Interactive evolution for cochlear implants fitting

Pierrick Legrand [1,2], Claire Bourgeois-Republique [4], Vincent Péan [5], Esther Harboun Cohen [6], Jacques Levy-Vehel [2], Bruno Frachet [6], Evelyne Lutton [2], Pierre Collet [3]

[1] *IMB, Institut de Mathématiques de Bordeaux, UMR CNRS 5251, Université de Bordeaux 2, 146 rue Léo Saignat, 33076 Bordeaux cedex, France*
Legrand@sm.u-bordeaux2.fr

[2] *COMPLEX Team - INRIA Rocquencourt, B.P. 105, 78153 Le Chesnay cedex, France*
Pierrick.Legrand@inria.fr, Evelyne.Lutton@inria.fr, Jacques.levy-vehel@inria.fr
`http://fractales.inria.fr`

[3] *Laboratoire d'Informatique du Littoral, ULCO BP719, 62100 Calais cedex, France*
Pierre.Collet@univ-littoral.fr
`http://lil.univ-littoral.fr`

[4] *LE2I, UMR 5158 CNRS, 9 avenue A. Savary, B.P. 47870, 21078 Dijon cedex, France*
Claire.bourgeois-Republique@u-bourgogne.fr

[5] *CRT Innotech, 1 Promenade Jean Rostand, 93005 Bobigny cedex, France*
vincent.pean@innotech.fr

[6] *Hôpital Avicenne, Service ORL, 125 rte de Stalingrad, 93000 Bobigny, France*
bruno.frachet@avc.ap-hop-paris.fr

Abstract. Cochlear implants are devices that become more and more sophisticated and adapted to the need of patients, but at the same time they become more and more difficult to parameterize. After a deaf patient has been surgically implanted, a specialised medical practitioner has to spend hours during months to precisely fit the implant to the patient. This process is a complex one implying two intertwined tasks: the practitioner has to tune the parameters of the device (optimisation) while the patient's brain needs to adapt to the new data he receives (learning). This paper presents a study that intends to make the implant more adaptable to environment (auditive ecology) and to simplify the process of fitting. Real experiments on volunteer implanted patients are presented, that show the efficiency of interactive evolution for this purpose.[1]

[1] This work has partially been funded by the French ANR - RNTS HEVEA project 04T550

1 Introduction

Cochlear Implants (CI) [Nih89] allow totally deaf people to hear again provided their auditory nerve and cochlear are still functional: a computer processes sounds picked up from a microphone, to stimulate directly the auditory nerve through several electrodes inserted inside the cochlea (cf. fig. 1).

As one can imagine, there are hundreds of parameters that can be tuned, and in the same time the patient has to learn to "hear" using new informations provided to his auditory nerve. The tuning of such a device is thus extremely complex, and highly dependent on the patient. This process is currently done "by hand" by medical practitioners, and looks like an optimisation process based on "trial and error." This process is so delicate that sometimes, no satisfactory fitting can be found for some patients.

Hence, it seems interesting to use an interactive evolutionary algorithm (IEA) to help finding the best values for implant parameters. This is the main topic of the HEVEA project, which is a collaboration between computer scientists, signal processing experts and medical researchers. The aim is actually twofold: to facilitate the initial fitting of cochlear implants, and to automatise the adaptation of cochlear implants to various sound environments. A simple IEA was developed with this in mind, and tested on a very basic feature, the range of intensities that a specific electrode can take when stimulating the auditory nerve. The IEA has been implemented on a PDA and tests have been performed on volunteering patients with satisfying results.

The paper is organised as follows: section 2 presents cochlear implants, and section 3 describes how they are currently tuned by medical practitioners. The approach of the HEVEA project is developed in section 4, and a first implementation of an IEA is detailed in section 5. Experiments on several patients are reported in section 6, yielding good results as well as important conclusions on manual fitting procedures. This first validation step is important: an analysis of the success and failures raises new questions that are developed in section 7, related to the well-known "user fatigue" problem of IEAs, and to the fact that different sound environments have an important influence on the fitting of implants. Automatic adaptation of the device to sound has been investigated, based on a sound signal classification scheme, which is detailed in section 7. Conclusions and perspectives are described in section 8.

2 Cochlear Implants

A cochlear implant is a surgically implantable device [GFM+98] that provides hearing sensations to individuals with severe to profound hearing loss, who cannot benefit from hearing aids. In a normal ear, sound energy is converted to mechanical energy by the middle ear, which is then converted to electrical impulses by the inner ear (see figure 1). In order to perform this last stage, the cochlea (part of the inner ear) contains a fluid which is set into motion by the oval window which is connected to the middle ear. Within the cochlea, sensory cells (inner and outer hair cells) are sensitive transducers that convert the mechanical fluid motion into electrical impulses conveyed to the brain by the auditory nerve. Cochlear implants are designed to be a substitute for the middle ear, cochlear mechanical motion, and sensory cells, transforming directly sound energy into electrical energy that will initiate impulses in the auditory nerve [Moo95], [Coh89] thanks to a digital signal processor.

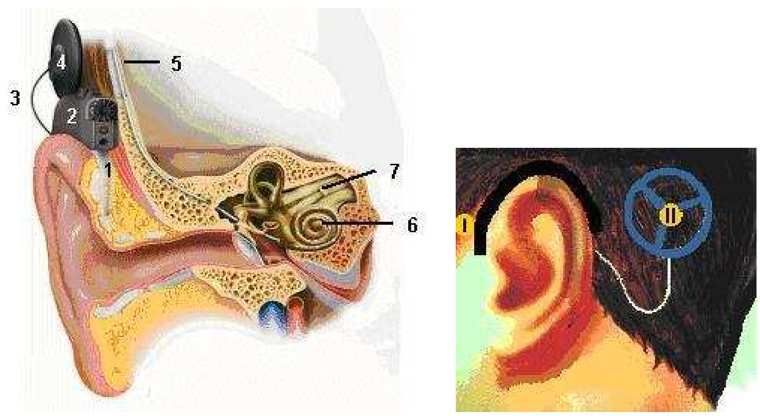

Fig. 1. All implant devices have the following features in common : sound is collected by a microphone (1) and sent to electronic components within a speech processor (2). The speech processor analyzes the input signal (sound) and converts it into an electronic signal (electrical). This code travels along a cable (3) to the transmitting coil (4) and is sent across the skin via frequency modulated (FM) electro-magnetic waves to the implant package (5). Based on characteristics of the code transmitted to the internal device, electrode contacts within the cochlea (6) provide electrical stimulation to the spiral ganglion cells and dendrites extending into the modiolus. Electrical impulses then travel along the auditory nerve (7), ascending auditory pathways to the brain.

Cochlear implants have been very successful in restoring partial hearing to profoundly deaf people [ALM95], [Osb97]. In 2006, around 70 000 deaf people are implanted with such devices around the world. Efficiency is quite variable, ranging from totally deaf patients that have fully recovered their audition and are capable to follow telephone conversations and enjoy music, to others who hear strange sounds they can't benefit from, to a point where they prefer to switch off the implant [COM94], [GTBVC01], [BTE04], [Rom98].

For many people, it is still difficult to fully take advantage of the device because it is not easy to tune the parameters of digital signal processor and adjust them for the characteristics for each patient, since all patients are different (cause of deafness, number of years between total deafness and implantation, age, depth of electrode insertion,. . .).

Research has been going on since nearly 50 years ago on how to electrically stimulate the auditory nerve to give a totally deaf patient sound sensations

[LPD00,Loi01]. Even though the early devices stimulated the auditory nerve with one electrode only, some lucky patients managed to hear again and even understand speech. Nowadays, it is technologically possible to use more than one electrode, in order to stimulate more of the thousands of neurons the auditory nerve is made of [PCMF79] [CFML83]. However, the more electrodes, the more parameters to tune.

The cochlea is used to interface electrodes and the auditory nerve. The cochlea is a biological device that mainly allows to map different sound frequencies onto different neurons. It is shaped like a snail shell. Only long wavelengths (low frequency sounds) can reach the far end of the cochlea, while short wavelengths (high frequency sounds) are stopped at the entrance of the cochlea. The idea is then for surgeons to use this frequency discriminator and insert into the cochlea a thin silicon wire, bearing an array of ring-shaped electrodes.

The array of electrodes (cf. fig. 1/6) is then connected to an antenna inserted under the skin of the patient, in a cavity created by the surgeon in the skull bone of the patient, just above his external ear (cf. fig. 1/5). On the outer side of the skin, the patient wears another antenna (centered over the inner antenna thanks to a powerful magnet (cf. fig. 1/4)) that is itself connected to a digital signal processor (DSP) which uses a microphone as input(cf. fig. 1/2).

When a sound is received by the microphone, the DSP processes it and sends electrical impulses to the external antenna, that are received by induction by the implanted antenna. when the microphone picks up a low frequency sound, the DSP will stimulate electrodes introduced deeply in the cochlea (that will make the patient hear a low pitch sound) while on the contrary, high pitch sounds received by the microphone will have the DSP stimulate electrodes closer to the entrance of the cochlea (that will make the patient hear a high pitch sound).

3 Cochlear Implant fitting

3.1 Complexity of the problem

Being able to use more than one electrode to stimulate different neuron areas is indeed a great improvement, but the number of parameters to tune increases drastically. Concerning electrodes only, many questions arise, among which:

- Which frequencies should be mapped to which electrodes ?
- Which range of intensities should be applied to which electrodes ?
- How many electrodes should be stimulated simultaneously ?
- Should the processor prohibit neighbour electrodes to be stimulated simultaneously in order to avoid diaphony (crosstalk between nearby electrodes) ?

Finding good answers to these questions is a difficult optimisation problem. This not only due to the extremely large size of the search space but to several other reasons. First of all, the quality of a fitting is a two stage process where subjectivity plays a large role: the practitioner has to interpret the quality of the fitting (second subjective process) from the answers given by the patient (first subjective process). The disparity of patient behaviour with respect to language and sensitivity to various thresholds, as well as the character of the practitioner deeply influences the results. For example the well known psychological "Pygmalion" effect biases answers of the patient, who often unconsciously tries to satisfy the practitioner's expectations.

