Automatic Learning and Extraction of Multi-Local Features

Oscar Danielsson, Stefan Carlsson and Josephine Sullivan
School of Computer Science and Communications, Royal Inst. of Technology, Stockholm, Sweden
(osda02,stefanc,sullivan)@csc.kth.se

1. Motivation
What image properties are common to different examples of an object class?

- Discriminative local features are often not shared among the exemplars of the class
- Intra-class shape variation too large for efficient template-based representation

- Sparse, global shape features are shared among exemplars

2. Definition of Multi-Local Features

$\Theta^{(k)} = \{ (y_1, d_1), (y_2, d_2), \ldots, (y_m, d_m) \}$

Coordinates and orientations of edges (in Bookstein coordinates)
Parameters of bounding box

An image contains a given multi-local feature if there is a set of edges in sufficiently similar constellation.

Some examples of multi-local features learnt from images ...

3. Learning a Single Multi-Local Feature

Features are built incrementally from (weighted) positive training examples ...

1. Find the most common pair:
   $\hat{x} = \arg \max_{x} \{ \theta(x) \}$

2. Greedily add edges:
   $(\hat{x}, \hat{d}) = \arg \max_{x, d} \{ \theta(x, d) \}$

4. Learning an Ensemble of Multi-Local Features

An ensemble is learnt using a boosting scheme ...

1. Start with a uniform weighting over training examples.
2. Learn one multi-local feature as described above.
3. Down-weight training exemplars containing the new multi-local feature.
4. Repeat from step 2 until a sufficient number of multi-local features have been learnt.

5. Category Representation

The resulting features are shared by different subsets of exemplars and each exemplar contains several features:

Giraffe Features

6. Detecting Multi-Local Features

Feature detection is invariant to similarity transform ...

1. Input image.
2. Detect edges.
3. Extract edges.
4. Loop over all ordered pairs of edgels and check for potential correspondence to first two edges of feature.
5. Are the rest of the edges there?
6. If yes, signal detection.

7. Object Detection using Multi-Local Features

1. Each detected feature votes for a bounding box.
2. Votes are clustered.
3. Clusters are scored based on the number of features they contain.

8. Detection Results

Results were obtained using 5-fold cross-validation on the ETHZ Shape Classes database.

Some false positives ...

...