Proceedings of the 28th IEEE
EMBS Annual International Conference
New York City, USA, Aug 30-Sept 3, 2006



# Simulation and Assessment of Bioinspired Visual Processing System for Epi-retinal Prostheses

Milena Vurro, * *Student Member, IEEE*, Giuseppe Baselli, Francesco Orabona, and Giulio Sandini

**Abstract – Retinal prosthesis represent the best near-term hope for individuals with chronic blinding disease of the outer retina. However the small number of stimulating electrodes produces a poor, low resolution image. We propose a new pre-processing method for epi-retinal implants and validate it through a novel simulation of the implanted blind perception. Twenty-one normally sighted, untrained subjects performed a face recognition test. Three different electrodes grids were simulated: rectangular, hexagonal and log-polar. The results show that the proposed pre-processing method has a good and statistically significant performance improvement.**

## I. Introduction

According to the World Health Organization statistics, approximately 40 million people are blind all over the world [1]. Since several studies have demonstrated that the sensation of a spot of light, called "*phosphene*" [2], can be produced by electrically stimulate a point of the visual pathway, researchers have been working on the development of visual prostheses for blinds.

Recent studies show that photoreceptor degenerative diseases, as the retinitis pigmentosa (RP) and age-related macular degeneration (AMD), destroy only the photoreceptor layer leaving a functional neuronal network in the retina. This has leaded to an increase interest in retinal prosthesis, which have actually developed in two different approaches: sub-retinal and epi-retinal [3]. In this work we will focus on epi-retinal implants. They consist in several subsystems that perform distinct functions, including units for image acquisition, image-processing electronics and a high-density microelectrode array, attached to the inner retinal surface, for charge delivery to the retina.

Although unlikely to recreate the full experience of vision due to the small number of electrodes, visual prostheses may provide enough visual clues to blind people to perform every-day tasks such as navigation, recognition, and reading. Yet few groups are searching effective image processing methods for enhancing viewer perception when using visual prosthesis prototypes [4-7].

In our opinion image degradation can be partially recovered by introducing some kind of pre-processing stage; it is still reasonable that an elaboration which mimics some aspects of the retina processing, as contrast enhancement and equalization, could improve the performance of a prosthetic visual device and reduce the information to meet the limited number of electrodes.

This research supports quality-of-life improvements for blind people by proposing a new pre-processing technique to enhance the quality of image for epi-retinal prosthesis, inspired by the retina layers processing.

To assess the proposed pre-processing method we have developed a new simulation approach to the perception of the implanted blind. Twenty-one subjects performed a face recognition task using this simulation. Also we evaluate the effect of the proposed method on different electrodes grids: rectangular, hexagonal [5] and log-polar [8]. This last one is a bio-inspired configuration we suggest.

## II. Methods

Every image went through two processing steps: first it was enhanced by a nonlinear algorithm and then "pixelized" to simulate prosthetic vision mask.

### A. Subjects

The subjects were twenty-one volunteers, ranging from 20 to 55 years of age. They had all normal or corrected visual acuity of 20/20 for both eyes and were instructed on the purpose of the tests.

### B. Images

The image database was freely given by the University of Stirling, Psychology Department[1]. We have chosen two photos of 24 individuals (12 women and 12 men), one in frontal position with normal illumination, and the other in a partially averted facial view or illumination. The age of the subjects shown was equally distributed in young, adult and middle age. An example of the two sets of images is shown in Fig. 1.


Fig. 1 - Example of images in reference set (A) and test set (B).

### C. Pre-processing model

The pre-processing model used is based on two computational retina models published in [9, 10].

M. Vurro (*corresponding author*) and G. Baselli are with the Department of Biomedical Engineering, Polytechnic of Milan, 20133 Milan Italy (e-mail: mv650280@lau.polimi.it).

F. Orabona and G. Sandini are with Department of Communication Computer and System Science (DIST), University of Genoa, 16145 Genoa, Italy.

[1] http://pics.psych.stir.ac.uk, last access June 2005.

The proposed elaboration was applied to the chromatic domain by considering independent filters for the basic colors. Differently to the original model, we tuned the contrast gain parameter to obtain better recognition results, in fact our aim was not to fit biological data.

