Auditory and Visual-based Signal Processing with Interactive Evolutionary Computation

Takagi, Hideyuki
Kyushu Institute of Design

出版情報：2004-04-01. World Scientific
バージョン：
権利関係：

Auditory and Visual-based Signal Processing with Interactive Evolutionary Computation

Hideyuki Takagi

Kyushu Institute of Design, Fukuoka, 815-8540 Japan
takagi@kyushu-id.ac.jp, http://www.kyushu-id.ac.jp/~takagi

Abstract. We discuss new type of signal processing approach, auditory and visual-based signal processing, and show some examples of filter design based on this approach. This approach allows us to design filters based on our auditory or visual evaluation of the filter outputs, while the conventional signal processing is based on mathematical expression. Interactive evolutionary computation is a main technology of this approach, and the evolutionary computation optimizes the target filter based on our auditory or visual inspection to the processed signal. Recovering speech from distorted speech, compensation of hearing impairment, and enhancing images are discussed as examples of this new approach of signal processing.

1 Introduction

Signal processing has been based on mathematics and dramatically progressed in these several decades. According to the rapid increase of computer performance, not only signal processing experts but also ordinary people receive the benefit of the advanced signal processing in our daily lives. We came to be able to handle multimedia daily thanks to the data compression technology used for mobile phones, DVD, MP3 and others; it became easy for us to retouch photo images taken by a digital camera or a scanner and make personal greeting cards; nobody is surprised to see photos from Mars or the vegetation images on the earth taken from an artificial satellite at home.

The cases that a domain expert is not an expert of signal processing and vise versa have increased according to the expansion of the application fields of the signal processing, for example, medical doctors who make diagnosis using medical images are not experts of the signal image processing. It is difficult for the approach of the conventional signal processing to solve this gap.

Developing new approach that domain experts conduct signal processing without the knowledge on signal and signal processing but with only the domain knowledge is a way to solve this problem. In this paper, we discuss this new approach of signal processing using interactive evolutionary computation (IEC). We shortly introduce the IEC and discuss IEC-based filter design method for speech processing and image enhancement in the following sections.

2 Interactive Evolutionary Computation

One of technologies that optimizes a target system to obtain better system outputs based on our auditory or vision, such as Fig. 1 is Interactive Evolutionary Computation (IEC). The IEC can be said as system optimization based on human subjective evaluation.

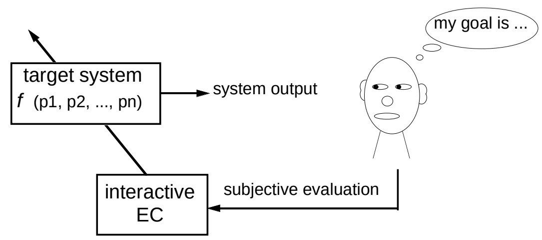

Fig. 1. General framework of an IEC system: system optimization based on subjective evaluation.

For example, suppose to design a building or bridge that well matches to the scene around, to fit a hearing aid to user's hearing, or to retrieve an image from an image database that matches given design concept. The IEC system displays the outputs of CAD, hearing aid sound, or images to an IEC user and make the user evaluate the system outputs. The IEC inputs the user's subjective evaluation as fitness values and optimizes the parameters of the CAD system, hearing aid, and retrieval engine (see Fig. 1). As a concrete EC technology, genetic algorithms (GA) or genetic programming (GP) has been frequently used.

From the optimization search point of view, the IEC searches out the coordinate that maximizes user's subjective evaluation in the space of parameters of the target system. An IEC user compares the real outputs from the target system with the ideal target in mind and evaluates the system outputs with the psychological distance between two. The EC searches the global optimum that corresponds to the ideal target in mind in a searching space consisting of parameter axes of the target system and one axis of the user's psychological distance (see Fig. 2).

