Contour-based Object Detection

Joseph Schlecht schlecht@uni-heidelberg.de

Björn Ommer ommer@uni-heidelberg.de

University of Heidelberg

Introduction

Goal: Detect objects using a local contour representation of shape

Finding object contours exactly is difficult

- Binford, Clowes, Brooks, Lowe, Pentland, ...

Recent approaches focus on local appearance descriptors

- HoG, SURF, SIFT, Shape context, GB

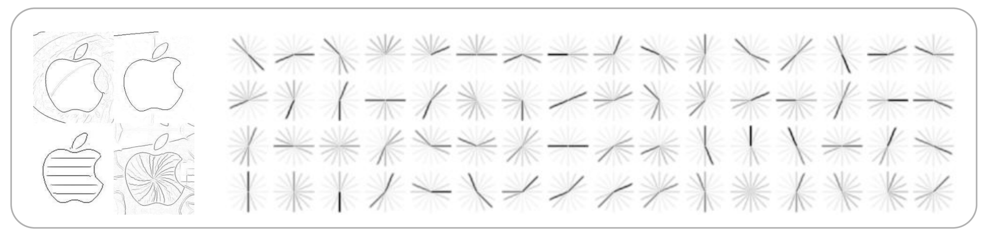
Contour curvature and junctions provide crucial shape information

- Attneave 1954, Biederman 1987

Investigate contour representation alongside appearance-based descriptors

Approach

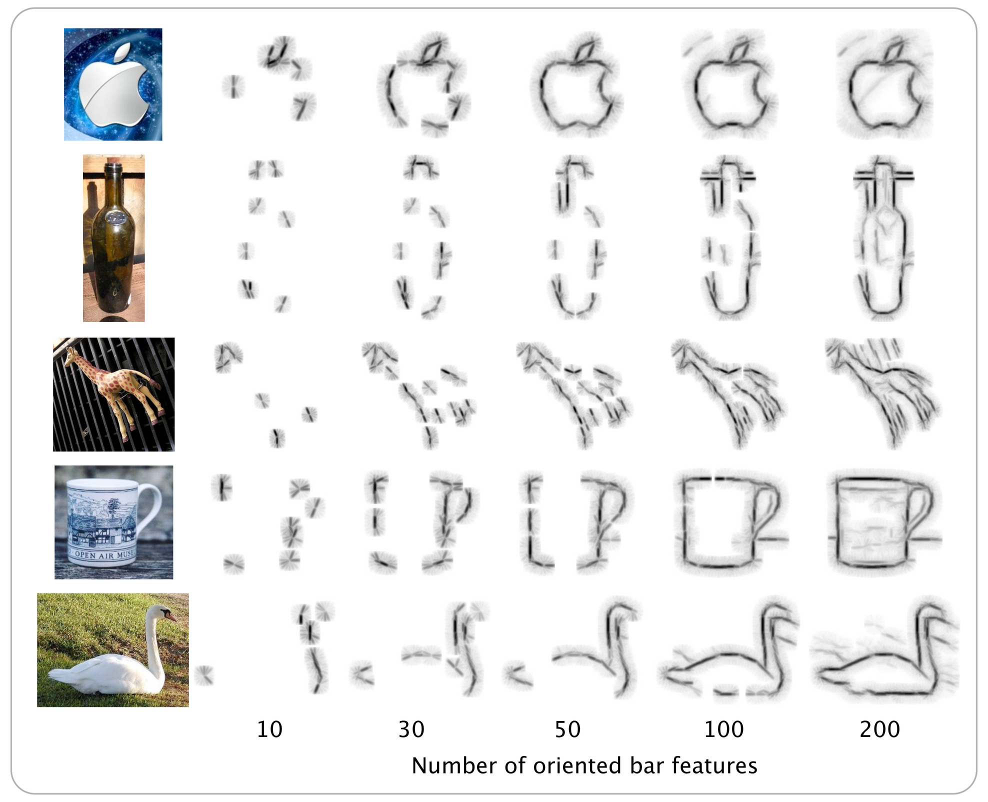
Represent local curvature and junctions with a distribution over oriented line segments at contour points



