Selective Search for Object Recognition

Introduction

Introduction

- Capture all scales
- Diversification
- Fast to compute

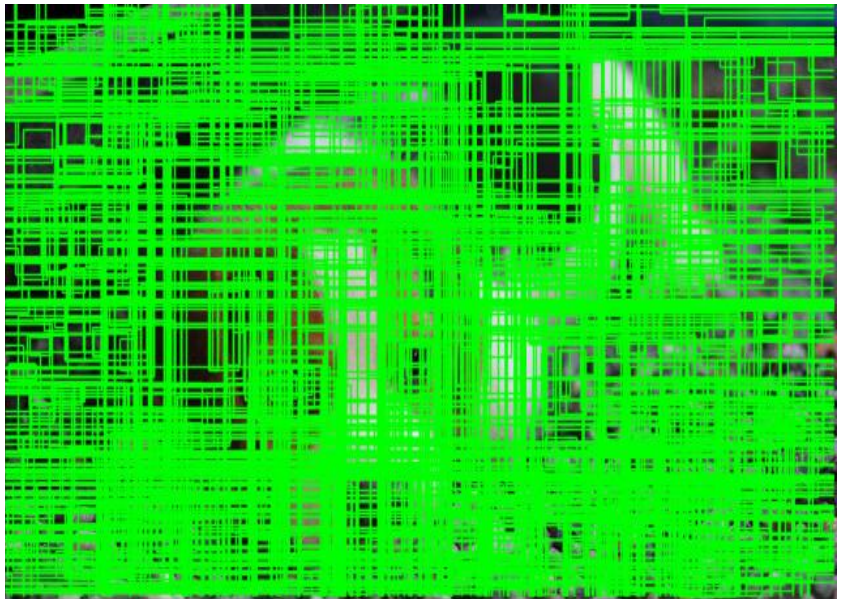
(a) (b)

(c) (d)

Related works

Exhaustive search

- Coarse search grid with weak classifier
- Linear classifier SVM and HOG
- Regular grid and fixed aspect ratio
- Branch and Bound technique



Segmentation

- Class Independent Object Hypothesis
- Pixel wise image classification
- Generate a part of hypotheses using grouping method

Other Sampling Techniques

- Train a classifier and randomly sample boxes
- Region wise image classification
- Jumping window strategies to predict objectness

Selective Search

Selective search by Hierarchical Grouping

- Bottom up grouping
- Regions Based Feature
- Fast method of efficient graph based method

Input: (colour) image
Output: Set of object location hypotheses L

Obtain initial regions $R = \{r_1, \cdots, r_n\}$
Initialise similarity set $S = \emptyset$
foreach *Neighbouring region pair* (r_i, r_j) **do**
    Calculate similarity $s(r_i, r_j)$
    $S = S \cup s(r_i, r_j)$

while $S \neq \emptyset$ **do**
    Get highest similarity $s(r_i, r_j) = \max(S)$
    Merge corresponding regions $r_t = r_i \cup r_j$
    Remove similarities regarding $r_i : S = S \setminus s(r_i, r_*)$
    Remove similarities regarding $r_j : S = S \setminus s(r_*, r_j)$
    Calculate similarity set S_t between r_t and its neighbo
    $S = S \cup S_t$
    $R = R \cup r_t$

Extract object location boxes L from all regions in R

Diversification Strategies

- Variety of color spaces with different invariance properties
- Different similarity measures
- varying starting regions

»»Complementary Similarity measures

★ **Color similarity measures**

- » Create a color histogram for each region
- » 1 dimensional color histogram for each color channel using 25 bins total 75 dimension
- » Measure similarity with histogram intersection

$$s_{colour}(r_i, r_j) = \sum_{k=1}^{n} \min(c_i^k, c_j^k)$$

★ **Texture similarity measures**

- » SIFT like Features
- » Take gaussian derivatives of the image in 8 orientation for each channel

★ Size similarity measures

- Merging larger regions to smaller ones to create balanced hierarchy
  - Adding a size component to our similarity metric that ensures small regions are more similar to each other.

