Segmentation as selective search for object recognition

Elie Cattan

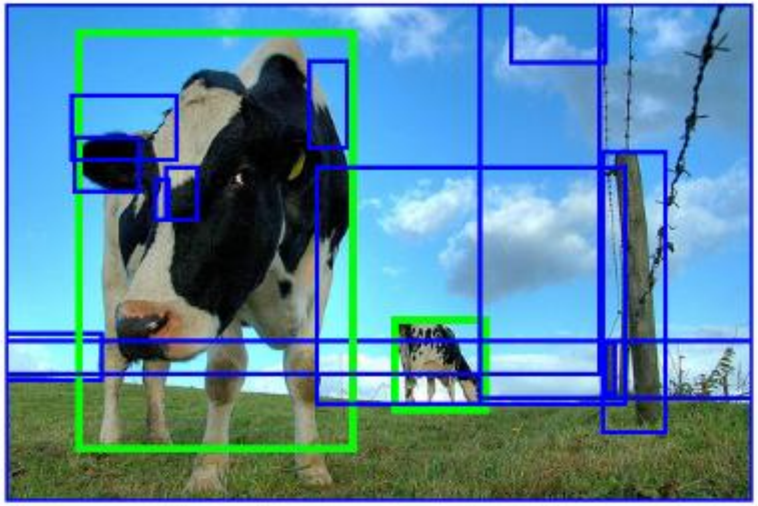
6/12/2013

Introduction

- Object recognition
- Exhaustive search
  - Quick computation needed
- Selective search

Introduction

- This paper
  - Coarse location
  - Emphasizing recall
  - Fast to compute



State of the art – exhaustive search

- Search object and part of the objects (Felzenszwalb et al.)
- Branch and bound (Lampert et al.)
- Use of random (Alexe et al.)

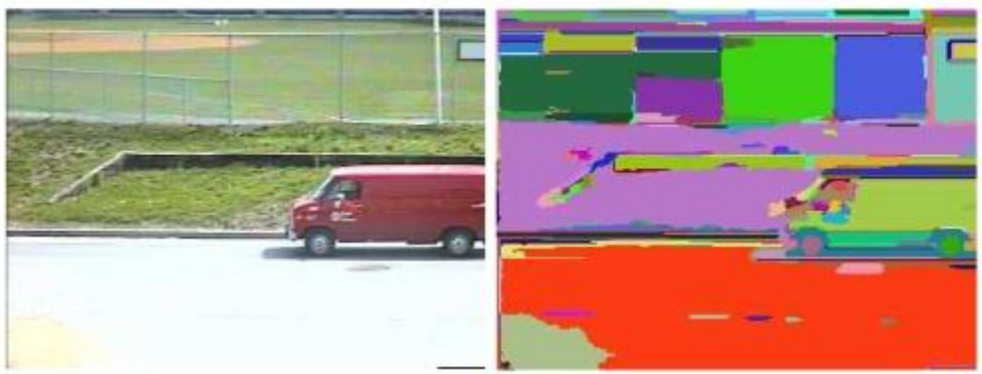
- Class dependent vs class independent

State of the art – selective search

- Gu et al. Work
  - But only a single hierarchy
- Foreground/Background segmentations (Carreira et al.)
  - With precise object delineations

Algorithm

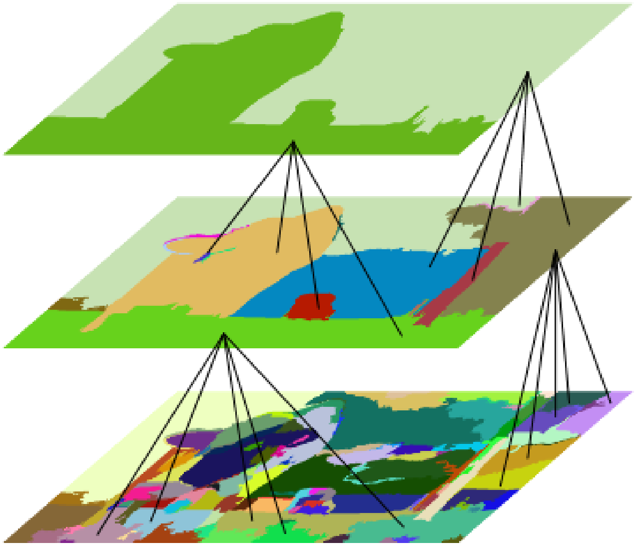
- The oversegmentation
  - Felzenszwalb et al.



