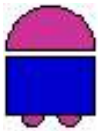
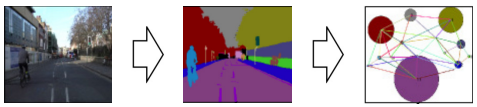
Automated Place Detection Based on Coherent Segments

Mahmut Demir and H. Isil Bozma

Intelligent Systems Laboratory
Electrical and Electronics Engineering, Boğaziçi University, Istanbul, Turkey

January 28, 2018





Introduction

- Goal: Automated appearance-based place detection
- Place is a specific spatial unit or area
- Place detection is a prior step to
  - Place recognition
  - Topological mapping
  - Semantic scene understanding

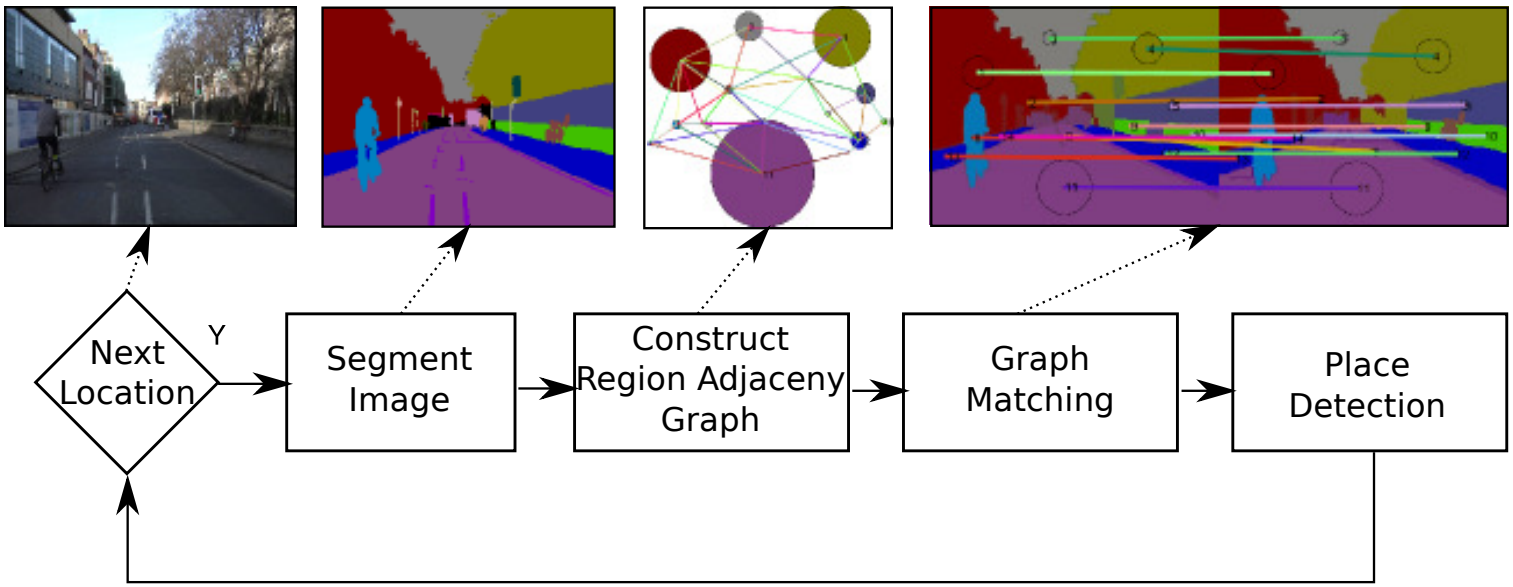
Introduction

- Why appearance-based approach?
  - Geometric or odometric data may not be avaiable
  - Suitable for scene content analysis
- Challenges
  - Appearance variability
  - Perceptual aliasing
  - Indiscriminate boundaries

