Discriminative Learning of Contour Fragments for Object Detection

Peter Kontschieder
kontschieder@icg.tugraz.at

Hayko Riemenschneider
hayko@icg.tugraz.at

Michael Donoser
donoser@icg.tugraz.at

Horst Bischof
bischof@icg.tugraz.at

Institute for Computer Graphics and Vision
Graz University of Technology
Austria

In this work we propose a simple and effective solution for contour-based object detection. We show that shape fragments of fixed length, when discriminatively learned in combination with their spatial location on the object contour (see Figure 1), are sufficient to effectively localize previously unseen category instances in test images. We first introduce a novel shape fragment descriptor, abstracting spatially connected edge points into a matrix consisting of angular relations between points. These descriptors are discriminatively learned in a Hough Forest [1] framework to be able to localize instances in a generalized hough voting manner by only considering shape as a cue. We finally verify and rank the obtained center hypotheses similar to related methods [3]. In experiments conducted on the ETHZ shape database, we obtain an average detection rate of 87.5% at 1.0 FPPI only from Hough voting, outperforming the currently highest scoring methods by almost 8%.

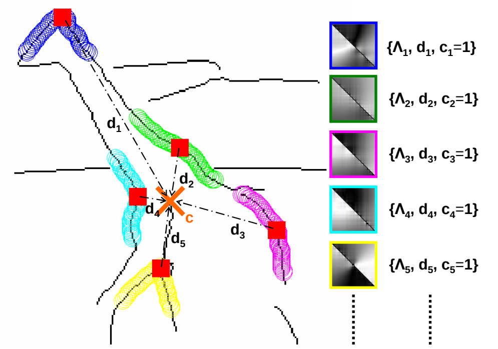

Figure 1: Illustration of our object detection method. For training, a local edge fragment i (color-coded) is abstracted in a triple consisting of our novel fragment descriptor matrix Λ_i, its corresponding center voting vector $\mathbf{d}_i$ and its class label $\mathbf{c}_i$. During testing, descriptor matches vote for object centroid $\mathbf{c}$ to hypothesize for object locations in the image.

Core contribution In order to discriminatively learn the local shape of object categories, we designed a novel descriptor based on angles between chords connecting sampled points of local edge fragments. In particular, our descriptor fulfills a series of important properties concerning distinctiveness, efficacy, invariance and computational efficiency, where in contrast existing shape descriptors like shape context fall short as demonstrated in our conducted experiments in Section 3 of the paper.

Our descriptor is based on contours obtained from the Berkeley edge detector [2] which we denote as *edges* after linking. Each edge is divided into severely overlapping parts, so-called *fragments*, of the same length N. Given a fragment, we can define a descriptor matrix as a function $\Lambda : N \times N \to [0, \pi]$ where $(\alpha_{ii}) = \Lambda(i,i) = 0$ and off-diagonal entries are defined as angles between lines connecting points $\overline{\mathbf{b}_i\mathbf{b}_j}$ and $\overline{\mathbf{b}_j\mathbf{p}_0}$, where $\mathbf{b}_i$ and $\mathbf{b}_j$ are two different sampled points on the currently analyzed fragment and $\mathbf{p}_0$ is a reference point fixed relatively to the fragment. Thus, an entry of our descriptor matrix is defined as

$$\alpha_{ij} = \measuredangle(\overline{\mathbf{b}_i\mathbf{b}_j}, \overline{\mathbf{b}_j\mathbf{p}_0}) \quad \forall i,j = 1,\ldots,N \text{ and } i \neq j.$$

The reference point $\mathbf{p}_0$ ensures encoding of the orientation of the considered fragments and is defined by the top- and leftmost coordinates of its enclosing bounding box (and an offset for numerical stability reasons). For the training process, we additionally provide the offset vector pointing from the fragment center to the object center. This process is repeated for all fragments that can be extracted from the obtained edges.

Results at a glance We provide a performance comparison of our shape fragment descriptor to state-of-the-art as well as a quantitative evaluation of our method on the ETHZ shape and INRIA horses databases. The ETHZ shape database scores are listed in Table 1 for the respective stages of our method and exemplary results are shown in Figure 2. On the INRIA horses database we achieve a detection rate of 85.5% at 1.0 FPPI.

Evaluation Stage	Apples	Bottles	Giraffes	Mugs	Swans	Average
Voting (FPPI=1.0)	94.4	90.9	86.7	92.3	73.3	87.5
Ranking (FPPI=1.0)	100	95.5	93.3	88.5	93.3	94.2
Verif. (FPPI=0.3/0.4)	94.4/100	100/100	91.1/93.3	80.8/87.2	100/100	93.3/96.1

Table 1: Detection scores of our method for ETHZ shape database in %, using the PASCAL50 criterion.

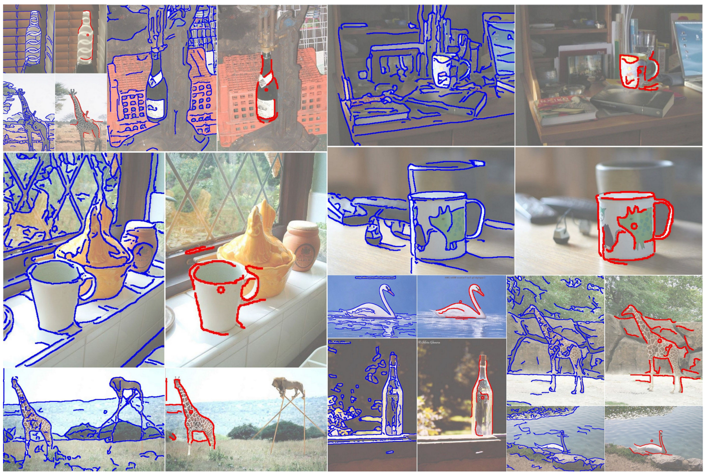

Figure 2: Exemplary results on ETHZ shape database. Extracted edges (blue) and reprojected fragments (red) voting for the centroid of the localized object are shown.

Conclusion We showed a simple way to tackle the contour-based object detection problem. Our novel method discriminatively learns a number of local contour fragment descriptors in combination with their spatial location relative to the object centroid in a Hough Forest classifier. We designed a fragment descriptor that abstracts spatially connected edge points into angular relations in a matrix form and demonstrated that our proposed descriptor shows distinctive patterns for differently shaped fragment primitives, while tolerating small perturbations and intra-class variabilities. In our conducted experiments we showed the superiority to state-of-the-art descriptors when trained in different machine learning frameworks and achieved promising scores on the ETHZ shape and INRIA horses benchmark databases.

[1] Jürgen Gall, Angela Yao, Nima Razavi, Luc Van Gool, and Victor Lempitsky. Hough forests for object detection, tracking, and action recognition. *(PAMI)*, to appear, 2011.

[2] David R. Martin, Charless C. Fowlkes, and Jitendra Malik. Learning to detect natural image boundaries using local brightness, color, and texture cues. *(PAMI)*, 2004.

[3] Hayko Riemenschneider, Michael Donoser, and Horst Bischof. Using partial edge contour matches for efficient object category localization. In *(ECCV)*, 2010.