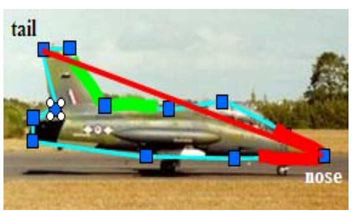
Discriminative Object Class Models of Appearance and Shape

Presented by
Priyanka Godbole

Points to cover:

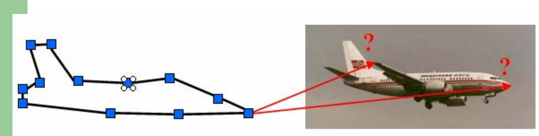
- Shape model
- Joint model of shape and appearance
- Correlograms vs Histograms
- Experimental results

Shape model

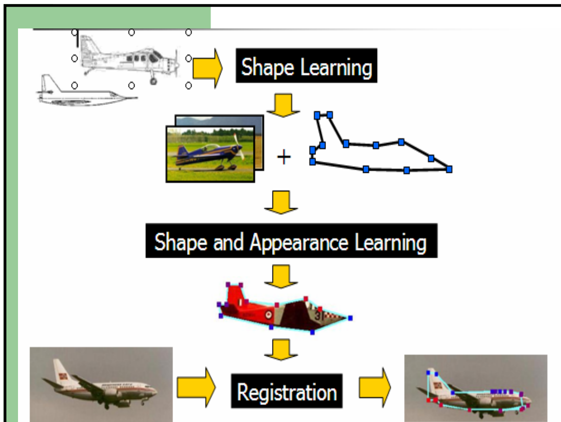


- Defined by set of landmark points

Shape + Appearance model

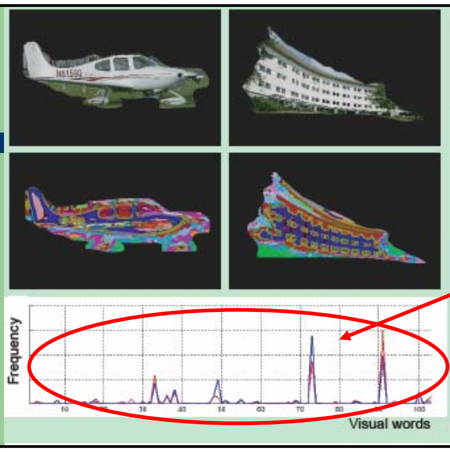


- Assigns landmark to each pixel



Concept of Histogram

- Only captures distribution of pixel labels
- Does not take in account the spatial relationship between pixels
- Liable to false positives (see next 2 slides)
- Not robust to large appearance changes



False Positives

Similar histogram

False positives

Similar histogram

Concept of Correlogram (color)

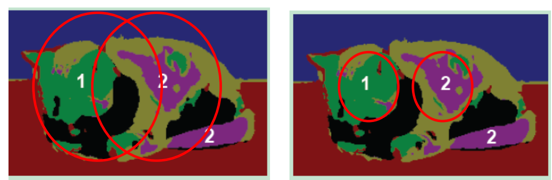
- Tolerates large changes in appearance and shape
- Table indexed by color pairs
  - K th entry for (i,j)
  - Probability of finding a pixel of color j at a distance k from pixel of color i in the image

Use of Circular Kernels

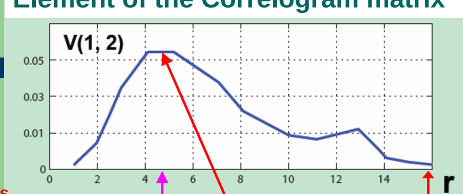
- Used to compute spatial correlation between pairs of visual words
- Correlation is computed as a function of kernel radius

At radius = distance between regions, kernel gives greatest correlation.

At all other radii, correlation is less.