Algorithm

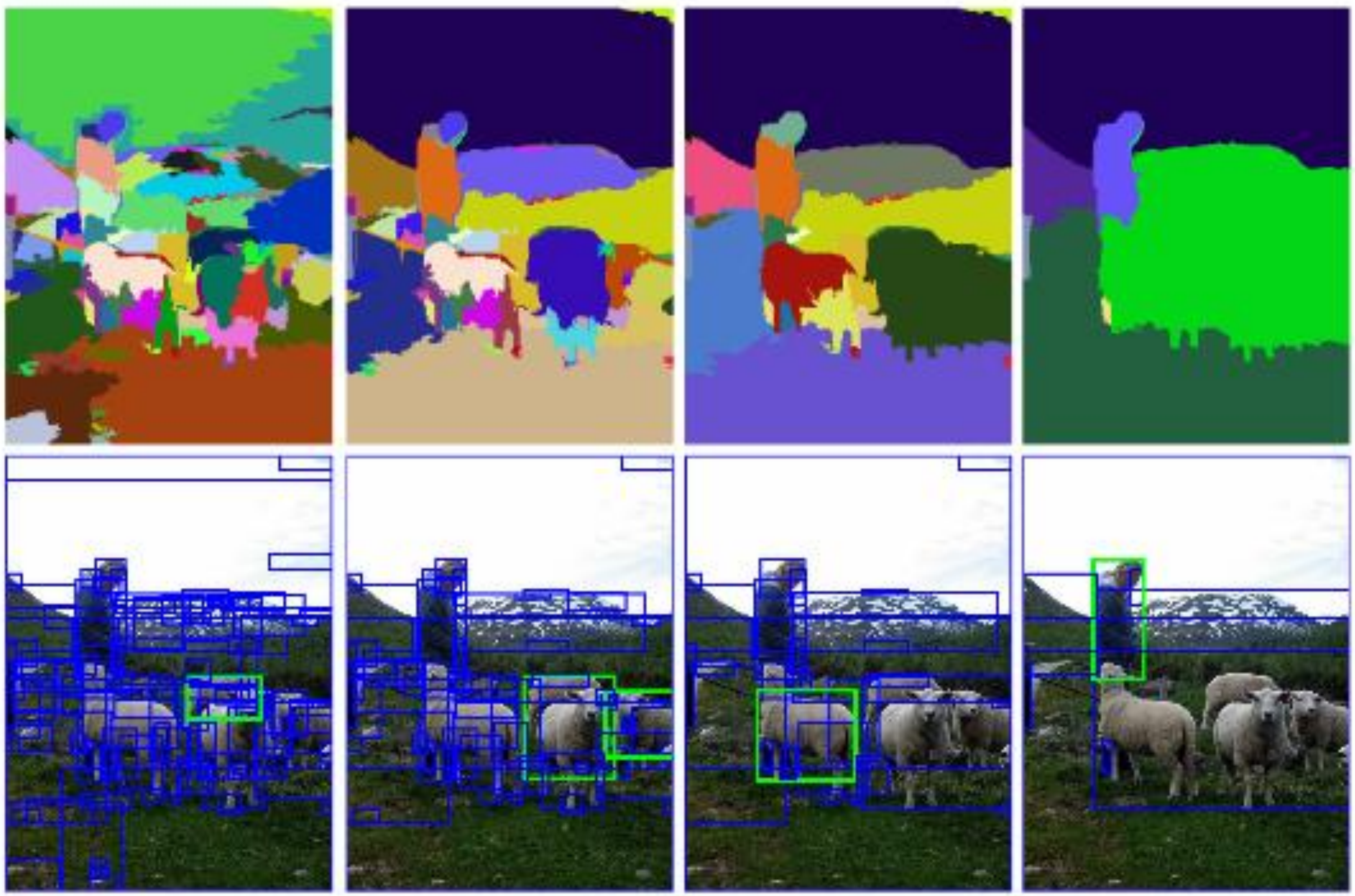
- Group similar regions
  - S = Ssize + Stexture

- Multiple color spaces



Algorithm

- Results :



Object recognition system

- Bag of feature
  - SIFT + OpponentSIFT + RGB-SIFT
  - 4096 words
- Training + retraining

