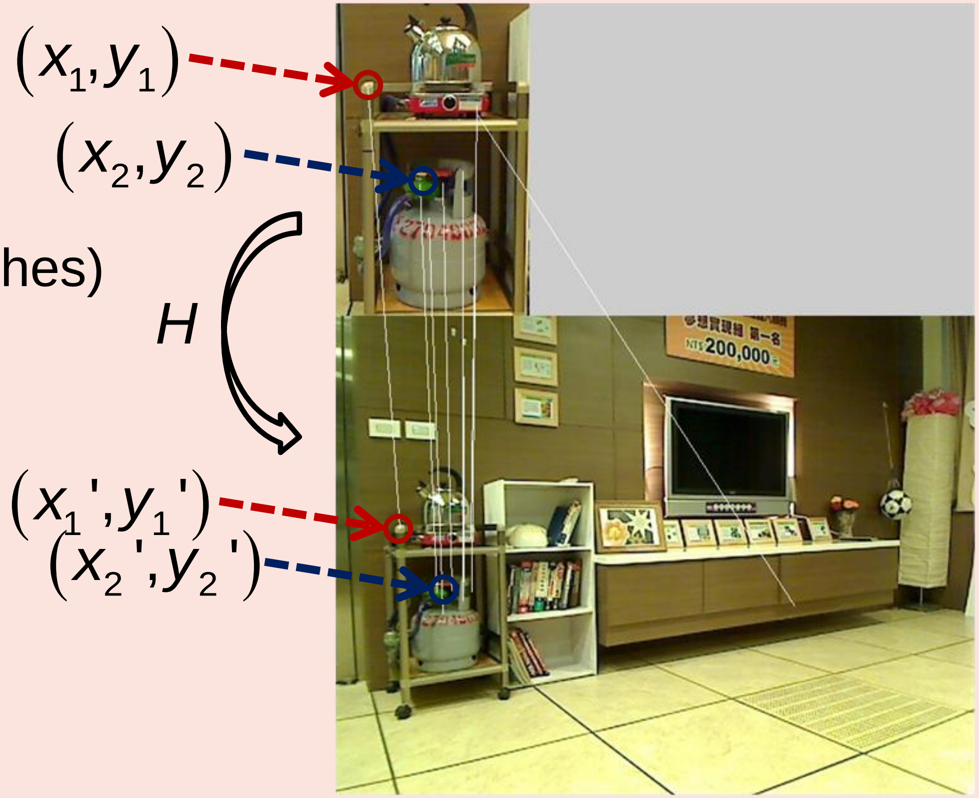
- Communication with the robot
  - Pointing with intelligent devices

Feature Selection - Top down approach

- Robot redetects known objects process
- Based on an object recognition algorithm
  - *RANSAC*

