Advances and Application of Visual Attention Mechanism

Li Gun, Xu Fei, Yu Lei, Zhang Liang

Department of Biomedical Engineering, School of Electronic Information Engineering, Xi'An Technological University, Xi'An, China

Email address:

zhangliang@xatu.edu.cn (Zhang Liang)

Abstract: Senses and cognition of humans are mainly done by the visual nervous system. Most of the information people absorb from the world all conducted by the visual system. Therefore, visual attention mechanism is very important for exploring the visual system. In this paper, several basic problems of visualization of cellular electrical activity and visual information processing in the central nervous system are reviewed; then, models of visual attention mechanism are systematically reviewed. Finally, application of the visual attention mechanism in medical image segmentation is discussed.

Keywords: Visual Attention, Object Extraction, Image Retrieval, Object Detection, Visual Nervous System

1. Introduction

People understand and feel the world through the visual organs, where the eye as a prism to refract light of the outside image, retinal vision is caused when the visual information refracted to the retina, so that the external material causes the feeling. Rhodopsin or rhizome conducts the treatment and conversion of these external optical signals and energy for the physiological signal. Some investigations have shown that perceived location and the representation of distance information can be interpreted by the attention modulations, but it is still obscure that what influence spatial attention on the underlying structure of visual space [1]. Visual attention is influenced by many factors when a scene is concerned, for instance, the number of objects [2], emotion [3], event-related potentials [4], textural properties [5] and so on. Theoretically methods had been conducted recently to explore the visual working mechanisms. Jifan Zhou et al. used sub-object elements to reveal that how visual working memory forms intact perceptual representations. They tent to think that the object fragments can be integrated into objects within visual working memory in a manner similar to that of visual perception [6]. Many application have been put forward both in theoretical and technological area just as visual tracking algorithm [7-8] color image compression algorithm[9], regions of interest extraction[10], content-based image retrieval [11], image registration[12], etc.

As a fundamental sense, vision is humans assess to interact with the outer environments. The development of accurate visual attention model should take for the direction and distance of an object from the perspective of biology. In human visual system, when the light stimulates ganglion cells, it will produce a large action voltage potential; its frequency is proportional to the degree of membrane depolarization, so that the potential can transfer to other neurons. Obviously, the visual information from the neurons to anther needs the visual system itself and the central nervous system and to deal with. Visual cortex is located in the inner sides of the cortex; it can distinguish the shape of the object, brightness and color. Hubel and Wiesel argued that visual cortical neurons are not excited by the light spot, they are only sensitive to linear light stimulation, and can sense the direction of the stimulus, the direction and velocity of the movement, and the spatial changes [13]. So, the eyes can detect the distance of the target. According to the different characteristics of visual field in cerebral cortical neurons, the neurons can be divided into simple, complex and super complex ones. Cortical neurons as simple neurons, they can identify specific directions of stimulation. The neurons that react to a particular orientation of the rod or edge of any part of the field are called more complex neurons. From this perspective, this paper will discuss the several mechanisms on the visual attention, and then study the visual selective attention mechanism and its application of in the medical image segmentation.

2. Models of Visual Attention Mechanism

Visual perception is very fast and its speed makes many

people to think that it is very easy to get the information. In fact, it is need to conduct complex calculations and analyses for visual system to see the world. It is very important for visual system to discriminate the relevant information from irrelevant, and the process is called as selective visual attention, in this process the brain could decide the preferable information to process firstly. Corrado Corradi-Dell'Acqua, et al., tended to think that the ability for us to select the information relevant is influenced by two factors, the one is internal settings, and the other is the salient features of external stimuli. Their results highlight the role of the inferior frontal junction and posterior parietal regions in integrating bottom-up and top-down attention control signals [14].