Experiments

- Flat vs hierarchical
- Object recognition
- Object delineation
- Accuracy

Experiment 1

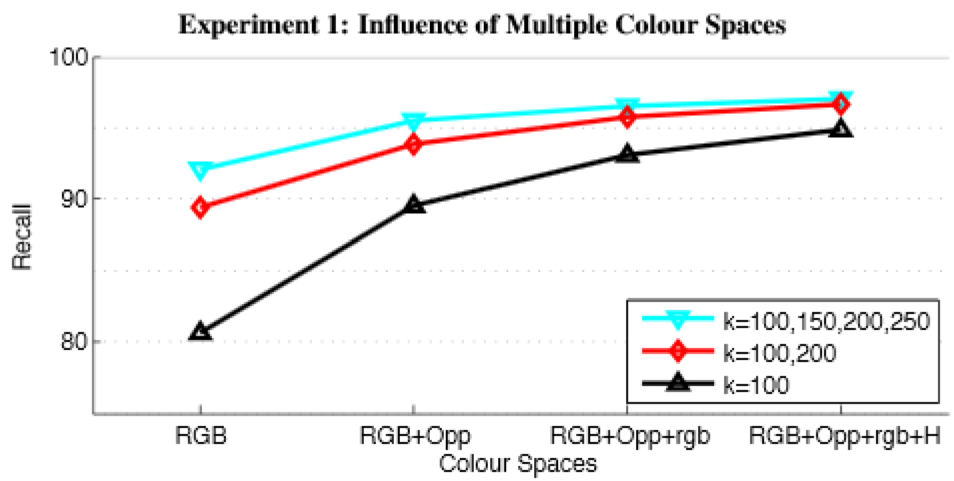
- Flat vs Hierarchical

Experiment 1: Multiple Flat segmentations versus Hierarchy

	Max. recall (%)	# windows
[10] $k = 100, 150 \dots 1000$	84.8	665
[10] $k = 100, 110 \dots 1000$	87.7	1159
Hierarchical $k = 100$	80.6	362
Hierarchical $k = 100, 200$	89.4	511

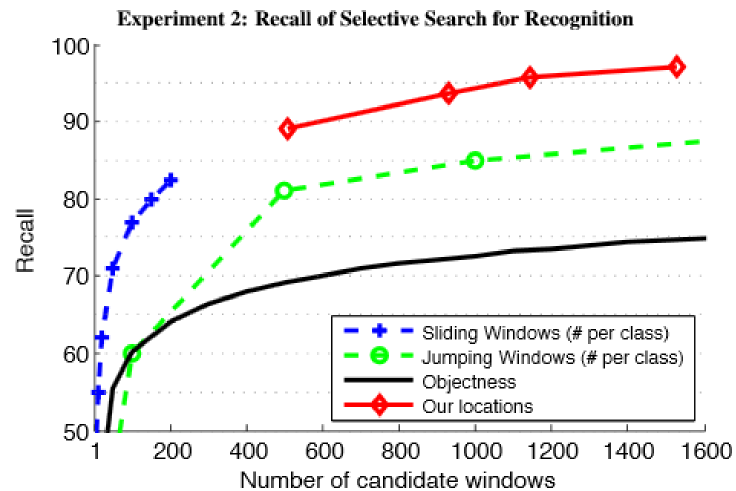
Table 1. Comparison of multiple flat segmentations versus a hierarchy in terms of recall and the number of windows per image.

- Multiple colour spaces



Experiment 2

- This paper vs State of the art – object recognition



Experiment 2: Maximum Recall of Selective Search for Recognition

	Max. recall (%)	# windows
Sliding Windows [13]	83.0	200 per class
Jumping Windows [27]	94.0	10,000 per class
'Objectness' [1]	82.4	10,000
Our hypotheses	96.7	1,536

Table 2. Comparison of maximum recall between our method and [1, 13, 27]. We achieve the highest recall of 96.7%. Second comes [27] with 94.0% but using an order of magnitude more locations.