3 Auditory-based Filter Design

3.1 Recovering Speech from Distorted Speech

One of our IEC researches in early time was auditory-based design of a distortion recovery filter that we aimed to evaluate how the IEC works for time-sequential display task which cannot be spatially compared and how IEC works in signal processing field that there was few IEC applications [12]. Auditory check in

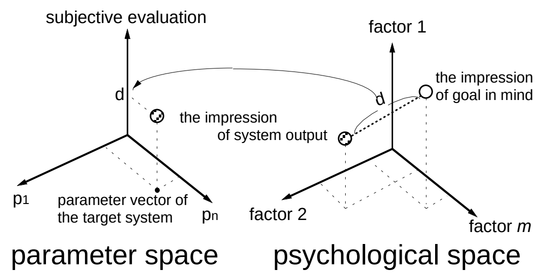

Fig. 2. Psychological distance between target in our psychological spaces and actual system outputs becomes the fitness axis of a feature parameter space where EC searches the global optimum in an IEC system.

the final stage is necessary for speech processing filter design when the filter is designed using only physical indexes because, for example, noise reduction to increase the signal-noise ratio of speech sometimes decreases speech articulation. This fact means that there are many IEC tasks that are optimized based on auditory evaluation.

The input speech in out experiment was made by suppressing the low frequency band where there is important formant information. We use an eighth order FIR filter, $f(z) = \sum_{i=0}^{7} \alpha_i z^{-i}$, as a recovering filter for distorted speech and the coefficients, α_i, are optimized by GA. An IEC user hears the filtered speech and subjectively evaluate its quality from the view point of balance between distortion and articulation. The GA uses the evaluation values as fitness values and optimizes the filter coefficients (see Fig. 3.)

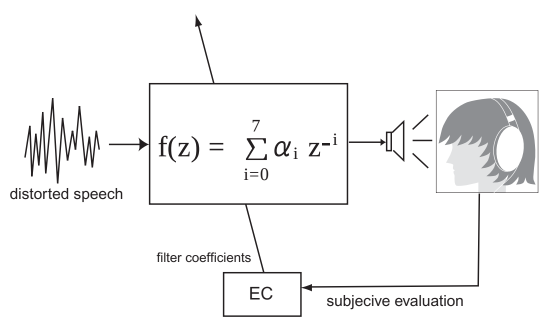

Fig. 3. IEC-based filter design recovering speech from distorted speech.

Statistical test for a subjective test showed that the recovered speech through the filter designed by the IEC was significantly better that the original speech for not only the IEC user who designed it but also other subjects who did not designed the filter [12]. This result showed that the filter was not designed as self-satisfaction, but this IEC approach was a useful technology for speech processing.

We also investigated the IEC interface for speech, music, or other sound that must be displayed to an IEC user time-sequentially. Unlike images, it is almost impossible to evaluate sounds by comparing them at same time, which causes user's fatigue problem. We proposed three methods to reduce the fatigue of an IEC user who deals with sound signal and evaluated them using subjective tests [11].

The first method decreases human fatigue by using multiple sound sources. We evaluated two systems based on this method: a system that assigns a different sound source to each individual and a system that assigns a different sound source to each generation. Subjective tests showed that the first system was not as effective as the second system, which significantly decreases fatigue. The second method displays the previous best individual as a reference to compare subsequent individuals. Subjective tests showed that this method significantly improved the IEC operability and had a tendency to improve the IEC convergence. The third method sequentially displays individuals that cannot be spatially compared. Subjective tests showed that this method significantly improved the IEC operability.

3.2 Compensation for Hearing Impairment

Wide spread of hearing aids like as glasses is necessary to help the activity of aged people in coming aging society. Not only the cost and outlooking of hearing aids but also their insufficient performance is main reason why the hearing aids have not widely spread yet. Digital hearing aids which can realize complex signal processing that the conventional analog hearing aids cannot have become popular, and the latter technical problem will be gradually solved according to the research and development of the digital hearing aids.

However, even if perfect performance of hearing aids would be realized, there is no fitting method which adjusts the hearing aid parameters for a certain user to bring the technical capability of the hearing aids into full play. The essential reason is that no one can know how other persons hear and only a hearing aid user can evaluate how his or her hearing aid is matched to his/her hearing. Unfortunately, no hearing aid user can adjust the parameters of the hearing aid for him or her unless the user is a fitting expert.

