ELSEVIER

Prediction of visual perceptions with artificial neural networks in a visual prosthesis for the blind

Cédric Archambeau[a,*], Jean Delbeke[b], Claude Veraart[b], Michel Verleysen[a]

[a]Microelectronics Laboratory, Université catholique de Louvain, Place du Levant 3, B-1348 Louvain-la-Neuve, Belgium
[b]Neural Rehabilitation Engineering Laboratory, Université catholique de Louvain, Avenue Hippocrate 54, B-1200 Bruxelles, Belgium

Summary Within the framework of the OPTIVIP project, an optic nerve based visual prosthesis is developed in order to restore partial vision to the blind. One of the main challenges is to understand, decode and model the physiological process linking the stimulating parameters to the visual sensations produced in the visual field of a blind volunteer. We propose to use adaptive neural techniques. Two prediction models are investigated. The first one is a grey-box model exploiting the neurophysiological knowledge available up to now. It combines a neurophysiological model with artificial neural networks, such as multi-layer perceptrons and radial basis function networks, in order to predict the features of the visual perceptions. The second model is entirely of the black-box type. We show that both models provide satisfactory prediction tools and achieve similar prediction accuracies. Moreover, we demonstrate that significant improvement (25%) was gained with respect to linear statistical methods, suggesting that the biological process is strongly non-linear.
© 2004 Elsevier B.V. All rights reserved.

1. Introduction

The European project OPTIVIP (Optimisation of the Visual Implantable Prosthesis) has recently been launched. Its primary aim is to further investigate the feasibility and the prospects of an optic nerve based visual prosthesis in order to restore partial vision to the blind. In this paper, an attempt is made to understand, decode and model the neurophysiological process that links the stimulation parameters to the visual sensations described by a blind volunteer. Preliminary results of this work have been published in [1].

We propose to use artificial neural networks (ANNs) for predicting the features of the visual sensations. ANNs are adaptive methods, able to evolve when new data are gained from experiments on the blind volunteers. Two alternatives are investigated. The first one is a hybrid neural network. It combines a neurophysiological model with a black-box ANN model. The second one is entirely of the black-box type.

The performances of both models are compared to classical linear statistical methods. We show that the non-linear statistical ANN methods possess much better prediction capabilities than the linear ones. This result suggests that the neurophysiological

*Corresponding author. Tel.: +32-10-47-25-40; fax: +32-10-47-25-98.
E-mail address: archambeau@dice.ucl.ac.be (C. Archambeau).

process is essentially non-linear. Besides, we demonstrate that the prediction accuracy is not reduced when a complete black-box model is used.

In the next section, we describe the principles of the optic nerve based visual prosthesis and we present briefly the overall electrical system. In Section 3 we clarify the notion of phosphene, i.e. visual sensations, and we introduce their prediction problem. In Section 4, we describe the structure of the hybrid ANN and review some of the most popular artificial neural networks, as well as an enhanced training procedure. Finally, in Section 5, we present our prediction results.

2. The optic nerve based visual prosthesis

Since the late eighties cochlear implants have rehabilitated patients with hearing loss for whom there is no other potential treatment. Further to this success in an analogous manner several multidisciplinary teams were established during the past decade with the goal to restore partial vision to the blind and improve their quality of life. The principle consists in implanting a neural prosthesis either intraocularly or intracranially and bypass, by electrical stimulation, neurons that have become non-functional.

To date three types of visual implants have arisen: cortical [2], retinal [3,4] and optic nerve based [5] implants. Cortical prostheses have the advantage of being able to treat blindness secondary to retinal or optic nerve diseases. Nevertheless the approach needs to deal with the complex geometry of the brain and requires performing an intracranial surgical procedure with high risks. Retinal and optic nerve based approaches, on the other hand, are ocular prostheses, and, as a result, avoid the high surgical risks. They though only apply to diseases where the optic nerve is still intact; retinitis pigmentosa (RP), one of the leading causes of blindness, is an example of such disease.

In a previous European project, MIVIP (MIcrosytem based VIsual Prosthesis), the feasibility of the optic nerve based visual prosthesis was investigated: the electrical stimulation of the optic nerve was demonstrated [6], a microelectronic prototype was built [7] and a considerable amount of data was gathered concerning the visual sensations evoked by a blind volunteer.

The neural prosthesis aims at capturing images and transforming them into electrical stimulations applied to the optic nerve. The system includes an artificial retina, an external processor, a transcutaneous antenna and its antenna driver, an implanted stimulator and a cuff-electrode (Fig. 1).

The external processor is a portable device processing the image and coding it into a restricted data stream. The transcutaneous link is of the inductive type in order to avoid wires through the skin. It transmits the data, the power supply and the clock to an internal circuit, the stimulator, which in turns decodes the data stream into waveforms. Those are applied to the optic nerve through a cuff-electrode with four contacts wrapped around it.

Lately the OPTIVIP project has started [8]. The main purpose of the project is to further develop the optic nerve based visual prosthesis designed during MIVIP: the electrical components are improved in terms of size and consumption, the
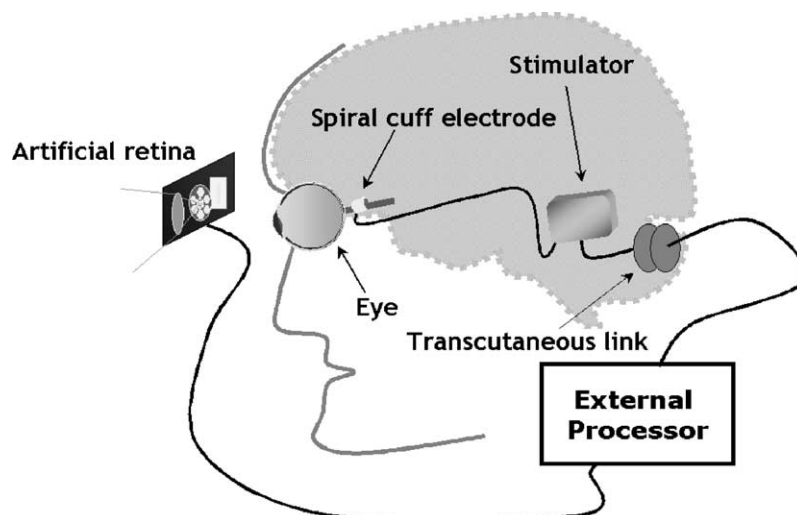

Figure 1   The optic nerve based visual prosthesis.

surgical procedure involving the electrode implantation is adapted and eased, the underlying neurophysiological process linking the stimulation parameters to the evoked visual sensations is being decoded and modelled, physiological data are further collected, new RP volunteers are recruited and the cosmetic aspects are better integrated in the design.

