ABSTRACT

KEYWORDS: Saliency; Generic Object Segmentation; Feature Rarity; Background Prior; Objectness.

Saliency is an important property of the human visual perception. Most often the focus of interest of the human eye gets attracted to a region or salient object appearing distinctly in the foreground of a scene. The ability to automatically segment the objects of interest an image is useful for many computer vision tasks such as, shape-based formulations, object recognition, indexing and retrieval, image-retargeting, object tracking in videos and so on. This pre-processing step helps to reduce the search space for feature-extraction, matching, enhancement, compression etc., reducing the computational time.

There are two primary approaches of such algorithms. First, there are methods that find any object of interest based on visual stimuli, without any prior knowledge about its category. These are known as bottom-up methods. Second class of methods find category specific objects which are known and learned a priori. The former methods attempts to find regions or objects in an image that are prominent and vividly stand out from the rest of the image. This is a subjective perceptual quality in human visual system that has the ability to select regions with important visual information from its bottom-up stimuli. This quality is known as visual saliency and the objects which stand out are considered salient. We concentrate on bottom-up methods of finding salient objects.

We present two different approaches for bottom-up salient object segmentation. The first one relies on basic perceptual cues alone. Whereas, the second one uses generic objectness features along with saliency to segment salient objects from complex natural scenes. In either of the approaches, we first segment the image into small homogeneous patches, known as superpixels. In the first method, we utilize the low-level perceptual cues such as, rarity of feature, center-bias, boundary prior, and mathematically model them to generate a probability map

depicting saliency in an unsupervised framework. Rarity of feature is computed by exploiting graph-based spectral feature rarity and the spatial compactness. Graph-based rarity is computed by finding the uniqueness of the spectral features of the Laplacian of the graph over superpixels. Spatial compactness is obtained using distribution of similar colors over the image. Boundary prior is obtained by statistically modeling (in color space) the set of superpixels near to the boundary of an image. Our method produces a full resolution saliency map where each pixel is assigned a probability value of being salient.

The second formulation addresses the complex issues in many natural images where the object cannot be segmented using the low-level perceptual cues alone. Natural images exhibit spatial interactions, e.g., neighboring superpixels are likely to belong to the same object unless disarticulated by image edges, spatially bounded superpixels generally together represent an object part. These dependencies can be captured by a graphical model approach which again helps in good spatial propagation of saliency and different prior information. Hence, to solve this problem of generic object segmentation, we construct a conditional random field (CRF) over superpixels. Since, saliency alone is not sufficient, we exploit saliency in conjunction with different objectness criteria and appearance features, to formulate an energy function. In our algorithm. the edge-cost produces a sub-modular CRF. Thus, we perform an exact inference using graph cut. The CRF parameters are learnt by formulating a max-margin optimization. As the energy cost is a linear function of its parameters, we take a primal-dual approach to estimate the CRF parameters. Hence, both learning and inference are done efficiently so as to work on practical purpose.

The performance of both the propose methods are shown using illustrations and Precision-Recall-Fmeasure metric is used to compare the same with recent state-of-the-art techniques using challenging benchmark real world datasets.