- $$s_{size}(r_i, r_j) = 1 - \frac{\text{size}(r_i) + \text{size}(\text{r}_\text{j})}{\text{size}(im)}$$

★ Fill measures

- Measures how well region ri and rj fit into each other
- Idea is to fill gaps (avoid holes)

  bij is the tight bounding box

$$fill(r_i, r_j) = 1 - \frac{size(BB_{ij}) - size(r_i) - size(r_i)}{size(im)}$$

★ Final Similarity Measures

- measure the similarity between two patches as a **linear combination** of the four given measures:

$$s(r_i, r_j) = a1s_{colour}(r_i, r_j) + a2s_{texture}(r_i, r_j) + a3s_{size}(r_i, r_j) + a4s_{fill}(r_i, r_j)$$

Complementry Starting regions

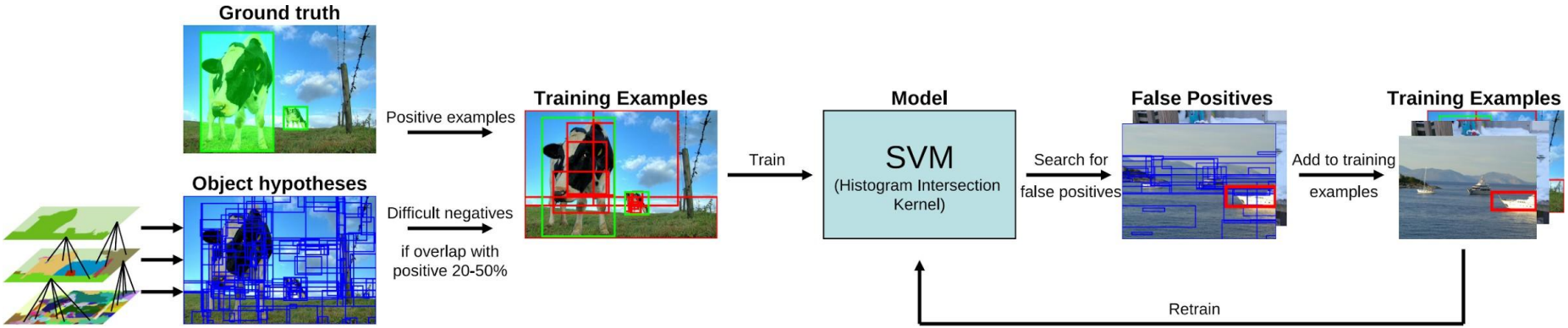
- Oversegmentation
- we vary the threshold additional to the various similarity measures

Object Recognition

Object Recognition

- Bag of Words for Object Recognition
- Employ variety of color SIFT descriptors and a finer spatial pyramid division
- Jumping window strategies to predict objectness
- Visual codebook
- Total feature of vector length of 360,000

- we employ a support vector machine with histogram intersection kernal



Evaluation

Diversification Strategies Results

★ **Individual Diversification strategies**

» Full hierarchy is more natural than using multiple flat partitionings

threshold k	MABO	# windows
Flat $k = 50, 150, \cdots, 950$	0.659	387
Hierarchical (this paper) $k = 50$	0.676	395
Flat $k = 50, 100, \cdots, 1000$	0.673	597
Hierarchical (this paper) $k = 50, 100$	0.719	625

★ Individual Diversification strategies

Similarities	MABO	# box
C	0.635	356
T	0.581	303
S	0.640	466
F	0.634	449
C+T	0.635	346
C+S	0.660	383
C+F	0.660	389
T+S	0.650	406
T+F	0.638	400
S+F	0.638	449
C+T+S	0.662	377
C+T+F	0.659	381
C+S+F	0.674	401
T+S+F	0.655	427
C+T+S+F	0.676	395

Colours	MABO	# box
HSV	0.693	463
I	0.670	399
RGB	0.676	395
rgI	0.693	362
Lab	0.690	328
H	0.644	322
rgb	0.647	207
C	0.615	125

Thresholds	MABO	# box
50	0.676	395
100	0.671	239
150	0.668	168
250	0.647	102
500	0.585	46
1000	0.477	19

★ Combination of Diversification strategies

» Combining strategies improves performance even more:

Version	Diversification Strategies	MABO	# win	# strategies	time (s)
Single Strategy	HSV C+T+S+F $k = 100$	0.693	362	1	0.71
Selective Search Fast	HSV, Lab C+T+S+F, T+S+F $k = 50, 100$	0.799	2147	8	3.79
Selective Search Quality	HSV, Lab, rgI, H, I C+T+S+F, T+S+F, F, S $k = 50, 100, 150, 300$	0.878	10,108	80	17.15

Thank you!