The first step consists in splitting the image *I(x,y)* in the three color channels $s_i(x,y)$. Then the stimulus $s_i(x,y)$ is convolved with a Difference of Gaussians (DoG) filter (1) for contrast enhancement:

$$k(x,y)=\frac{\alpha_+}{2\pi\sigma^2}\bullet e^{-\frac{(x^2+y^2)}{2\sigma^2}}-\frac{\alpha_-}{2\pi\beta\sigma^2}\bullet e^{-\frac{(x^2+y^2)}{2\beta\sigma^2}} \qquad (1)$$

$$h(x,y)=s(x,y)\otimes k(x,y) \qquad (2)$$

where $\alpha_+$ and $\alpha_-$ represent, respectively, the relative weights of center and surround, $\sigma$ the standard deviation and $\beta$ their diameters' ratio. The output was multiplied by a non-linear function $c_\gamma(x,y)$ called ***contrast gain control function***, expressed in (3),

$$c_\gamma(x,y)=\begin{cases}\dfrac{1}{1+\gamma\bullet h(x,y)^4}, & \mathrm{h}>0\\ 1, & \mathrm{h}\le 0\end{cases} \qquad (3)$$

$$u(x,y)=h(x,y)\bullet c_\gamma(x,y) \qquad (4)$$

It is finally rectified by a static nonlinear function *F(u)*.

$$F(u)=\begin{cases}0, & u\le 0\\ u, & u>0\end{cases} \qquad (5)$$

The parameters values were $\alpha_-/\alpha_+$ =0.8, $\sigma$=80 µm and $\beta$=3, as in [9, 10].

As said above, to obtain the value of the model parameter $\gamma$, the ***contrast gain***, a pilot test was performed on five of the twenty-one volunteers; they viewed 120 images through a rectangular grid in one session. As it will be said in Section IV the best $\gamma$ value was 4 (Table I).

Each channel could be differently weighted, summed in a single image and displayed in 256 gray levels. In this work we chose a weight of 1/3 for all them, i.e. averaging them; in the future a different combination could be investigated.

*D. Simulation of blind implanted vision*

Some assumption were made on the simulated device based on the existing ones [3]. Likewise in other studies [6, 11] the center of the simulated implant was considered to be on the optic axis, and the gap interface problem (stimulating array-ganglion cell layer) was ignored.

Almost all the epi-retinal prosthesis prototypes have electrodes forming a rectangular grid; however, recently, Suaning et al. [12] have proposed a hexagonal configuration demonstrating a performance advantage over the rectangular grid for correct identification of a test symbol [6]. Guided by the studies of Sandini and Tagliasco [8] we tested also a new retina-like configuration, the log-polar grid.

Hence we have simulated an array of 100 electrodes on a stimulating area of 4.7x4.7 mm$^2$, considering three grid types: rectangular (Rect), log-polar (LP) and hexagonal (Hex) [12]. The electrodes of the rectangular sampling grid were arranged in a 10x10 square; in the log-polar grid the centers of the electrodes are placed on circumferences with exponential increasing radius [8]. This arrangement is based upon the photoreceptors' distribution of the retina and it has the advantage of having a larger size of the input image being the number of electrodes the same.

The electrode spacing, used in rectangular and hexagonal grid, was fixed at 220 µm edge-edge (14 pixels on screen - see next paragraph), which covers a 0.75° angle in the visual space, and their diameter at 250 µm (16 pixels). Instead log-polar grid electrode diameter ranged from 120 µm to 500 µm (8-32 pixels). The electric stimulus of each electrode was supposed linearly proportional to the mean luminance intensity of the image inside a rectangular, hexagonal or circular area centered in the electrode, as shown in Fig. 2, and then quantized to 5 levels. So it implicitly applies a low pass filter on the image, necessary when sub-sampling.

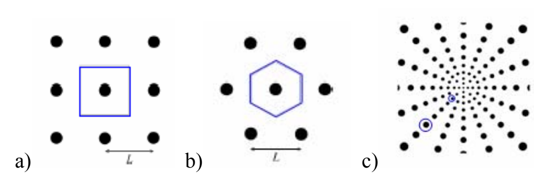


Fig. 2 – Mean area in a) rectangular, b) hexagonal and c) log-polar case.

Evidence suggests that ***phosphene*** size and brightness may be modulated [13, 14] by the intensity an duration of the electric stimulus. To modulate the visual characteristics of phosphenes in the simulation a 'phosphene strength' value was calculated. As shown in Fig. 3 there are two different approaches to simulate phosphenes appearance:

1. grayscale monochromatic dots with the gray level value proportionally dependent on the electric stimulus [7, 11] (Fig. 3-B);
2. grayscale circular Gaussian luminance profiles with fixed intensity value of center and phosphene radius dependent on the electric stimulus [4, 12] (Fig. 3-C).