In this paper we concentrate on the modelling and the approximation of the neurophysiological process when single electrode contacts are activated. As a matter of fact, additionally to the four electrode contacts, a fifth reference contact is placed relatively far from the cuff-electrode. Each of them may be used as anode or cathode for the electrical stimulations. The experiments taken into account in this work are those where the cathode is chosen as one of the four electrode contacts and the anode as the reference contact.

3. The visual perceptions

The electrode and the stimulator have been implanted on a blind volunteer and a number of experiments have been conducted in order to establish the features of the visual sensations. The stimulation principle relies on the selective response of the human optic nerve to adequately chosen electrical pulses and an adequate combination of the electrode contacts.

Brindley's seminal work [9] and subsequently Dobelle's experiments [10] have already demonstrated that:

- Current intensities delivered to the visual cortex evoke perceptions of bright spots of light, called phosphenes.
- Multiple spots of light could be perceived simultaneously following multiple points of stimulation.
- There is a perceptual alignment roughly correlated with the spatial organisation of the visual cortex.

Similarly, when current pulses are delivered to the optic nerve, so-called phosphenes are perceived by the patient. The visual perceptions are spatially organised in the visual scene of the volunteer according to a retinotopic map (Fig. 2). In other words, each contact around the optic nerve activates fibres located in a certain area of the optic nerve cross-section, which in turns corresponds to a well-defined area in the visual field of the blind [5]. In addition, spatial and temporal summation of phosphenes was observed when different electrode contacts were combined [11].
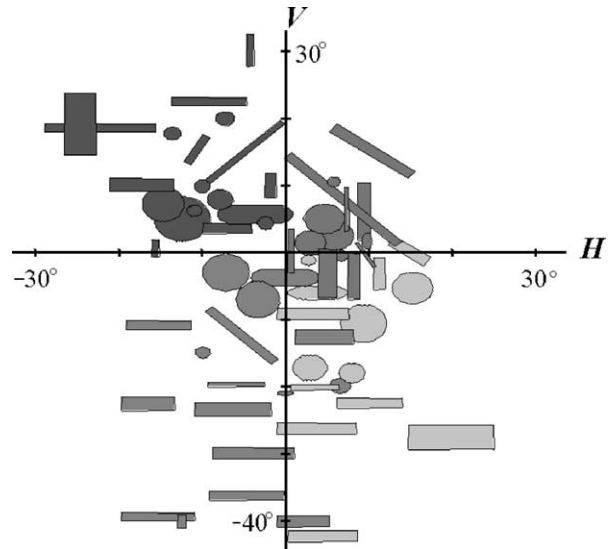
Figure 2 Examples of phosphene mappings in the visual field of the blind volunteer and their corresponding apparent shapes. Each grey level corresponds to one of the four electrode contacts.

3.1. The prediction problem

The prediction problem is depicted in Fig. 3. The complexity of the neurophysiological process, whereby the electrical pulses applied to the optic nerve generate phosphenes, makes it difficult to study the entire process on a biological level: some unknown parameters influence it to a large extent. Furthermore, it must be stressed that the optic nerve of RP patients is probably damaged up to some unknown degree. Finally, the characteristics of the phosphenes correspond to the description, by a blind volunteer, of subjective perceptions, clearly subject to human inaccuracies and errors. As a consequence, the collected data set can be expected to be very noisy.
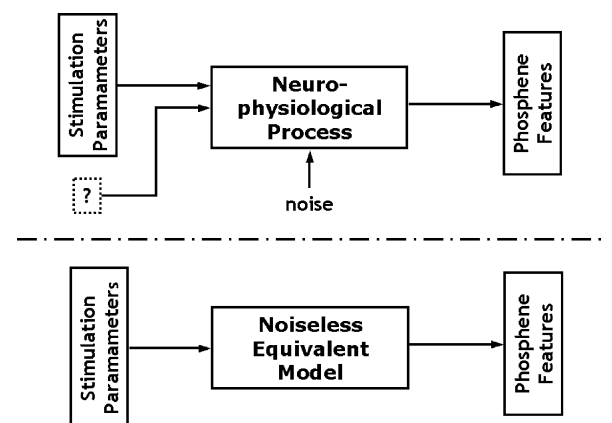
Figure 3 Prediction problem.

For these reasons, even if partial decoding can be achieved by using physiological knowledge (see Section 4), a mathematical identification of the undecoded part of the process is needed in order to obtain a model that can be used to predict the stimulation results and plan new tests efficiently. As a long-term goal, the model will allow us to derive algorithms that will be used to convert an incoming image into an adequate set of electrical pulses and give rise to a well-reconstructed image.

Consider single electrode contact stimulations. The contacts can be identified by their angular position around the optic nerve:

$$C\,(^\circ) = \{0, 90, 180, 270\}.$$

Let us define for each contact C a set of stimulating parameters X_C, which are the current pulse shape, its amplitude I, the pulse duration D, the number of pulse repetitions N and the frequency λ between successive pulses in the pulse train. Rectangular pulses with charge recovery are used in our experiments. Besides, let us define a set of features Y_C describing the perceived phosphenes, including their location in the visual field of the volunteer, their area, colour, shape and intensity. The input—output relationship can be described by a physiological process g, defined as

$$g : X_C \subseteq \Re^n \to Y_C \subseteq \Re^m : g(\mathbf{x}) = \mathbf{y},$$

where n is the dimension of the stimulation parameters space and m the dimension of the feature space.

The process g can be decomposed into two biological sub-processes g_1 and g_2 of lower complexity:

$$g(\mathbf{x}) = (g_1 \circ g_2)(\mathbf{x}).$$

At this stage g_1 and g_2 can be considered as any reasonable and meaningful split of the overall physiological process g. In Section 4.1, a particular decomposition will be addressed.