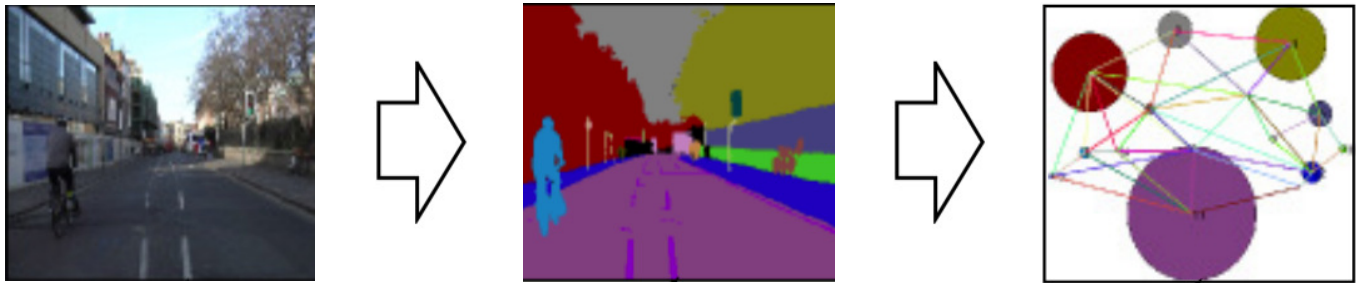
Introduction

- Related work
  - Partioning of incoming sensory data
  - Clustering
  - Feature types:
    - Global: Intensity, Histograms, Optical Flow, GIST ✗ Sensitive
    - Local: SIFT, SURF ✗ Low level, Matching
    - Hybrid: BoW, Bubble Space ✗ Low level
  - Identifying transition regions (i.e. doors, passages, corridors)
    ✗ Fails if transitions are not obvious
- Contribution:
  - Content based place detection
  - Segments Summary Graphs representation

General Approach



Region Adjacency Graphs

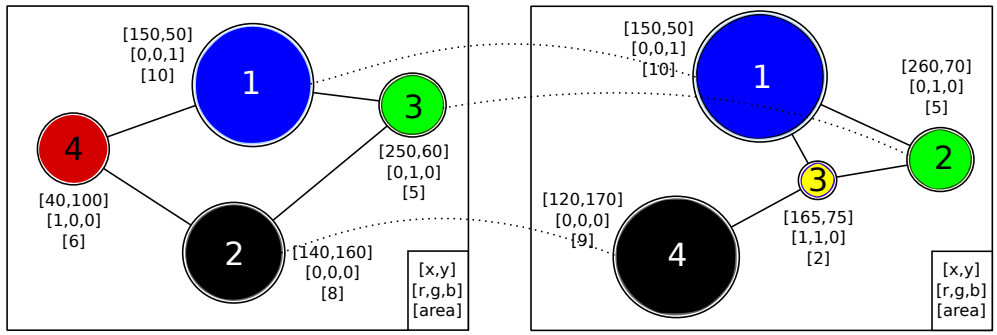


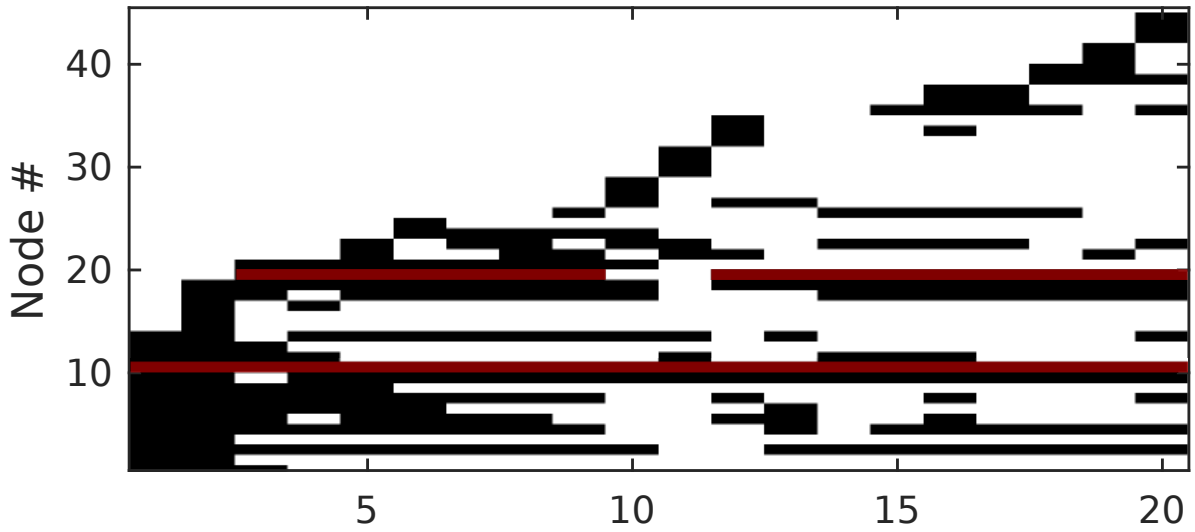
- Graph based segmentation method [Felzenszwalb, Huttenlocher, 2004]
- Segmentation ⇒ Segments ⇒ Nodes
- Neighboring segments ⇒ Edges

