Description of Objectives and Results of Paper on:

Visual Saliency Estimation through Manifold Learning

(by Richard M. Jiang & Danny Crookes)

Salman Mohammad
10630

Visual salience (or **visual saliency**) is the distinct subjective perceptual quality which makes some items in the world stand out from their neighbors and immediately grab our attention. This paper presents a **better algorithm for detecting salient visual features in complex images** than the previously available methods.

OBJECTIVES/MOTIVATION:

- Most early work makes more effort to build saliency models on low-level image features based on local contrast. These methods investigate the rarity of image regions with respect to (small) local neighborhoods.
- Recent efforts have been made toward global contrast based saliency estimation, where saliency of an image region is evaluated at the global scale with respect to the entire image.
- From previous research it is known that producing a correct salient map is sensitive to the size of regions, and a manual fine tuning of the segmentation is a prerequisite for some challenging images.
- In summary, it has been observed that most global approaches depend on either regions from image segmentation, or blocks with specified sizes.
- While image structures at global scale are usually an important factor for producing salient stimuli, this paper presents a novel approach for saliency modeling(that too manifold based saliency estimation)

Description of The New Algorithm:

In the proposed approach, an image is considered as a manifold of visual signals (stimuli) spreading over a connected grid, and local visual stimuli are compared against the global image variation to model the visual saliency. With this model, manifold learning is then applied to minimize the local variation while keeping the global contrast, and turns the RGB image into a multi channel image. After the projection through manifold learning, histogram based contrast is then computed for saliency modeling of all channels of the projected images, and mutual information is introduced to evaluate each single channel saliency map against prior knowledge to provide cues for the fusion of multiple channels. In the last step, the fusion procedure combines all single channel saliency maps according to their mutual information score, and generates the final saliency map. In our experiment, the proposed method is evaluated using one of the largest publicly available image datasets.

CONCLUSION:

Thus a robust manifold-based saliency estimation method has been proposed for robotic vision to capture the most salient objects in the observed scenes. An image is considered as a manifold of visual signals (stimuli) spreading over a connected grid, and projected into a multi-channel format through manifold learning. Histogram-based saliency estimation is then applied to extract the saliency map for each single channel, respectively, and a fusion scheme based on mutual information is introduced to combine all single-channel saliency maps together according to their mutual information score.

CONTRIBUTION:

The proposed method is evaluated using a well-known large image dataset. The experimental results validated that our algorithm attained the best prevision and recall rates among several state-of-art saliency detection methods.

Furthermore, the proposed method has been demonstrated on a test video to show its potential use for robotic applications, such as tracking a moving target in an arbitrary scene while the camera is moving with a robot. The experimental results show that the proposed approach can successfully accomplish this sort of tasks, revealing its potential use for similar robotic applications.