Neural processing model were used widely in the area of interpreting visual attention mechanism [15]. Roman Borisyuk et al. put forward a selective attention module based on the famous Hodgkin-Huxley neuron equation in 2009. The module is designed as a neuron system with star-like connections among central unit and peripheral elements. The module is regarded as a system that can automatically scan a visual scene, and select regions of interest sequentially, and then this information is analyzed for the following object segmentation. The object segmentation module is used for precise determination at the next stage as a network of phase oscillators of noise suppression [16]. Quoc Do, et al., emphasized on the question of visual landmark recognition to exploring the functions of the human visual system, put forward a theoretical model for recognizing visual landmarks in real time. The model combined the pre-attentive and attentive stages within human visual system. The results showed that the selective attention adaptive theory based on neural network and its derivatives are very useful in object recognition. The author concluded that the architecture could recognize objects in cluttered backgrounds, and its application in navigation was verified through many trials in different environments (both indoor and outdoor) [17].

All scenes have their own distinct characteristics. Selective visual attention responds for the scrutiny on the selective part of a scene. Attention may be focused on a particular object (object-based attention), which requires the recognition of a specific target. From this perspective, Alcides X. Benicasa, et al., consider early visual features combined with the object-based visual selection modulations to put forward a new visual selection model. The proposed model not only can handle the segmentation of the scene, but the average of saliency of each object can be computed too to provide visual attention for features. The results show that when a memorized object is selected, a stronger effect of the top-down modulation than bottom-up saliency [18]. In order to timely generate full-resolution saliency map, Jufeng Zhao, et al., tend to think that the optical computing is a practical and efficient method. They build a system to achieve fast visual saliency extraction to generate full resolution saliency maps based on optical information processing method. And the result shows that the system could produce full resolution saliency maps [19]. In 2005, T Nathan Mundhenk and Laurent Itti proposed a very famous computational model for detecting visual saliency. A biologically plausible device was used to simulate the elements aligned collinearly with a contour in an image. The authors tent to think that the model can be used to investigate its usage in different images, such as real and artificial ones. Then the model was found to be very useful for discovering processing mechanisms of complex visual features. The test results showed that the system is very good in human vision, and the visual cortex may be more efficient [20].

Tony Lindeberg tended to think that the visual field is consisted of receptive fields, and then the author presented a theory for an idealized vision system. The author also presents a general framework for relating and interpreting receptive fields possibly, which can predict new receptive field profiles. The system assumes structural requirements of visual processing in the first stages reflect symmetry properties of surrounding world. The results of the paper showed that the concerning receptive field profiles regarding the mathematical foundation outline how the proposed theory could be used in vision, and which way the receptive field responses be interpreted [21]. Duzhen Zhang, et al., established a general computational salient objects detection system beyond top-down and bottom-up mechanism. The system had 3 parts, the first is selection history, the second is current goal and the third is physical salience. In the system, the first part is an integrative framework composed of current goal and physical salience and an image window saliency in the system could be defined as the objectness score of the window. The author conclude that the main effect of selection history is to concentrate on salient objects, the current goal has strong effect to detect correct salient objects [22].

3. Application of Visual Attention Mechanism in Medical Image Segmentation

When the medical image is concerned, It requires an in-depth knowledge of visual system in terms of computational mechanisms [23]. The visual attention with the model could be interpreted from the perspective of the characteristics of the image: color, brightness and direction. In the feature graph, the value of each feature pixel can be obtained by Matlab. The characteristic of the various pixel values is very important for the following segmentation, from which a suitable pixel value p should be chosen. That is, it is need to pay attention to the visual attention area. Here a medical image (figure 1) as an example is used to interpret the usage of the visual attention mechanism on the problem of processing images.