The sound environment is another cause of variability of results, as the fitting session usually takes place in a small room at hospital with the practitioner. However the cochlear implant must also be used in real life, and a correct fitting at hospital may reveal very uncomfortable or unuseful when in the street, or in a restaurant.

Fatigue and brain adaptation are also other sources of trouble: it is impossible to test many possible parameter sets during a single session, so the process is very long and needs sometimes weeks to obtain a satisfying result. In the same time, a fitting that may not appear immediately as satisfying, may improve when testing it on a longer period (brain has a plasticity that cannot be neglected).

There are many factors that make this problem highly irregular. However, it has been proved that an acceptable or even good fitting is reachable by a manual search conducted by an experienced practitioner. We describe below this manual fitting technique, which is mainly a human-guided "trial and error" process, resembling a local search.

3.2 Manual fitting

Nowadays, depending on the manufacturer, the number of electrodes varies between 8 and 22. Cochlear implant "fitting" is performed by an expert practitioner, who proceeds in the following way:

- Right after the surgical intervention, the practitioner tries to determine which electrodes are functional (an electrode is functional if the patient hears a sound when current is applied to the electrode).
- For each functional electrode, the practitioner tries to determine the range of intensities that can be used. The lowest intensity above which the patient perceives a sound is called T (for Threshold). The maximum comfortable intensity (loudest sound the patient can bear for a reasonable amount of time) is called C (for Comfort threshold).
  Determining the T and C values for each electrode takes time (communication with a deaf patient, a young child, or with an old patient can be difficult), and due to the increasing number of electrodes, some manufacturers now advise to determine T and C values for one every three or four electrodes, and extrapolate the values for the other electrodes. See [Rou01], [Hes02] for more informations on this topic.
  Other manufacturers even set average values for T and C, based on neural response or even statistics.
- Then, once the $C - T$ range is maximised for all the electrodes, the "real" fitting begins. The practitioner uses his expertise to map frequency bands logarithmically onto the different functional electrodes, and starts to tune

the gain and sensitivity depending on sound frequencies, then tunes the number of simultaneously active electrodes,... while at the same time asking the patient whether they understand better or worse, whether the sound quality is comfortable or not, a.s.o.. Of course, misunderstandings between the practitioner and the deaf patient may occur here (elderly patients, small children, ...) that may affect the quality of the fitting. In certain cases, the practitioner will slightly reduce the $C-T$ range for some electrodes, when he has the feeling that the "neurologic" bandwidth is limited, and that the neurons facing the electrode are getting saturated at only moderate auditory levels.

Results are variable, but often good. Usually, a fitting session starts with the practitioner asking whether the current fitting is better or worse than the previous one. The best of the recent fittings is taken as a basis that the practitioner will try to improve, resulting in some sort of hill climbing process.

The patient tries to describe the quality of his audition, and the practitioner tries to modify some parameters to help solving the problems. Two or three parameters can be changed during a 30 to 90 minutes fitting session. Then, the patient leaves with the new settings that he keeps for a couple of months, before he comes back for another fitting session. The whole process is therefore very long (several years for problematic patients).

4 Description of the Problem

As seen above, fitting cochlear implants is done through a set of correlated parameters [LPD00], and perception and comfort thresholds are linked to histopathological factors specific to the patient [KSC+98]. In most cases, the fitting strategy simply consists in maximising the number of electrodes and maximising their dynamic range [BPG+92]. This often gives good results, but for some patients this approach does not work. Moreover, the following observations have also been reported:

– Better results might be obtained by decreasing the dynamic range [FXP03].
– Only using a subset of electrodes might improve speech recognition [ZCW97].
– Holes in spectral representation can exist in tonotopic representation (mapping of the sound frequencies on the electrodes) and spectral information redistribution around the holes does not increases results [SGD02].

Moreover:

– Most of the patients do not use all the information given by the electrodes [Fis96].
– All the electrodes are not necessary to obtain maximal speech perception performance in silent [DDML89,LWZF96,Fis96,KVR+00] and noisy environments [FSBW01] (part of this could be due to electrical interaction between channels [SLM+06]).

These published observations show that choosing a good subset of electrodes can have an influence on speech understanding, as well as the dynamic range on the electrodes. Finally, taking into account a real sound environment could increase speech understanding for some patients.

The work presented in this paper will try to address both the problems of choosing a good subset of electrodes, and taking into account a real life sound environment.

5 Description of the Interactive Evolutionary Algorithm

Before this work was started, several fitting sessions were observed, with patients who were not satisfied with their cochlear implant. During these experimentations, it really seemed that the fitting was stuck in a local optimum, since the expert's heuristics looked quite like what is known in computer science as a local search (trial of neighbours of the current best fitting) that would not bring any improvement.

This triggered the idea to use evolutionary algorithms, that are both quite good at optimising parameters and not easily trapped in local optima. The genetic loop is the following: the EA "suggests" a set of parameters that are directly uploaded into the Cochlear Implant's processor, and waits for an evaluation.

Other works have been conducted on interactively fitting hearing aids with evolutionary algorithms, [Dur02,Tak01], but they concern only conventional hearing aids, with a relatively small number of parameters that can be tuned. To our knowledge, nobody has tried to apply evolutionary algorithms to Cochlear Implants fitting.

5.1 Managing the runs

In an interactive evolutionary algorithm, a human user evaluates the different individuals proposed by the algorithm.

Thomas Bäck's results ([Bae05]), suggest that an evolutionary algorithm may do as well (if not better) than a human expert on a number of evaluations of the same order than the number of real parameters to optimise. Therefore, if the problem has around 100 parameters to tune, performing only 100 evaluations may allow to obtain interesting results. If it is possible to find an evaluation procedure that takes around 5mn, a run would last around 8 hours.

However, it is also important to take psychology and human fatigue into account: a well tuned convergence speed over 100 evaluations could seem discouraging for a human patient, who may think that improvement is too slow. Besides, since it is not possible to have an 8 hour run in one go, an elegant solution consists in fractioning the experimentation into several partial fast-converging runs, with a restart at the end of each run [Jan02]. Dividing the 8 hour run into 5 makes for 5 1h30 runs, that are quite manageable.

Rather than finding ways to avoid premature convergence, it is on the contrary a very fast convergence that is sought on these short runs of approximately

20 generations. This feature is easily obtainable with evolutionary algorithms, since they are known to converge quite fast, if no counter-measures are taken.

This policy allows to use a very fast converging algorithm trying to exploit local minima, rather than a slow converging algorithm trying to widely explore the search space, looking for the global minimum. The consequences of premature convergence are dealt with thanks to the periodical restarts. During the last run, one can restart the algorithm with the best individuals found in the 4 first runs, so as to benefit from the results previously found.

Population size and number of children per generation. For an identical number of evaluations, two possibilities exist: either many children per generation and a small number of generations, or a small number of children per generation and many generations.

Out of these two possibilities, it is the algorithm that maximises the number of generations that will favour most convergence. This suggests a *SteadyState* replacement policy, or a $(\mu+\lambda)$ with a very reduced λ (number of children) [Bae95]. Then in order not to spend too many evaluations in the initial population, one can also reduce it as is done in *micro-GAs* [Kri89].

Extremely low values can be used, such as 3 to 6 individuals for the initial population, with 1 to 3 children per generation. For the fifth run, 4 individuals could be used for the initial population, taken from the best individuals of the 4 previous runs.

The algorithm chosen for this specific interactive optimisation will therefore be a modern evolutionary algorithm, in the sense that it does not take after any of the four usual paradigms (Evolution Strategies, Genetic Algorithm, Genetic Programming, Evolutionary Programming) [DJ05].

According to Bäck [Bae05], using an Evolution Strategy paradigm for 100 evaluations should allow to optimise up to 100 real variables. In Cochlear Implants fitting, however, one can start with trying to find the best T and C values for each electrode. With the MXM 15 electrodes CI used for this experiment, the genome is therefore an array of only 30 real values, meaning that the chances to find a good fitting are much higher.

5.2 Initialisation

One hard constraint needs to be respected: the algorithm should not go beyond the maximum intensity for each of the electrodes for fear of destroying some of the patient's auditory neurons. Therefore, for each new patient, a first session with a practitioner is realised to determine the maximum admissible intensity for each electrode, that is called a *psychophysical test*. In order to reduce the search space, a minimal intensity below which the patient does not hear anything is also determined.

The initialisation of each individual therefore simply consists, for each of the 15 electrodes, to pick up two random values within the $[min, max]$ interval determined during the psychophysical test, and to take the lower value as a T threshold, and the higher value as a C threshold for the each of the 15 electrodes.

5.3 Selection of the parents

Parents selection is different from the replacement stage, in that it can select an individual several times. Whenever a child must be created, two different individuals are selected among the parent's population, that can be selected again to create another child.

Since the selection pressure of proportional selection depends on the fitness landscape of the problem to be solved (which is unknown), a stochastic tournament is selected [BT97], with a 90% probability, that consists in randomly selecting 2 individuals and to take the best of the two with a 90% probability.

5.4 Crossover

The genes are real values, which could have suggested some kind of barycentric crossover (such as used in Evolution Strategies), where each gene of the child is an average between the two genes of his parents. But since it is intervals that must be evolved, this type of crossover would have led to reducing the intervals progressively.

The chosen crossover is that of genetic algorithms, which exchange the parent's genes after a crossover point (*locus*) chosen randomly. A mono-point crossover was chosen, as a multiple crossover would have had a tendency to break efficient genomes, and would have turned the crossover in a kind of macro-mutation.

In this same attempt to not break good configurations, the determination of the locus is made electrode by electrode (the two T and C values are not separated). Since we are using a $(\mu + \lambda)$ evolutionary engine, with a number of children smaller than the size of the population, the crossover is called to create each child (100% probability).

5.5 Mutation

Mutation is also called with a 100% probability on each created child. In the proposed algorithm, each gene has a 10% probability to be mutated. Since there are 30 genes, each child undergoes 3 mutations in average. This may seem important, but due to the large epistasis, modifying a threshold on the global genome only has a limited influence on the global evaluation. This high mutation rate allow to keep a reasonable exploratory character to the algorithm, in spite of the very small number of evaluations.