Element of the Correlogram matrix



Y axis (Correlation = pixel match between 2 regions covered by kernel)

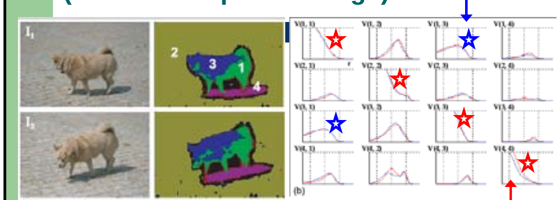
Distance between regions

Maximum Correlation (distance = radius of kernel)

Correlation decreases as difference between radius and distance increases

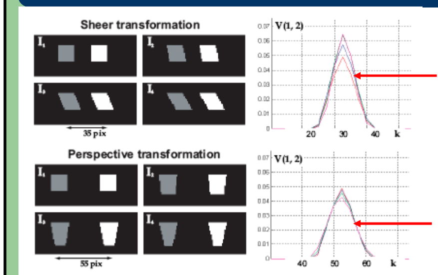
Correlogram Matrix (invariant to pose change)

Observation: Regions 1 and 3 are close by. So V(1,3)=V(3,1)



For similar regions, distance = 0. So the element peaks at r=0.

Invariance to Geometric transformations (affine / perspective)



Different amount of shear produce similar profiles of correlogram elements.

Perspective deformations produce similar profiles of correlogram elements.

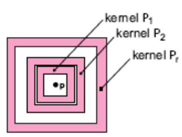
Properties of Correlograms

- Translational / rotational invariance
- Invariant to affine, perspective transformations
- Invariant to general object pose changes.

- Correlograms are not invariant with respect to scale.
- However, scale invariance can be learnt if multiple training images at multiple scales are available.

Use of Square Kernel

- Circular kernels induce rotational invariance.
- However, square kernels are computationally more efficient at partial expense of rotational invariance

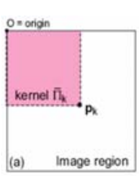


Integral Histogram

- Computational complexity of Correlogram = $O(N^4)$
- Computational complexity of Integral Histogram = $O(N^2)$

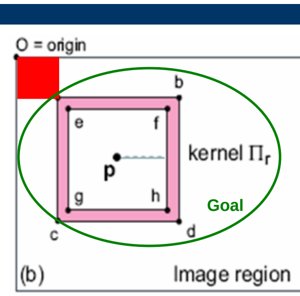
- Using integral histogram approach, Computational complexity of Correlogram = $O(N^3)$

Computational efficiency using square kernels (Integral Histogram approach)



Kernel (shown in pink) :
A rectangle defined by origin of the image region and pixel p^k

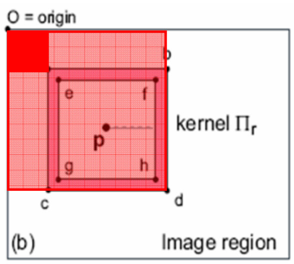
Computational efficiency using square kernels (Integral Histogram approach) contd..



Computing Integral Histogram for "a" =

$$H(\bar{\Pi}_a)$$

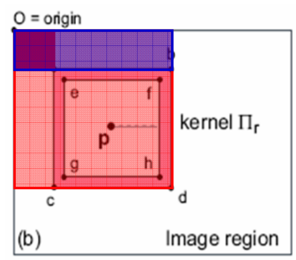
Computational efficiency using square kernels (Integral Histogram approach) contd..



Computing Integral Histogram for “d” =

$$\mathrm{H}(\bar{\Pi}_d)$$

Computational efficiency using square kernels (Integral Histogram approach) contd..

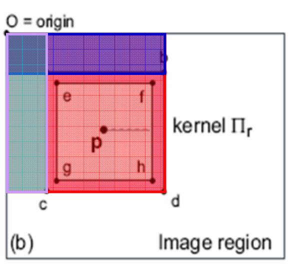


Computing Integral Histogram for “b” =

$$\mathrm{H}(\bar{\Pi}_b)$$

Need to subtract this

Computational efficiency using square kernels (Integral Histogram approach) contd..

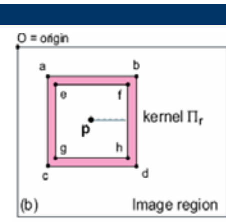


Computing Integral Histogram for “c” =

$$\mathrm{H}(\bar{\Pi}_c)$$

Need to subtract this

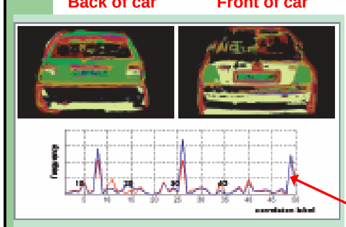
Computational efficiency using square kernels (Integral Histogram approach) contd..