Experiment 3

- This paper vs State of the art – object delineation

Experiment 3: Recall of Selective Search for Segmentation

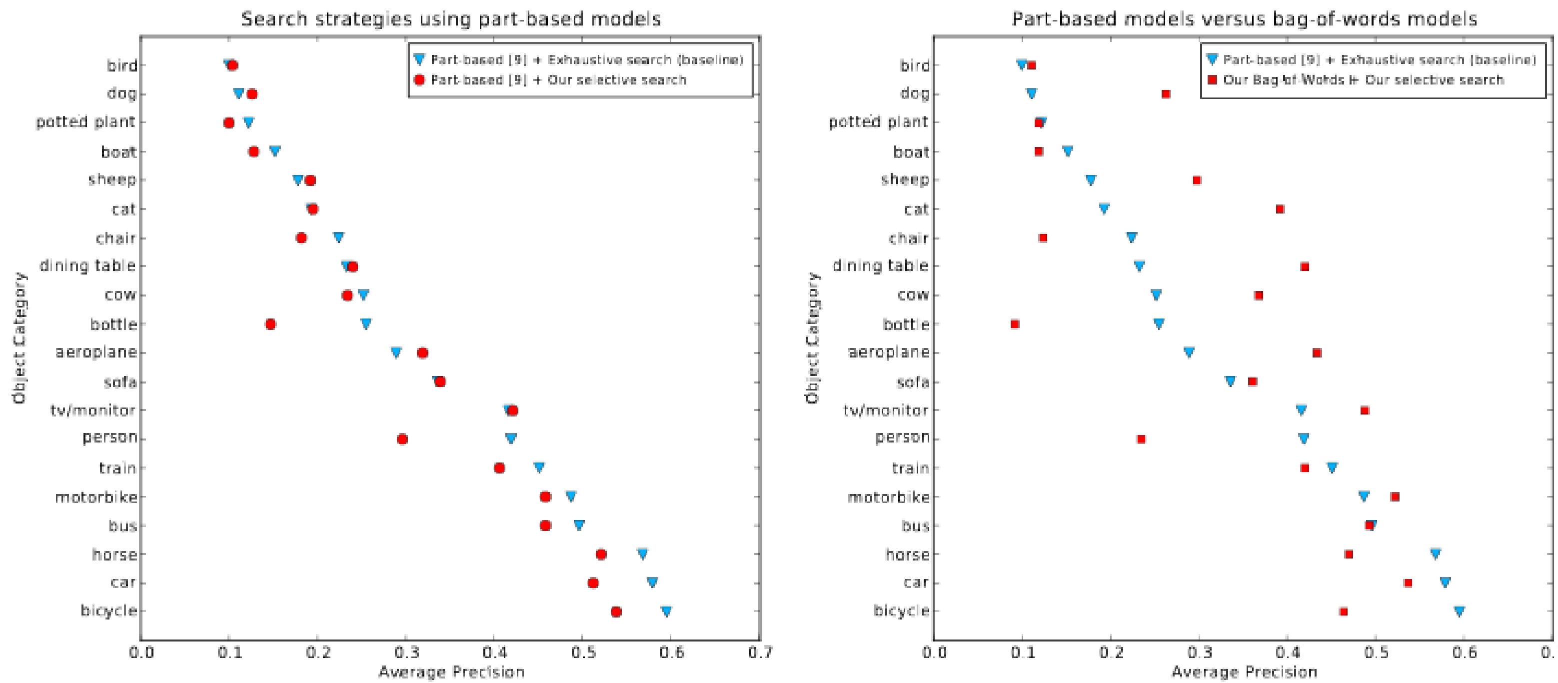
	Max. recall (%)	# windows	Time (s)
Carreira [3]	78.2	697	432
Endres [7]	82.2	1,989	226
Our hypotheses	79.8	1,973	8
Combination	90.1	4,659	666

Table 3. Comparison of our paper with [3, 7] in terms of recall on the Pascal VOC 2007 segmentation task. Our method has competitive recall while being more than an order of magnitude faster.

Experiment 4

- Accuracy

Experiment 4: Object Recognition Accuracy on VOC2007 Test Set



Experiment 4: Object Recogntion Accuracy on VOC2010 Test Set

System	plane	bike	bird	boat	bottle	bus	car	cat	chair	cow	table	dog	horse	motor	person	plant	sheep	sofa	train	tv
NLPR	.533	**.553**	**.192**	**.210**	.300	.544	.467	.412	**.200**	**.315**	.207	.303	.486	.553	.465	.102	.344	.265	**.503**	.403
MIT UCLA [29]	.542	.485	.157	.192	.292	**.555**	.435	.417	.169	.285	.267	.309	.483	.550	.417	.097	.358	.308	.472	.408
NUS	.491	.524	.178	.120	.306	.535	.328	.373	.177	.306	.277	.295	**.519**	**.563**	.442	.096	.148	.279	.495	.384
UoCTTI [9]	.524	.543	.130	.156	**.351**	.542	**.491**	.318	.155	.262	.135	.215	.454	.516	**.475**	.091	.351	.194	.466	.380
This paper	**.582**	.419	**.192**	.140	.143	.448	.367	**.488**	.129	.281	**.287**	**.394**	.441	.525	.258	**.141**	**.388**	**.342**	.431	**.426**

Table 4. Results from the Pascal VOC 2010 detection task test set, comparing the approach from this paper to the current state-of-the-art. We improve the state-of-the-art up to 0.085 AP for 8 categories and equal the state-of-the-art for one more category.

Conclusion

- Many approximate locations
- Set of complementary segmentations
- Very effective for object recognition

Questions