5.6 Replacement

A *Steady State*-like replacement is used, i.e. with a very small number of children per generations, in order to promote a fast convergence. During a strict Steady State replacement, only one child would be created, that would replace the worst of both parents. Since we decided to have several children per generation, it is a $(\mu + \lambda)$ replacement scheme that is used, with only 2 or 3 children per generation (where Evolution Strategies usually create more children than there are individuals in the population).

5.7 Evaluation

It is possible to memorize 2 or 3 fittings on modern cochlear implant processors (called $P1, P2, P3$). Until this research was conducted, the evaluation of the patient's understanding was done by two different ways. Either the patient was sent home with the new fitting on $P1$ and the previous fitting on $P2$, which allowed him to compare both fittings in his environment, or an evaluation was done by an orthophonist with intensive tests during more than one hour.

Even though an interactive evolutionary algorithm requires a reduced number of evaluations [Tak98] none of these methods were suitable for an interactive evolutionary algorithm, so various evaluation protocols have been devised and will be described in details in section 6.

5.8 Execution

The evolutionary algorithm has been implemented both on a regular Personal Computer and on a PDA so that it is possible for a patient to tune his cochlear implant in a real environment (in a train station, for instance, if the patient works there and really needs a specific fitting for this particular environment).

The first versions have been implemented using the EASEA[2] [CLSL00] language in combination with the GALib library [Wal]. Later versions have been completely re-implemented from scratch in C++, because the GALib library has a bug that makes it ignore the evaluation of the initial population. Although this is not very important for applications where evaluations are easy to obtain, losing n interactive and tiring human evaluations was too high a cost in this special case.

6 Experiments

The first three sub-sections present results obtained with patient A, that were conducted by Claire Bourgeois-République, as part of her PhD. thesis of the Université de Bourgogne. These results have already been presented in several papers [BR04,BRVC05,BRFC05].

The following experiments have been conducted by Vincent Péan and Pierrick Legrand within the RNTS HÉVÉA project, funded by the French Ministry of Health.

6.1 Presentation of Patient A

Patient A has received an MXM cochlear implant 10 years ago in 1994. Unfortunately, he has not recovered a perfect audition (he understands some words quite well, but not others), although he is able to hold a conversation over the telephone, which is already quite a feat.

[2] http://sourceforge.net/projects/easea or http://complex.inria.fr/cgi-bin/twiki/view/Complex/SoftwareEASEA

He was initially given a waist processor (called *Boîtier*) to be carried attached to his belt, until MXM recently came up with a tiny "Behind The Ear" BTE processor. In 2003, patient A has received a BTE with the hope that new technology would allow him to hear better.

Unfortunately, this is not the case. After many disappointing fitting sessions with an expert practitioner, he still feels uncomfortable with the BTE and apparently cannot follow a conversation with it. He therefore keeps the BTE in a drawer and uses the old *Boîtier* for every day life.

The automatic fitting algorithm described in this paper was developed with the latest MXM technology, i.e. BTEs. It was thought that *Patient A* could be a good patient to test the evolutionary algorithm, with the remote hope to find parameters that would allow him to hear with his state of the art BTE at least as well as with his old *Boîtier*.

To start with, *Patient A* came to the hospital for yet another fitting session with a practitioner, with the aim to determine the minimum and maximum (C and T) intensity values for each of the electrodes for his BTE, to feed the evolutionary algorithm (cf. table 1).

Electrodes 10, 11 and 12 have C and T values of 0 because the auditory neurons they face have apparently been damaged (Patient A does not hear anything whatever intensity is applied to these electrodes).

Electrode	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Min	6	6,5	6,5	9	9	9	8	8	8	0	0	0	7	6	5
Max	9,5	13	13	18	20	21,5	21,5	18	16,5	0	0	0	12	10	9

Table 1. Minimum and maximum intensity (C and T values) for each electrode for *Patient A*. A pulse stimulation is defined by two parameters: Pulse height = intensity in mA (between about ten microA and 2 mA). Pulse width = duration in microseconds with a step of 0,5 microsecond. The min and max units are the min and max of the pulse width (in microsecond) of the stimulation for a given pulse height.

In order to be able to compare fittings, evaluations were done with the best fittings on the *Boîtier* and the BTE. The results corresponded to his claims. With the 78%/22% evaluation described above:

- The *Boîtier* obtained an evaluation of 53% (slightly more than 50% of the 78 words were understood).
- The BTE obtained an evaluation of 48.5% (fewer words were understood and the BTE is less confortable).

Although the results above are not statistically significant (only one evaluation for the *Boîtier* and the BTE), they matched well with the patient's feelings.

6.2 First set of experiments

Evaluation for the Patient A. A new evaluation protocol have been devised, using calibrated sentences extracted from a list of "cochlear" sentences elaborated by Pr. Lafon [Laf64], that are supposed to contain representative syllable of the French language allowing to evaluate pathological cochlea. Ten sentences were selected, for a total of 78 words, that would give 78 points if all words were correctly understood.

A comfort mark between 0 and 10 completes the evaluation, as an uncomfortable fitting will not be used by the patient. The comfort mark is multiplied by 2.2 so that the global evaluation is made of 78 points coming from the recognised words + 22 points coming from the comfort of the tested fitting.

Tests have shown that this evaluation procedure takes slightly less than 4 minutes. This is clearly not enough to obtain a fine evaluation of the audition of the patient, but it allows to perform 100 evaluations in 6h40mn only (i.e. 1h20mn per run if the 100 target evaluations are decomposed in 5 runs). If this reduced protocol is enough to guide the evolutionary algorithm and allow it to improve the fitting over 100 such evaluation, the aim is reached.

Such an aim is different from the aim of the complete evaluation of a standard practitioner, because due to the very small number of fittings they can perform in a year (about 10 fitting sessions per year and per patient), they need a very precise evaluation procedure in order to test the quality of the audition of the patient.

The experiments below were made in order to find the good values for the parameters of the interactive evolutionary algorithm (size of the population, number of children per generation, mutation rate, ...), hence the different tested values.

Experiment 1 and results. For the first experiment with patient, the size of the population was limited to 3 individuals and the evolutionary algorithm was asked to create 3 children per generation. Mutation rate was 0.1 and crossover rate was 1.

On the first evaluation (of a randomly created individual) 42 words were understood on a total of 78. Patient A gave an evaluation mark of only 1 (over 10) because even though he could understand more than half of the words, the BTE sound was resonating and feeling uncomfortable. The global evaluation was therefore of $42+1\times2.2=42.2$.

On this first experiment, 12 evaluations were performed, which is a large number, knowing that preparation and evaluation of one fitting takes between 15 and 20 mn for an experienced practitioner. With the evolutionary algorithm, only 4 mn were needed per fitting.

The result of the evaluation is given in the table .

The first three evaluations (44.2, 21.2, 9.2) correspond to random individuals. Artificial evolution starts on fitting number 4, with 3 children per generation (generations are marked with a double vertical bar).

Fitting	1	2	3	4	5	6	7	8	9
Evaluation	44,2	21,2	9,2	31,4	55,6	46,4	74,8	74,8	58,4
Fitting	10	11	12						
Evaluation	81	81	79,8						

Table 2. Experiment 1 -patient A

From the 5th evaluation onwards, obtained results are better or equivalent to the best fitting performed by the medical practitioner (48.5).

Fittings 7 and 8 are nearly identical, as well as fittings 10, 11 and 12. These results have never been approached by the expert neither with the BTE nor with the *Boîtier*.

Patient A is enthusiastic, and a second experiment is started with 6 individuals, to avoid premature convergence.

Experiment 2 and results. The only changes that have been made are a population size of 6 individuals and 4 children per generation (generations are marked with double vertical bars).

Fitting	1	2	3	4	5	6	7	8	9	10
Evaluation	24	17	30	19	53.2	37.4	22.6	24	33.4	32
Fitting	11	12	13	14	15	16	17	–	–	–
Evaluation	9	27.4	34	34.5	12	27	32	–	–	–

Table 3. Experiment 2 - patient A

The first four random individuals get poor results. Then, crossover and mutations have difficulties creating better individuals, with some really poor individuals (fittings 11 and 15).

Patient A gets tired and disappointed. The test is stopped after the 17th fitting.

Experiment 3 and results. For the 3rd test, the population is reduced back to three individuals, but with 2 children per generation. Mutation rate is increased to 0.6 and roulette-wheel is used as a selector in order to increase the selective pressure when choosing parents.

The three initial individuals obtain great values (54, 33 and 26.5). The second generation obtains values near 50. Then evaluations increase towards 60s and 70s without dropping below 50 again.

Fitting	1	2	3	4	5	6	7	8	9	10	11
Evaluation	54	33	26.5	48	52	51.6	54.6	62.8	59.6	65.6	60.1
Fitting	12	13	14	15	16	17	18	19	20	21	22
Evaluation	60	72	69.4	53.4	73	67	50.1	62	68.3	67.3	65

Table 4. Experiment 3 - patient A

Around generation 10 or 11 (fittings 20, 21, 22), evaluations seem to stabilise near 70 without beating value 73 of fitting 16.

Experiment 4 and results. For the fourth experimentation, population size is set to four individuals and four children per generation. Mutation rate is brought back to 0.1 and parents selection is set back to Tournament.

Fitting	1	2	3	4	5	6	7	8	9	10	11	12
Evaluation	59.4	62.2	57.3	58.9	57	62.3	65	73	75.3	65.2	83.1	68
Fitting	13	14	15	16	17	18	19	20	21	22	23	24
Evaluation	75.4	91	91.5									

Table 5. Experiment 4 - patient A

In average, the first four individuals present an average evaluation of 59.5 and all subsequent values are above 56.5.

Values of 91 and 91.5 are obtained at the end of generation 4. *Patient A* is tired but extremely satisfied and surprised by such results. He leaves for lunch with the BTE.

However, when he returns a couple of hours later, he says that the fitting is not very efficient in noisy environments, and feels like he still prefers his *Boîtier*, as it feels much more comfortable to wear, as he has used it for the past 10 years.

Experiment 5 and results. Population is now of 5 individuals, with two children per generation, a tournament selector and a mutation probability of 0.1.

Among the first five random individuals, two show a surprising evaluation of 70.1 and 71.9, which raises questions on the original fitting of the expert for the BTE, which only gets 48.5 on the exact same data. Some answers to this will be provided in third and fourth sets of experiments below.

However, evolution does not seem to find any better individuals.