Next, consider the approximation model $\hat{g}_1$. It can be defined as:

$$\hat{g}_1 : Z_C \times W_C \times \Theta_C \subseteq \Re^{k \times l \times 1}$$
$$\to Y_C \subseteq \Re^m : \hat{g}_1(\mathbf{z}; \mathbf{w}, \theta) = \hat{\mathbf{y}},$$

where Z_C is an intermediate parameter space of dimension k:

$$g_2 : X_C \subseteq \Re^n \to Z_C \subseteq \Re^k : g_2(\mathbf{x}) = \mathbf{z}.$$

In Eq. (1) Θ_C is the model selection parameter space and l the dimension of the model parameter space W_C. The model selection parameter θ enables us to select the complexity of an approximation model $\hat{g}_1$. The complexity of $\hat{g}_1$ is reflected by the number of parameters in the model. The model parameter vector $\mathbf{w}$ is thus a function of θ and it allows us to define a model of a chosen complexity θ.

Suppose that we have identified and decoded the biological process g_2. The prediction problem then consists in finding a model $\hat{g}_1$ that can approximate the undecoded part g_1 of the physiological process g as close as possible:

$$\hat{g}_1^{(\text{opt})} = \operatorname*{argmin}_{\hat{g}_1} E(g_1, \hat{g}_1),$$

where E is a well-defined error criterion.

As linearity of the whole physiological process g cannot be expected, non-linear models, i.e. artificial neural networks, are preferred for estimating $\hat{g}_1^{(\text{opt})}$. In addition, the adaptation abilities of neural networks allow us to build a model that evolves as information from new experiments is gained. Moreover, we expect in the future to reach different, but similar models for other patients. The global structure of the model, for example the number of nodes (see below), found with one patient may be expected to be appropriate for new ones.

4. Prediction tools

The aim of the ANN training is not to learn an exact representation of the training data itself, but rather to build a statistical model of the process that generates the data, such that it achieves a good prediction for new inputs.

In this work, we focus on the prediction of the horizontal and vertical position of the visual sensations. A correct prediction of the phosphenes location seems to be essential in order to reconstruct a visual scene. Indeed, let us consider an incoming image. This image can be pixelised and each pixel can be identified by its position in the image grid. If the position of the phosphenes can be predicted with a sufficient accuracy, we could associate each pixel or set of pixels to the corresponding phosphene, and, as a result, reconstruct the image.

Subsequently, we will first describe a grey-box hybrid ANN using neurophysiological information. Secondly we will review the principles of linear regression, as well as those of two classical neural networks: the multi-layer perceptron (MLP) and the radial basis function network (RBFN), which are both non-linear data prediction models. Finally, we will present the training procedure of neural networks and a parameter estimation method based on K-fold cross-validation.

4.1. A hybrid artificial neural network

In [12], a model was proposed to describe the perception threshold of phosphenes by direct stimulation of the human optic nerve. The model shows that the classical strength—duration equation is valid for describing the perception threshold in the optic nerve. The strength—duration equation states that the minimum current intensity I_v required in order to activate the nerve fibres is related to the pulse duration D. In addition, the perception threshold is linked to the other stimulation parameters, namely the frequency λ and the number of pulses N. Therefore, when fitted on the data set of each contact C, resulting from experiments that were conducted in order the measure the perception threshold, the following model has been obtained [13]:

$$g_2 \equiv I_v = K_C \times 10^{-6}$$
$$\times \frac{8.6 + 141.4/S}{(1 - 0.47/S)(1 - \exp(D/164.9 \times 10^{-6}))}, \quad (2)$$

where K_C is a constant depending on the electrode contact:

$$K_0 = 0.636, \qquad K_{90} = 0.812,$$
$$K_{180} = 1.178, \qquad K_{270} = 1.374.$$

The factor S is defined as

$$S \equiv \sum_{j=1}^{N} \exp\left(\frac{j - N}{39 \times 10^{-3} \times \lambda}\right).$$

We propose a hybrid ANN structure illustrated in Fig. 4. Before entering into the classical ANN, the current amplitude is pre-processed according to the perception threshold model (2). The known physiological information enables us to filter the valuable fraction of the current intensity. The information contained in the current amplitude

is thus adjusted according to the biological process g_2:

$$I_{\text{useful}} = I - g_2(D, N, \lambda).$$

4.2. Artificial neural networks

In this section we recall several statistical prediction tools and their training procedure. Both are used for grey-box and black-box modelling.

In the following, we have renamed the approximation model as $\hat{f}$. Referring to Section 3, when we are dealing with a grey-box model, $\hat{f}$ corresponds to $\hat{g}_1$, whereas when we are dealing with complete black-box model, $\hat{f}$ corresponds to the approximation of entire physiological process g.

4.2.1. Training procedure
The training procedure of ANN can be decomposed in three steps. At first the unknown physiological process is learned by using a set of input—output pairs. Then, we generalise, that is we try to make the best prediction for a new set of inputs. At last the performances of the approximation model are tested on a third data set.

Consider the data set X_C corresponding to the electrode contact C. We can, respectively, define a learning set of size N_{L_C}, a validation set of size N_{V_C} and a test set of size N_{T_C}, which are mutually exclusive:

$$L_C = \{(\mathbf{x}^{(p)}, \mathbf{y}^{(p)}) \in X_C \times Y_C, 1 \le p \le N_{L_C} : \mathbf{y}^{(p)} = f(\mathbf{x}^{(p)})\},$$
$$V_C = \{(\mathbf{x}^{(q)}, \mathbf{y}^{(q)}) \in X_C \times Y_C, 1 \le q \le N_{V_C} : \mathbf{y}^{(q)} = f(\mathbf{x}^{(q)})\},$$
$$T_C = \{(\mathbf{x}^{(r)}, \mathbf{y}^{(r)}) \in X_C \times Y_C, 1 \le r \le N_{T_C} : \mathbf{y}^{(r)} = f(\mathbf{x}^{(r)})\},$$
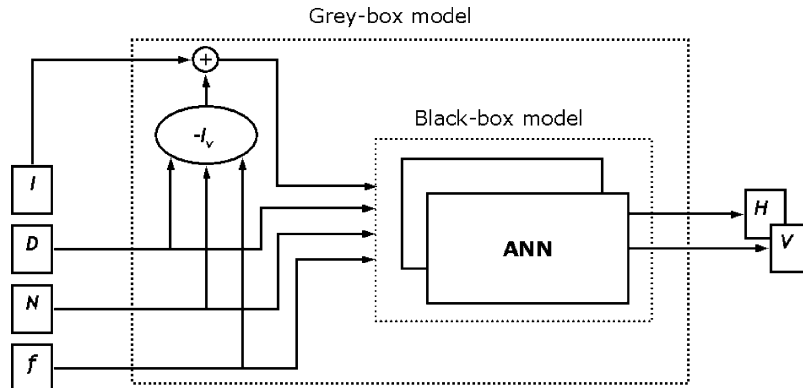

Figure 4   Hybrid artificial neural network.

such that: $L_C \cap V_C = \emptyset$, $V_C \cap T_C = \emptyset$, $T_C \cap L_C = \emptyset$ and $L_C \cup V_C \cup T_C = X_C$.