For image of a significant area derived from an image, the significant area of the image may be relatively large and it is difficult to obtain the effective result. So it is necessary to solve the problem and to get a more detailed area map. Form this perspective; the problem by getting the boundary of the image is analyzed as seen in figure 2.
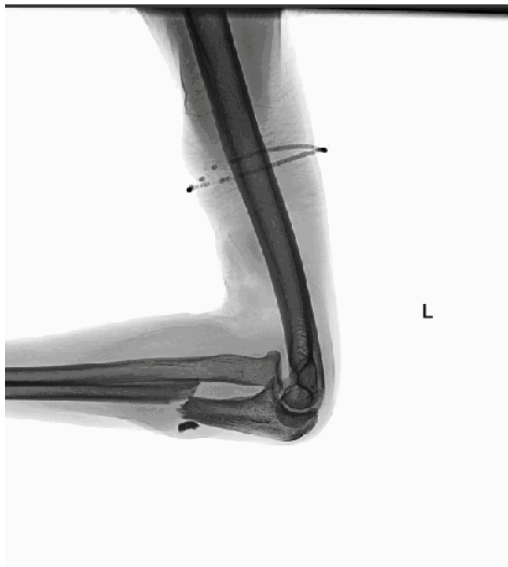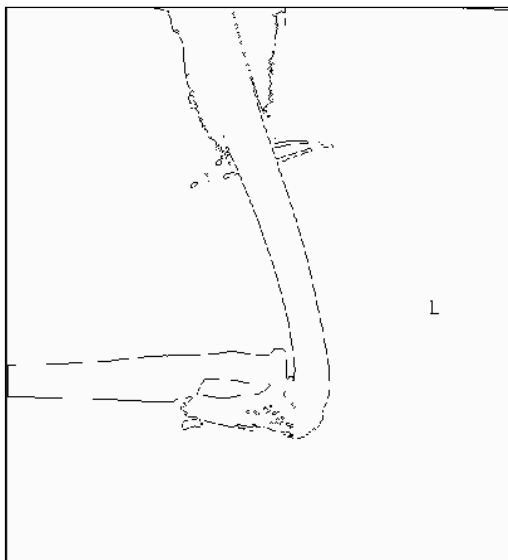
Figure 1. The original medical image.

Figure 2. Image boundary.

Figure 3. Significant area of image.

Figure 4. Identification results of the attention area.

Generally, the pixel region is obtained by using the pixel comparison method and the appropriate pixel values for a particular image is exactly where the attention need be pay to. Here, in order to get a relatively bright area of a medical image, it can be set a automatically query program to get the pixels and their specific location of a specific percentage range (figure 3).

In figure 4, the area of the fracture in the image and the area of the problem in the original image are separated. Three parts of them are the character 'L', the arm ring, and the fracture area. It is clear that the desired results are get in the paper, and the regional recognition result is most people may pay their attention to. Although for some specific images, this method works well to find the region of interest, manual analysis is required for medical diagnosis. It is difficult to use the visual selection attention mechanism to solve the difficult problems.
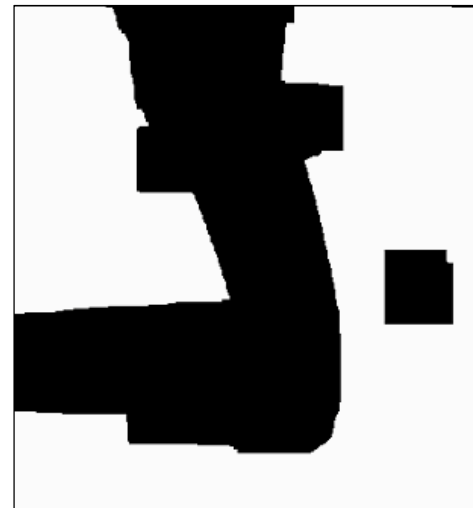
The existing visual significance detection model is mainly based on the theory of feature fusion. It is normalized into a global significant area by extracting the basic visual primaries such as color, brightness and so on, comparing them into a significant feature, generating a variety of significant target segmentation and other visual processing areas.

4. Conclusion

Due to the lack of understanding of the visual attention mechanism and the difference between the calculation method and the actual physiological structure, the existing visual attention model has many shortcomings, which cause it is not suitable for practical application. Finding the target of interest in the image automatically is the basic problem of computer vision, it attracted a lot of scholars' attention because of its extensive application. Based on the characteristics of human visual mechanism, this paper builds a model that can simulate the visual selective attention mechanism to segment the obvious goals in the medical image and the results show that it is can be applied in processing medical image based on their content.

Acknowledgements

References

[1] Francesca C. Fortenbaugh, Lynn C. Robertson, Michael Esterman. Changes in the distribution of sustained attention alter the perceived structure of visual space. Vision Research, 2017, 131: 26-36.

[2] Halely Balaban, Roy Luria. The number of objects determines visual working memory capacity allocation for complex items. NeuroImage, 2015, 119: 54-62.

