Volume : 2 | Issue : 11 | Nov 2013 • ISSN No 2277 - 8160

Research Paper | Engineering

Review paper of Visual Attention and Saliency Detection Methods

Geet kiran Kaur	Chandigarh Engg. College, Landran, Mohali
Parvinder Kaur	SUSCET Tangori
Yasmeen Kaur	Chandigarh Engg. College, Landran, Mohali

ABSTRACT

Computer vision is the science of extracting information from images This paper is an attempt to present basic idea of Saliency detection methods. Saliency is the word derived from "Salient" which means to stand out. Human perception is complex, thorough and in many cases highly variable and hence difficult to model for computers to replicate. The working of the human vision system is such that it shifts the attention from one salient object in the scene to another. It is fascinating and challenging jobs for the researcher to simulate the working of the Human visual System. This paper highlights some of the work done in the field of Visual Attention using saliency detection methods.

KEYWORDS :

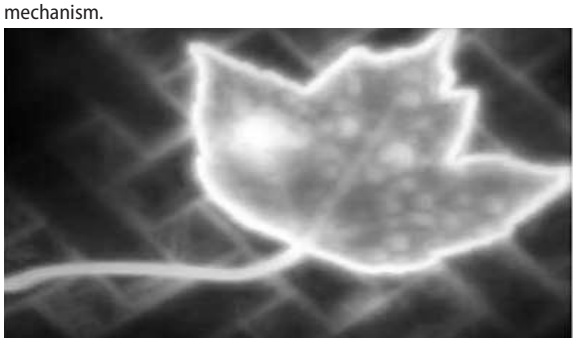
INTRODUCTION

Human perception can be thought of as responses to visual cues presented in the image. In other words, there are regions in an image that attract visual attention Selective attention is what the evolution has favored for the human need to deal with high amount of sensory data every moment. The Data is in range of 10^8 bits per second in general to high to be processed completely at one go and, so the possible action at one and the same time that is parallel processing is restricted, so the brain has to prioritize, Many modern technical system also faces the same problem .

SALIENCY

Saliency is something that's "you know it when you see it" .Human are capable of detecting, what pops-out of the imaging data with respect to the background. When we look in the night sky the moon appears Salient that is inspite of many other stars in the sky our attention goes to moon first, likewise red rose amidst other flowers is salient. Now the question arises why some parts of the scene are more salient then others, or in other words we can say that what guides our attention.

The answer comes from neuroscience, its states that brain works serially and not in parallel. This is due to the fact that the amount of information that our optic nerve receives, is something in order of 10^8 bits per second that far exceeds our brain processing capacity. So our brain processes selective information from the environment. The information that is registered by our visual system and affects our behavior appears as salient .We can say when we watch a scene or a image, our attention shifts from one salient object to another. Depending on their saliency quotient, saliency detection is the process of detecting the interesting visual information in an image It is fascinating to know that human visual system can adapt to wide range of environmental changes automatically and extract only useful information needed from the complex scene quickly and this has led to a wide spread view that such an efficient information processing capability relates a lot to human attention mechanism.



SALENCY DETECTION METHODS

Saliency detection is field where we try to find out where the human visual system will view in the freely viewing environment .Saliency map is the output of the saliency detection algorithms that depicts the salient regions of the image In this field saliency has been relevant for feature detection, foreground prediction and compression as costs of these methods are very high and have problem in stability also it is quite helpful in application like Image retargeting ,Collage making to name some. Saliency estimation methods have been classified in many ways some of the classifications are given below.

3.1 CLASSIFICATION BASED ON TECHIQUES

Depending upon the techniques used, Saliency Detection can be classified as biologically based, purely computational, or a combination. The goal of any method, in general, is to detect properties of contrast, rarity, or unpredictability in images, of a central region with its surroundings, either locally or globally, using one or more low-level features like color, intensity, and orientation. Biological model based methods attempt to mimic known models of the visual system for detecting saliency.

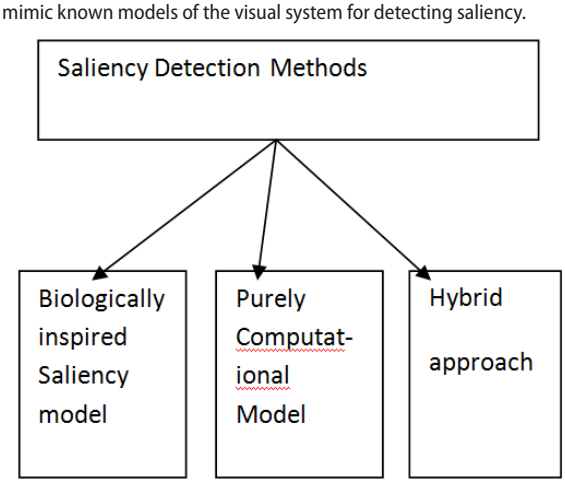


Fig.3: Classification of Computational Models

3.1.1 Biologically Inspired Saliency detection Methods

Most computational models of attention have a structure that is adapted from the first algorithmic model of attention by Itti and Koch[1]. The main idea is that several features are computed in parallel and then their values are fuse in a representation which is called a saliency map. The models generally include the following steps: First, the computation of one or several image pyramids from the input image takes place to enable the computation of features at different scales. Then, image features are computation takes place. Commonly used features include

intensity, color, and orientation, depth. Every feature channel is subdivided into several feature types (for example, r, g, b maps for color). Differences of Gaussians or center surround mechanisms are used to collect within map contrast into feature maps. The feature maps are summed up to forms feature dependent maps called conspicuity maps. Finally, the conspicuity maps are normalized, weighted and combined together to form the saliency map. The saliency map is usually visualized as gray-scale image in which the brightness of a pixel is proportional to its saliency

3.1.3 Hybrid approaches

The third category of methods is those that incorporate ideas that are partly based on biological models and partly on computational ones. For instance, Harel al. [2].Creates feature maps using Itti's [1] method but perform their normalization using a graph-based approach. Bian and Zhang [3] explain the biological plausibility of their spectral domain saliency detection method. Similarly, for the techniques of Torralba et al. and Zhang et al. [4] which rely on global scene context and natural image statistics, respectively, the authors allude towards the biological plausibility of their models.

