HIERARCHICAL SHAPE MODEL FOR OBJECT DETECTION

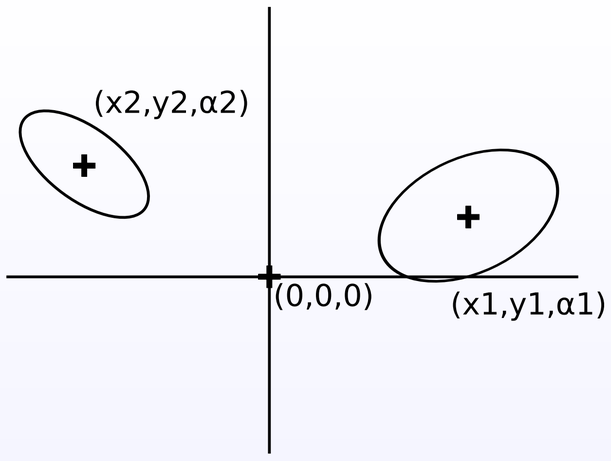
Mačák J., Drbohlav O. - Czech Technical University in Prague
{macakj1,drbohlav}@cmp.felk.cvut.cz

ICVSS 2011
Sicily ~ 11-16 July
International Computer Vision Summer School

Abstract

Hierarchical models are very suitable for detection and classification of objects from high number of classes because of their natural ability to model the exponential growth of the complexity of image scenes as the resolution and the number of objects in scene increases. The model is build in layers and each layer consists of a set of compositions that are frequent in the learning data. Each layer describes relations between compositions from the preceding layer. The lowest layer describes small areas of objects like line segments, arcs, etc. The highest layer can describe a whole object or a big part of an object. In this framework, compositions of all but especially lower layers can be easily shared within categories which leads to an effective description of a high number of classes. This approach is based on the work of Leonardis et al.

Compositions



Each component is assigned average location (x_i, y_i, α_i) and covariance matrix Σ_i describing its uncertainty.

Learning compositions

Most frequent and statistically significant compositions are learned

$$p(A, B) > p(A) \cdot p(B).$$

Layers are build incrementally from the lowest one to the highest one.
After learning a new layer, reprojection is done in order to locate all instances of compositions in the data.

Reprojection

Each component of the composition can vote for a location of the center of the composition. A composition is considered as present at a certain location, if a sufficient number of its components is found in their positions in the local neighbourhood. Mahalanobis distance

$$d = \sqrt{(\mathbf{x_l} - \mathbf{x_d})^T \Sigma^{-1} (\mathbf{x_l} - \mathbf{x_d})}$$

is used for determining distances.

Illustrative experimental results

The framework was used for windows detection using very small set of artificial hand-drawn window samples; six layers of compositions were learned.

Learning data

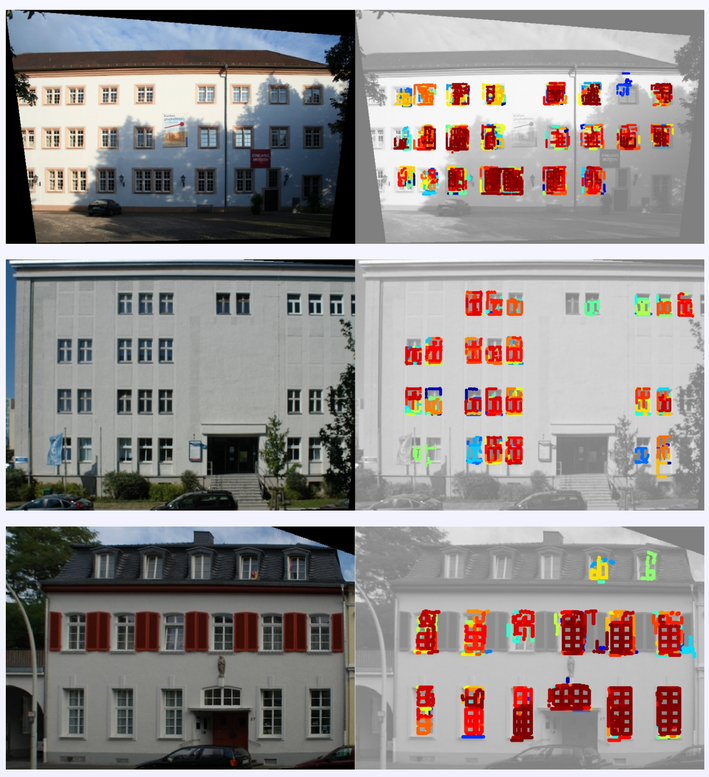
Compositions (5th layer)

Compositions (3th layer)

Compositions (6th layer)

Window detections

Rectified



Non rectified

References

[1] Sanja Fidler, Gregor Berginc, and Ales Leonardis. Hierarchical statistical learning of generic parts of object structure. In *Proc. CVPR*, pages 182–189, 2006.

[2] Sanja Fidler and Ales Leonardis. Towards scalable representations of object categories: Learning a hierarchy of parts. In *Proc. CVPR*, 2007.

[3] Jan Mačák and Ondřej Drbohlav. Hierarchical shape model for windows detection. In *ÖAGM/AAPR*, 2011.