We propose a novel simulation, called "phosphene gaussian" representation (Fig. 3-D), where phosphene appearance is simulated as a grayscale circular Gaussian luminance profile with the center luminance value and standard deviation modulated by phosphene strengths, as many neurophysiological studies show [14, 15]. The max gray scale value was fixed at 250 and the gaussian profile was truncated when reached the 20% of it. This minimum represents the activating threshold, which is the minimum value under which there is no neuronal activation.

Note that in the gaussian phosphene representation two or more phosphenes can sum their values and fuse in one blurred site, as in real electric stimulation [16]. On the contrary other simulations, as shown in Fig. 3, ignore such physiological event.

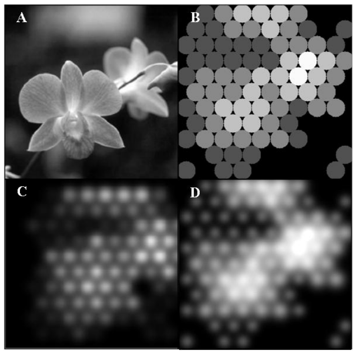


Fig. 3 – Approaches of simulation of epi-retinal prosthesis stimulation: A) original figure. B) Monochromatic dots. C) "Electrodes gaussian pixel". D) "Gaussian phosphene" proposed in this work.

## III. Procedure

Using Matlab 6.5, we developed a user friendly software program to present the images for the facial recognition task.

At the start of each experimental session, each of the twenty-one subjects was set at 30 cm distance from a 15" monitor (resolution 1024x768).

During each trial, subjects first viewed a reference set of four phosphene simulated facial images of different individuals covering a visual angle of around 15.45° each. The four faces in each trial contained images representing each age group and both gender. While watching these images, a fifth one was shown, the **test image**, that was a partially averted facial view of one of the individuals in the reference set, altered with the phosphene simulation.

On determining the identity of the test face, subjects digited the number of the picture of the reference set that matched the identity of the test face. They performed 144 trials; in each trial the images were viewed under pre-processing or original condition in one of the three grid configuration (rectangular, log-polar and hexagonal). Each combination was presented 24 times in a pseudorandom, counterbalanced order, so that any learning effects would be evenly distributed across the test conditions. No maximal response time was imposed. An example of the original image and after simulation in all configurations with and without pre-processing is showed in Fig. 4.

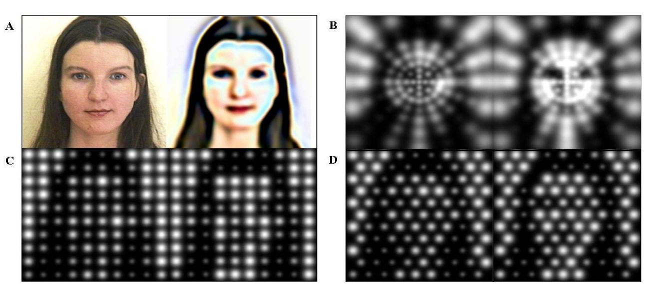


Fig. 4 – Pre-processing image (right) against base image (left) for: A) original image; B) log-polar grid simulation; C) rectangular grid; D) hexagonal grid

## IV. Results

### A. Data analysis

For each test condition were evaluated mean, percentage mean and standard deviation (SD) of subjects' correct responses.

Due to the great inter-subject variability two performance gains were evaluated for each subject *i*, the ***processing gain***:

$$g_p(i) = \frac{\text{correct responses with pre-processing}}{\text{correct responses without pre-processing}} - 1 \qquad (6)$$

and the ***configuration gain***:

$$g_e(i) = \frac{\text{correct responses, choosen grid}}{\text{correct responses, rectangular grid}} - 1 \qquad (7)$$

The one-tailed *t*-test was used to determine whether the percentage of correct responses for each trial condition significantly exceeded that expected by chance alone (25%).

### B. Gamma parameter

We have tested this value of $\gamma$: 1, 2, 4, 8 obtained on some preliminary tests, and 78, taken from [9], on five of the twenty-one volunteers. The best resulted $\gamma$=4, as shown in Table I.