Let us next consider a prediction model $\hat{f}_d(\mathbf{x}; \mathbf{w}_d | \theta_d)$. The *learning step* consists in finding the model parameters $\mathbf{w}_d$ minimising a well-defined error criterion computed according to the learning data set L_C, while the complexity θ_d of the model is fixed. For the sake of simplicity, let us consider the mean square error:

$$MSE_{L_C,d} = \frac{1}{N_{L_C}} \sum_i^{N_{L_C}} (f_d(\mathbf{x}^{(i)}) - \hat{f}_d(\mathbf{x}^{(i)}; \mathbf{w}_d | \theta_d))^2,$$

where d denotes either the horizontal or the vertical direction:

$$d \in \{H, V\}.$$

During the training both directions are considered independently. The neural networks for both outputs are therefore trained separately.

The learning step can then be rephrased as follows:

$$\mathbf{w}_d^{(opt)} = \underset{\mathbf{w}_d}{\operatorname{argmin}} \, MSE_{L_C,d}.$$

Accordingly, during the learning step, we compute the optimal model parameters $\mathbf{w}_d^{(opt)}$ for a selected model $\hat{f}_d$, θ_d being fixed. Therefore, $\mathbf{w}_d^{(opt)}$ depends on θ_d. The search for the optimal $\mathbf{w}_d^{(opt)}$ is repeated for all θ_d values considered. In the following we omit the dependency in θ_d in order to ease the mathematical notations.

In the *validation step*, we determine the most appropriate model, i.e. we search for the model $\hat{f}_d(\mathbf{x}; \theta_d | \mathbf{w}_d^{(opt)})$ that shows the best generalisation on a new data set. This step consists thus in selecting the appropriate model $\hat{f}_d$ by minimising an error criterion, the optimal model parameters $\mathbf{w}_d^{(opt)}$ being known for each complexity θ_d. The error criterion is computed on the validation data set V_C. Again, let us consider the mean square error:

$$MSE_{V_C,d} = \frac{1}{N_{V_C}} \sum_i^{N_{V_C}} (f_d(\mathbf{x}^{(i)}) - \hat{f}_d(\mathbf{x}^{(i)}; \theta_d | \mathbf{w}_d^{(opt)}))^2.$$

Here the mean square error can be viewed as a prediction error, since it measures to which extend a model $\hat{f}_d$ predicts the response value of a future, not yet learned observation.

The model selection can be formulated as

$$\theta_d^{(opt)} = \underset{\theta_d}{\operatorname{argmin}} \, MSE_{V_C,d}.$$

During the last stage of the training algorithm, *the test step*, the performances of the selected model are estimated by calculation of the error criterion on the test data set T_C:

$$MSE_{T_C,d} = \frac{1}{N_{T_C}} \sum_i^{N_{T_C}} (f_d(\mathbf{x}^{(i)}) - \hat{f}_d(\mathbf{x}^{(i)} | \mathbf{w}_d^{(opt)}, \theta_d^{(opt)}))^2.$$

The procedure we have exposed here is of general interest. However, in our work we have omitted the last step for practical reasons. As a matter of fact, the data acquisition of the phosphene characteristics turns out to be fastidious and slow. Hence, we have favoured the approach that avoids defining a data set for the test. This enables us to have a greater amount of data at our disposal during the learning step and therefore enhance the approximation of the unknown process. Moreover, in practice there is only a slight difference between the $MSE_{V_C,d}$ and the $MSE_{T_C,d}$ for the optimal model parameters $\mathbf{w}_d^{(opt)}$ and the optimal model selection parameter $\theta_d^{(opt)}$, since the prediction error on the validation set V_C is a good estimator of the test error.

For further improvement of the prediction error estimates cross-validation is used. A practical problem often encountered with ANNs, is that they appear to be unstable predictors [14]. Indeed, during training they get easily trapped into local minima.

4.2.2. K-fold cross-validation
Cross-validation is a standard statistical tool for approaching the true value of prediction errors. It allows us to obtain a more realistic guess of the model selection by applying the training procedure to a number of different data splits and by averaging out the calculated prediction errors.

Suppose we split the data set at random in K roughly equal sized parts. The method can be described as follows [15]:

(1) Repeat for $k = 1, \ldots, K$:
   ○ define the validation data set V_C as the kth part of the split and the learning data set L_C as the remaining $K - 1$ parts (the test set T_C is empty in our case);
   ○ calculate the prediction error $E_k^{(P)}$ on V_C of the model fitted on L_C.
(2) Combine the K prediction errors by computing the cross-validation error:

$$CV^K = \frac{1}{K} \sum_{k=1}^K E_k^{(P)}.$$

Such a method makes it possible to use a high proportion $(1 - 1/K)$ of the available data to train the networks, while also making use of all the

data points when evaluating the cross-validation error. A disadvantage of the approach is that it requires the training process to be repeated K times, which can lead to large processing times.

4.2.3. Linear regression
A linear regression with $n + 1$ inputs can be defined as follows:

$$\hat{y}_d^{(i)} = \sum_{j=1}^{n+1} w_{d,j} x_j^{(i)}, \quad i = 1, 2, \ldots, N_{L_c},$$

where n is the dimension of the stimulation parameter space. In this equation N_{L_c} is the number of training data and $x^{(i)}$ the input vector. By convention, $x_1^{(i)} = 1$ and the corresponding parameter $w_{d,1}$, called the bias, represents the independent term in the model.

This linear method is exposed as reference, to validate the approximation improvement achieved by the non-linear ones, as discussed in Section 5.