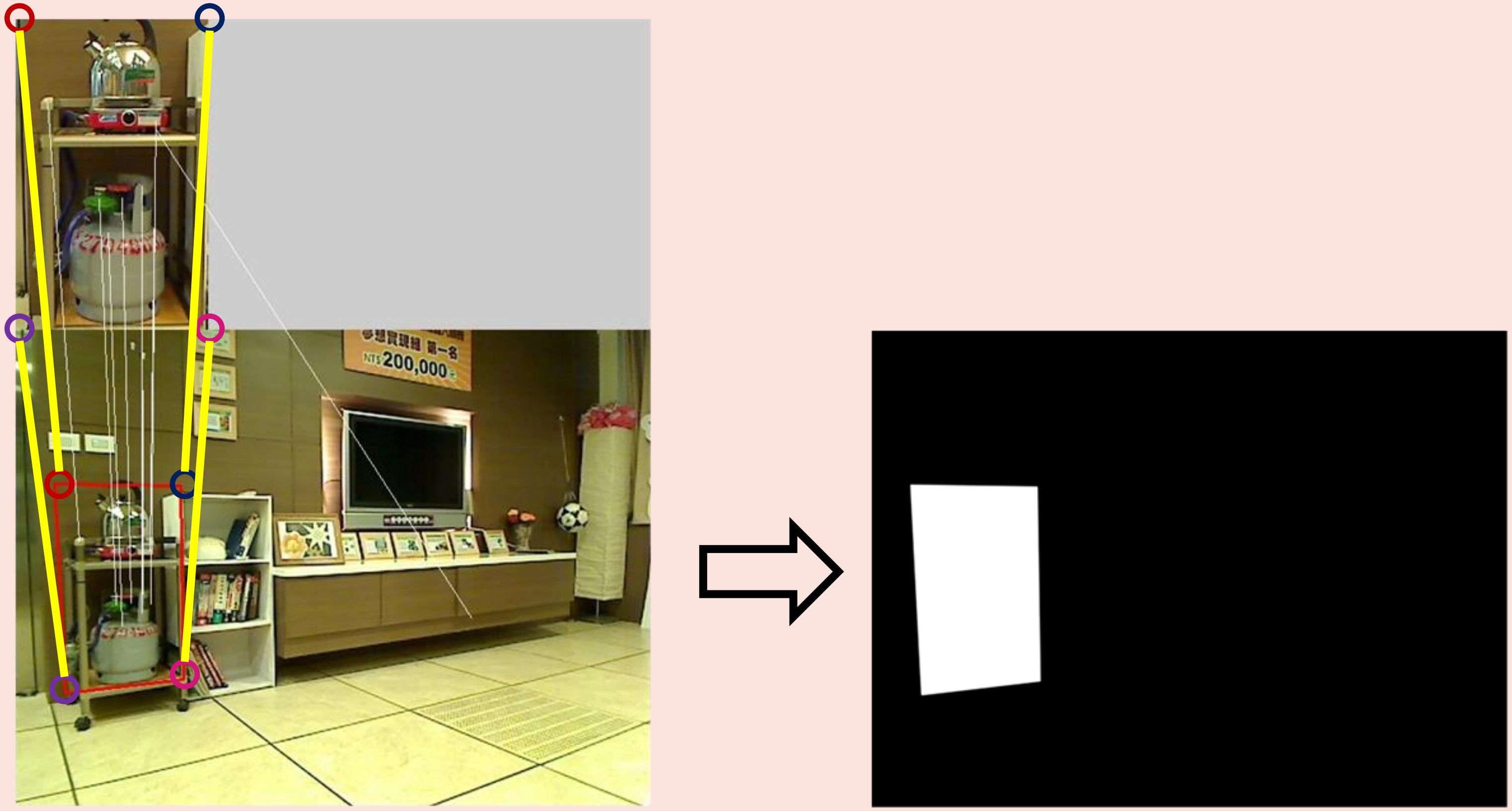
    (Reject inconsistent matches)
- Compute *Homography*

$$s_i \begin{bmatrix} x_i' \\ y_i' \\ 1 \end{bmatrix} = H \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix}$$



