Fusing shape and appearance information for object category detection

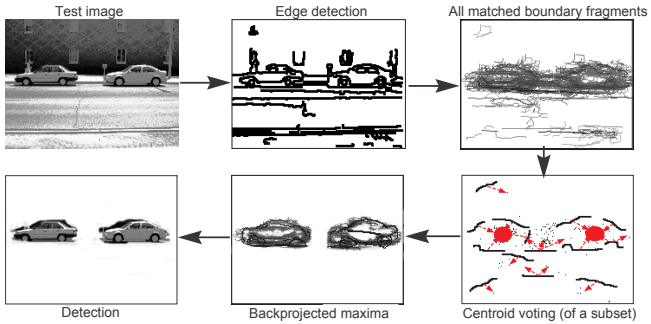
A. Opelt*, A. Pinz* and A. Zisserman**

* Vision-based Measurement Group, EMT, Graz University of Technology

* * Visual Geometry Group, University of Oxford

Basic Approach

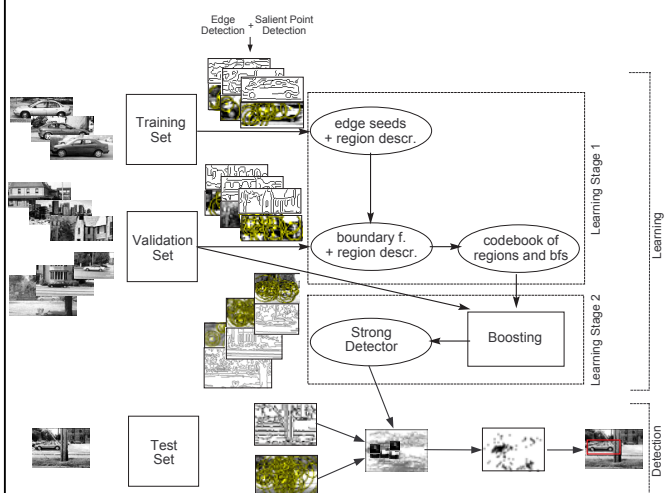
We build on our shape based Boundary Fragment Model from ECCV 2006.
Detection is shown here.



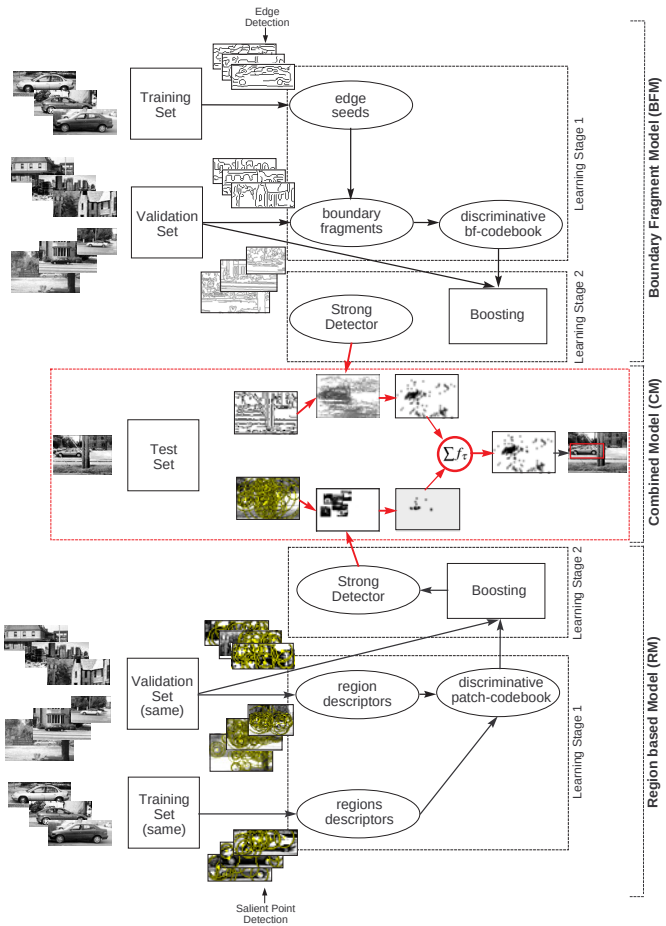
The central question is:

How should different feature types (appearance and shape) be combined in a centroid based category detector?