4.2.4. Multi-layer perceptron
Linear regressions can be considered as single-layer networks, which can be extended to multiple-layer networks. Whereas single-layer networks have limited applications, MLPs have the universal approximation property [16] and, therefore, can approach any kind of process provided the network contains a sufficient number of parameters and the number of data available tends to infinity. Since a restricted number of layers reduces the complexity as well as the training time, a minimum number of layers is suitable.

A 2-layer perceptron can be described as follows [17,18]:

$$\hat{y}_d^{(i)} = h\left(\sum_{l=1}^{\theta_d+1} w_{d,l}^{(2)} \cdot g\left(\sum_{j=1}^{n+1} w_{d,lj}^{(1)} x_j^{(i)} \right) \right),$$

$$i = 1, 2, \ldots, N_{L_c},$$

where θ_d is the number of nodes, $h(\cdot)$ and $g(\cdot)$ the so-called activation functions and $w_{d,lj}^{(1)}$ and $w_{d,l}^{(2)}$ the weights. The activation function of the output units, $h(\cdot)$, is chosen, in our case, as the identity function in order to avoid saturation of the outputs; the internal one, $g(\cdot)$, is a sigmoidal function.

The MLPs used in our experiments have been trained according to the Levenberg–Marquardt algorithm [19]. This algorithm is specifically designed for minimising a sum-of-squares error and is an extension of gradient descent methods. Unfortunately, since the error surface is a highly non-linear function of the weights, we get easily trapped into local minima. Therefore, a good initialisation of the weights is important.

4.2.5. Radial basis function network
While an MLP approximates any type of function by combining sigmoids, an RBFN approximates functions by combining radial basis kernels. The universal approximation property holds for RBFN as well [20].

Suppose we want to approximate the physiological process by a set of θ_d basis kernel $\varphi(\cdot)$. The corresponding RBFN can be characterised by [17,18]:

$$\hat{y}_d^{(i)} = \sum_{j=1}^{\theta_d} w_{d,j} \varphi(\| \mathbf{x}^{(i)} - \mathbf{c}_{d,j} \|), \quad i = 1, 2, \ldots, N_{L_c},$$

(3)

where $\mathbf{c}_{d,j}$ is the centre of the radial basis kernels j. The most commonly used type of basis function is the Gaussian kernel:

$$\varphi(\| \mathbf{x}^{(i)} - \mathbf{c}_{d,j} \|) = \exp\left(-\frac{\| \mathbf{x}^{(i)} - \mathbf{c}_{d,j} \|^2}{2 q_d \sigma_{d,j}} \right),$$

$$i = 1, 2, \ldots, N_{L_c},$$

where $\sigma_{d,j}$ is the standard deviation of kernel j, whereas q_d the width-scaling factor [21].

The training of an RBFN is decoupled in a three-stage procedure. Firstly the centres $\mathbf{c}_{d,j}$, of the radial basis function are computed by a vector quantisation method. The vector quantisation method we have selected is a competitive learning scheme (see, for example [22]). Secondly the standard deviation $\sigma_{d,j}$ of each cluster is estimated by computing the standard deviation of the distance between the data and their corresponding centres. In addition an overlapping of the Gaussian kernels is forced by learning the width-scaling factor q_d in order to improve the generalisation process by smoothing the approximation model. Indeed, there is no reason for $q_d = 1$ to come out to be the best choice. Finally the last step consists in computing the weights $w_{d,j}$ by solving the set of equations (3). Those equations are linear since the radial basis functions are fixed after the two first stages of training.

5. Results

At first, consider the hybrid prediction models, i.e. a hybrid linear regression, a hybrid MLP and a hybrid RBFN. Their input vectors are the stimulation parameters, the current intensity being modified according to Eq. (2):

$$\mathbf{x}^{(i)} = [I_{useful}^{(i)}, D^{(i)}, N^{(i)}, \lambda^{(i)}], \quad i = 1, 2, \ldots, N_{L_c}.$$

Both hybrid ANNs are trained using a 10-fold cross-validation scheme. The cross-validation is repeated 15 times with randomised data sets, in order to

reduce the variance of the results due to sampling. The parameters computed according to this procedure are the number of nodes for the MLP and the number of Gaussian kernels as well as the width-scaling factor for the RBFN.

Separate models were trained for each electrode contact and in each direction. As we mentioned before, both directions were assumed independent. The four contacts were also assumed independent since they were separately electrically activated. Additionally, separating the prediction models makes it possible to view in which areas of the blind's visual field phosphenes are likely to be generated when a certain electrode contact is active.

In order to demonstrate our approach, let us consider for example the electrode contact 90°. In Figs. 5 and 6, we have illustrated the corresponding estimated prediction error curve for the hybrid MLP and the hybrid RBFN respectively. Remark that the learning and the validation data sets L_C and V_C are normalised according to the learning data set L_C, the prediction error being thus unitless.

The estimated prediction error was computed as follows:

$$\hat{E}^{(P)} = \frac{1}{15}\sum_{j=1}^{15} CV_j^{10},$$

where CV_j^{10} is the jth of the 15 10-fold cross-validations.

As discussed in the previous sections, the estimated prediction error $\hat{E}^{(P)}$ depends on the number of nodes θ for the MLP; it depends on the number of Gaussian kernels θ and on the width-scaling factor q_d for the RBFN. Therefore, in the case of the RBFN, the plotted error is the one obtained for the optimal width-scaling factor q_d.

For each electrode contact, the optimal prediction error is comparable for both ANN. Nevertheless, the prediction accuracy is better in the vertical direction. This is probably due to the fact that the observed range of the non-normalised phosphene positions is much larger in the vertical than in the horizontal direction. The measurements in horizontal direction are thus more sensitive to relative inaccuracies.

