3Tu3E.3.pdf Imaging and Applied Optics 2018 (3D, AIO, AO, COSI, DH, IS, LACSEA, LSC, MATH, pcAOP) © OSA 2018

Light-field background de-cluttering for visual prostheses

Jae-Hyun Jung* and Eli Peli
The Schepens Eye Research Institute, Massachusetts Eye and Ear, Department of Ophthalmology, Harvard Medical School, Boston, MA
jaehyun_jung@meei.harvard.edu

Abstract: Object recognition is challenging with current visual prostheses, especially with background clutter. We have developed an imaging system to remove the background clutter in the visual prostheses using the light-field camera and bipolar edge filtering.
OCIS codes: (100.6890) Three-dimensional image processing; (330.0330) Vision, color, and visual optics

1. Introduction

An estimated 260,000 individuals in the U.S. are functionally blind [1]. Various prosthetic vision systems, such as retinal or cortical implants and visual sensory substitution devices (SSDs) [2], have been proposed to restore vision. Most visual prostheses capture scenes using a head-mounted video camera and convert them into a format appropriate for the device. However, the low resolution (<1,000 electrodes), limited dynamic range (2-4 gray levels), and narrow field of view (<20°) of current devices severely restrict their utility. The cluttered images at the low resolution and dynamic range created by current prostheses are difficult to interpret [3]. We have demonstrated the impact of background clutter on object recognition [3-5] and developed an imaging system, active confocal imaging system, which will enable a blind user of any visual prosthesis to efficiently scan, focus on, and see only an object of interest (OI) while suppressing interference from background clutter (background de-cluttering) [3, 6, 7].

2. Light-field for background de-cluttering in visual prostheses

The active confocal imaging system [3] captures three-dimensional scene information using a light-field sensor and generates confocal images (i.e., computational integral imaging [8]) at depth planes selected by the user from potential depth planes that may contain OIs. There is a blur difference between the objects at a selected depth plane and other planes. Edge detection of confocal images removes the blurred background clutter and prepares the features of OIs for presentation within the limited dynamic range. Allowing the user to select OIs from a small subset of depth planes benefits from the user's intent, familiarity with the environment, and situational awareness. In preliminary experiments, we verified the positive impact of confocal-based background clutter removal on recognition of objects in simulated binary edge image (Fig. 1a). However, the recognition rate with the background de-cluttering in the resolution of current available visual prostheses is still too low in static images, and thus the various benefits of live motion imaging need to be evaluated.

Motion parallax can serve as an additional background clutter removal method. In human vision, observers keep fixation on an OI while moving the head, aided by the ocular-vestibular reflex that rotates the eyes in response to head movement. The motion is transferred only to the background clutter and thus the OI can be perceptually separated. However, the head-mounted camera in visual prostheses continues to point straight during the lateral head movement and does not rotate as the eyes do during similar head movements. This causes the OI to move as well, and possibly out of the narrow prosthetic field of view. We proposed a compensatory computational (cropping) method as a fixation aid to simulate the action of eyes' rotation [6]. The OI remains fixed at the center of the prosthetic field during head movement.

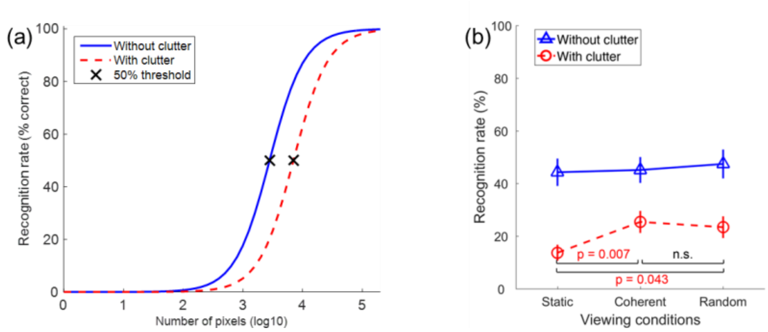

Fig. 1. Impact of background de-cluttering on object recognition. (a) Object recognition rate in static binary edge image with/without background clutter, as a function of resolution [3], (b) Object recognition rate in 20×20 resolution 8-bit grayscale image under three viewing conditions: static, coherent and random motion parallax [4].

