Simulated prosthetic vision: object recognition and localization approach

Grégoire Denis, Marc J.-M. Macé*, Christophe Jouffrais.

IRIT, UMR5505, Université de Toulouse & CNRS - Toulouse
* CA: marc.mace@irit.fr

Keywords
Blindness, Visual Prosthesis, Simulation, Object recognition, Object localization.

ABSTRACT

Introduction

Currently, there is no efficient treatment against eye diseases such as retinitis pigmentosa (RP) or age related macular degeneration (ARMD). They affect millions of people worldwide and may result in blindness after a few years of evolution [1]. Within the last decades, a large collection of assistive devices have been designed to assist people with visual impairment and enhance their autonomy. However, assistive devices didn't prove to be very useful in complex tasks such as navigation or object recognition and grasping. In the meantime some research groups aim to restore vision via neural interfaces, and have designed varied neuroprosthesis that elicit perceptible spots of light (phosphenes). Today, the most advanced prostheses are implanted in the retina, and are based on two different strategies. In the first one, an image is captured from a micro camera mounted on a pair of goggles, then transmitted to a video processor and finally converted into electric pulses by an electrode array implant. In the second one, an array of light sensors implanted in the retina directly converts the incoming light into electrical signals [2]. Both devices have been designed to restore an image of the whole field of view through a point-to-point display of phosphenes (sometimes called "scoreboard" approach). The first clinical trials with arrays of 60 electrodes are encouraging: implanted blind subjects are able to perceive simple visual stimuli in a highly contrasted and controlled environment. However, more complex visuo-motor tasks, such as grasping an object among other objects, are still very limited with these neuroprosthesis.

Material and methods

To overcome the limits of current visual prosthesis, we developed an alternative approach based on object recognition and localization. Our hypothesis is that we may restore visual scene perception and complex spatial behavior on the basis of a sparse representation of the surrounding space (a few representative phosphenes only are highlighted). In this study we tested this hypothesis with a simple object localization and grasping task. We developed a simulator of prosthetic vision (SPV) following the recommendations published by Chen et al [3]. The SPV was made of a Head Mounted Display with a stereo camera that captured the scene in real-time. We used a bio-inspired algorithm [4] to perform real-time recognition and localization of target objects in the image. In this experiment, twelve subjects were seated on a chair in front of a table. In each trial, the subjects were asked to localize and grasp an object which was placed among nine other objects randomly spread over the table. Four conditions were systematically assessed: (SC1) scoreboard approach with a simulated 6*10 electrode array, (SC2) scoreboard approach with a simulated 15*18 electrode array, (SC3) scoreboard approach with a simulated 32*38 electrode array, and (LOC) localization approach with a simulated 6*10 electrode array. Each subject performed 24 trials per condition. Performance was measured in terms of time (time to grasp an object) and accuracy (% of correctly grasped objects). In

the SC conditions, the camera image was resized to fit with the number of electrodes available in each condition (6*10, 15*18 & 32*38) and reduced to 8 grey levels. The luminance of each pixel in this low resolution image was then used to display a round shape phosphene with a Gaussian profile. Concerning the LOC condition, the aspect of the phosphene was the same as in the SC condition, but we only displayed the nearest phosphene related to the localization of the target object in the image.

Results

The average time to grasp the target object (correct trials only) was 22.8s (SE=4.8s) in SC1, 26.1s (SE=3.6s) in SC2, 22.4s (SE=3.3s) in SC3, and 17.4s (SE=3.0s) in LOC. There were no significant differences between any of these conditions (Friedman ANOVA, $\chi^2=7.3$, df=3, p=0.06). The accuracy in SC1 was slightly above chance level (15.6%, SE=3.9% - Chance level was 10% correct), and better for SC2 (37.5%, SE=9.6%). The accuracy was very good and comparable for the two remaining conditions (SC3: 71.9%, SE=10.8%; LOC: 80.9%, SE=9.4%). The effect of the condition on accuracy was significant ($\chi^2=32.2$, df=3, $p < 0.001$). Pairwise comparisons indicated that the difference between SC1-SC3 and SC1-LOC was highly significant ($p < 0.001$). Another significant difference was revealed between SC2 and LOC ($p < 0.01$) with a better accuracy in the localization condition.

Conclusion

We showed that performance in object localization and grasping is very good in the LOC condition, even with 60 electrodes only. Indeed it is similar to the performance reached with the classical “scoreboard approach” but with 1216 electrodes (32*38). Currently implants with 1000+ electrodes have been tested, but they didn’t prove to be more efficient. In fact, apparent resolution with 1500 electrodes seems comparable to apparent resolution with 6*10 electrodes, probably because of cross-talks between electrodes. Altogether, these results suggest that the effectiveness of the actual and near future small electrode arrays may be enhanced if a contextual object localization algorithm was included in the visual neuroprosthesis. In addition to object localization and grasping, the interleaved artificial vision may subserve other perceptual [5] and visuo-motor tasks (pointing, reaching, heading, etc.).

[1] WHO, “Visual Impairment and blindness Fact Sheet N° 282,” World Health Organization, 2012.

[2] E. Margalit, M. Maia, J. D. Weiland, R. J. Greenberg, G. Y. Fujii, G. Torres, D. V. Piyathaisere, T. M. O’Hearn, W. Liu, G. Lazzi, G. Dagnelie, D. A. Scribner, E. De Juan, and M. S. Humayun, “Retinal Prosthesis for the Blind.,” *Survey of Ophthalmology*, vol. 47, no. 4, pp. 335–356, 2002.

[3] S. C. Chen, G. J. Suaning, J. W. Morley, and N. H. Lovell, “Simulating prosthetic vision: I. Visual models of phosphenes.,” *Vision Research*, vol. 49, no. 12, pp. 1493–1506, Jun. 2009.

[4] F. Dramas, S. Thorpe, and C. Jouffrais, “Artificial Vision For The Blind: A Bio-Inspired Algorithm For Objects And Obstacles Detection,” *International Journal of Image and Graphics*, vol. 10, no. 4, pp. 531–544, 2010.

[5] R. Parlouar, F. Dramas, M. J.-M. Macé, and C. Jouffrais, “Assistive device for the blind based on object recognition: an Application to Identify Currency Bills,” in *SIGACCESS conference on Computers and accessibility*, 2009, pp. 227–228.