Discriminative Training for Object Recognition Using Image Patches

Thomas Deselaers – Daniel Keysers – Hermann Ney
deselaers@cs.rwth-aachen.de

PASCAL Workshop – VOC– 11. April 2005

Human Language Technology and Pattern Recognition
Lehrstuhl für Informatik VI
Computer Science Department
RWTH Aachen University, Germany

Introduction

object recognition in cluttered scenes:



question: **What objects are contained in an image?**

several approaches known, active field of research

Idea

promising approach:

- **objects consist of parts which are modelled independently**

advantages:

- **changes in geometrical relationship can be modelled**
- **partly occlusion can be handled**
- **maybe: objects share parts**
  learning instances of an object might improve performance for other objects

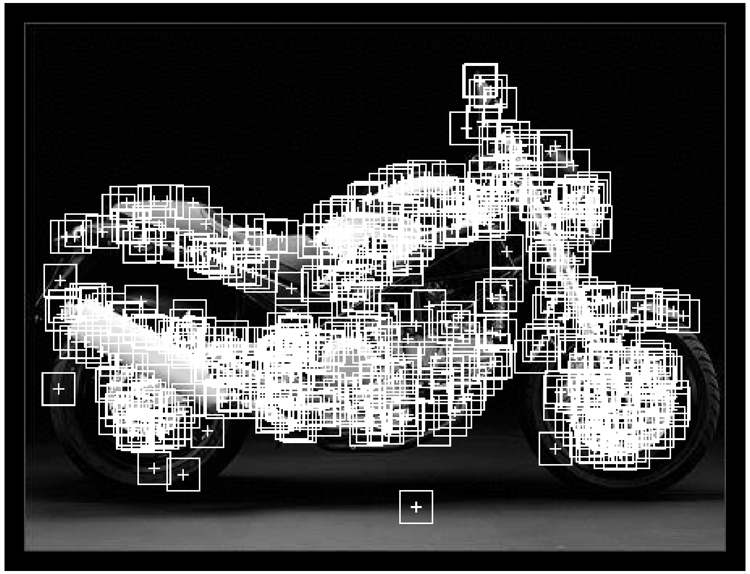
Proposed Approach

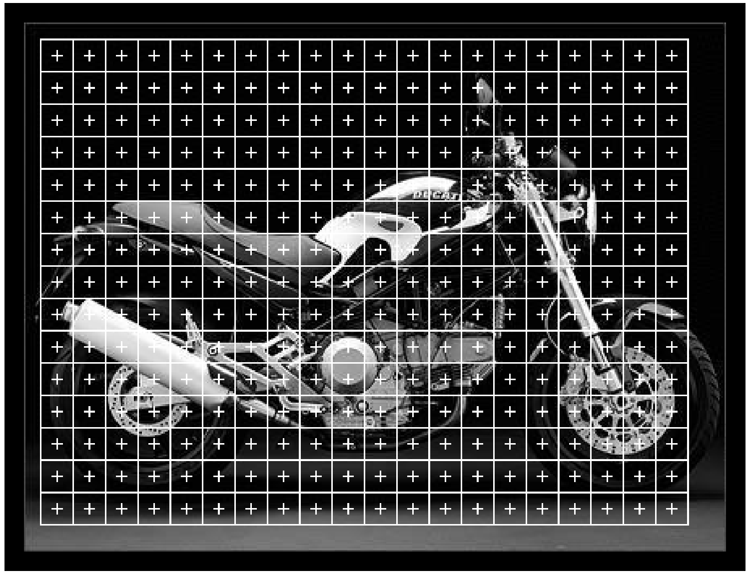
4 steps:

- **salient point detection and patch extraction**
- **PCA transformation**
- **clustering and histogramization**
- **discriminative training (maximum entropy approach)**

Feature Extraction

use feature extraction points from interest detector and from a regular grid:





at each of these points extract patches of various sizes

scale the extracted patches to a common size

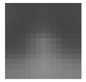
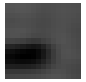
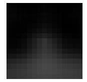
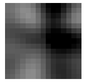
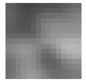
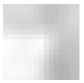
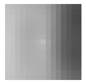
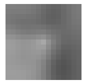
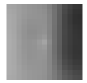
salient point detector from [Loupias & Sebe$^+$ 00]

PCA transformation

the extracted patches from all training images are pca transformed keeping 40 coefficients

PCA transformations offers possibility for brightness normalization:

– first PCA coefficient encodes overall image brightness

	1st PCA comp.	w/ 1st PCA component		w/o 1st PCA component	
		bright	dark	bright	dark
airplanes					
faces					
motorbikes					

discarding first PCA coefficient yields invariance wrt. image brightness

Clustering

use Linde-Buzo-Gray Clustering

1. **start with an initial Gaussian density**
2. **split this density by slightly disturbing mean**
3. **reestimate until convergence is reached**
4. **repeat steps 2 & 3**

this method yields $2^{\#\text{splits}}$ clusters

Histogramization

discard all information for each patch except its closest cluster identifier

create histogram of patches for each image:

$$h_c(X) = \frac{1}{L_X} \sum_{l=1}^{L_X} \delta(c, c(x_l))$$

L_X	**number of patches for image** X
x_l	l**-th patch for image** X
$c(x_l)$	**closest cluster identifier for** x_l
$h_c(X)$	c**-th bin of histogram for image** X

