Object Recognition using a Region Detector Based on Hierarchies of Partitions

Anonymous Author(s)
Affiliation
Address
email

Abstract

This article proposes the use of a novel region detector based on hierarchies of partitions, so-called Hierarchy-Based Salient Regions (HBSR), as a part of an approach for object recognition. HBSR enables to combine the clues given by a high quality contour detector with a custom salient region detection procedure. The evaluation of HBSR with a standard feature detection assessment framework shows that HBSR outperforms the state-of-the-art methods, in average. Finally, we obtain 94.17% of accuracy applying HBSR for object recognition.

1 Introduction

The extraction of local image features is a conventional approach for providing compact image descriptors that can be used to solve many computer vision tasks, like image stitching, tracking, reconstruction, image retrieval. Some examples of local features are edges, corners, ridges and blobs. The desirable qualities of image features (*e.g.* repeatability, distinctiveness, accuracy) (9) are tightly linked to the invariance properties of the detector (*e.g.* invariance to viewpoint, to luminosity, and to compression). Some of the best-known feature detectors are SIFT(5), SURF(1), ORB(8), MSER(6), Harris-Affine and Hessian-Affine(7). In this article, we present a local region detector based on hierarchies of partitions.

Existing feature detection methods based on hierarchies, like MSER (6), TBMR(10), or TOS-MSER (2), rely on component trees (min-tree, max-tree, and level-line tree) and thus on the study of the lightness of the image, seen as a topographical relief. Here, we propose to replace the use of component trees by hierarchies of partitions whose construction rely on the gradient of the image. Actually, this approach allows us to take advantage of machine learning based contour detectors to obtain a high-quality multiscale representation of the image from which we select salient nodes.

2 HBSR: Hierarchy-Based Salient Regions

Ideally, in a hierarchy of partitions of an image, the scene is iteratively refined in its objects, parts of the objects, parts of the parts, and so on. Thus, each region (also called node) of the hierarchy should represent a *salient* element of the scene. However, in practice, hierarchical representations are not perfect and generally contain artifacts (regions that do not correspond to any meaningful element of the scene) and redundancy (several nodes representing the same region with slight variations). The proposed method aims at selecting nodes from a hierarchy of partitions of an image by determining the *salient nodes* of the hierarchy and then filtering redundancy among them (see Fig. 1). Finally, each selected node of the hierarchy is represented by its best fitting ellipse.

Table 1 shows the results of repeatability and matching scores. We can see that our method obtains the best average score, with a repeatability very close to the best method and with a matching score significantly higher than all other methods.

Submitted to 33rd Conference on Neural Information Processing Systems (NeurIPS 2019). Do not distribute.

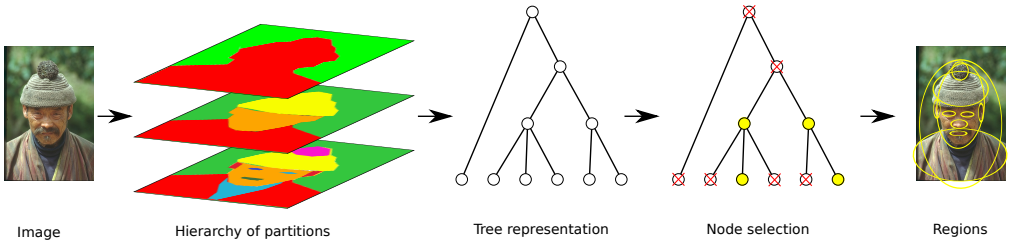

Figure 1: Main steps of the proposed region detector HBSR.

Table 1: Scores for HBSR detector and some detectors of the state-of-the-art

Measure	Harris	Hessian	MSER	TBMR	HBSR
Repetability	54.74%	64.22%	55.46%	51.96%	61.32%
Matching score	28.49%	35.03	39.27%	32.24%	53.26%
Average score	41.62%	49.62%	47.36%	42.10%	**57.29%**

3 HBSR for Object Recognition

These promising results may lead to improvements in many computer vision tasks, for example, object recognition. We are proposing an approach for object recognition using HBSR. This approach consists of detecting representative regions using HBSR, then describe the regions using CNN (Convolutional Neural Network)(4) and finally classify using SVM (Support Vector Machine) (3).

4 Experiments & Results

In our first experiment, we have used: two classes of the PASCAL VOC dataset (car and dog) with 1200 samples per class, and k-fold validation ($k = 5$).

We obtain **94.17%** of accuracy as preliminary results. Figure 2 shows an example of regions detected by HBSR method in an image of the chosen dataset.

Figure 2: Salient regions from HBSR method in an image from Pascal VOC dataset.

Acknowledgments

Not for this version.

References

[1] Bay, H., Ess, A., Tuytelaars, T., Gool, L.V.: Speeded-up robust features (surf). CVIU **110**(3), 346–359 (2008)

[2] Bosilj, P., Kijak, E., Lefèvre, S.: Beyond MSER: Maximally Stable Regions using Tree of Shapes. In: BMVC. Swansea, United Kingdom (2015)

[3] Cortes, C., Vapnik, V.: Support-vector networks. Machine learning **20**(3), 273–297 (1995)

[4] Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. In: Advances in neural information processing systems. pp. 1097–1105 (2012)

[5] Lowe, D.G.: Object recognition from local scale-invariant features. In: ICCV. pp. 1150–1157 (1999)

[6] Matas, J., Chum, O., Urban, M., Pajdla, T.: Robust wide-baseline stereo from maximally stable extremal regions. IVC **22**(10), 761–767 (2004)

[7] Mikolajczyk, K., Schmid, C.: Scale & affine invariant interest point detectors. IJCV **60**(1), 63–86 (2004)

[8] Rublee, E., Rabaud, V., Konolige, K., Bradski, G.R.: ORB: an efficient alternative to SIFT or SURF. In: ICCV. pp. 2564–2571. IEEE Computer Society (2011)

[9] Tuytelaars, T., Mikolajczyk, K.: Local invariant feature detectors: A survey. Found. Trends Comput. Graph. Vis. **3**(3), 177–280 (2008)

[10] Xu, Y., Monasse, P., Géraud, T., Najman, L.: Tree-Based Morse Regions: A Topological Approach to Local Feature Detection. IEEE TIP **23**(12), 5612–5625 (2014)