- Nodes
  - Color
  - Position
  - Size
- Edges
  - Mean color difference

Temporal RAG Tracking

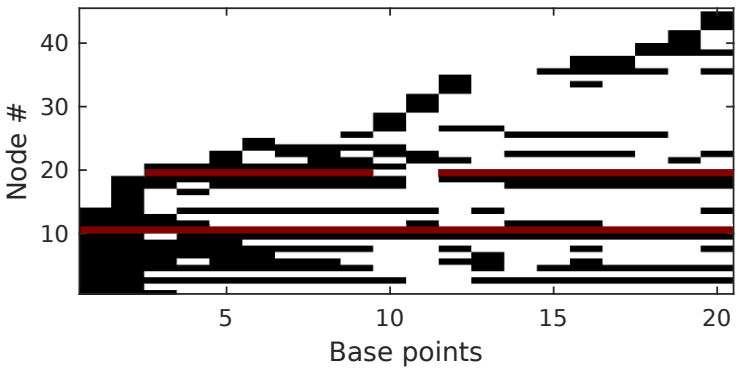
- Matching consecutive RAGs:
  - Cost matrix C^{kl} with $c_{ij} = \delta(s(\mathcal{N}_i^k), s(\mathcal{N}_j^l))$
  - Optimal match by Hungarian method
  - Remove nodes with matching cost $> \tau_m$
- Nonmatched nodes - Matching via backtrack





Coherency score calculation

- Coherency over temporal window
- Parameters:
  - τ_w - window size
  - # appearing nodes
  - # disappearing nodes
  - node weights - ρ_i^l



$$\varphi^k = 1 - \sum_{l=k-\tau_w}^{k} \sum_{i=1}^{|n^l|} \rho_i^l (a_i^l + b_i^l) \quad (1)$$

