Information Sciences 415–416 (2017) 1–18

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

A real-time image optimization strategy based on global saliency detection for artificial retinal prostheses

Heng Li[a], Tingting Han[a], Jing Wang[b], Zhuofan Lu[a], Xiaofei Cao[a], Yao Chen[a], Liming Li[a], Chuanqing Zhou[a], Xinyu Chai[a,*]

[a] *School of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China*
[b] *College of Information Technology, Shanghai Ocean University, Shanghai 201306, China*

ARTICLE INFO

Article history:
Received 21 November 2016
Revised 12 June 2017
Accepted 16 June 2017
Available online 17 June 2017

Keywords:
Retinal prostheses
Simulated prosthetic vision
Saliency detection
Eye-hand coordination

ABSTRACT

Current retinal prostheses can only generate low-resolution visual percepts constituted of inadequate phosphenes which are elicited by a limited number of stimulating electrodes and with unruly color and restricted grayscale. Fortunately, for most retinal prostheses, an external camera and a video processing unit are employed to be essential components, and allow image processing to improve visual perception for recipients. At present, there have been some studies that use a variety of sophisticated image processing algorithms to improve prosthetic vision perception. However, most of them cannot achieve real-time processing due to the complexity of the algorithms and the limitation of platform processing power. This greatly curbs the practical application of these algorithms on the retinal prostheses. In this study, we propose a real-time image processing strategy based on a novel bottom-up saliency detection algorithm, aiming to detect and enhance foreground objects in a scene. Results demonstrate by verification of conducting two eye-hand-coordination visual tasks that under simulated prosthetic vision, our proposed strategy has noticeable advantages in terms of accuracy, efficiency, and head motion range. The study aims to help develop image processing modules in retinal prostheses, and is hoped to provide more benefit towards prosthesis recipients.

1. Introduction

Retinal degenerative diseases, such as retinitis pigmentosa (RP) and age-related macular degeneration (AMD), are caused by progressive damage of the rod and cone photoreceptor cells. They can be expected to result in gradual loss of vision and eventually profound blindness, without any established therapies [45]. Retinal prostheses have proven to partially restore vision in patients suffering from advanced stages of above diseases by a surgical implanted electrode array to electric stimulate the inner neurons in the retina (e.g. bipolar and ganglion cells. Most of them are intact and remain functional) [47]. The elicited partial percepts are composed of a set of discrete spots of light called as “phosphenes”, which matches the stimulation pattern generated by encoding visual information into the implanted electrode array under ideal conditions [7].

Currently, there are several types of retinal implants that have been developed or under development, on the basis of their implanted anatomical location, such as the inner (epiretinal) surface of retina, the outer (subretinal) surface of retina, and the choroid plexus (suprachoroidal) [47]. Most retinal implant designs employ an external camera and an external video

* Corresponding author.
E-mail address: xychai@sjtu.edu.cn (X. Chai).

http://dx.doi.org/10.1016/j.ins.2017.06.014

processing unit (VPU) to drive the implanted electrode array. For example, the Argus II epiretinal implant (Second Sight Medical Products, Sylmar, California, USA) requires recipients to use head movements via the camera for inspecting scenes and permits the VPU to make image processing and individual electrode signal adjustments [39]. An exception is the Alpha IMS subretinal implant (Retina Implant AG, Reutlingen, Germany), which instead of an external camera uses a microphotodiode array dependent on natural eye movements for scene inspection and allows globally gain/offset adjustments [37]. Although there are some sensible distinctions between the two implant systems from a technological perspective, the patients either implanted with Argus II or Alpha IMS can obtain stable, retinotopically organized visual precepts. At present, both the Argus II (approved for implantation by the FDA in the USA and EMA in the European Community) and Alpha IMS (has EMA approval) have been successfully commercially available [11]. Furthermore, the number of implantable electrodes has also been developed to 60 in Argus II and 1500 in Alpha IMS compared with the early 16 electrodes [33].