The technology that has potential to solve this fitting problem is IEC Fitting [3, 7, 4, 8, 9]. The IEC Fitting is an IEC-based hearing aid fitting method that automatically optimizes the hearing aid parameters based on how a user hears the sound processed by his/her hearing aid (see Fig. 4). This approach has unique features that the conventional hearing aid fitting methods do not have. Conventional fitting method is an approach that a fitting expert measures auditory characteristics of a user such as audiogram, observe the difference between the characteristics of the user and a hearing-normal people, and adjusts fitting parameters based on the difference and user's oral explanation on how the user hears. Problems lain in this approach are that measurable auditory characteristics are just a part of whole human auditory characteristics, it takes time to measure the characteristics, and linguistic interpretation from the user to the fitting expert is based on guessing. On the other hand, the IEC Fitting does not request to previously measure the partial auditory characteristics (though previously known characteristics help the IEC Fitting), and optimizes the parameters of a hearing aid based on user's evaluation of his/her hearing which is a result of auditory processing from auditory peripherals to the central.

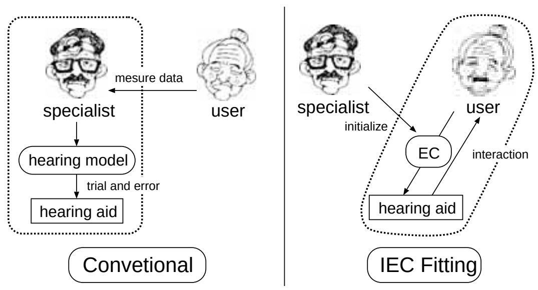

Fig. 4. Left: conventional fitting that a fitting expert adjusts a hearing aids based on his/her knowledge, experience, and user's linguistic expression on how the user hears. Right: IEC Fitting that EC optimizes a hearing aid based on how the user hears.

We have installed the IEC Fitting system on a Personal Digital Assistant (PDA). To take advantage of the biggest feature of the IEC Fitting, i.e. fitting at any time at anywhere with any source, it is important to realize the IEC Fitting system in portable equipment. Fig. 5 is an overlook photo of the IEC Fitting system on a PDA and user interface. The dots on the widest window are fitness values assigned to individuals mapped from an n-D parameter space to a 2-D space for visualization. This is our Visualized IEC letting an IEC user actively intervene in EC search by pointing the estimated the global optimum in the mapped 2-D space [1, 2, 6].

4 Visual-based Filter Design

Image enhancement is a process to visually enhance certain parts of images as domain experts desire, for example, enhancement of disease parts in medical images or certain features in geological images. The domain experts are not always experts of image processing and cannot design such image enhancement filters; at the same time, image processing experts may not have domain knowledge.

We come to be able to visually design image filters by introducing the IEC. Genetic Programing (GP) is used as an EC and generates math equations that

Fig. 5. Photo and illustration of the user interface of an IEC Fitting system on a PDA.

determine the characteristics of the filters. Concretely speaking, for example, numerical value assigned to each pixel is input to the generated math equation, and calculated value becomes a numerical value of a pixel in an enhanced image.

Some of this kind of IEC applications are designing a color filter that enhances MRI images to help the decision making of a medical doctor and designing an image synthesis filter that make it easy to see when two echo-cardiograms are mapped onto one [5].

We experimentally design an image filter based on our vision to investigate how the IEC-based approach can design filters whose characteristics are similar to those of an conventional filter that is mathematically specified. The mathematical equations generated by the IGP input numerical values of 3×3 pixels and output a new value at the center pixel among the input pixels. Images enhanced by the generated math equations are displayed to an IGP user and the user returns his/her evaluation to the GP. This process is iterated until the user finds satisfied images.

Fig. 6 shows the IGP-based filter design process and its user interface. Original input image is shown at the right upper corner on the IGP interface. An IGP user compares 12 images process of 12 math equations generated by the GP. According to his/her intention of how to emphasize the input image, subjective evaluation to each is given by clicking the five levels of input buttons under each image. The evaluation values are fed-back to the GP, and the GP creates new math equations, i.e. image enhancement filters. This process is iterated until satisfied image is obtained.