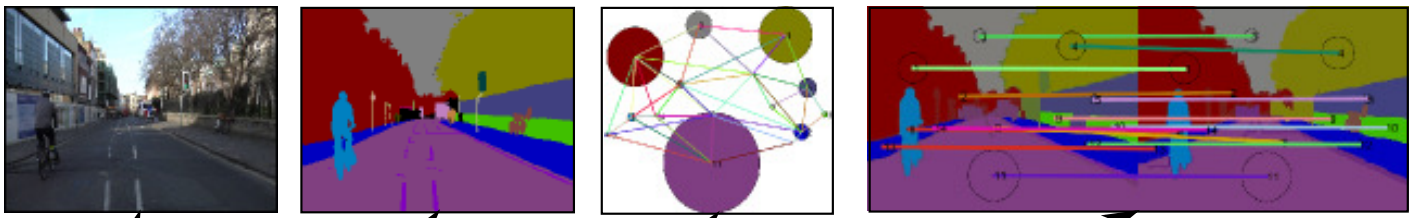
where

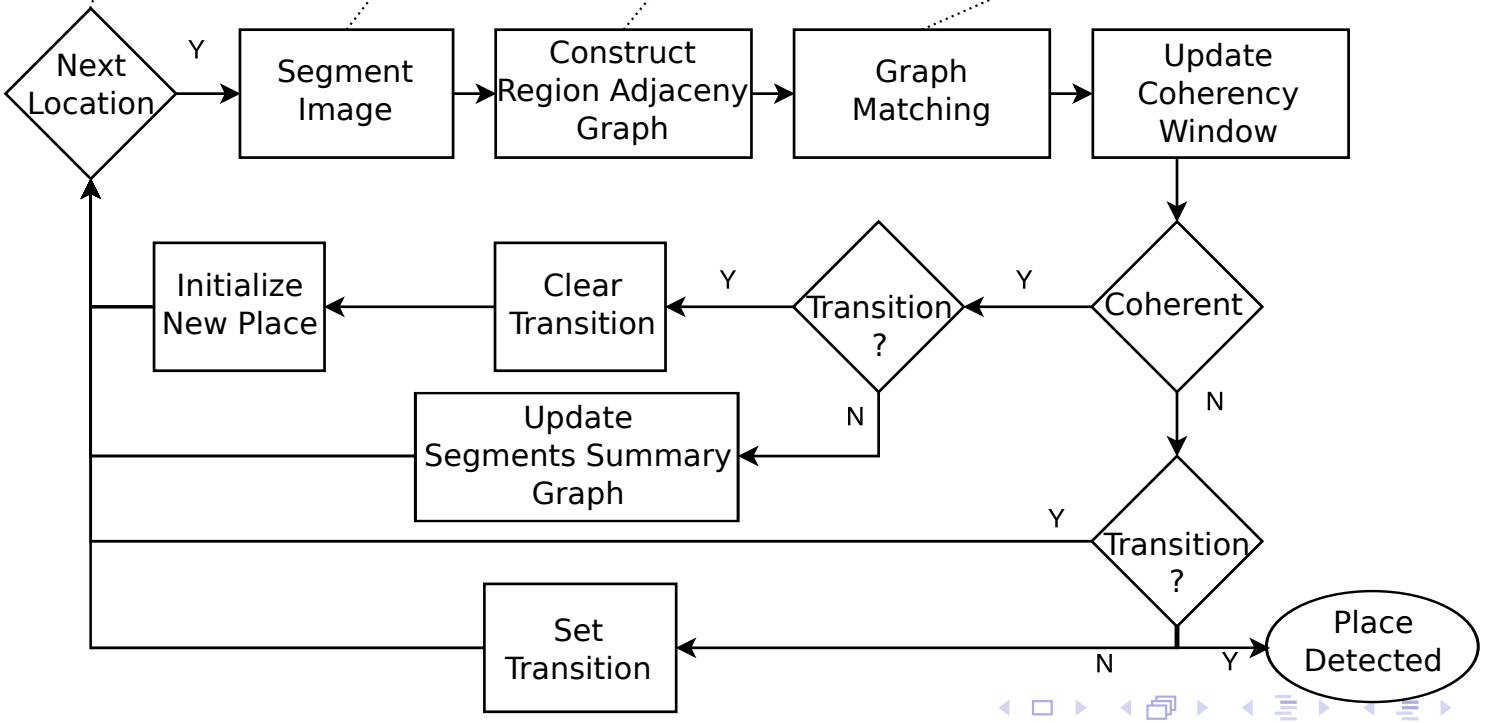
$$a_i^l = \begin{cases} 1 & \text{if } M_{li} > 0,\ M_{l-1,i} = 0 \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

$$b_i^l = \begin{cases} 1 & \text{if } M_{li} = 0,\ M_{l-1,i} > 0 \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

$$\rho_i^l \propto s_3(N_i^l) \times \sum_{k=l-\tau_w}^{l} M_{ki} > 0 \quad (4)$$