Such encouraging development notwithstanding, the visual acuity reported by retinal implant recipients is still poor under the current number of the implantable electrodes. The latest clinical trials showed that the best visual acuity of patients implanted with the Argus II was 20/1262 and that the Alpha IMS could deliver a visual acuity of 20/546 in the best case, which are both still lower than limit for legal blindness (20/200) [47]. Although the recipients with such visual acuity had a significant improvement of activities of daily living and mobility by assessing variety of daily visual tasks, such as letter and large geometric shapes recognition, word reading, object localization, and outdoor movement detection [10,38], it's extremely difficult for them to perform more sophisticated visual tasks (e.g. object and face recognition, independent locomotion in unknown surroundings) [35]. This implies a need to increase the number of the implantable electrodes in future designs. Currently, there are some studies on high-density retinal implant designs, but they are in a stage of testing in animals or humans, awaiting further headway [21,28,29,40]. Moreover, recipients report that their visual features are mostly lost. For instance, colors of the phosphenes are elicited unruly, as well as the elicited brightness is generally confined within 10 levels [7]. Therefore, based on these different clinical trials, we can draw a conclusion that the prosthetic visual perception evoked by current retinal implants with the limited number of electrodes has the characteristics of low resolution and scarce visual features, causing recipients' poor visual understanding.

For most retinal implants, an external VPU is an important component of them, and allows image processing to potentially improve understanding of visual perception for recipients [11]. Based on this, researchers have developed various image optimization strategies, and assessed their performances by conducting daily visual tasks under simulated prosthetic vision (SPV) [5,6,26,43,49]. These studies suggest that recipients may benefit from image processing algorithms, and that the utility of feature integration may be conductive to improve their visual perception, such as an "importance mapping" [5]. In essence, enhancing certain image features or integrated features, to some extent, implies to compensate partially for the loss of visual features towards the recipients. This limited compensation may be significant to recipients carrying out visual tasks in daily life. Since many visual tasks involve object detection or locating in which the selective visual attention mechanism dependent on visual features plays an extremely important role and can help humans detect rapidly the location of objects of interest in their visual scenes.

In view of the advantage of visual attention mechanism, some researchers have investigated the application of various attention models on SPV. Li et al. [19] proposed a pixelization model that assigned higher resolution on where prominent features including contrast, edge, orientation, and symmetry. The model proved effective for detecting areas where humans are interested. Boyle et al. [5] applied important maps or region-of-interest (ROI) processing to binary images, and their results indicated that the method used in zoom processing improved scene recognition performance. Ashley et.al [36] developed a salient information processing system for obstacle avoidance under SPV, and the results demonstrated that the system could effectively identify obstacles. Parikh et.al [30] evaluated the benefits provided by a saliency-based cueing algorithm [31] to subjects performing mobility and search tasks under SPV and suggested that using a saliency model might help retinal prosthesis recipients detect important objects in an unknown environment. Recently, Wang et al. [42] proposed two image processing strategies using a saliency segmentation method based on the Itti saliency model and Grabcut for object recognition under SPV. Their results indicated that using the saliency model subserved object detection and that the proposed strategies improved recognition accuracy of objects. The above studies have suggested that image feature integration or saliency models can be used to emphasize the regions with significant features in an image, to such an extent as to assist prosthesis recipients in locating objects and improving their visual perception [17].

Although a variety of sophisticated image processing algorithms have been proven beneficial for recipients to improve their visual perception, most of them cannot achieve real-time processing due to the complexity of the algorithms and the limitation of platform processing power. This greatly limits the practical application of these algorithms on the retinal implant system. For example, [19,42,43,49] were able to only use static images as experimental materials for object recognition under SPV. For some simple visual tasks, a few studies report their algorithm processing time, but all of them are unable to meet the real-time requirement. For instance, Parikh et.al [31] developed a computationally efficient model with three information streams (color saturation, intensity, and edge information) based on the Itti saliency model and embedded this model in a digital signal processor. Results showed the execution speed was only close to 1.2 fps. Ashley [36] developed a saliency-based system that could achieve processing speed of approximately 5 fps under the design of 100 phosphenes. Chris McCarthy reported that, for avoidance obstacles, the proposed strategy based on the augmented depth was able to realize the dynamic display of approximately 20 fps under the simulation of only 20 phosphenes [26].

In this paper, we fully consider the superiority of the saliency model on prosthetic vision and the real-time requirement of image processing algorithms when in practical application, which strongly motivates us to develop a real-time image