Table I
Mean and standard deviation of correct response over 24 faces for different $\gamma$ values

| | Base $\gamma$=0 | $\gamma$=1 | $\gamma$ =2 | **$\gamma$=4** | $\gamma$=8 | $\gamma$=78 |
|---|---|---|---|---|---|---|
| Mean | 17.7 | 19.4 | 19 | **20.1** | 18.5 | 17 |
| SD | 1.17 | 1.64 | 1.72 | **1.06** | 1.21 | 1.33 |

### C. Face recognition: pre-processing and grid performance

The other sixteen of the twenty-one subjects performed, instead, the final recognition test.

Table II summarizes the results for different experiment conditions.

Table II
Mean, standard deviation and percent mean (random=25%) of correct response over 24 faces in all condition

| | Rect Base | Rect Enh | LP Base | LP Enh | Hex Base | Hex Enh |
|---|---|---|---|---|---|---|
| Mean | 18.1 | 19.2 | 16.7 | 17.4 | 17.1 | 19.1 |
| SD | 1.99 | 2.29 | 3.18 | 2.53 | 1.17 | 3.85 |
| % Mean | 75.5 | 79.9 | 69.5 | 72.4 | 71.3 | 79.6 |

In all test combinations the subjects achieved good facial recognition accuracy, between the 69.5% and the 79.9%, against a random case of 25% ($p<0.001$). The use of the pre-processing enhancement (Enh) algorithm improved face recognition in *all* configurations, with a gain of 4.12% for the log-polar, 5.86% for the rectangular and 11.69% for the hexagonal grid. The tests were statistically significant ($p<0.05$, paired t-test) except for the log-polar configuration.

Note that the standard deviation is rather high, meaning an elevated intra-individual difference in the responses. For this reason and to see the real subject improvement when using the pre-processing model, we calculated the inter-subject gain $g_p$. The processing gain percent is 5.81% for the log-polar, 6.66% for rectangular and 12.1% for hexagonal grid, which means a good performance improvement due to the pre-processing algorithm.

The hexagonal grid obtained the best benefit from pre-processing, while rectangular configuration, starting from a higher recognition base and getting the highest post-enhanced recognition rate, improved less.

As shown in Table III we estimated the inter-subject configuration gain $g_c$. The rectangular grid results the best both with and without pre-enhancement.

TABLE III
CONFIGURATION GAIN: COMPARING RECTANGULAR GRID TO HEX AND LOG GRIDS IN BOTH PROCESSING CONDITIONS

| | LP base/ Ret base | Hex base/ Ret base | LP Enh/ Ret Enh | Hex Enh/ Ret Enh |
|---|---|---|---|---|
| Mean gain | -7.61% | -6.2% | -8.91% | -1.27% |
| SD | 1.5 | 1.3 | 1.2 | 1.5 |

## V. DISCUSSION AND CONCLUSIONS

Visual prosthesis research is steadily progressing, and medium-scale implantation of safe and useful prostheses may occur soon. In the designs under development, electrode numbers are limited, and there is a need to efficiently use the electrodes and corresponding perceived phosphenes. It is therefore timely to suggest means of increasing the information content of artificial vision systems. In this work we accomplish three objectives: 1) developed a pre-processing enhancement algorithm to provide a less degraded visual information; 2) developed a biologically motivated simulation of the implanted blind vision; 3) tested the performance of three different electrodes configurations.

The pre-processing model proposed in this work improves the performance of facial recognition in all configuration grids, with max gain of 12% for the hexagonal grid. The rectangular grid was the best configuration for facial recognition. However the hexagonal grid allows a higher electrode density compared to the rectangular grid, so it is possible to increase the number of electrodes from 100 to 112 on the same simulated array space, allowing a higher resolution of the phosphened image. Considering in addition that the hexagonal grid has the best performance after pre-processing, it appears to be the best choice for static images. Log-polar grid, instead, was penalized by the fixed position of the high resolution area at the image center; this determined that some important details, such as the eyes, were misinterpreted. In future works, consenting to the subject to move this high resolution area around the image, the potentiality of this configuration could be fully exploited.

These experiments provide a basis from which more complex vision prostheses simulation tests may be derived. Also the tests involved presentation of static (still) images to viewers and a logical extension to this work is to conduct similar experiments on image sequences (video).

## REFERENCES


[1] B. Thylefors, A. D. Negrel, R. Pararajasegaram, and K. Y. Dadzie, "Global data on blindness," Bulletin of World Health Organization, vol. 73, pp. 115-121, 1995.

