Segmentation as selective search for object recognition.

Current State-of-the-art

- Exhaustive search.
- Cheap Features.
- Few, precise locations.

Related Work

Localization

Selective Search → Exhaustive Search

Selective Search
approximate locations (this paper) 1k - 10k
coarse locations 100k - 10M

Object Delineation - - - - -> Object Recognition
Precise Locations 10 -100

Proposed Work

- selective search over proposed regions →
- More expensive features
- Many, approximate locations

- high recall
- faster to compute
- coarse locations are sufficient

Contributions

- Use segmentation to select best locations for recognition
- Accounting for scene conditions through invariant color spaces results in high recall.
- Selective search enables use of expensive features like BoW.

Approach

- use many scales, multiple invariant color spaces.
1. Over segmentation
2. Greedily group together similar adjacent regions, until the entire image is in the same region.
3. Resulting heirarchy of regions contains rich information that can be used to generate proposal regions

Intuition: ① if we miss a location, we can never detect the object
② appearance and near context are most effective for recognition

Exhaustive Search
~ sliding window
(coarse grid, fixed aspect ratios ...)

Color Spaces
Standard RGB
Normalized RGB
HSV
↑
most invariant

Similarity
$$S = S_{size} + S_{texture}$$
Execution a Vb
Total image
SIFT