[3] Helen F. Dodd, Julia Vogt, Nilgun Turkileri, et al. Task relevance of emotional information affects anxiety-linked attention bias in visual search, Biological Psychology, 2017, 122: 13-20.

[4] Sven-Thomas Graupner, Sebastian Pannasch, Boris M. Velichkovsky. Saccadic context indicates information processing within visual fixations: Evidence from event-related potentials and eye-movements analysis of the distractor effect, International Journal of Psychophysiology, 2011, 80: 54-62.

[5] Kenji Fujii, Shinofu Sugi, Yoichi Ando. Textural properties corresponding to visual perception based on the correlation mechanism in the visual system. Psychological Research, 2003, 67: 197-208.

[6] Jifan Zhou, Haihang Zhang, Xiaowei Ding, et al. Object formation in visual working memory: Evidence from object-based attention. Cognition, 2016, 154: 95-101.

[7] Wanyi Li, Peng Wang, Hong Qiao. Top-down visual attention integrated particle filter for robust object tracking. Signal Processing: Image Communication, 2016, 43: 28-41.

[8] Yayun Ren, Benlian Xu, Peiyi Zhu, Mingli Lu, et al. A multiCell visual tracking algorithm using multi-task particle swarm optimization for low-contrast image sequences. Applied Intelligence, 2016, 45(4): 1129-1147.

[9] Yao Juncai, Liu Guizhong. A novel color image compression algorithm using the human visual contrast sensitivity characteristics. Photonic Sensors, 2017, 7(1): 72-81.

[10] Huang Chaobing, Liu Quan, Yu Shengsheng. Regions of interest extraction from color image based on visual saliency. The Journal of Supercomputing, 2011, 58(1): 20-33.

[11] Qiang Zhou, Limin Ma, Mehmet Celenk, et al. Content-based image retrieval based on ROI detection and relevance feedback. Multimedia Tools and Applications, 2005, 27(2): 251-281.

[12] Feng Jing, Ma Long, Bi Fukun, et al. A coarse-to-fine image registration method based on visual attention model. Science China Information Sciences, 2014, 57(12): 1-10.

[13] DH Hubel, TN Wiesel. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. Journal of Physiology, 1962, 160(1): 106-154.

[14] Corrado Corradi-Dell'Acqua, Gereon R. Fink, Ralph Weidner. Selecting category specific visual information: Top-down and bottom-up control of object based attention. Consciousness and Cognition, 2015, 35: 330-341.

[15] Yuhua Zheng, Yan Meng, Yaochu Jin. Object recognition using a bio-inspired neuron model with bottom-up and top-down pathways. Neurocomputing, 2011, 74: 3158-3169.

[16] Roman Borisyuk, Yakov Kazanovich, David Chik, et al. A neural model of selective attention and object segmentation in the visual scene: An approach based on partial synchronization and star-like architecture of connections. Neural Networks, 2009, 22: 707-719.

[17] Quoc Do, Lakhmi Jain. Application of neural processing paradigm in visual landmark recognition and autonomous robot navigation. Neural Computing and Applications, 2010, 19(2): 237-254.

[18] Alcides X. Benicasa, Marcos G. Quiles, Thiago C. Silva, et al. An object-based visual selection framework. Neurocomputing, 2016, 180: 35-54.

[19] Jufeng Zhao, Xiumin Gao, Guang Lin, et al. An optical information processing-based idea for visual attention analysis. Optik, 2016, 127: 3556-3559.

[20] T Nathan Mundhenk, Laurent Itti. Computational modeling and exploration of contour integration for visual saliency. Biological Cybernetics, 2005, 93(3): 188-212.

[21] Tony Lindeberg. A computational theory of visual receptive fields. Biological Cybernetics, 2013, 107(6): 589-635.

[22] Duzhen Zhang, Chuancai Liu. A salient object detection framework beyond top-down and bottom-up mechanism. Biologically Inspired Cognitive Architectures, 2014, 9: 1-8.

[23] Anna Schubö, Hermann J. Müller. Selecting and ignoring salient objects within and across dimensions in visual search. Brain Research, 2009, 1283: 84-101.