Fitting	1	2	3	4	5	6	7	8	9	10	11	
Evaluation	18.6	53	70.1	9	71.9	58.4	60.3	58	51	57.3	48.2	
Fitting	12	13	14	15	16	17	18	19	20	21	22	23
Evaluation	36	36.2	50	29	33.5	50.3	40.2	44.5	48.3	49.3	45.2	50

Table 6. Experiment 5 - patient A

Discussion on obtained results

Fitness evolution: The evolution of the best individual for the five runs The evolution of the best individual for each of the runs is shown figure 2. Fitness increases on all experiments but exp. 2, which is a nice result for such a small number of evaluations, meaning that the educated guesses made on the IEA implementation were probably good. It seems that the correct population size is 3 or 4 individuals, with 2 to 4 children per generation.

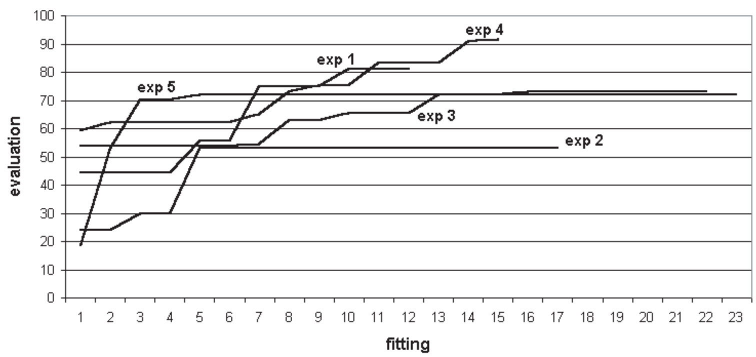

Fig. 2. Evolution of the best individuals per evaluations, for each experimentation.

Analysis of the best obtained individual: Analysis of the T/C values of the best individual is intriguing (Figure 3: Electrodes 10, 11 and 12 have been omitted as they are not functional). Sometimes, experts reduce the C - T range for some electrodes when they feel that the neural "bandwidth" is too narrow and there is a possibility of saturation if the auditory information is too important. In the fitting found by the IEA, however, many of the C - T ranges are reduced down to 1.5, 1, 0.5 and even 0. In fact, only electrodes 1, 7 and 9 have significant ranges (over 2.5, cf. circled arrows). Other good fittings show wider ranges for electrodes 7 and 9 and narrower ranges for the other electrodes, which raises a hypothesis: What if, for this precise patient, some electrodes had a negative influence on speech understanding ? If this were the case, the current practice

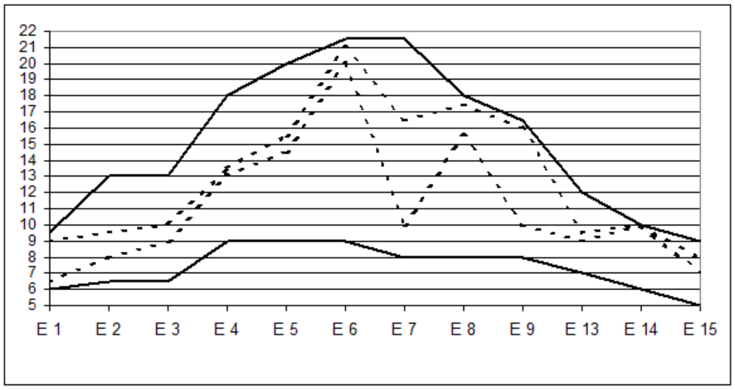

Fig. 3. X-axis: Electrodes, Y-axis: Intensity. The bold curves represent the maximum allowed envelope (T and C) for each electrode and the curves in dotted lines represent the best obtained individual (T and C).

(that has been going on for many years) of maximising the range of as many electrodes as possible would also maximise the range of "wrong" electrodes that prevent the patient of understanding speech. After this first evolutionary fitting session, the patient went back home with the original settings in his CI.
This experiment raises several questions:

– Is minimising the $T - C$ interval equivalent to shutting down an electrode ?
– Could there be a diaphony problem (crosstalk) between the electrodes ?
– Could the problem be combinatorial ?

6.3 Second set of experiments.

A second set of experiments has been conducted in order to verify some hypotheses that arose after the first set of experiments. The tests have been conducted with the same patient and with the same evaluation protocol, but one month later. It is important to note that between the two sets of experiments, the patient has used his old *Boîtier* and has resumed his normal life, meaning that it is very probable that no neuronal plasticity has had any chance to occur. The evaluation basis are therefore comparable. In the text below, the first set of experiments is noted C_1 while the second set is noted C_2.

Experiment 6. Surprisingly enough, the best individual obtained during the fourth run was virtually using only three of the 12 functional electrodes: Electrodes 1, 7 and 9. Each electrode corresponds to a given frequency band:

– Electrode 1: 4089 - 5798 Hz,
– Electrode 7 : 671 - 916 Hz
– Electrode 9 : 427 - 549 Hz

Since electrode 1 was mapped onto very high frequency sounds that are not discriminant for speech, the number of functional electrodes could be reduced to only 2. In order to confirm this strange result, the first deterministic test maximises electrodes 7 and 9 only (using the maximum C - T range of table 1), giving only a small range to electrode 1 figure 5. For all the other electrodes, T and C values are set to 1 and 1.5, i.e. much below the threshold, in order to cancel them totally. This setting obtains an evaluation of 82, which is much better than with all activated electrodes (best fitting of 48.5 obtained by the expert). Nearly 90% of the words were understood, and the fitting was rated as not very comfortable. This allows to conclude that for this patient, using only three electrodes out of 15 allows him to understand speech better than with all functional electrodes set to nearly maximum range.

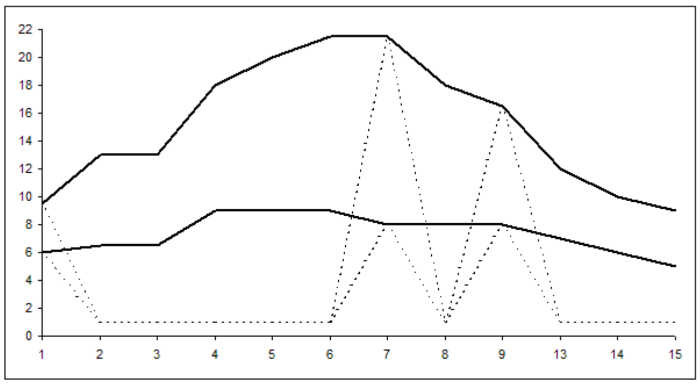

Fig. 4. Experiment 6. X-axis: Electrodes, Y-axis: Intensity. Testing with electrodes 1, 7 and 9 only. The bold curves represent the envelope (T and C) for each electrode. The dotted lines correspond to the manually tested T and C values for each electrode.

Experiment 7: On the influence of electrode 8. In the C_1 set of experiments, the evolutionary algorithm seems to hesitate a bit on electrode 8. In order to test its real contribution, the electrode 8 is added to the 1, 7 and 9 electrodes, by maximising its $C - T$ interval (using the values of table 1). The obtained evaluation is 81, and the patient finds that the fitting is slightly less comfortable than the previous one. Speech understanding is comparable. The electrode 0 does not seem to have an important role in speech understanding.

Experiment 8: Is there any diaphony between the electrodes ? In order to explore this hypothesis, even electrodes are suppressed (by setting T and C

values below the T liminary values for the patient), and the odd electrodes are maximised (using the values of table 1), so as to space active electrodes (cf. figure 5).

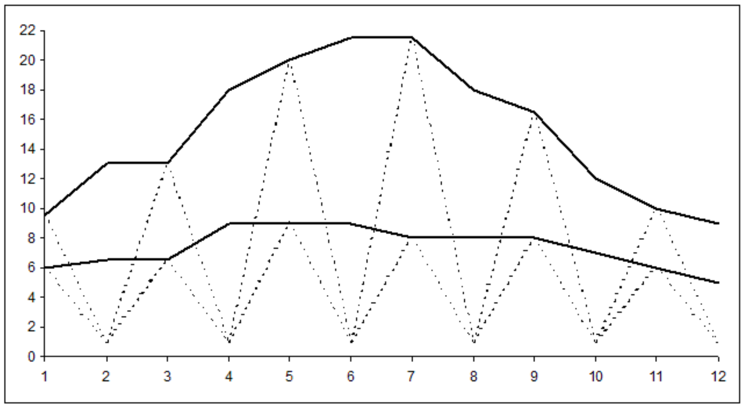

Fig. 5. Experiment 8. X-axis: Electrodes, Y-axis: Intensity. Checking for diaphony.

This fitting obtains an evaluation of 78.8, and is judged less comfortable by the patient. The result is therefore not as good as those obtained during experiments 6 and 7. Adding other electrodes does not seem to add much. The result is however still much better than the one obtained by the practitioner with the BTE (48.5).

Experiment 9: Spacing electrodes even more. This time, 2 electrodes out of 3 are canceled, by setting their T and C values to 1 and 1.5 (cf. figure 6). Therefore, electrodes 1, 4, 7 are activated. It was chosen to keep electrode 9 active, so as to keep a common comparison basis with the previous experiments. Finally, electrode 15 is maximised figure 6. This fitting obtains an evaluation of only 58.5, i.e. clearly not as good as the previous ones, and the patient rates it as quite uncomfortable. This is very surprising, as the only difference with the first test (that had obtained an evaluation of 82) is that electrodes 4 and 15 have been added. Clearly, not only is there no diaphony problem (spacing active electrodes did not improve evaluation), but it can be concluded that for this patient, electrodes 4 and 15 contribute negatively to speech understanding. The fact that functional electrodes can contribute negatively to speech understanding is a totally new concept in the cochlear implant medical field.