Fig. 7 (b) is an image processed by a conventional Laplacian filter using an original image of Fig. 7(a), and Fig. 7 (c) is an image processed by a filter visually designed based on IGP approach. From their comparison, the edges detected by the IGP filter is smoother than those of the Laplacian filter, which implies the IGP approach created a better edge detection filter in this case.

We also compared the characteristics of filters designed by the IGP with those of conventional high-pass filter (see Fig. 8), vertical component detection filter

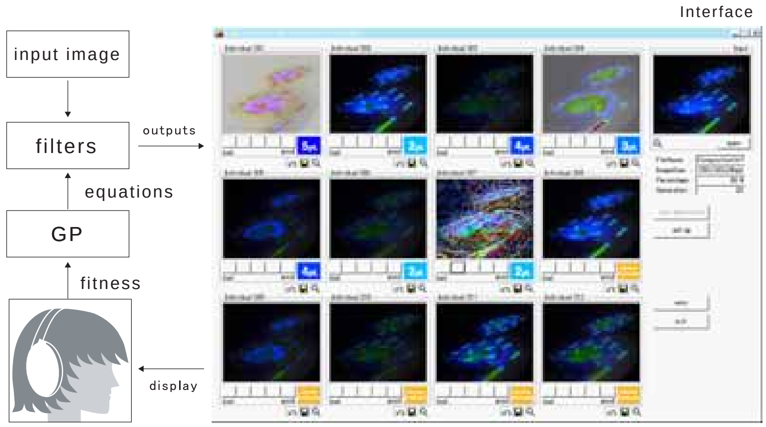

Fig. 6. IGP-base filter design process and its user interface

(a) (b) (c)

Fig. 7. (a) original image, (b) edge detection by a Laplacian filter, and (c) that by a IGP filter.

(see Fig. 9(a)(b)), and horizontal component detection filter (see Fig. 9(c)(d)). In these cases, the performance of filters visually designed by IGP do not reach to the perfect level of the conventional filters, but they are similar.

5 Discussion

From the experimental result showed that the speech distortion was significantly recovered, we can say that speech processing filters can be designed without any a priori knowledge on signal processing and the physical characteristics of the target speech but with only auditory evaluation. Although detail experimental conditions and results are not shown in this paper, we compared recovered speeches at the 10th, 20th, and 40th generations and found that the quality of the speech at the 10th generation was significantly improved and there were no significant difference among speeches at the 3 generations [12]. The population size of this experiment was 20. This result implies that 10th generation at most

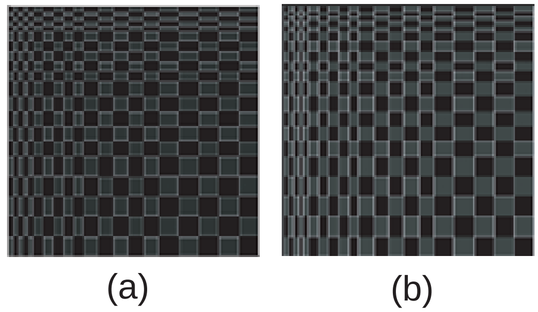

Fig. 8. (a) image processed by a linear high-pass filter, and (b) that processed by a filter generated by IGP.

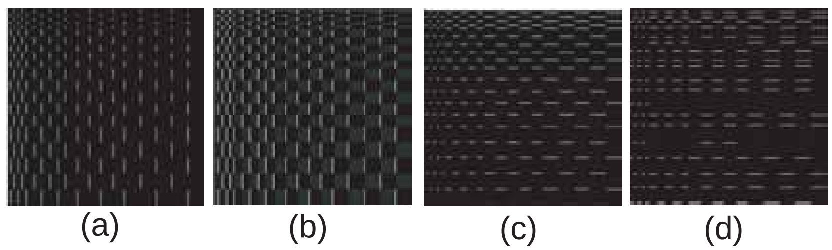

Fig. 9. Images processed by an x-axis component extraction filters: (a) Prewitt operator filter and (b) filter generated by the IGP. Images processed by a y-axis component extraction filters: (c) Prewitt operator filter and (d) filter generated by the IGP.

is enough to reach to the saturation level of its recovering performance for this task.