[2] O. Foester, "Beirtrage zur pathophysiologie der sehbahm und der spehphare," J. Psychol Neurol, vol. 39, pp. 435-463 (Abstract in English), 1929.

[3] E. Margalit, M. Maia, J. D. Weiland, R. J. Greenberg, G. Y. Fujii, G. Torres, D. V. Piyathaisere, T. M. O'Hearn, Wentai Liu, G. Lazzi, G. Dagnelie, D. A. Scribner, E. d. Juan, and a. M. S. Humayun, "Retinal prosthesis for the blind," Survey of Ophthalmology, vol. 47, pp. 335–356, 2002.

[4] J. S. Hayes, V. T. Yin, D. V. Piyathaisere, J. D. Weiland, M. S. Humayun, and G. Dagnelie, "Visually guided performance of simple tasks using simulated prosthetic vision," International Society of Artificial Organs, vol. 27, pp. 1016-1028, 2003.

[5] L. E. Hallum, G. J. Suaning, D. Taubman, and N. H. Lovell, "Towards Photosensor Movement-Adaptive Image Analysis in an Electronic Retinal Prosthesis," Proceedings of the 26th Annual International Conference of the IEEE EMBS, San Francisco, CA, USA • September 1-5, pp. 4165-4168, 2004.

[6] S. C. Chen, L. E. Hallum, N. H. Lovell, and G. J. Suaning, "Visual acuity measurement of prosthetic vision: a virtual-reality simulation study," J Neural Eng, vol. 2, pp. S135-S145, 2005.

[7] J. R. Boyle, A. J. Maeder, and W. W. Boles, "Image enhancement for electronic visual prosthesis," Australasian Physical Engineer Sciences Medicine, vol. 25, tecnical note, 2002.

[8] G. Sandini and V. Tagliasco, "An anthropomorphic retina-like structure for scene analysis," Computer Graphics and Image Processing, vol. 14, pp. 365-372, 1980.

[9] S. D. Wilke, A. Thiel, C. W. Eurich, M. Greschner, M. Bongard, J. Ammermuller, and H. Schwegler, "Population coding of motion patterns in the early visual system," J Comp Physiol A, vol. 187, pp. 549-558, 2001.

[10] M. J. Berry, I. M. Brivanlou, T. A. Jordan, and M. Meister, "Anticipation of moving stimuli by the retina," Nature, vol. 398, pp. 334-338, 1999.

[11] R. W. Thompson, G. D. Barnett, M. S. Humayun, and G. Dagnelie, "Facial recognition using simulated prosthetic pixelized vision," Invest Ophthalmol Vis Sci, vol. 44, pp. 5035-5042, 2003.

[12] L. E. Hallum, D. Taubman, G. J. Suaning, J. Morley, and N. H. Lovell, "A filtering approach to artificial vision: phosphene visual tracking task," IFMBE Proceed, World Congress on Medical Physics and Biomedical Engineering, August 24-29, Sydney, 2003.

[13] K. Nakauchi, T. Fujikado, H. Kanda, T. Morimoto, J. S. Choi, Y. Ikuno, H. Sakaguchi, M. Kamei, M. Ohji, T. Yagi, S. Nishimura, H. Sawai, Y. Fukuda, and Y. Tano, "Transretinal electrical stimulation by an intrascleral multichannel electrode array in rabbit eyes," Graefe's Archive for Clinical and Experimental Ophtalmology, vol. 243, pp. 169-174, 2005.

[14] J. Rizzo, J. Wyatt, J. Loewenstein, S. Kelly, and D. Shire, "Perceptual efficacy of electrical stimulation of human retina with microelectrode array during short-term surgical trials," Invest Ophthalmol Vis Sci, vol. 44, pp. 5362-5369, 2003.

[15] M. Wilms, M. Eger, T. Schanze, and R. Eckhorn, "Visual resolution with epi-retinal electrical stimulation estimated from profiles in cat visual cortex," Visual Neuroscience, vol. 20, pp. 543-555, 2003.

[16] M. S. Humayun, J. D. Weiland, G. Y. Fujii, R. J. Greenberg, R. Williamson, J. Little, B. Mech, V. Cimmarusti, G. Van Boemel, G. Dagnelie, and E. de Juan Jr, "Visual perception in a blind subject with a chronic microelectronic retinal prosthesis," Vision Res., vol. 43, pp. 2573-2581, 2003.