The selective features Model (SF)



The combined Model (CM)



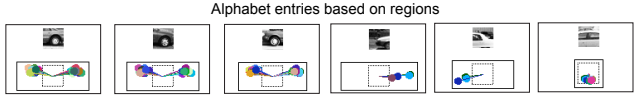
$$conf(x_n) = \frac{\sum_{\tau \in Q} f_\tau \sum_i^T p(c, h_i^\tau)}{\sum_\tau f_\tau}$$

Combining the confidences of each Hough voting space.

x_n ... a certain point in the voting space
τ ... a feature type (e.g. shape)
Q ... the set of all used feature types
f_τ ... a specific weight for the influence of that feature type
T ... the number of weak detectors
$p(c, h_i^\tau)$ ... the probability of a weak detector h_i^τ matching for class c

Our codebook model based on regions (RM)

Alphabet entries based on regions



Weak detectors based on regions

Results

RPC-equal error rate on the UIUC cars-side

Method	RM	CM	SF	BFM	Fergus *et al.* 03	Leibe *et al.* 04	Amores *et al.* 05	Shotton *et al.* 05
RPC-EER	10.5	6.2	7.0	15.0	11.5	9.0	10.0	7.2

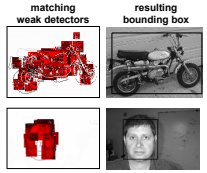
RPC-equal error rate on the Caltech dataset (detection)

Cat.	RM	CM	SF	BFM	Leibe *et al.* 04	Shotton *et al.* 05
Cars-rear	2.9	0.0	0.5	2.3	6.1	-
Airplane	22.5	4.2	13.4	7.4	-	-
Motorbikes	4.0	2.0	3.7	4.4	6.0	7.6
Faces	2.4	1.0	3.2	3.6	-	6.0

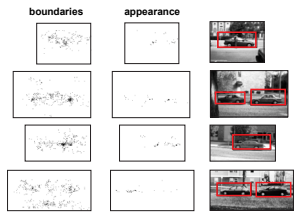
ROC-equal error rate on the Caltech dataset (classification)

Cat.	RM	CM	SF	BFM	Fergus *et al.* 03	BarHillel *et al.* 05	Zhang *et al.* 05
Cars-rear	1.7	0.5	0.5	0.5	9.7	2.3	-
Airplane	10.8	2.9	7.1	2.6	7.0	10.3	5.6
Motorbikes	0.0	0.0	2.3	3.2	6.7	6.7	5.0
Faces	0.7	0.3	0.7	1.9	3.6	7.9	0.3

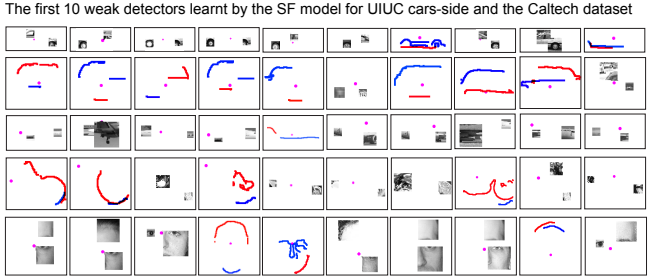
Detection examples



Detection examples with CM Showing both Hough Spaces



The first 10 weak detectors learnt by the SF model for UIUC cars-side and the Caltech dataset



This work was supported by the Austrian Science Foundation FWF, project S9103-N04, ECVision and Pascal Network of Excellence