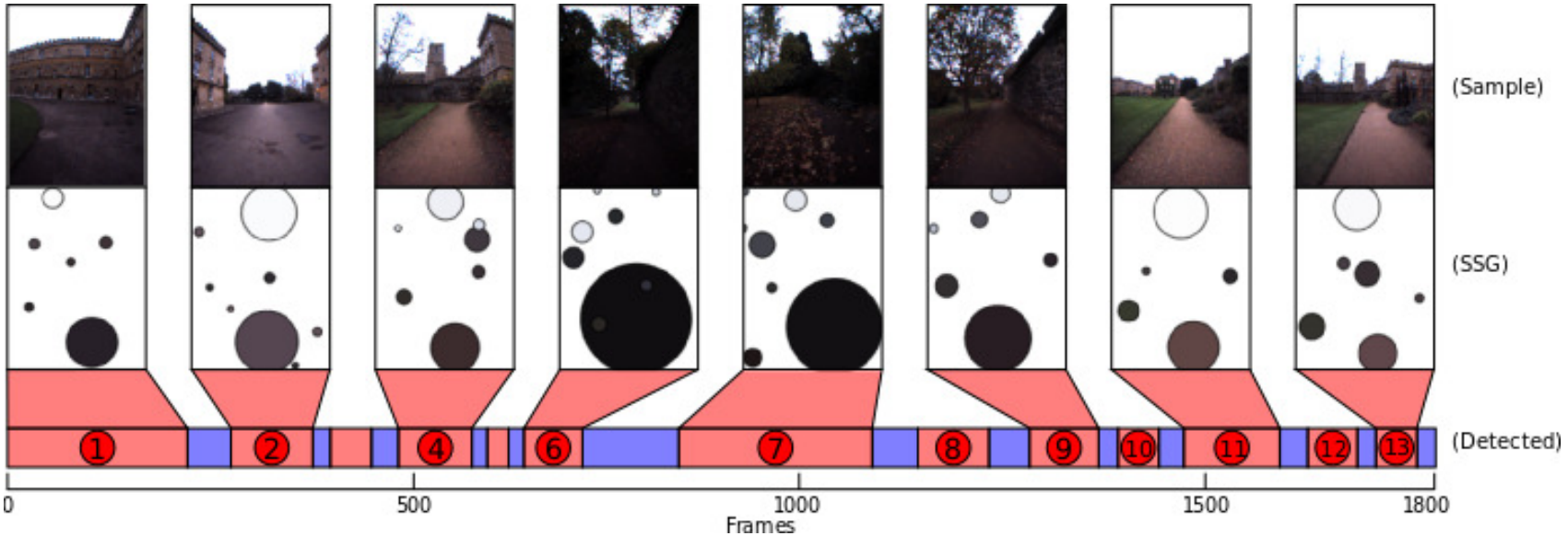
Place Detection





Segments Summary Graphs

- Contains apparent segments only
- Encodes spatial relations



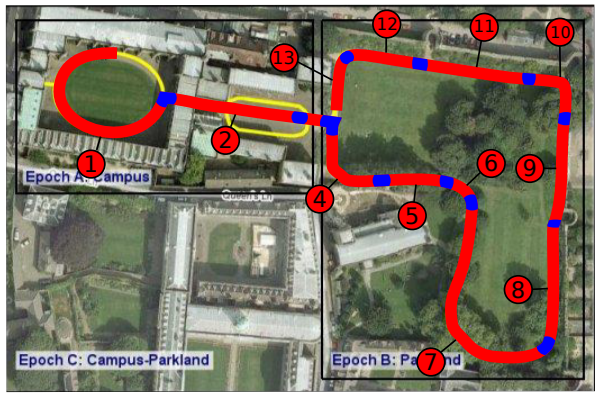
Experiments

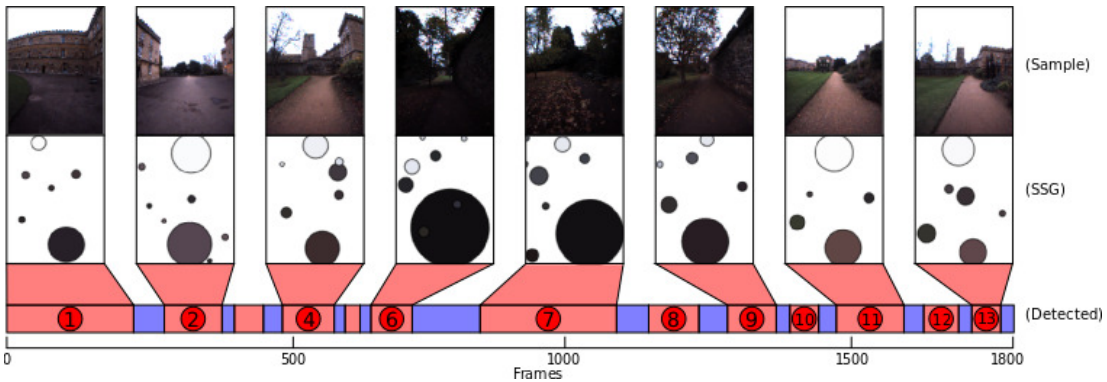
- Outdoor [New College Dataset]
- Indoor [COLD Dataset]
- Comparative study [VPC2009]