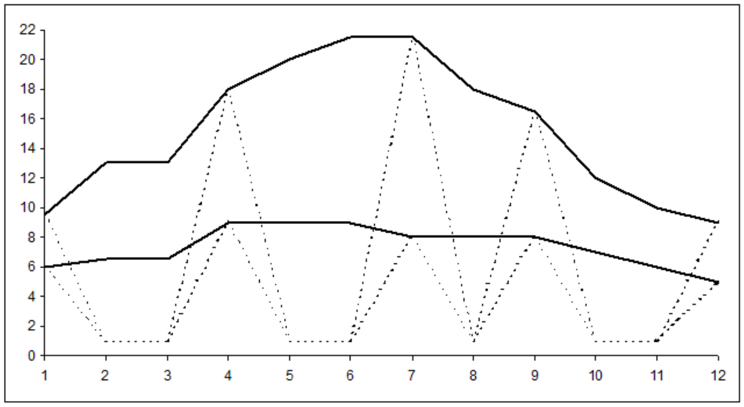

Fig. 6. Experiment 9. X-axis: Electrodes, Y-axis: Intensity. Checking for diaphony by selecting only one every 3 electrodes, and keeping electrode 9.

Experiment 10: Evaluation of the best individual of C_1. In order to test the evaluation procedure, the best individual of the set of experiments C_1 is tested again, one month later, and without telling anything to the patient.

The speech understanding test is again very good (94% of the words are understood, which is even better than one month before) but the comfort mark (over 22, cf. section 6.2) is not as good, resulting in a slightly lower evaluation of 86.2%. All in all, this value is slightly lower than the one obtained during C_1, but it is the best value obtained during C_2.

Experiment 11: Evaluation of the practitioner's fitting. This time, it is the practitioner's original fitting that is tested again (the one that more or less maximised all electrodes, and that had obtained 48.5 during C_1).

Here again, the number of recognised words is very low (only 33%) and comfort gets a bad 4/10 evaluation. The global evaluation is 41.8, which is also slightly worse than during C_1.

All in all, in one month, the best fitting found by the IEA went down from 91.5 to 86.2, while at the same time, the practitioner's fitting also went down from 48.5 to 41.8. This suggests that the proposed quick 4mn evaluation is quite reliable, as the results seem to be reproducible one month later, while the patient used his old *Boîtier* in the meantime.

Other tests. In order to verify that values obtained by the evolutionary algorithm are better than random ones, other experiments have been conducted with random values for T and C for all electrodes. Evaluations range from average to bad, although often greater than those obtained by the practitioner (48.5). The patient finds that these fittings are not comfortable.

6.4 Third set of experiments with others patients

In order to verify the gain obtained with computer-aided CI fitting, and develop its use at hospital, new experiments have been carried out with others patients. This set of experiments C_3 is conducted with 2 new patients: Patient B and patient C. For these experiments, the parameters of the IEA are the following:

Population	3
Children	2
Mutation rate	0.1
Crossover rate	1

The new population is obtained by taking the best individuals of the intermediate population consisting of the 3 parents and the 2 children (i.e. in the style of a (3+2)ES).

Corpus and methodology. Both patients have received MXM cochlear implants some years ago, but they are not satisfied with their devices and have no good results (general evaluation by the practitioner is less than 50%). The IEA has been used to try to determine optimal C (Comfortable) and T (Threshold intensity) values for each of the electrodes of the CI.

To start with, the patients came to the hospital for a fitting session with a practitioner, and minimum and maximum intensity values for each electrodes of their BTE have been determined, to give boundaries to the evolutionary algorithm.

For these 2 patients (B and C), the same procedure that was used for patient A (a set of calibrated sentences) has been tested. Unfortunately, the results are disappointing as patients B and C recognise but a few words, meaning that this test is too hard for them.

Therefore, a new evaluation procedure was set up, based on a weighted evaluation of the results of:

- A discrimination test (ASSE) on 7 items. The ASSE test consists in emitting a sound n times (an [i] for instance), and within these occurrences, replacing one of the [i] with an [a] (for the following sequence: i i i i a i i). The patient needs to detect that one of the sounds was different. The ASSE test counts for 20% of the evaluation.
- A VCV (Vowel/Consonant/Vowel) test ([APA], [ATA], ...), where the patient must recognise the consonant between the two vowels. In one VCV test, each VCV is repeated 3 times, meaning that 48 VCVs are proposed to the patient (because in French, there are 16 different phonetic consonants). This test counts for 50% of the evaluation.
- A comfort evaluation with a mark from 0 to 10, that counts for 30% of the evaluation.

Unfortunately, the complete evaluation takes a long time (much more than 4 minutes), and the patients are less compliant than patient A, so it is impossible to get around 100 evaluations (as for patient A).

After the first sessions, the P1 and P2 settings of the CI were loaded with respectively the fitting obtained with the IEA, and the manual fitting of the practitioner, after which the patients were sent home with the instruction to use the best fitting of P1 or P2.

After two weeks, the patients came back for new tests:

1. a discrimation test with P1 and with P2,
2. a VCV recognition test with P2 and with P1,
3. a sentence recognition test with 10 sentences using the P1 setting (IEA).

Third set of experimentations with patients B and C

– First session for Patient B:

Eval Nb	Manual	1	2	3	4	5	6	
ASE Result	4/7	5/7	5/7	5/7	6/7	5/7	7/7	
VCV Result	33%	31%	25%	18%	29%	31%	31%	fatigue
Comfort	7/10	6/10	7/10	5/10	5.5/10	6/10	8/10	
Evaluation		5	5	4	5	5	6	

– Second session for Patient B three days later:

Setting	Manual	1	2	3	4	5	6	7	
ASE Result		6/7	7/7	7/7	7/7	6/7	5/7	5/7	
VCV Result	35%	25%	27%	10%	18%	18%	20%	27%	fatigue
Comfort		5/10	6/10	6/10	5/10	5/10	5/10	5/10	
Notation		4	5	4	4	4	4	4	

The best obtained fitting (6th fitting of the first session) was loaded in memory P1 of the BTE and the patient was sent home for two weeks for a long evaluation of the new fitting.

– First session for Patient C:
A first set of independent random tests has been performed, to be compared to the manual fitting results, in the table below:

Setting	Manual	Random								
ASE Result	5/7	6/7	5/7	5/7	5/7	5/7	4/7	5/7	6/7	5/7
VCV Result	45%	33%	29%	22%	39%	31%	18%	29%	39%	35%
Comfort		4/10	5/10	5/10	5/10	5/10	4/10	5/10	5/10	5/10
Notation		5	4	4	5	4	3	4	5	4

Then the IEA is used, but only based on a VCV evaluation, to shorten the time of evaluation.

Setting	1	2	3	4	5	6	7	Fatigue
VCV Result	35%	41%	39%	33%	20%	43%	37%	
Notation	4	4	4	3	2	4	4	

As for Patient B, Patient C is sent back home for a long evaluation of fitting obtained in setting 6 (VCV result = 43%).

After two weeks, patients B and C come back to hospital with the following results for patient B:

Test	ASE	VCV	Words/list	Comfort
Auto	3/7	33%	7	n/a
Manual	5/7	27%	10	n/a

And for Patient C:

Test	ASE	VCV	Words/list	Comfort
Auto	3/7	52%	1	8/10
Manual	4/7	37%	2	8/10

Remarks:

1. Both patients preferred to use the P1 fitting, i.e. the one found by the interactive evolutionary algorithm (although the very small number of evaluations due to the long evaluation makes it more like a random search than an evolutionary search).
2. Random fitting can do really well, sometimes slightly better than what the practitioners do when they maximise the number of electrodes and their dynamic.
3. Comfort is too difficult to evaluate accurately for the patients.
4. As for the interactive evolutionary algorithm, each of these non word based evaluations is much too long to obtain, meaning that patients get tired very rapidly and the run is stopped before a significant number of evaluations could be done that could allow the evolutionary algorithm to suggest other fittings than random ones.

The results obtained by this random search again question the maximisation of the number of electrodes and the maximisation of their dynamic range.

These random tests also show that the ranges of possible parameters values is well chosen, providing a search space having many "average good" solutions, but with a rather "flat" search landscape. In these conditions, and considering the parameter setting of the IEA (a (3+2) Evolution Strategy), time for convergence is too short to really obtain the beginning of a convergence. The problem of user fatigue makes it impossible to obtain meaningful results. Additionally it can be argued that the evaluation is not discriminant enough to provide an efficient fitness landscape to the IEA.

New tests have been designed, taking these results into account.

6.5 Fourth set of experiments

The same patients (Patient B and C) were tested. The parameters of the IEA are the following:

Population	4
Children	3
Mutation rate	0.8
Crossover rate	1

The new population is obtained by an elitist binary tournament between a population made by the parents and the children. The elitism is "soft," in the sense that it is the best individual of the 7 individuals that is taken to be part of the next generation (and not the best of the parents only). The three other individuals are selected by a standard binary tournament.

Corpus and methodology. Each trial was based on the results of a VCV recognition test such as [APA], [ATA] ... The patient has to recognise the consonant in the VCV. Each VCV is proposed once, meaning that there are only 17 items in a test. The result over the 17 VCV counts for 100% of the evaluation.

Experiments

- Patient B:
  The IEA fitting tested over two weeks obtains an evaluation of 2 over the 17 tested VCVs. The expert fitting is tested again, and here again, only 2 of the 17 tested VCVs were recognised. A new run gives the following results:

Evaluation	1	2	3	4	5	6	7	8	9	10
VCV Result	2	3	3	2	3	4	3	4	4	3

  And after one hour break and restart:

Evaluation	1	2	3	4	5	6	7	8	9
VCV Result	4	2	3	2	4	2	2	3	3

  Several fittings were found where 4 VCVs were recognised rather than only 2 previously, but it must be noted that these fittings were found at random. Patient B was satisfied with this result.
- Patient C:
  The IEA fitting tested over two weeks obtains an evaluation of 8 over the 17 tested VCVs. A new run gives the following results:

Evaluation	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
VCV Result	6	5	5	5	4	7	5	7	7	8	7	4	6	4	4	7	5	4

After a lunch break, the algorithm is restarted, initialised with two individuals which are the IEA fitting the patient had been using for the previous week and the best fitting of the previous run (fitting 10).

Evaluation	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
VCV Result	3	5	6	6	7	8	7	7	6	3	5	8	6	6	6	4

Then after another break, the algorithm is restarted again to produce the following results:

Evaluation	1	2	3	4	5
VCV Result	6	3	7	6	8

Remarks:

– The IEA was working fine, although no real improvement could be seen, even during the longest runs (like the first run of patient C, i.e. 18 evaluations, i.e. evolution during five generations).
– The probable explanation is that the chosen VCV evaluation is too difficult for both patients, and the algorithm cannot find any fitting leading to a stable improvement of the audition of the patients.