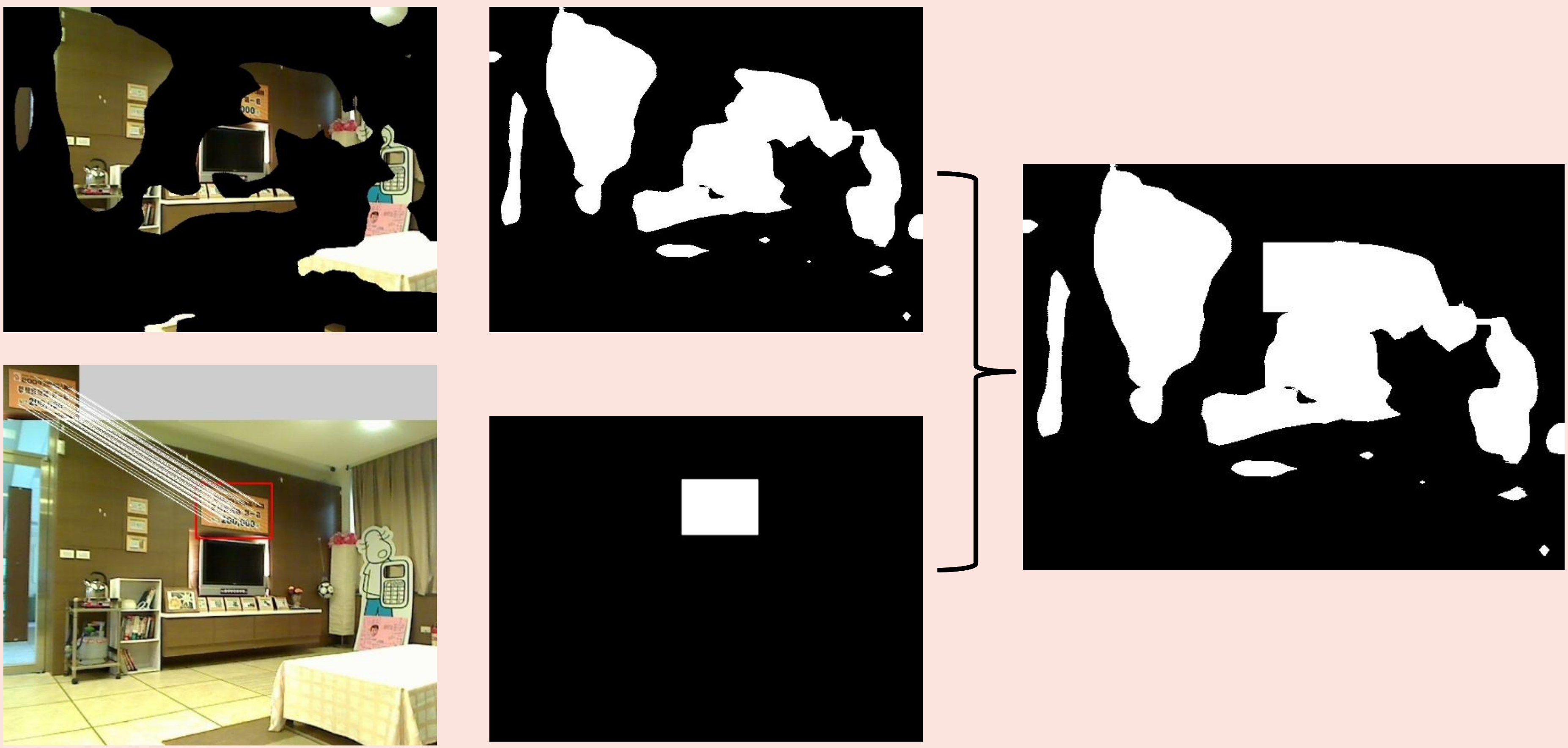
Feature Selection - Top down approach

- Solving the homography matrix, then we project the four end-points to determine top-down ROI



Feature Selection

- Merge two ROIs to obtain versatile features



Outline

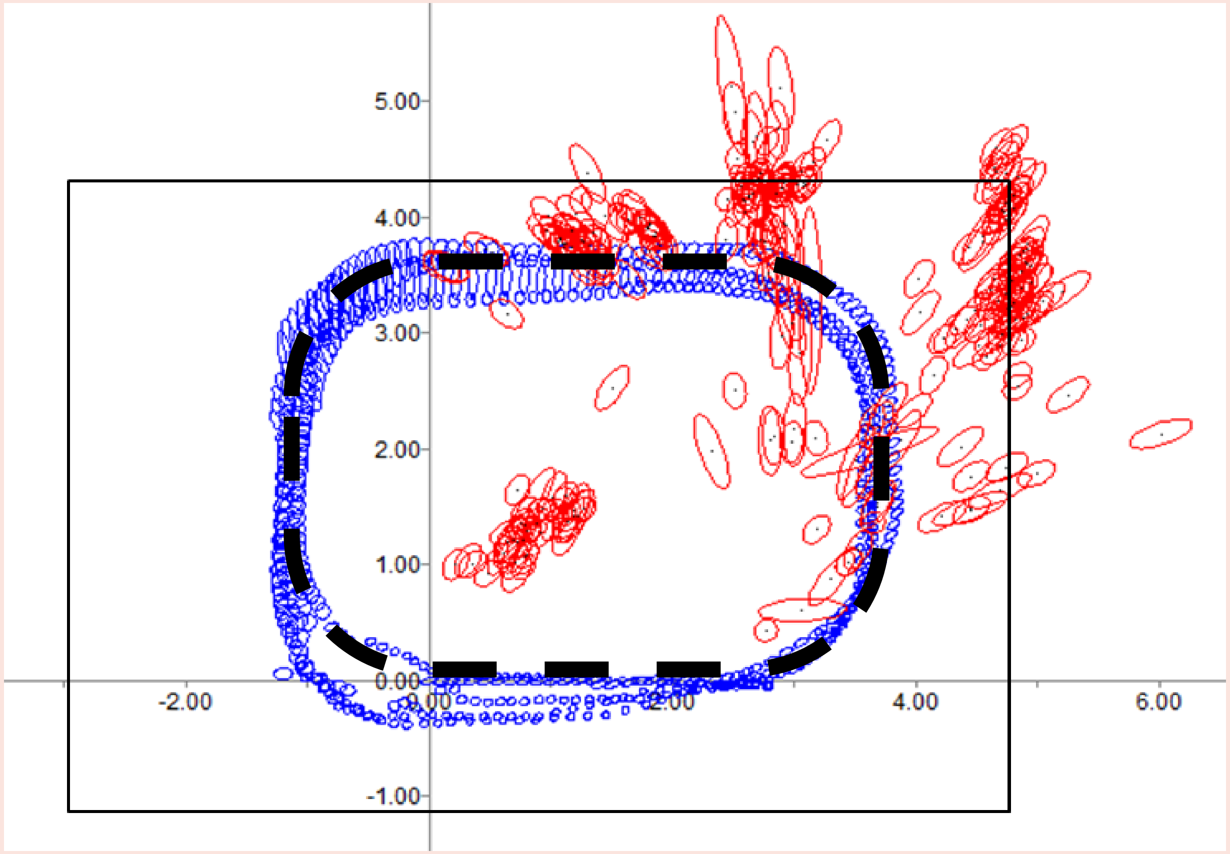
- Introduction
- Feature Extraction
- Bottom-up Feature Selection
- Top-down Feature Selection
- Experimental Results
- Conclusions