$$\mathrm{H}(\Pi_r) = \mathrm{H}(\bar{\Pi}_a) + \mathrm{H}(\bar{\Pi}_d) - \mathrm{H}(\bar{\Pi}_b) - \mathrm{H}(\bar{\Pi}_c)$$
$$-(\mathrm{H}(\bar{\Pi}_e) + \mathrm{H}(\bar{\Pi}_h) - \mathrm{H}(\bar{\Pi}_f) - \mathrm{H}(\bar{\Pi}_g))$$

Intra-class variations

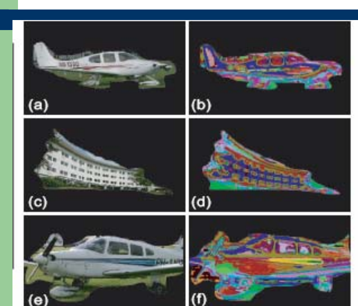
Back of car | Front of car



- Invariant enough to incorporate intra-class variations
- Discriminative enough to differentiate among different classes

Similar correlatons

Discriminative Object class models

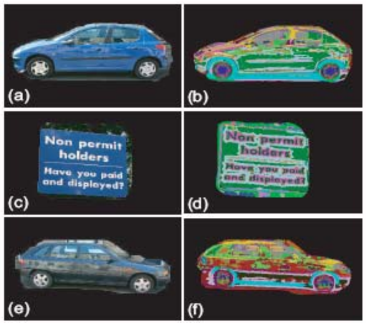


Query Image

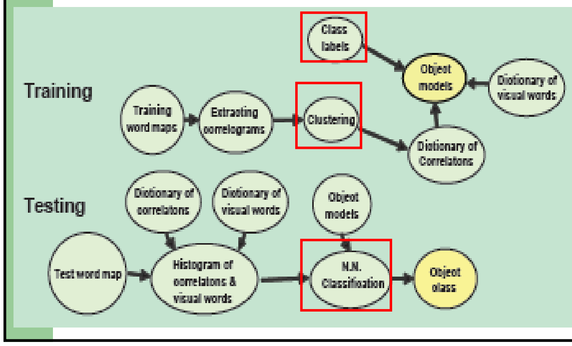
Using histogram of visual words

Using histogram of visual words + Correlations

Another example



Combining appearance and shape



Experiments

Reduction of vocabulary size using method defined in [2]

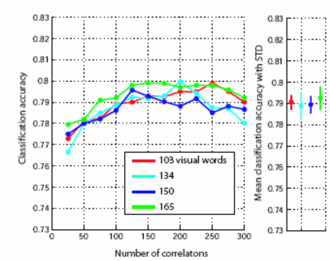
Original size of vocabulary

- Database with 15 classes

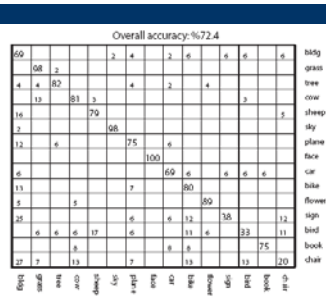
Model	Number of visual words			
	103 (200)	134 (400)	150 (600)	165 (1000)
V. words	72.4%	73.6%	74.6%	74.4%
Correl.	47.1%	37.3%	33.9%	36.9%
Joint	79.1%	78.9%	79.0%	79.3%
Joint +	80.0%	80.5%	81.1%	80.7%

Models learnt using the joint model approach

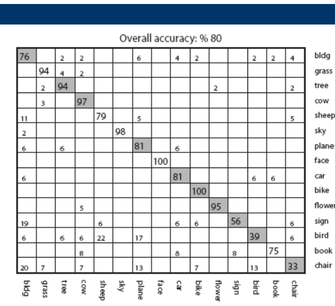
Classification Accuracy



Confusion Matrix: Word Size 103



Confusion Matrix: Word Size 150



Conclusions

- Introduced correlations to capture spatial correlation between visual words.
- Efficiency of computation achieved by integral image processing.
- Introduced object models, based on joint distributions of visual words and correlations.
- Proposed models are shown to outperform appearance-only based models.

References

- Discriminative Object Class Models of Appearance and Shape by Correlations (Savarese, Winn, Criminisi) [1]
- Object Categorization by Learned Universal Visual Dictionary (Winn, Criminisi, Minka) [2]
- Image Indexing Using Color Correlograms (Huang, Kumar, Mitra, Zhu, Zabih) [3]

A video demo for Object Classification (testing Correlatons approach)