Discriminative Training

training of log-linear model

- **take into account information of competing classes**
- **maximize class posterior probability**

$$\prod_{k=1}^{K} \prod_{n=1}^{N_k} p(k|X_{kn})$$

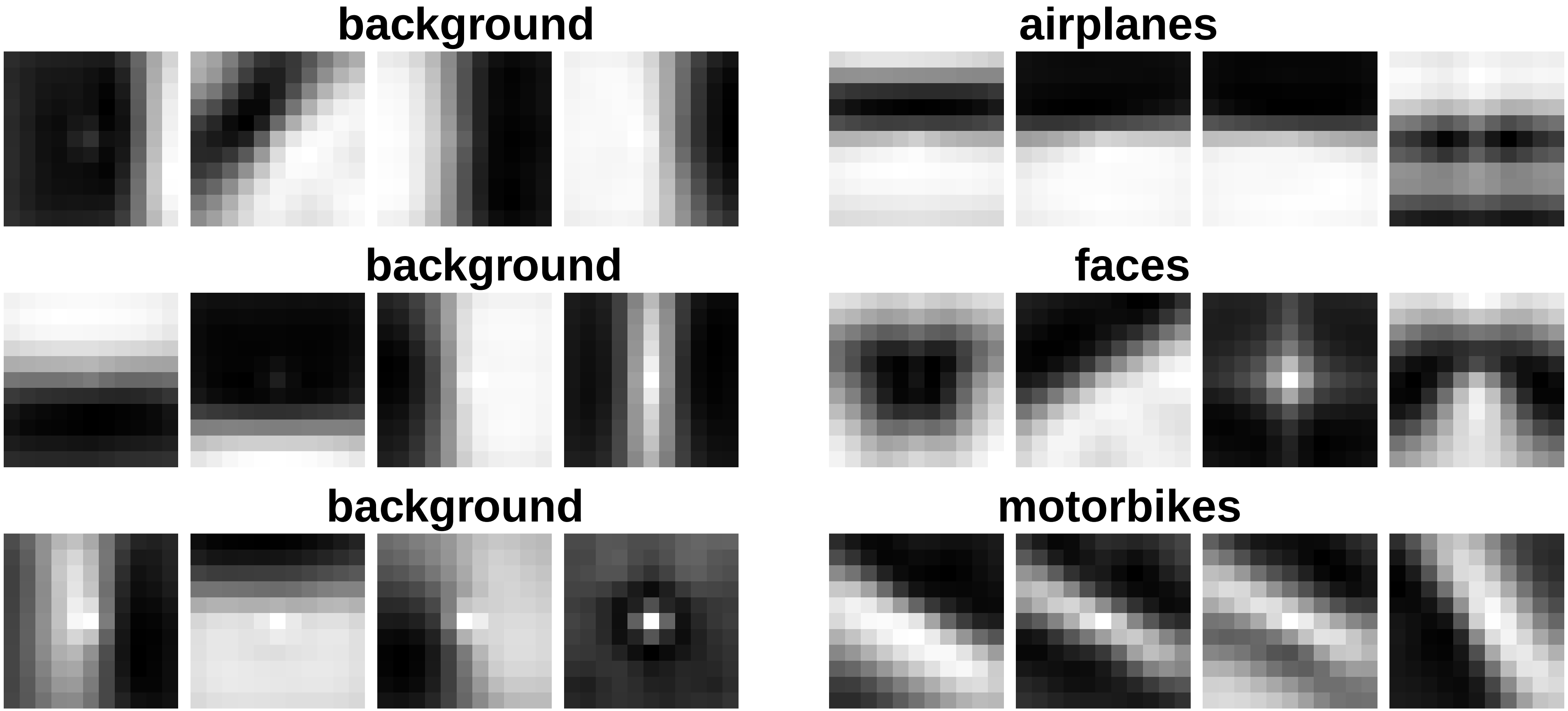
- **maximize parameters of log-linear / maximum entropy model:**

$$p(k|h) = \frac{1}{Z(h)} \exp\left(\alpha_k + \sum_{c=1}^{C} \lambda_{kc} h_c\right),$$

- **maximizing model is unique**
- **can be obtained using generalized iterative scaling**

Which features are discriminative?

(examples for Caltech database)



Results

Results on Caltech database: [Deselaers & Keysers$^+$ 05]
very good on motorbikes and airplanes, quite good on faces

Results on medical radiographs:[Deselaers & Keysers$^+$ 05]
quite good, specialized approaches are better

Ranks in PASCAL Visual Object Classes Challenge:

	motorbikes	bicycles	people	cars
Task 1	6,8 (17)	5,8 (15)	5,6 (15)	4,5 (17)
Task 2	2,3 (11)	2,3 (9)	2,3 (9)	2,3 (10)

Conclusion & Outlook

combination of 4 well understood steps yields good results:

1. **interest point detection and patch extraction**
2. **PCA transformation**
3. **clustering and histogramization**
4. **discriminative training**

approach generalizes well to other data

incorporation of spatial data for this approach not straight forward

visually meaningful parts are learned to be discriminative

SVM not necessary to obtain good results ;-)

SIFT not necessary to obtain good results ;-)

Thank You for Your Attention

References

[Deselaers & Keysers^{+} 05] T. Deselaers, D. Keysers, H. Ney. Discriminative Training for Object Recognition using Image Patches. Proc. *CVPR 05*, in press, San Diego, CA, June 2005. 10

[Loupias & Sebe^{+} 00] E. Loupias, N. Sebe, S. Bres, J. Jolion. Wavelet-based Salient Points for Image Retrieval. Proc. *International Conference on Image Processing*, Vol. 2, pp. 518–521, Vancouver, Canada, Sep. 2000. 4