While the averaged optimal prediction error is comparable for both networks, the complexity of the MLP is smaller than the complexity of the RBFN. Though, the standard deviation of the MSE is greater in the case of the MLP (Figs. 7 and 8); this suggests a greater instability in the learning. Moreover, it should be stressed that some results were left out for the MLP: in some cases the network did not converge properly and reached a bad local minimum;
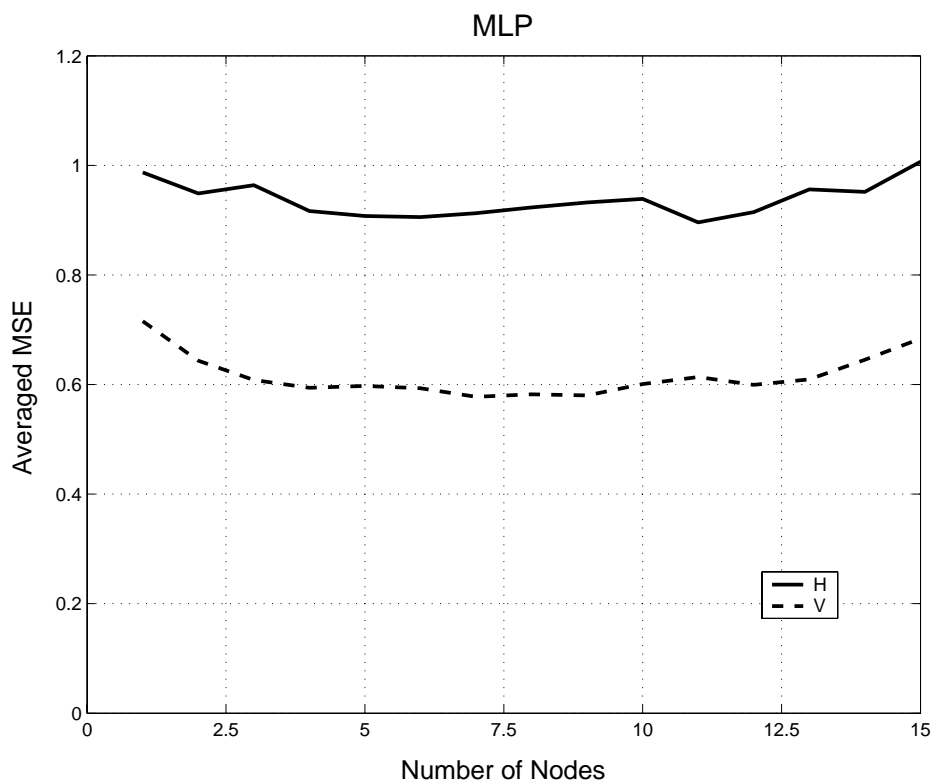

Figure 5 Averaged mean square error of the hybrid MLP for the electrode contact 90° after 15 10-fold cross-validations. H and V stand, respectively, for the horizontal and vertical directions.
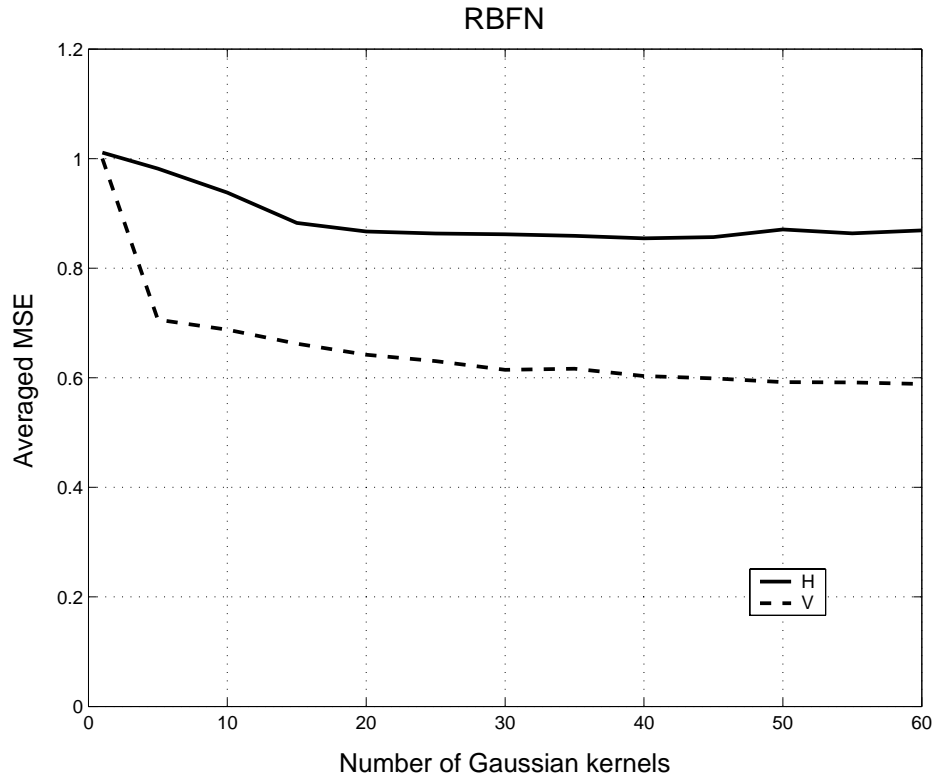
Figure 6 Averaged mean square error of the hybrid RBFN for the electrode contact 90° after 15 10-fold cross-validations. *H* and *V* stand, respectively, for the horizontal and vertical directions.
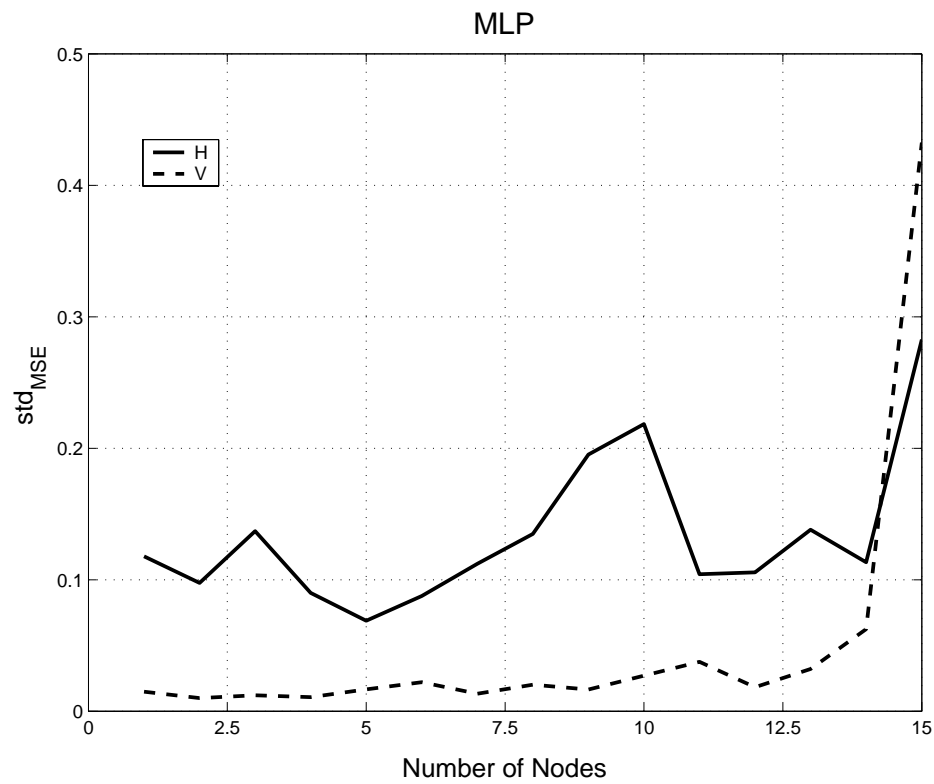

Figure 7 Standard deviation of the mean square error of the hybrid MLP for the electrode contact 90° and 15 10-fold cross-validations. *H* and *V* stand, respectively, for the horizontal and vertical directions.