Experiment Setting

- Pioneer 3DX
- Logitech webcam V-UBH44

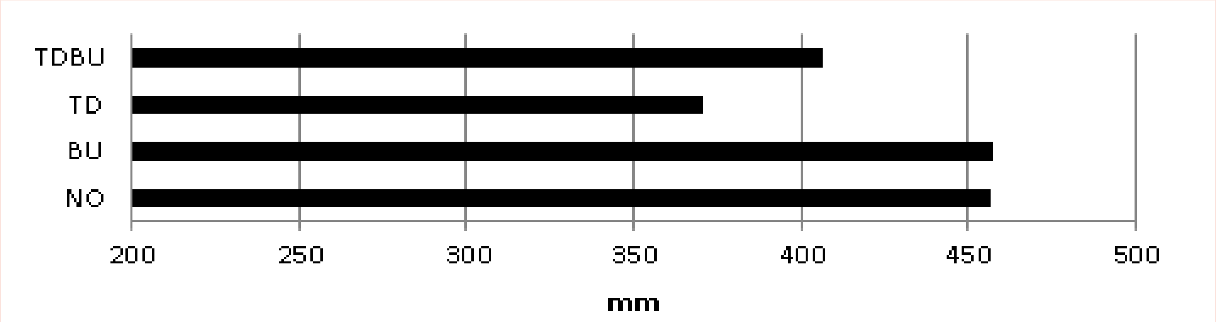
Room size: 6m*8m

SLAM Map

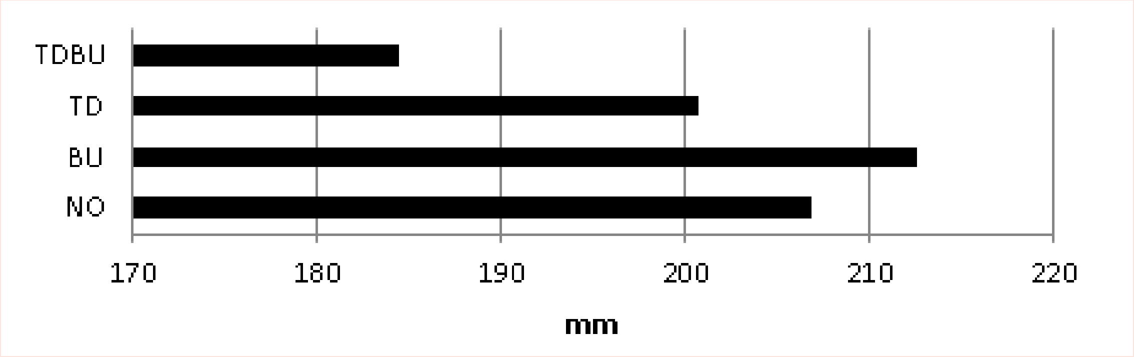


Localization Error Comparison

Average