Explain object shape with a combination of repeatable local contour features

Shape cues concentrate at points of high contour curvature

- Order by intensity and curvature $\sum b_{ij}\, e^{-|\theta_i - \theta_j|}$

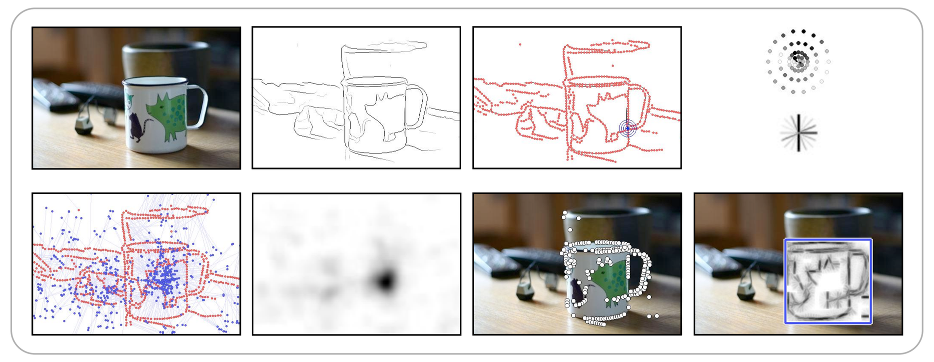


Combine local contour and texture features (GB) in a standard Hough voting framework for detection

Show that contour information complements appearance and significantly improves detection

Detection

Hough transform for object detection



Extract edge contours from images

Sample interest points uniformly along edges

$x_1, \ldots, x_N$

Compute oriented bar and geometric blur at interest points and concatenate

$$b_1, \ldots, b_N \qquad g_1, \ldots, g_N \quad \Longrightarrow \quad f_i = (b_i, g_i)$$

Match f_i to nearest neighbor $\hat{f}_i$ in training data to get object shift vector v_i

Cast votes for object position and scale

$$H(x, \sigma) = \sum_{i=1}^{N} w_i\, \delta(\|x_i + \sigma v_i - x\|)$$

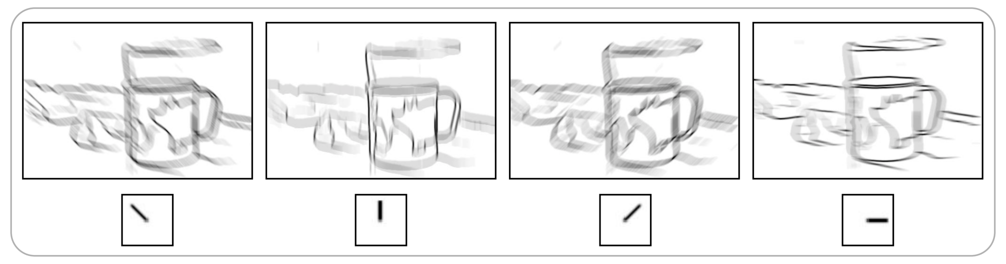
Estimate discrete density and extract strong hypotheses

$$h(x, \sigma) = c \sum_{\mathsf{x},\mathsf{s}} H(x - \mathsf{x}, \sigma - \mathsf{s})\, \eta(\mathsf{x}, \mathsf{s}; \omega_x, \omega_\sigma)$$

Representation

Contours

Describe local curvature and junctions with a non-parametric distribution over oriented bars at interest points



Create oriented bar filters F_d and convolve with edge map E

$$B_d(x) = \sum_{\mathsf{x}} E(x - \mathsf{x}) F_d(\mathsf{x})$$

Sample each channels at an interest point, normalize by magnitude

$$\hat{b}_i = B_1(x_i), \ldots, B_D(x_i)$$

Texture

Describe local texture near interest points with summary of edge signal under all affine transformations, i.e., geometric blur [1]

Descriptor centered at x is a convolution with spatially varying Gaussian kernel

$$G_x(y) = \sum_{\mathsf{x}} E(x + y - \mathsf{x})\, \eta(\mathsf{x}; \alpha\|x - y\| + \beta)$$