7 Actual work and perspectives

Concerning the evolutionary runs, the evaluation function is very important. If for these patients, the VCV test is really too hard, the IEA will not be able to find any improvements (the fitness landscape is too flat to give a direction for improvement to the algorithm).

It seems important to spend some time to set up an evaluation function specific to each patient, that can return an average value, neither too low, like 3/17 or 5/17 (because this would mean that the test is too difficult) or too high, like 15/17 or 16/17 (because this would not leave any room for improvement).

The evaluation function must be quick. If it is too slow, the patient will get tired before any significative number of evaluations are done (in the set of experiments 3).

Finally, until the IEA procedure is routinely giving good enough results, it may be interesting to choose "easier" patients, i.e. patients for whom the cochlear implantation works slightly better ...

Even though the sets of experiments 3 and 4 have not been really satisfying evolutionary-wise, the results are very interesting on a medical point of view,

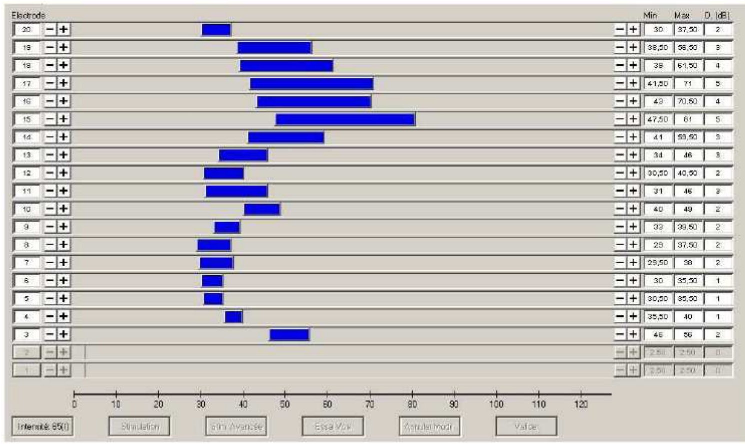

Fig. 7. Best fitting found by the practitioner for patient C: each rectangle represents the $[T, C]$ interval for each electrode.

since it has confirmed that narrower intervals (or even removal of one or several electrodes) can lead to better speech understanding.

In all tested patients (of which A B C were a subset), it was possible to find fittings that were working at least as well as manual fittings maximising the dynamics for all electrodes, and in many cases, these fittings were simply random fittings !

In order to have a visual example, fig. 7 shows the intervals for all the electrodes of patient C on the best fitting found by the practitioner, while fig. 8 shows the best fitting obtained . . . randomly, that gives better results than the practitioner's. Please note the skinny intervals compared to those of fig. 7. In some cases, some electrodes are virtually cancelled (electrodes 5, 8, 11, 12 and 13), which goes against reason (and against what is advocated by the cochlear implants manufacturers).

7.1 Classification of sound environments

Many users of cochlear implants or hearing aids find that the parameter setting of their device is not perfectly adapted to all situations of their everyday life: in restaurants, they find clicking cutlery aggressive, and they have a hard time following a conversation, in the street, some noises are nearly unbearable, and so on. Some patients may need a setting for a quiet environment (such as home) but may work in a noisy environment (metal industry, garage, other noisy environments) so there is no miracle solution.

The aim of the HEVEA project is to improve hearing with cochlear implants by several means. One is to help the expert find good fittings using an interactive evolutionary algorithm [BRC05], and another is to integrate into the processor

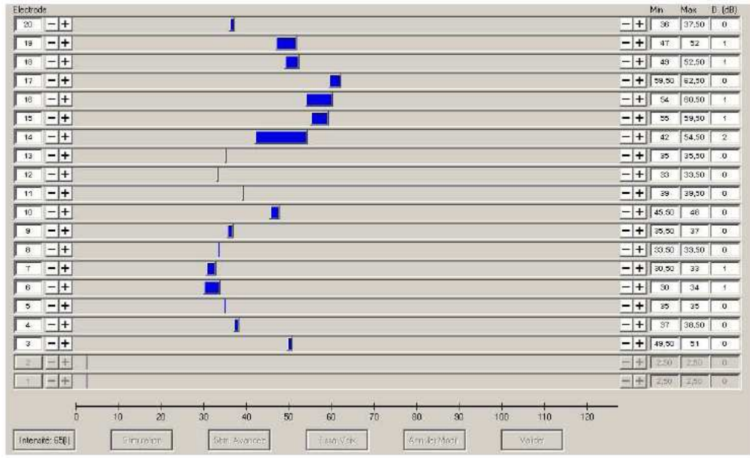

Fig. 8. Best fitting found at random for patient C, that beats the best fitting found by the practitioner: each rectangle represents the $[T, C]$ interval for each electrode.

a small signal analysis software that would be able to recognise the sound environment and automatically select a fitting accordingly, among a set of available fittings corresponding to different situations.

In order to achieve this second task, several stages must be performed :

1. The medical team must determine with the patient a number of common environments for which the patient would need a specific fitting, for instance: home, work, supermarket, cinema, …
   The number of specific environments should be limited, because for each of the specified environments, a special set of parameters needs to be found for the cochlear implant, and finding a good set of parameters can be a long and difficult task (even with the help of an evolutionary algorithm).
2. For each of the specified environments, the patient must take a number of sound samples to bring back to hospital.
3. Specific parameters must be found, to deal with each of the specified environments (possibly with the help of an interactive evolutionary algorithm).
4. In parallel, the different samples must be analysed to extract some common features, so that a classifying algorithm can determine them in which category falls the sound environment that is surrounding the patient.
5. Finally, the characteristics and parameters for the different environments must be uploaded into the cochlear implant processor, along with a signal processing program that will automatically choose the correct parameters to match the environment in which the patient is evolving.

The result is an "intelligent" cochlear implant that can automatically switch between potentially different sets of parameters, depending of the sound environment surrounding the patient.

This section presents the sound sampling, characterization and classification stage. It starts with a description of the specific sound sampler developed for this application, followed by a sub-section recalling the wavelet theory on which the scientific work is based. Then, a third sub-section describes how the energy content of a sample can characterise a sound environment. Finally, results are presented on the classification of different environments using a standalone piece of software.

Development of an *a posteriori* sound sampler. In this application, sound sampling is essential to provide accurate data for two orthogonal needs:

1. The sound environment must be accurately recorded so that it can be recognised in the future by the processor with sufficient confidence to switch between different sets of parameters.
2. Particularities must be also recorded so that a specific fitting can be found that will help to cope with the current environment.

This distinction must be made because it is necessary to tune the Cochlear Implant (CI) on possibly transient sounds that are not representative of the general sound environment. For instance, one patient currently switches off his cochlear implant whenever cycling to work, because the sound of a motorbike passing by is too stressful to be bearable with his usual CI fitting. Choosing to switch off his CI (and becoming totally deaf) in a street environment is quite radical, but shows how much an adaptive and "intelligent" CI would be needed for this patient.

So it would be necessary for the adaptive CI to recognise a street environment, in order to choose for a fitting that would allow to cope with passing motorbikes, although passing motorbikes are exceptional in a street. One must therefore find a fitting adapted to an exceptional event, that should be selected when a sound environment (that has nothing to do with the exceptional event) is detected.

Sampling the regular environment for characterization. The sampling must be as accurate as possible, so that the processor can select the correct parameters without making any mistakes. Therefore, recording a sound environment on an old tape recorder may not be sufficient. A small jack plug has been added to the processor of the CI so that it could output directly the sound picked up by the microphones of the CI to a digital sampler.

Then, a sampling software has been developed on a PDA (Personal Digital Assistant) that the patient plugs directly onto the CI processor in order to sample the exact sound that is received by the processor (cf. fig. 9).

Sampling the exceptional event for CI fitting. Then, another problem arises: whenever an exceptional event occurs for which the CI should be tuned, it is often too late (the unbearable motorbike sound has vanished before the patient could record it, or in a crowded restaurant, the words that have not been understood cannot be repeated in exactly the same manner). A solution could be to sample

Fig. 9. A sampling software has been developed on a PDA that the patient plugs directly onto the CI processor in order to sample the exact sound that is received by the processor.

the street (or the restaurant) for a long enough time, but here again, it is difficult do predict when the right motorbike will appear (or when the waiter will speak in an unintelligible way), and this could result in hours of recording, and hours to replay the records to find the relevant information.

A special sampling software has therefore been developed that constantly records the current sound for a period of n seconds. When the patient hits the `record` button, whatever happened during the previous n seconds is stored in a file, for future use. 30 seconds seems to be a correct period, so that when the patient uses the PDA to record precise sounds, he has 30 seconds to press on the button after he noticed that some interesting sound has occurred.

These very transient sounds samples (motorbike) have a different content than the samples that are used to characterise the general environment ("standard" street noise).

Characterisation of a sound environment. We distinguish two steps in the problem of "sound environment classification". The first step is the extraction of the characteristics, in order to build the representation's space. The second step is to find a classification method which allows to fit each point of this space with a probability of being in a specified family. We can extract a lot of information from a sound in order to make a classification. For example, one can use the frequential content, the cepstral characteristics, the loudness, the pitch ...

From what is know from the human perception, spectral characteristics are discriminant for the recognition of all kind of sounds. This is the reason why we decided to use spectral measurements for artificial characterisation.

For this work we will analyse the frequential content at each dyadic scale because the implant performs the same kind of analysis. We will use a wavelet transform in order to perform a multiscale analysis (see [Dau92] and [Mey90]).

We could use a simple Fourier Transform but we prefer keep the possibility to use the time localisation provided by the wavelet transform for a future work. In fact, Wavelet analysis allows to adjust the width of analysis windows, and achieves a perfect localisation in time and frequency. Logically, temporally extended windows are used to study low frequencies, while narrower windows are used for higher frequencies. This localisation property makes wavelet theory predominant in several areas of signal processing.

Continuous Wavelet Transform (CWT). A wavelet is a "wave localised in time." More precisely, it is a function $\psi \in L2(\mathbf{R})$ such that $\int_{\mathbf{R}} \psi(t)dt = 0$.
If $\int_{\mathbf{R}} \psi^2(t)dt = 1$, then we use normalized wavelets.