We tested the benefit of the method in the simulated phosphene vision (20×20) using Oculus Rift head-mounted display and found that the motion parallax with the proposed system can improve object recognition with background clutter (Fig. 1b). A randomized background motion, control condition, where the multiple viewpoints of the object are not coherent (random) with user's head movement was also helpful to de-clutter the background. The results suggest that motion parallax alone may not provide sufficient de-cluttering effect.

3. De-cluttering based on multi-band bipolar edge filtering with local contrast reduction in light-field

We have developed a better edge filtering method to clearly de-clutter the background and enhance the visibility of the OI in limited dynamic range. We have applied bipolar edge filtering to visual prostheses that is motivated by a model of the human visual system [7, 9]. The polarity of edges and cusps (black or white features on a gray background) in 3 gray levels carries shape-from-shading information not available in the binary edge image.

For bipolar edge filtering, we applied one-octave wide bandpass filters separated by one octave. Binary phase congruence across a range of scales results in three gray levels to represent two different polarities of contrast (black and white lines over gray background), as shown in Fig. 3. Human visual system models perform multiscale bandpass filtering, and the measured contrast sensitivity function is a measure of the system detection threshold in each band.

However, the global threshold in each band misses the effects of local contrast adaptation (e.g., numbers around 12 o'clock in Fig. 2b) or leave high contrast background clutter. Since we generate confocal images (shallow depth of field) from the integration of shifted sub-aperture images (wide depth of field), the local contrast difference between two images from different apertures in the light-field can be used for local contrast thresholding (Fig. 2c).

At the background clutter, the local contrast reduction from the sub-aperture to confocal images (Fig. 2a) is higher than the ratio between their MTFs regardless of original contrast. The contrast reduction at the OI is minimal since the OI is in-focus in both sub-aperture and confocal images. With the proposed bipolar filtering, the active confocal imaging system can provide more details of OI with the limited dynamic range of visual prostheses.

Fig. 2. Light-field background de-cluttering. (a) Sub-aperture image (left) and confocal image (right) from the integration of shifted sub-aperture images. Background de-cluttering in confocal image using bipolar edge filtering: (b) conventional global thresholding loses low contrast details of the OI (e.g., numbers and clock legs) or leave high contrast background clutter while (c) local thresholding of contrast reduction based on MTF of light-field camera keeps all low contrast details of the OI (1024×1024 in left and 64×64 in right). Details of the OI are the more important cues for the object recognition in the low resolution of visual prostheses as shown in right images in (b) and (c).

4. Future works

We are developing a multiple camera array mounted on the spectacle frame to enlarge the de-cluttering range. In addition, we will further evaluate the impact of the systems on object recognition and search with the actual retinal implant and SSD users.

Acknowledgment

Supported in part by US Department of Defense grant W81XWH-16-1-0033.

References

[1] R. Leonard and A. R. Gordon, *Statistics on Vision Impairment: A Resource Manual*, 5th ed. (Lighthouse International, 2002).
[2] J. M. Ong and L. da Cruz, "The bionic eye: a review," Clin. Experiment. Ophthalmol. **40**, 6-17 (2012).
[3] J.-H. Jung, D. Aloni, Y. Yitzhaky, and E. Peli, "Active confocal imaging for visual prosthesis," Vision Res. **111**, 182 -196 (2015)
[4] C. Qiu, K. Lee, J.-H. Jung, R. Goldstein, and E. Peli, "De-cluttering using motion parallax improves recognition of new objects in simulated prosthetic vision," Transl. Vis. Sci. Technol. (to be published).
[5] S. Han, C. Qiu, K. Lee, J.-H. Jung, and E. Peli, "Word recognition: re-thinking prosthetic vision evaluation," J. Neural Eng. (to be published).
[6] E. Peli, J.-H. Jung, and C. Qiu, "Motion Parallax in Object Recognition," US provisional patent application 15/126,775 (2017).
[7] J.-H. Jung, T. Pu, and E. Peli, "Comparing object recognition from binary and bipolar edge for visual prostheses," J. Electron. Imaging **25** 061619 (2016).
[8] S. H. Hong, J. S. Jang, and B. Javidi, "Three-dimensional volumetric object reconstruction using computational integral imaging," Opt. Express **12**, 483-491 (2004).
[9] E. Peli, "Feature detection algorithm based on a visual system model," Proceedings of the IEEE **90**, 78-93 (2002).