Sample along concentric circles about x_i and normalize by magnitude

$$\hat{g}_i = G_i(y_1), \ldots, G_i(y_C)$$

Results

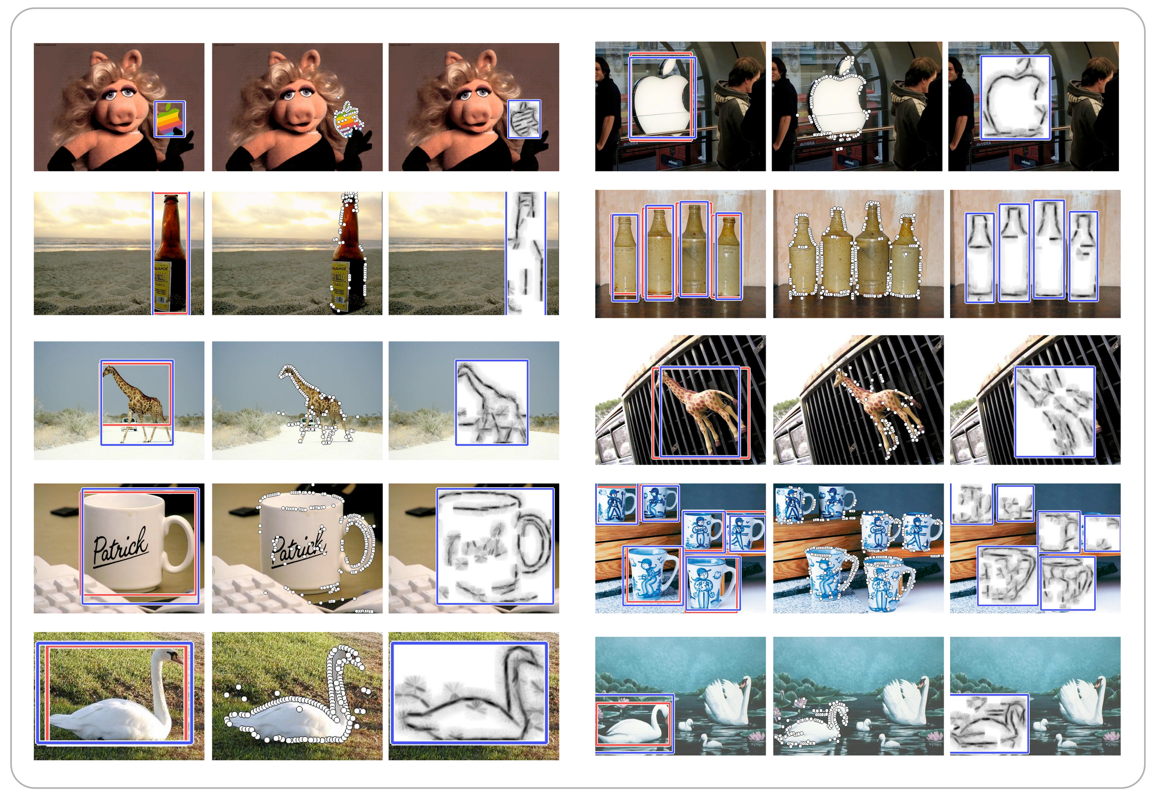
Evaluate approach in multi-scale object detection in cluttered scenes

- ETHZ shape dataset and INRIA horses

Average results over 10 random trials for each category

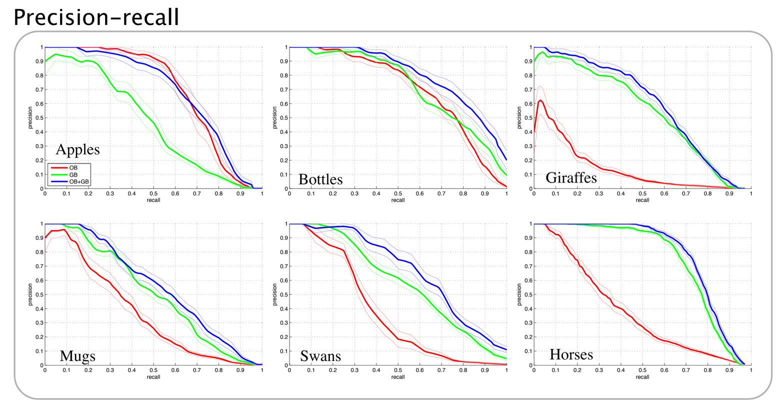
Compare against baselines of contour only (OB) and texture only (GB)