Figure 8 Standard deviation of the mean square error of the hybrid RBFN for the electrode contact 90° and 15 10-fold cross-validations. *H* and *V* stand, respectively, for the horizontal and vertical directions.
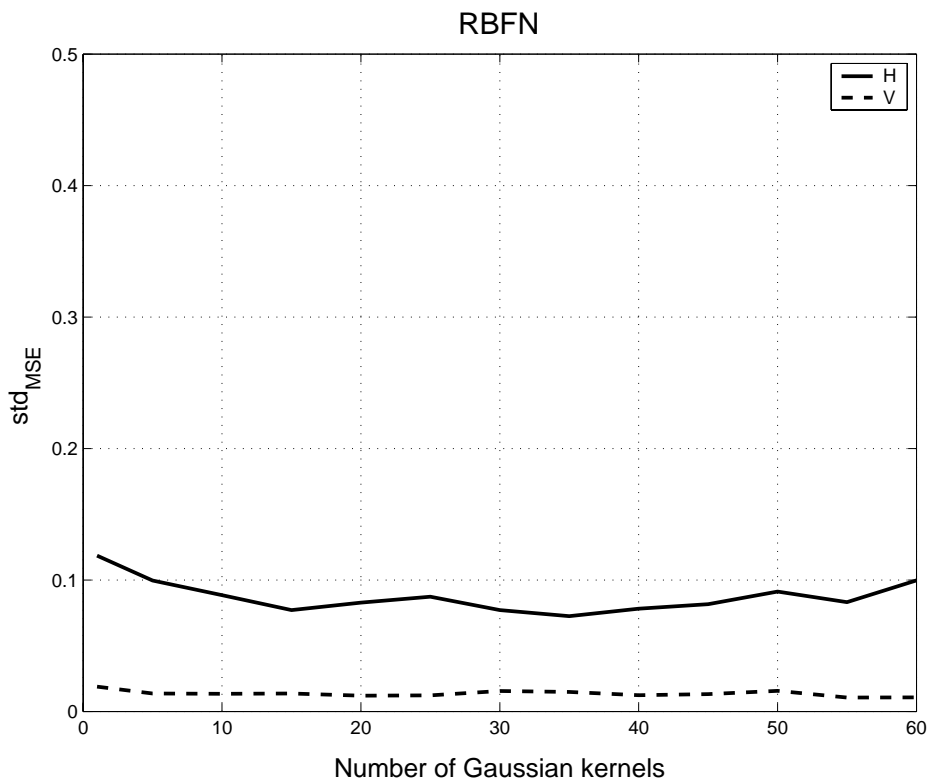
as a consequence the MSE was not representative of the performance that could be achieved by the MLP. For the RBFN no convergence problem was found.

Looking at Fig. 8, one can notice that for the RBFN the prediction error does not increase by overfitting after a certain number of kernels, as is usually the case in other applications. This can be rationalised as follows. The universal approximation property holds for MLP and RBFN provided that the number of data tends to infinity, as well as the number of hidden units, that is the nodes in MLPs or the number of Gaussian kernels in RBFNs. In practice, however, the number of available data is limited. In order to avoid overfitted data, the number of hidden units should be limited as well. Remarkably, in this case, overfitting does not seem to occur for the RBFN. Actually, when too many centres are used, the weights of the corresponding radial basis kernels are set to merely zero. As a result the performances of the RBFN are not sensitive to the number of Gaussian kernels, provided that a sufficient number of them is taken. A good choice is the number corresponding to the end of the elbow, since a further increase would not improve significantly the quality, whereas the complexity would grow drastically.

Let us now consider the black-box modelling, i.e. classical linear regression, classical MLP and RBFN. The input vectors are in this case:

$$\mathbf{x}^{(i)} = [I^{(i)}, D^{(i)}, N^{(i)}, \lambda^{(i)}], \quad i = 1, 2, \ldots, N_{L_c}.$$

The prediction performances of the three of them are very similar to the performances of their corresponding hybrid counterparts (Table 1).

Let us next discuss Table 1. In order to evaluate the performances of the prediction models we have to compare the prediction accuracy to the spreading of the phosphenes in their respective domain.

Table 1 Prediction accuracy (°) of the grey-box and black-box statistical ANN models

	Contact (°)			
	0	90	180	270
Dispersion of the data	15.6	15.3	16.6	14.7
Hybrid linear regression	15.4	14.0	15.8	14.8
Hybrid MLP	12.7	12.3	13.2	11.7
Hybrid RBFN	13.0	12.1	13.4	11.4
Linear regression	15.4	14.1	15.6	14.9
MLP	12.9	12.0	13.0	11.8
RBFN	12.8	12.1	13.1	11.2

Table 2 Coefficients of determination of the linear regressions between the measured phosphenes' positions and the predicted ones by means of the black-box statistical ANN models for the electrode contact 90°

	r_H^2	r_V^2
Linear regression	0.001	0.264
MLP	0.278	0.635
RBFN	0.345	0.575

The prediction accuracy can be measured by considering the root of the MSE computed on the position vectors **y**. Next, let us define the dispersion of the phosphenes belonging to the contact C as the root of the sum of the variances in horizontal and vertical direction. The dispersion measures the scattering of the phosphenes in the visual field of the blind volunteer.