Same result was obtained from the experiment of designing an image enhancement filter. We use completely no a priori knowledge on image and filter design for when we designed the filter using IGP. Without such knowledge, the obtained filter characteristics became similar to those of conventional filters. Even, the edge detection performance of a filter designed by the IGP was superior to that of a conventional filter.

These experimental evaluation showed the higher potential of the IEC filter design approach.

6 Conclusion

We experimentally evaluated new approach of signal processing that we do not use any a priori knowledge on the task signal or signal processing but our auditory or visual evaluation of the process signal. The experimental result showed that even if non-experts of signal processing can quickly design filters for domain fields by using this IEC-based signal processing approach. Since the domain experts are not always the experts of signal processing, this approach must be useful for the experts in several application fields.

References

1. Hayashida, N. and Takagi, H., "Visualized IEC: Interactive evolutionary computation with multidimensional data visualization" in *IEEE Industrial Electronics, Control and Instrumentation (IECON 2000)*, Nagoya, Japan, pp. 2738–2743 (Oct., 2000).
2. Hayashida N. and Takagi H., "Acceleration of EC Convergence with Landscape Visualization and Human Intervention," *Applied Soft Computing*, Elsevier Science, (will appear in 2002).
3. Ohsaki, M. and Takagi, H., "Application of Interactive Evolutionary Computation to Optimal Tuning of Digital Hearing Aids," in *5th Int'l Conf. on Soft Computing (IIZUKA'98)*, pp.849–852, Iizuka, Fukuoka, Japan: World Scientific, Singapore (Oct., 1998).
4. Ohsaki, M. and Takagi, H., "Design and Development of an IEC-based Hearing Aids Fitting System," in *4th Asia Fuzzy System Symposium (AFSS'00)*, Tsukuba, Japan, pp.543–548 (May, 2000).
5. Poli, R. and Cagnoni, S., "Genetic programming with user-driven selection: Experiments on the evolution of algorithms for image enhancement," in *2nd Annual Conf. on Genetic Programming*, 1997, pp. 269–277.
6. Takagi, H., "Active user intervention in an EC search," in *Int. Conf. on Information Sciences (JCIS2000)*, Atlantic City, NJ, USA, pp. 995–998 (Feb./Mar., 2000).
7. Takagi, H., Kamohara, S., and Takeda, T., "Introduction of Soft Computing Techniques to Welfare Equipment," in *1999 IEEE Midnight-Sun Workshop on Soft Computing Methods in Industrial Applications (SMCia'99)*, Kuusamo, Finland, pp.116–121 (June, 1999)
8. Takagi, H. and Ohsaki, M., "IEC-based Hearing Aids Fitting," in *IEEE Int. Conf. on Systems, Man, and Cybernetics (SMC'99)*, Tokyo, Japan, **III**, pp.657–662 (Oct., 1999).
9. Takagi, H. and Ohsaki, M., "IEC fitting: New framework of hearing aid fitting based on computational intelligence technology and user's preference for hearing," in *Int'l Hearing Aid Research Conference (IHCON2000)*, Poster session PB9, pp.49–50, Lake Tahoe, CA, USA (Aug., 2000)
10. Takagi, H., "Interactive Evolutionary Computation: Fusion of the Capacities of EC Optimization and Human Evaluation," in *Proceedings of the IEEE*, **89**(9), pp.1275–1296 (2001).
11. Todoroki, Y. and Takagi, H, "User Interface of an Interactive Evolutionary Computation for Speech Processing," in *6th Int'l Conf. on Soft Computing (IIZUKA2000)*, Iizuka, Fukuoka, Japan: World Scientific, Singapore, pp.112–118 (Oct., 2000).
12. Watanabe, T. and Takagi, H., "Recovering system of the distorted speech using interactive genetic algorithms," in *IEEE Int. Conf. on Systems, Man and Cybernetics (SMC'95)*, Vancouver, Canada, Oct. 1995, **1**, pp. 684–689.