Qualitative results explain which contours support object hypothesis

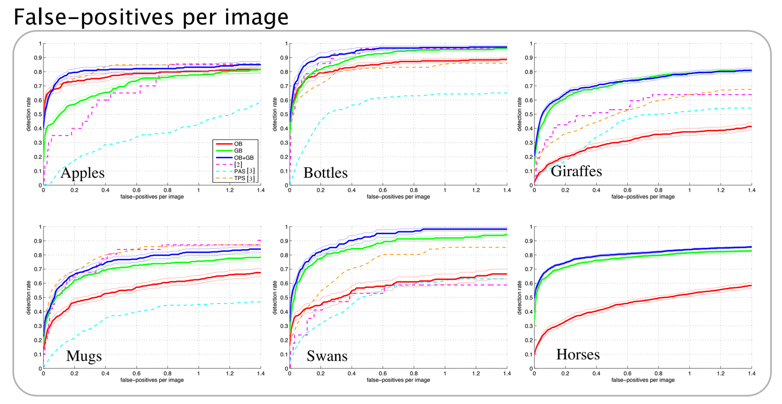


PR and FPPI results show significant gains in detection

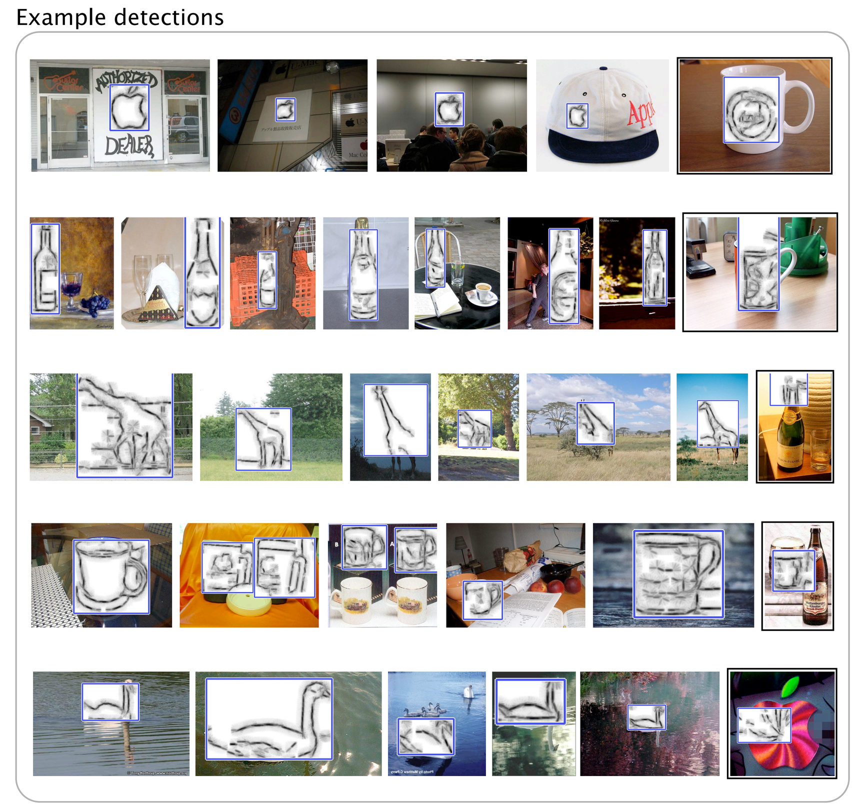
Precision-recall



False-positives per image



Example detections



Mean average precision

	OB	±	GB	±	OB+GB	±
Apples	67.4	2.6	45.5	2.9	67.0	2.9
Bottles	69.0	2.6	71.1	2.9	79.4	2.6
Giraffes	14.0	1.4	54.9	2.1	59.1	1.4
Mugs	35.7	2.2	50.2	2.0	55.3	2.3
Swans	34.5	2.1	56.6	1.7	65.2	2.4
Horses	37.6	2.0	74.4	0.9	78.1	0.7

Mean detection rate at 1.0 FPPI

	OB	±	GB	±	OB+GB	±	[2]	PAS[3]	TPS[3]	[4]*	[5]
Apples	76.1	2.5	65.5	2.7	81.4	2.5	60.0	28.9	83.2	85.0	75.0
Bottles	83.2	2.2	89.3	2.1	93.4	2.8	92.9	56.4	81.6	67.0	89.3
Giraffes	27.0	2.3	68.4	2.1	70.0	1.8	51.1	34.1	44.5	55.0	62.0
Mugs	52.8	2.5	69.2	2.2	74.6	2.6	77.4	35.5	80.0	55.0	65.0
Swans	54.8	3.5	84.5	2.3	90.2	3.4	52.9	44.9	70.5	42.5	53.0
Horses	40.4	2.1	77.4	2.4	79.2	0.8	76.9	45.0	68.0	52.0	–

References

[1] A. Berg and J. Malik, Geometric blur for template matching, CVPR, pp. 607–615, 2001.

[2] V. Ferrari et al., Groups of adjacent contour segments for object detection, PAMI, 30(1):36–51, 2008.

[3] V. Ferrari, F. Jurie, and C. Schmid, From images to shape models for object detection, IJCV, pp. 1–20, 2009.

[4] S. Maji and J. Malik, Object detection using max-margin Hough transform, CVPR, pp. 1038–1045, 2009.

[5] B. Ommer and J. Malik, Multi-scale object detection by clustering lines, ICCV, pp. 484–491, 2009.