3.2 BASED ON THE ATTENTION ASPECTS

Base on the Attention aspect used saliency detection methods can be classified into bottom up saliency detection methods and Top down Saliency Detection Method

3.2.1 Bottom up Saliency Detection Methods

When we observe our visual environment, we do not perceive all its components as being equally interesting. Some objects automatically and effortlessly "pop out" from their surroundings, that is, they draw our visual attention, in a "bottom up" manner, towards them. In a first approximation, focal visual attention acts as a rapidly shiftable "spotlight," which allows only the selected information to reach higher levels of processing and representation

3.2.2 Top down Saliency Detection Methods

What happens once our attention is focused onto a restricted part of our visual field. There is mounting experimental evidence that attention is far more sophisticated than a simple feed-forward spatially-selective filtering process. Indeed, visual processing appears to be significantly different inside the attentional spotlight than outside. That is, in addition to its properties as a feed-forward information processing and transmission bottleneck, focal visual attention feeds back and locally modulates, in a "top down" manner, the visual processing and representation of selected objects. This experimental dataset places strict constraints on our model of early pattern vision. The model, however, is eventually able to reproduce the entire dataset while assuming plausible neurobiological components.

3.3 BASED ON THE CONTEXT CONSIDERATION

On the basic of the context used i.e. whether the method is taking in considering the salience of an region/object locally, globally or Both Local and Globally.

3.3.1 Local Consideration

Many visual attention approaches have been proposed for detecting a few key locations with maximum local saliency and employ biologically motivated low-level features [5], [6]. They are driven by low-level stimulus such as intensity, color, orientation, and texture. While these approaches manage to find a few fixation points, they do not aim at finding regions of visual attention. For example, the saliency results of the motorcycle picture in Fig.4 (b) illustrate why this is insufficient: The high local contrast includes all the transitions between the background stripes

3.3.2 Global Consideration

These methods are based on finding regions in the image which imply unique frequencies in the Fourier domain. Therefore, they are able to quickly locate visual "pop-outs" that can serve as candidates for salient objects. These methods are very fast, but since they are based on global considerations, they do not detect object boundaries accurately, as illustrated in Fig. 4(c)

3.3.3 Local-Global Consideration

The approach we propose unifies local and global saliency by measuring the similarity between each image patch and other image patches, both locally and globally. In **fig,** the resemblance between patches was also used as a cue for saliency estimation; however, there only local neighborhoods were compared. In Fig. 4 (d) an image region is considered salient if it cannot be explained by anything similar in other portions of the image. Since their goal was to detect salient regions, they search for ensembles of patches that are unique globally. The lack of local considerations leads to coarse detection results. A consideration leads to coarse detection results. An attempt to incorporate local and global features for the purpose of detecting a single salient object is proposed in [7], [8]. This is done by using three features: a center surround histogram, color spatial distribution features, and local multistage contrast features.



INPUT



LOCAL METHOD



GLOBAL METHOD



Fig.4: d) LOCAL GLOBAL METHOD

NEED OF SALIENCY

In computer vision, robotics, human-computer interaction, computer graphics and design, there is a strong interest in having models, inspired by the human visual system that selects the most relevant parts within a large amount of visual data. The objective of these models is to improve artificial vision systems by computing a numerical value of the likelihood of attending to, or the saliency of, every location in an image.

CONCLUSION

In this paper, we conclude that lot of saliency detection models exists which are classified in many ways like based on techniques, aspects,

and other consideration. But irrespective of techniques used, all the methods try to detect the salient object or region from the image based on the differential quotient of the object or region with respect to the surrounding. Since using visual saliency techniques can tremendously reduce the amount of data to be processed in image and videos and we can detect the regions of interest from highly clustered background, we could apply these techniques in object recognition problems.

REFERENCES

1. L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 20, no. 11, pp. 1254–1259, November 1998. | 2. J. Harel, C. Koch, and P. Perona, "Graph-based visual saliency," Advances in Neural Information Processing Systems, pp. 545–552, 2007. | 3. P. Bian and L. Zhang, "Biological plausibility of spectral domain approach for spatiotemporal visual saliency," Advances in Neuro-Information Processing, vol. 5506/2009, pp. 251–258, and 2009. | 4. A. Torralba, A. Oliva, M. S. Castelhano, and J. M. Henderson, "Contextual guidance of eye movements and attention in real-world scenes: the role of global features in object search," Psychological Review, vol. 113, pp. 766–786, 2006. | 5. R. Achanta, S. Hemami, F. Estrada, and S. Susstrunk, "Frequency- Tuned Salient Region Detection," Proc. IEEE Conf. Computer Vision and Pattern Recognition, pp. 1597-1604, 2009. |