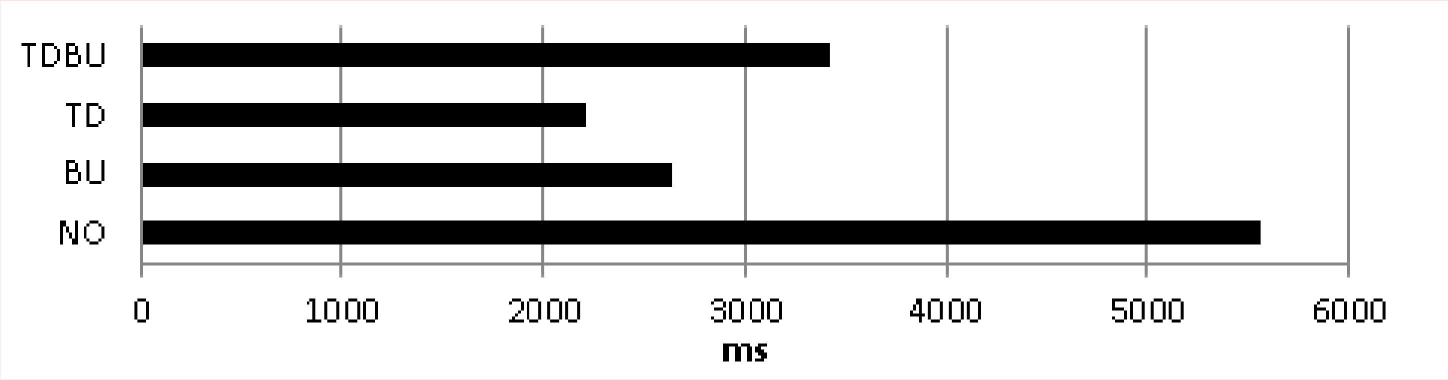


Variance



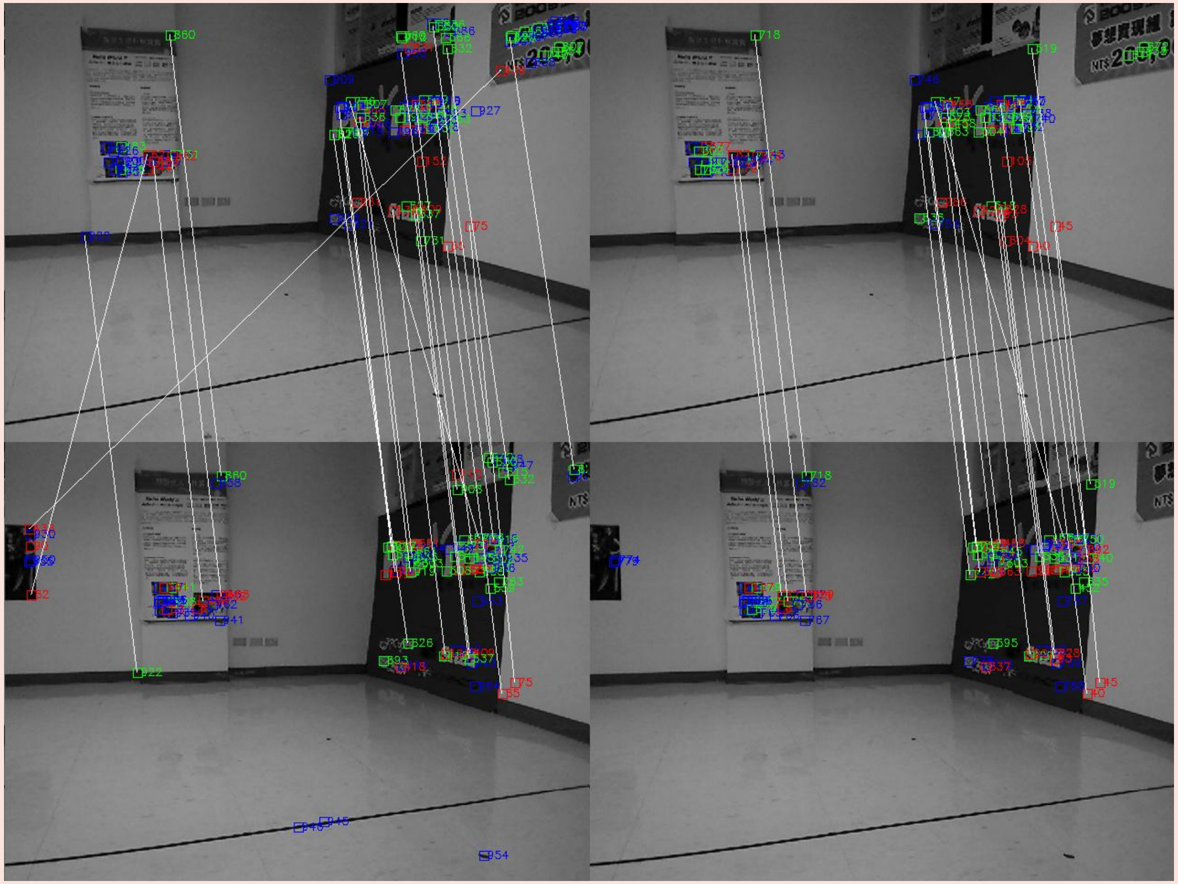
TDBU: the proposed integrated bottom-up and top-down selection
TD: top-down
BU: bottom-up
NO: without selection