When we look at the prediction accuracy of the linear statistical methods, we notice that their value is approximately equal to the dispersion of the phosphenes. This indicates that the linear methods are only able to predict the location of new phosphenes by the mean of the learning data sets. As a result, the linear methods are ineffective referred to a prediction by the mean. On the other hand, when considering the non-linear statistical methods, we clearly observe a prediction improvement of roughly 25% on average for the four electrode contacts. This result suggests, firstly, that the features of the visual sensations can be predicted with a certain accuracy and, secondly, that the generation process is non-linear.

Still, the prediction accuracy can be considered as relatively modest at first sight. However when computing a linear regression between the predicted and experimentally measured positions, substantial correlations are found. To illustrate this, consider again the electrode contact 90°. In Table 2, we have presented the coefficients of determination r_H^2 and r_V^2 of the linear regressions between the measured locations and the locations predicted by the black-box models, respectively in horizontal and vertical directions. Obviously, a significant improvement is achieved by using non-linear methods, and a relatively high proportion of variation is explained, leading to a fair relative prediction accuracy. For the other electrode contacts and/or the grey-box models we observe similar results.

6. Conclusion

Within the framework of the OPTIVIP project, we made an attempt to decode and model a complex neurophysiological process linking the stimulating parameters of an optic nerve based visual prosthesis to the visual sensations generated in the visual field of blind RP patients. Adaptive neural techniques were proposed to estimate the characteristics of the visual sensations.

A grey-box model, incorporating physiological knowledge, as well as a complete black-box model were investigated. They both provide tools for predicting the features of phosphenes and achieve a comparable accuracy. Reminding that the observed values to approximate result from human response, which is definitely not accurate, the prediction accuracy found with the non-linear models is satisfactory. As a result, it is expected that the predictions obtained with such models will be used for image reconstruction in the optic nerve based visual prosthesis.

Finally, we show that the prediction performances of the ANNs are superior to the performances achieved by classical linear statistical methods, suggesting that the physiological process is strongly non-linear.

In our future work, we plan to further develop neurophysiological aspects of the modelling in order to whiten the black-box parts of the model and acquire a finer understanding of underlying processes. Moreover, improved interactions within the grey-box models will be investigated in order to further develop the prediction capabilities.

Acknowledgements

References

[1] Archambeau C, Lendasse A, Trullemans C, Veraart C, Delbeke J, Verleysen M. Phosphene evaluation in a visual prosthesis with artificial neural networks. In: Proceedings of the European Symposium on Intelligent Technologies, Hybrid Systems and their implementation on Smart Adaptive Systems, Puerto de la Cruz, December 13—14, 2001. Spain: Tenerife; 2001, pp. 509—15.

[2] Norman RA, Maynard EM, Guillory KS, Warren DJ. Cortical implants for the blind. IEEE Spectr 1996;33(5):54.

[3] Humayun MS, de Juan E, Dagnelie G, Greenberg RJ, Prost RH, Phillips DH. Visual perceptions elicited by electrical stimulation of retina in blind humans. Arch Ophtalmol 1996;114(1):40—6.

[4] Rizzo JF, Wyatt J. Prospects for a visual prosthesis. The Neurosci 1997;3(4):251—62.

[5] Veraart C, Raftopoulos C, Mortimer JT, Delbeke J, Pins D, Michaux G et al. Visual sensations produced by optic nerve stimulation using an implanted self-sizing spiral cuff electrode. Brain Res 1998;813:181—6.

[6] Delbeke J, Wanet-Delfalque MC, Gérard B, Troosters M, Michaux G, Veraart C. The microsystems based visual prosthesis for optic nerve stimulation. Artif Organs 2002; 26(3):232—4.

[7] Doguet P, Mevel H, Verleysen M, Troosters M, Trullemans C. An integrated circuit for the electrical stimulation of the optic nerve. In: Proceedings of the Conference IFESS'2000, June 18—19, 2000. Denmark: Aalborg; 2000, pp. 18—20.

[8] Official website of the European project OPTIVIP. http://www.dice.ucl.ac.be/optivip/.

[9] Brindley GS, Lewin WS. The sensations produced by electrical stimulation of the visual cortex. J Physiol-Lond 1968;196:479—93.

[10] Dobelle WH, Mladejovsky MG, Evans JK, Roberts TS, Girvin JP. ''Braille'' reading by a blind volunteer by visual cortex stimulation. Nature 1976;259:111—2.

[11] Delbeke J, Parrini S, Glineur O, Vanlierde A, Veraart C. Phosphene perception thresholds to direct stimulation of a human optic nerve shows spatial and temporal summation. Soc Neurosci Abstr 1999;25:1040.

[12] Delbeke J, Parrini S, Michaux G, Vanlierde A, Veraart C. In: Proceedings of the Fifth Annual Conference of the International Functional Electrical Stimulation Society on Perception Threshold Changes in Phosphenes Generated by Direct Stimulation of a Human Optic Nerve, June 18—19, 2000. Denmark: Aalborg; 2000, pp. 152—155.

[13] Hill AV. The strength—duration relation for electric excitation of medullated nerve. Proc Roy Soc Lond Ser B 1936; 119(815):440—53.

[14] Cunningham P, Carney J, Jacob S. Stability problems with artificial neural networks and the ensemble solution. Artif Intell Med 2000;20:217—25.

[15] Efron B, Tibshirani RJ. An Introduction to the Bootstrap. Boca Raton: Chapman & Hall/CRC Press LLC; 1998.

[16] Cybenko G. Approximation by superposition of sigmoidal functions. Math Contr Signals Syst 1989;2(4):303—14.

[17] Bishop CM. Neural Networks for Pattern Recognition. Oxford: Oxford University Press; 1995.

[18] Haykin S. Neural Networks—A Comprehensive Foundation. Upper Saddle River: Prentice-Hall; 1999.

[19] Marquardt DW. An algorithm for least-squares estimation of non-linear parameters. J Soc Ind Appl Math 1963;11(2): 431—41.

[20] Poggio T, Girosi F. A theory of networks for approximation and learning. A.I. Memo No. 1140. Artificial Intelligence Laboratory, Massachusettes Institute of Technology; 1989.

[21] Benoudjit N, Archambeau C, Lendasse A, Lee JA, Verleysen M. Width optimization of the Gaussian kernels in radial basis function networks. In: Proceedings of the European Symposium on Artificial Neural Networks, April 24—26, 2002. Belgium: Bruges; 2002, pp. 425—32.

[22] Kohonen T. Self-organizing Maps. Berlin: Springer; 1995.