The continuous wavelet transform of a signal f is given by:

$$CWT(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} f(t)\psi\left(\frac{t-b}{a}\right) dt$$

In this expression, a is a scale factor and b is a translation parameter (temporal shift). Variable a represents the inverse of the frequency: the smaller a, the (temporally) narrower the wavelet (i.e. the analysing function).

Therefore, one can see this expression as the projection of the signal on a family of analysing functions:

$$\psi_{a,b} = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right)$$

constructed by widening and translation from the original ψ wavelet.

Discrete Wavelet Transform. In this work we use a discrete wavelet transform which is faster than the continuous transform. The Discrete Wavelet Transform can be obtained thanks to the discretization of the parameters of resolution (a) and position (b). Let $a = a_0^m$ with m an integer, a_0 a resolution step greater than 1 and $b = nb_0a_0^m$ with n an integer and $b_0 > 0$.

Furthermore, if $a = 2$ and $b = 1$, the transform is called "dyadic." One then has:

$$C_{j,k} = 2^{-\frac{j}{2}} \int_{-\infty}^{\infty} f(t)\psi(2^{-j}t - k)dt$$

If $\psi_{j,k} = 2^{-\frac{j}{2}}\psi(2^{-jt} - k)$ we get a tiling of the time-frequency space called a dyadic grid (see fig 10).

Energy of a signal. For a given scale, if we use a normalized wavelet, the energy of the signal can be obtained from the continuous wavelet transform. More precisely: one can compute the energy of the a scale by adding the squares of the wavelet coefficients of the continuous transform at this scale:

$$E_a^2 = \int [CWT(a,b)]^2 db \tag{1}$$

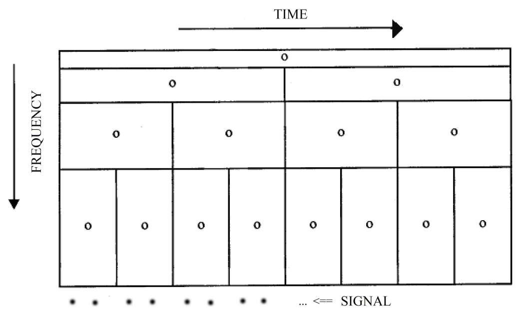

Fig. 10. Dyadic grid. X-axis: Time, Y-axis: Frequency. At the bottom, each point is a point of the signal. The matching discret wavelet coefficients are the circle in the grid. At low frequencies, the computation of the wavelet coefficient uses large windows in time, then we only have few coefficients. On the opposite, at high frequencies the computation uses small windows.

where E_a^2 is the energy at scale a. If we use the discrete wavelet transform, we get:

$$E_j^2 = \sum_{k=1}^{2^{j-1}} [C(j,k)]^2 \tag{2}$$

where E_j^2 is the energy at scale j.

Characterisation of a class by its energy content. As said above, a class will be characterised by its energy content. Let us consider a sound environment $S1$. The patient records a collection of `*.wav` files, that are chopped into a family of n_1 subsignals of 2^{14} points (almost 3 seconds for each subsignal, the 2^{14} number of points being chosen because it is a good compromise between quantity of information and computing speed). If one computes the discrete wavelet transform of theses signals and the energy of each of the obtained frequency bands during multi-resolution analysis, one then gets n_1 vectors of 14 coordinates. We choose to characterize a class by the mean value of these vectors. We obtain for each class a value at each dyadic bandwidth frequency (see fig 11).

Classification of sound environments. The aim is to create a class for a specific environment, by using a collection of `.wav` files as input. The set of sounds chosen below are part of a patient's environment.

When the patient is in a new environment, he uses the sound sampler and records a sample of this environment. A `.wav` file is imported and chopped into 2^{14} micro-samples. When clicking on `compute`, each of the mini-samples is associated with the family that matches the sample best.

A ratio is then displayed, that presents the number of samples that corresponded to each family, and the results are displayed in a bar-chart. The bar-

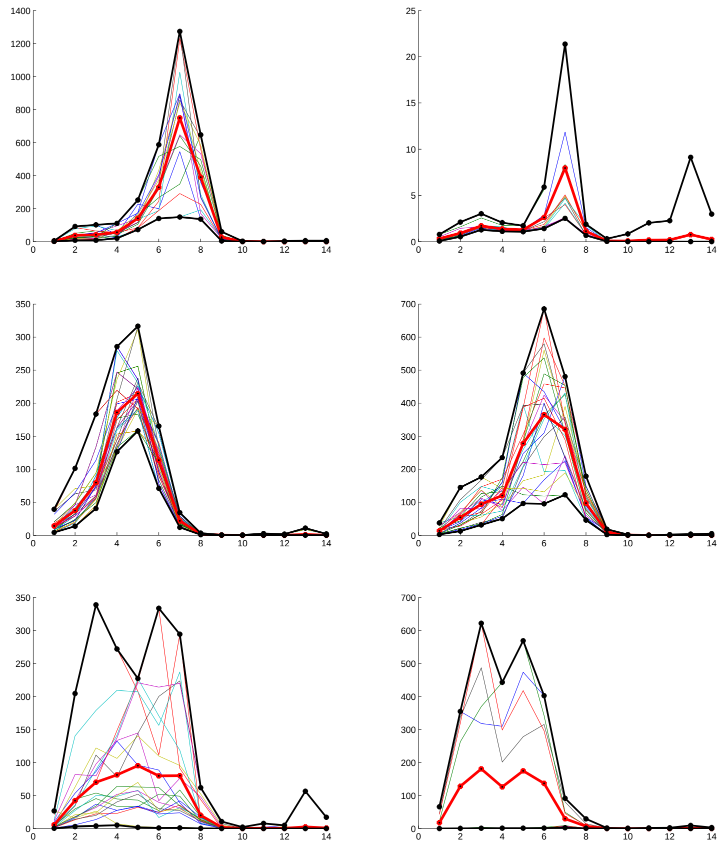

Fig. 11. X-axis: frequency, Y-axis: Energy. Left up: "Car-radio" environment. Right up: "Birds" environment. Left middle: "Supermarket" environment. Right middle : "road corner" environment. Left down: "School-yard" environment. Right down: "Lawn mower" environment. Set of values of the energy for each frequency (fine lines), envelope and mean criterion (thick lines).

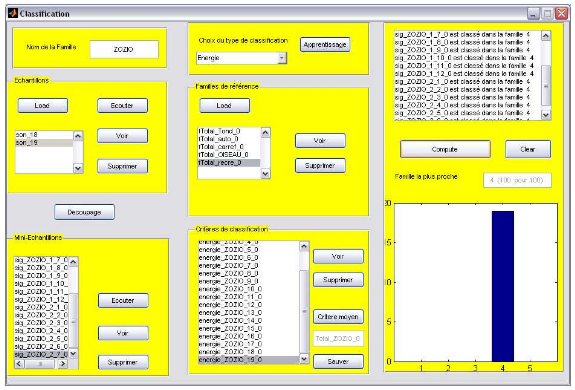

Fig. 12. Graphical Interface for the classification toolbox.

chart provides us the matching family with a certain confidence. For example if 80% of the micro-sample are classified in the class $S1$, then the sample will be classified in the class $S1$ with a confidence of 80%.

Results. For each family, available `.wav` files have been chopped into mini-samples of 2^{14} points. 66% of the mini-samples chosen randomly are used for the learning set, and 33% for the test set. The results are presented in the following table:

Family	Learning set	Test set	matching family	Confidence
Car-radio	16	8	Car-radio	100%
Cross-roads	24	13	Crossroads	84 %
Birds	12	7	Birds	100%
School-yard	22	11	School-yard	100%
Supermarket	35	15	Supermarket	100%
Lawn-mower	10	5	Lawn-mower	80%

This set of sound was chosen because they were part of the sound environment of the tested patient.

All samples have been correctly classified. For Car-radio, Bird, School-yard, and Supermarket environments we have 100% of confidence. The worst results are for the Crossroad and Lawn-mower environments, the sample have been correctly classified with a confidence of respectively 84% and 80% (on the 13 Crossroad test samples, one is identified as a Supermarket environment and another one as a lawn-mower, and on the lawn-mower, one out of 5 samples is classified as being a crossroad).

Future work. What needs now to be done for the scheme to be fully functional is to connect the PDA to the cochlear implant, so that if the PDA is able to classify an environment with a confidence rate greater than 50%, it selects automatically the corresponding CI fitting adapted to this sound environment and it uploads it into the CI.

If, on the contrary, the confidence rate is less than 50%, the sound environment is sampled and memorized, so that it can be classified later on (which may require to create a new sound class).

8 Conclusion

The problem of cochlear implants fitting belongs to a class of very difficult problems, impossible to solve in a deterministic way in a limited time, for at least two reasons:

- The function to be optimised cannot be modeled. It is extremely variable, because it is dependent on the patient and linked to a subjective evaluation of his auditive sensations.
- The search space is very large, therefore, strict optimality is out of reach.

The work presented in this paper describes an approach of this problem, based on an interactive evolutionary algorithm with a micro-population. The first results with patient A are promising: evolution has taken place (as the curves show in fig. 2) and the obtained results were far better than those obtained by an expert practitioner.

However, this experiment showed that it was possible to obtain good fittings by simply selecting values at random, which questions the usual aim, that is to maximise the number and range of electrodes to improve audition and comprehension. A number of other experiments has been conducted that shows that indeed, the strategy advocated by CI manufacturers may not be the best, which is a new result in the medical field.

But this work is obviously a preliminar one, that needs to be confirmed with additional experimental analysis on other patients, having various profiles. Moreover, the aim of this project is to make cochlear implants more adaptive to patients and to their environments: The adaptation to audio environment that has been sketched in section 7, needs now to be tested by patients in real environments.

Other points of improvements are more technical and relate to the heart of the interactive optimisation method. The real experiments presented in this paper actually prove the importance of user fatigue, which is a general problem in IEAs. But in the case of audio interaction this problem is even more crucial, for two reasons: only one signal can be evaluated at once (on the contrary to visual evaluations), and the attention needed to correctly evaluate a fitting is extremely demanding for implanted patients.