Outdoor experiments

- New College dataset
- 1800 basepoints 550 m
- Contains gradual changes

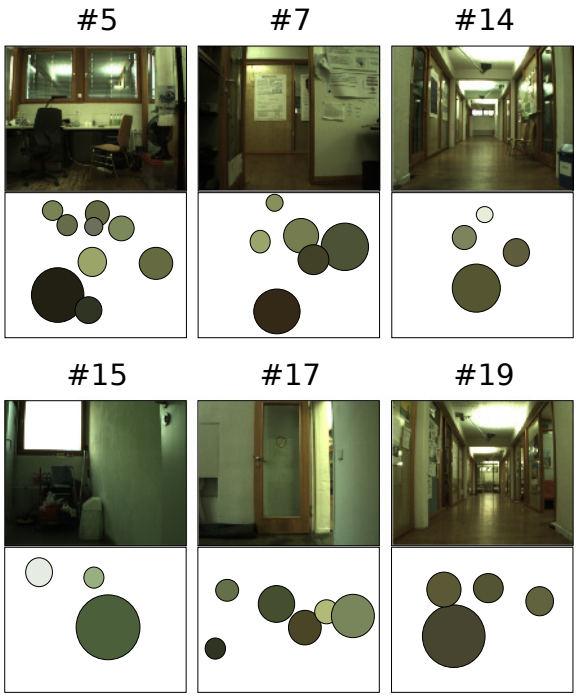
New College Map



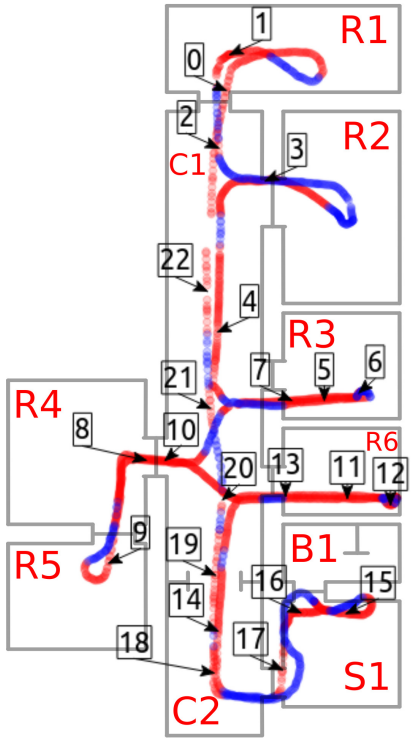


Indoor experiments

- Freiburg (Fr), Saarbrucken (Sa) and Ljubljana sites of COLD Dataset

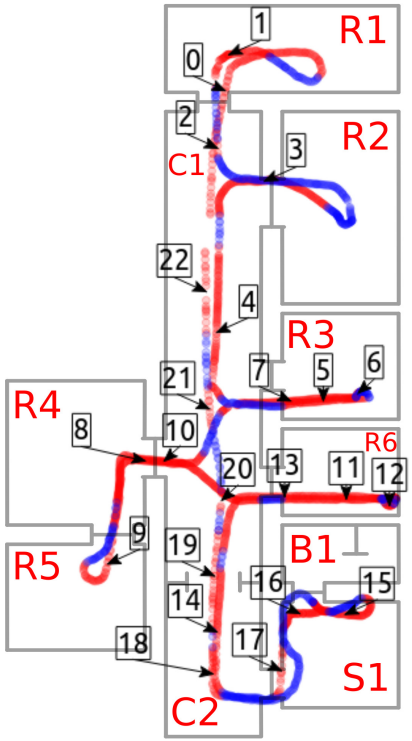


Freiburg site

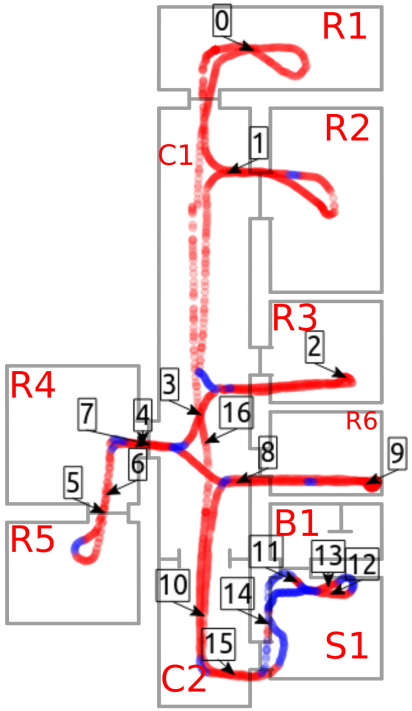


Comparison of Detected Places: SSG vs BD[1] (Bubble Descriptors)

SSG approach



BD approach



1 - Karaoguz & Bozma, ICRA 2014, Autonomous Robots 2015

VPC2009 Dataset

- 21019 images from three different homes
- Challenging dateset:
  - Unclear place boundaries
  - Visual content varies greatly with respect to the viewpoint due to small FOV
- Comparison based on 43 manually annotated transition regions
- Criteria: Minimum %30 overlap

Approach	SSG	BuS
Correct detection (%)	88.3	84.9

Conclusion

Segments based Place Detection

- Stable under wide range of view-points and dynamical changes compared to low-level descriptors
- Reliable place detection
- SSG enables semantic content analysis

Future work

- Use semantic segmentation
- Use SSG for place recognition and hierarchical place representation