Time Comparison

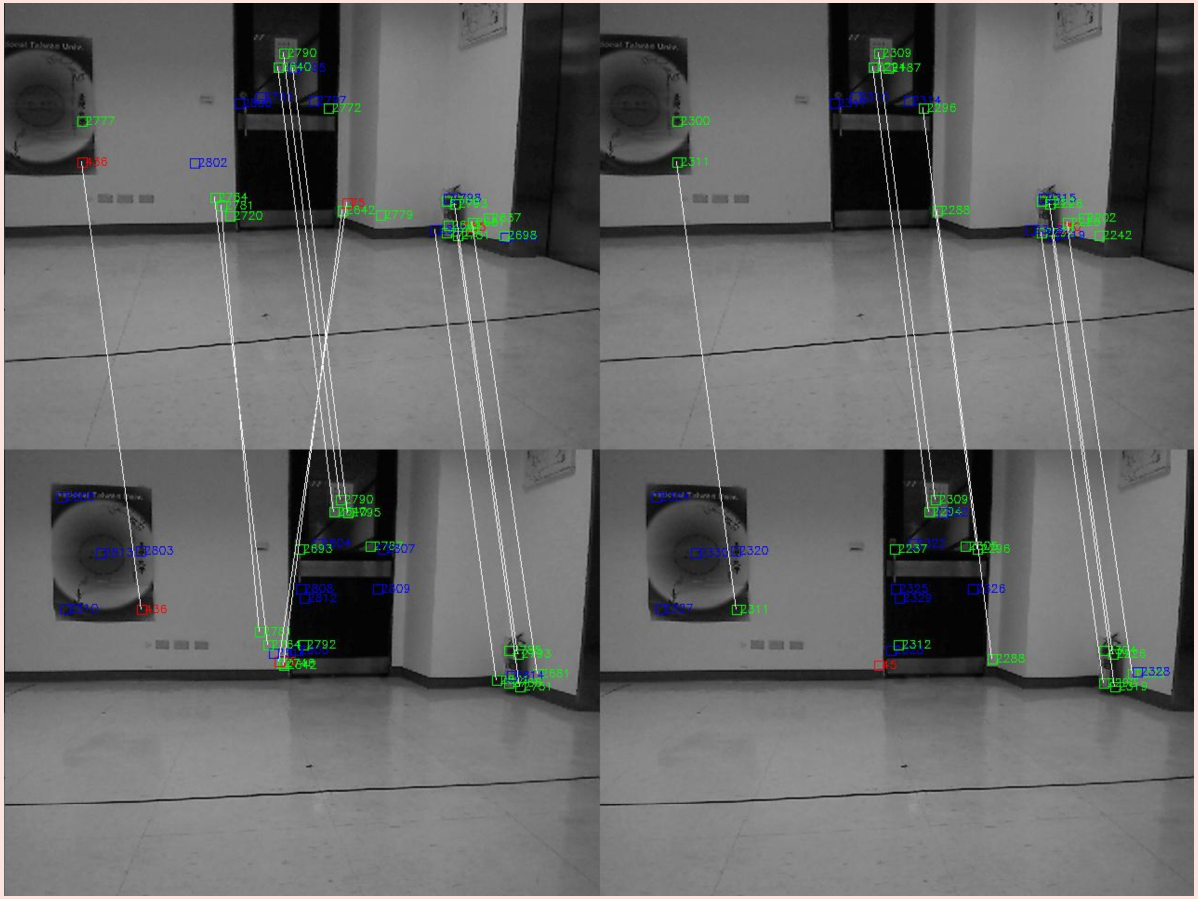


TDBU: the proposed integrated bottom-up and top-down selection
TD: top-down
BU: bottom-up
NO: without selection

Matching Comparison 1/2



Matching Comparison 2/2



Conclusions

- We propose an integrated feature selection strategy based on visual attention system for bearing-only SLAM with EKF
  - Reduce computation time to 62%
  - Reduce localization error to 89%
- Combining bottom-up and top-down approach to construct ROIs allow us to
  - Select salient and useful features
  - Improve data association