Usually, one copes with user fatigue in three ways: [PC97,Tak98,Ban97] :

- reduce the size of the population and the number of generations,
- choose specific models to constrain the research in a priori "interesting" areas of the search space,
- perform an automatic learning (based on a limited number of characteristic quantities) in order to assist the user and only present to him the most interesting individuals of the population, with respect to previous votes of the user.

In this paper we have used the first item, i.e. a micro-EA. The experimental analysis that has been presented proves the necessity to try other strategies. According to the third item above, experiments have been conducted on another application (image denoising) with a fitness map technique [LPLV05], where the fitness rating has been extended to individuals of a larger population via the analysis of the user judgment on a small sample of individuals. Future work on cochlear implants could use a similar strategy, in order to evolve a larger population of parameter settings while keeping a low number of user evaluations.

Additionally, other strategies to better exploit the user interactions should be considered, such as using partial evaluations (shorter audio tests), and refinements of audition, understanding and comfort evaluations only on areas of the search space that have been identified as promising by the IEA.

Acknowledgements

We would like to thank *Neurelec* (an *MXM* company, http://www.neurelec.com) who provided us with equipment that made this research possible.

References

[ALM95] S. Archbold, M. Lutman, and D. Marshall. Categories of auditory performance. In *Otol Rhinol Laryngol*, 1995.

[Bae95] T. Baeck. *Evolutionary Algorithms in Theory and Practice*. New-York: Oxford University Press, 1995.

[Bae05] T. Baeck. Tutorial on evolution strategies. Genetic and Evolutionary Computation Conference Gecco'05, 2005.

[Ban97] W. Banzhaf. Handbook of evolutionary computation. In *Chapter Interactive Evolution*. Oxford University Press, 1997.

[BPG+92] P.J. Blamey, B.C. Pyman, M. Gordon, G.M. Clark, A.M. Brown, R.C. Dowell, and R.D. Hollow. Factors predicting postoperative sentence scores in postlinguistically deaf adult cochlear implant patients. In *Ann. Otol. Rhino. Laryngol.*, volume 101(4), pages 342–348, Apr 1992.

[BR04] C. Bourgeois-République. *Plateforme de réglage automatique et adaptatif d'implant cochléaire par algorithme évolutionnaire interactif*. Phd thesis, University of Bourgogne, Fr, 2004.

[BRC05] C. Bourgeois-République and P. Collet. An interactive evolutionary algorithm for cochlear implant fitting : First results. In *ACM Symposium on Applied Computing Santa Fe, New Mexico*, 2005.

[BRFC05] C. Bourgeois-République, B. Frachet, and P. Collet. Using an interactive evolutionary algorithm to help fitting a cochlear implant. In *MEDGEC (GECCO)*, 2005.

[BRVC05] C. Bourgeois-République, G. Valigiani, and P. Collet. An interactive evolutionary algorithm for cochlear implant fitting: first results. In *SAC*, pages 231–235, 2005.

[BT97] T. Blickle and L Thiele. A comparison of selection schemes used in evolutionary algorithms. *Evolutionary Computation*, pages 361–394, 1997.

[BTE04] BTE. Implant cochléaire digisonic convex bte. 2004.

[CFML83] C.H. Chouard, C. Fugain, B. Meyer, and H. Lacombe. Indications and results of the multichannel cochlear implant. In *Annals of the NY Academy of Sciences*, volume 405:387-411, 1983.

[CLSL00] P. Collet, E. Lutton, M. Schoenauer, and J. Louchet. Take it easea. In PPSN2000editors, editor, *PPSN2000*, LNCS 1917, pages 891–901. sv, 2000. http://sourceforge.net/projects/easea.

[Coh89] L. Cohen. Time frequency distribution a review. In *Proceedings IEEE volume 77*, 1989.

[COM94] C.H. Chouard, M. Ouayoun, and B. Meyer. Speech coding strategies of the digisonic fully digitized cochlear implant. In *Acta Otolaryngol*, 1994.

[Dau92] I. Daubechies. *Ten Lectures on Wavelets*, volume 61. CBMS-NSF Regional Conference Series in Applied Mathematics, SIAM, Philadelphia, 1992.

[DDML89] M.F. Dorman, K. Dankowski, G. McCandless, and Smith L.M. Consonant recognition as a function of the number of channels of stimulation by patients who use the symbion cochlear implant. In *Ear Hear.*, volume 10(5), pages 288–291, 1989.

[DJ05] K.A. De Jong. *Evolutionary Computation: a Unified Approach*. MIT Press, 2005.

[Dur02] E.A. Durant. *Hearing Aid fitting with Genetic Algorithms*. Phd thesis, University of Michigan, USA, 2002.

[Fis96] G.S. Fishman. Monte-carlo: Concepts, algorithms and applications. In *Springer series in Operations Research*. Springler-Verlarg, 1996.

[FSBW01] L.M. Friesen, R.V. Shannon, D. Baskent, and X. Wang. Speech recognition in noise as a function of the number of spectral channels: comparison of acoustic hearing and cochlear implants. In *J. Acoust. Soc. Am.*, volume 110(2), pages 1150–1163, Aug 2001.

[FXP03] K.H. Franck, L. Xu, and B.E. Pfingst. Effect of stimulus level on speech perception with cochlear prostheses. In *J. Assoc. Res. Otolaryngol.*, volume 4(1), pages 49–59, Mar 2003.

[GFM+98] S. Gallego, B. Frachet, C. Micheyl, E. Truy, and L. Collet. Cochlear implant performance and electrically-evoked auditory brain-stem response characteristics. In *Electroencephalography neurophysiologie and clinical neurophysiology*, 1998.

[GTBVC01] S. Gallego, E. Truy, C. Berger-Vachon, and L. Collet. Electrically auditory brainstem responses in cochlear implant assessment : possibility and interest. 2001.

[Hes02] G. Hesse. *Programmation des seuils liminaires de l'implant cochléaire MED-EL Tempo+*. Thèse de doctorat de médecine, Université de Rennes, 2002.

[Jan02] T. Jansen. On the analysis of dynamic restart strategies for evolutionary algorithms. In *Parallel Problem Solving from Nature*, pages 33–43, 2002.

[Kri89] K. Krishnakumar. Micro-genetic algorithms for stationary and non-stationary function optimization. *SPIE: Intelligent Control and Adaptive Systems*, 1196, 1989.

[KSC+98] A. Kawano, H.L. Seldon, G.M. Clark, R.T. Ramsden, and C.H. Raine. Intracochlear factors contributing to psychophysical percepts following cochlear implantation. In *Acta Otolaryngol.*, volume 118(3), pages 313–326, Jun 1998.

[KVR+00] J. Kiefer, C. VonIlberg, V. Rupprecht, J. Hubner-Egner, and R. Knecht. Optimized speech understanding with the cis sampling speech coding strategy in patients with cochlear implants: effect of variations in stimulation rate and number of channels. In *Annals of Otology Rhinology and Laryngology*, volume 109(11), pages 1009–1020, 2000.

[Laf64] J.C. Lafon. Le test phonétique et la mesure de l'audition. In *Ed. Centrex*, Eindhoven, 1964.

[Loi01] P. Loizou. Url http://www.fing.org/index.php. In *Conference on Implantable Auditory Prostheses*, 2001.

[LPD00] P. Loizou, O. Poroy, and M. Dorman. The effect of parametric variations of cochlear implant processors on speech understanding. In *Journal of Acoustical Society of America*, 2000.

[LPLV05] E. Lutton, M. Pilz, and J. Levy-Vehel. The fitness map scheme. application to interactive multifractal image denoising. In *CEC2005, IEEE Congress on Evolutionary Computation is in Edinburgh, UK*, 2005.

[LWZF96] D.T. Lawson, B.S. Wilson, M. Zerbi, and C.C. Finley. Speech processors for auditory prostheses. In *Third Quaterly Progress report. NIH Contract N01-DC-5-2103*, 1996.

[Mey90] Y. Meyer. *Ondelettes et Opérateurs*. Hermann, Paris, 1990.

[Moo95] B.C.J. Moore. Perceptual consequences of cochlear damage. In *Oxford Medical Publication*, 1995.

[Nih89] Nih. Consensus development conference statement on cochlear implants. In *Arch otolaryngol head neck*, 1989.

[Osb97] M.J. Osberger. Cochlear implantation in children under the age of two years: candidacy considerations. In *Otolaryngol of Head and Neck Surgery*, 1997.

[PC97] R. Poli and S. Cagnoni. Genetic programming with user-driven selection: Experiments on the evolution of algorithms for image enhancement. In *2nd Annual Conference on Genetic Programming*, 1997.

[PCMF79] P. Pialoux, C.H. Chouard, B. Meyer, and C. Fugain. Indications and results of the multichannel cochlear implant. In *Acta Otolaryngol*, volume 87(3-4):185-9, Mar-Apr 1979.

[Rom98] S. Roman. *La réhabilitation des sourds profonds par implant cochléaire*. Thèse de doctorat de médecine, Université de Marseille, 1998.

[Rou01] G. Roux. *Synthèse et réalisation d'études cliniques sur l'implant cochléaire*. Thèse de doctorat de médecine, Université de Rennes, 2001.

[SGD02] R.V. Shannon, J.J. 3rd Galvin, and Baskent D. Holes in hearing. In *J. Assoc. Res. Otolaryngol.*, volume 3(2), pages 185–199, Jun 2002.

[SLM+06] G.S. Stickney, P.C. Loizou, L. Mishra, P.F. Assmann, R.V. Shannon, and J.M. Opie. Effects of electrode design and configuration on channel interactions. In *Hearing Research*, volume 211, pages 33–45, 2006.

[Tak98] H. Takagi. Interactive evolutionary computation : System optimisation based on human subjective evaluation. In *Proceedings of the IEEE Intelligent Engineering Systems (INES'98)*, Vienna, Austria, 1998.

[Tak01] H. Takagi. Interactive evolutionary computation: fusion of the capabilities of ec optimization and human evaluation. In *Proceedings of the IEEE, Vol.89, No.9*, 2001.

[Wal] M. Wall. Matthew's genetic library. http://lancet.mit.edu/ga/.

[ZCW97] T.A. Zwolan, L.M. Collins, and G.H. Wakefield. Electrode discrimination and speech recognition in postlingually deafened adult cochlear implant subjects. In *J. Acoust. Soc. Am.*, volume 102(6), pages 3673–3685, Dec 1997.