Design of Sign Sounds using an Interactive Genetic Algorithm

Mitsunori Miki, Hiroko Orita, Sanae H.Wake and Tomoyuki Hiroyasu

Abstract— In recent years, many kinds of sign sounds have come into use in our daily life. The purpose of sign sounds is to communicate a message. However the ease of message understanding and listening comfort have rarely been taken into consideration. Because people have different tastes and attach a different "image" to sign sounds according to their individual sensibility, it is important for each person to be able to make melodies freely and easily, and customize sign sounds. We propose a sign sounds design system using an Interactive Genetic Algorithm (IGA). The proposed system produces a melody of two bars based on this IGA. A melody is considered to be a sort of "organism," whose chromosome consists of genes indicating the length, pitch, and velocity of the notes of the melody. Experimental results showed that this system is effective for generating sign sounds. In creating auditory signals, it was found that the method of evolving part-by-part is better than the method of evolving the whole two bars.

I. INTRODUCTION

Auditory displays are one area of research which exemplifies the presentation of information using sound. Here, particular focus is placed on non-verbal sounds, and research has been conducted regarding techniques for effective use of such sounds as an information communication media, and the component technologies needed to achieve that. From this standpoint, Gaver and Blattner proposed using non-verbal sounds on PCs to communicate messages such as OS errors and file deletion[1], [2], and this had a significant effect on subsequent OS like Windows and Macintosh.

In this way, presentation of information using sound originated primarily in interactions with a PC, but in recent years, message communication using sound has appeared in various contexts in ordinary life, such as the ring tones of mobile phones. Sounds like these are called "sign sounds"[3]. Sign sounds are used in many contexts, particularly in Japan. They are widely used-in both public and private spaces-in sounds announcing train arrival, household appliances and crosswalk signals.

The purpose of sign sounds is to communicate information, but in reality, not much thought has been put into issues like pleasantness when hearing a sound, or how easy it is for users to understand the message [3]. While appliances in the home have diversified in response to diversifying uses, the sign sounds produced by each device do not function as part of an overall system, and this can confuse the user [4].

For this reason, the Japanese Standards Association decided to improve the functionality of message communication using sign sounds, and in 2002 issued " Guidelines for the elderly and people with disabilities-Auditory signals on consumer products " [5], a document stipulating frequencies, and describing matters like the patterning of melodies for starting and stopping. Also, in response to design issues like personal preferences and listening comfort for sign sounds, a new product called MUPASS (Music & Multidate pass) [6] has appeared on the market. With MUPASS, melodies downloaded via mobile phone are transferred to an appliance using infrared signals, so that users can switch to sign sounds they prefer.

In this way, research and trials have been conducted to ameliorate the problems of sign sounds, but today it is still hard to say that points like functionality and preferences have been adequately addressed. Also, people have preferences for melodies, and factors like " image " vary depending on the sensibility of the individual. Therefore, in addition to downloading existing melodies, it would be desirable if users could freely and easily create sign sounds, so they could customize them to coordinate with their tastes, or other appliances.

However, creating melodies is not easy for the ordinary person. For this reason, this research proposes a sign sound design system in which sign sounds based on optimal melodies are generated in response to evaluation of a person. An Interactive Genetic Algorithm (IGA) [7] is used as the optimization technique.

II. SIGN SOUNDS

"Sign sounds" is a general term indicating sounds intentionally added by a manufacturer to products, equipment or the environment-such as warning sounds produced by appliance products, ring-tones of mobile phones and departure bells for trains [3]. Sign sounds can be roughly classified into electronic sounds and melody sounds. Their respective features are given below.

- Electronic sounds
  These are sign sounds which communicate a message with a simple sound only. A typical example is the "beep-beep" sound of an electronic thermometer. Since the sound source is inexpensive, this type is frequently used in old electronic products. However, there are some problems. For example, there are individual differences, due to age and other factors, in the frequency band of

M. Miki is with Department of Knowledge Engineering and Computer Science, Doshisha University, 1-3 Tataramiyakodani, Kyotanabe-shi Kyoto `mmiki@mail.doshisha.ac.jp`

H. Orita is with Graduate School of Engineering,Doshisha University, 1-3 Tataramiyakodani, Kyotanabe-shi Kyoto `horita@mikilab.doshisha.ac.jp`

S.H. Wake is with Department of Information and Media, Doshisha Women's College, 1-3 Tataramiyakodani, Kyotanabe-shi Kyoto `swake@dwc.doshisha.ac.jp`

T. Hiroysau is with Department of Knowledge Engineering and Computer Science, Doshisha University `tomo@is.doshisha.ac.jp`

sounds which are easy-to-hear, and thus there are cases where the sound is hard-to-hear or jarring.

- Melody sounds
  These are sign sounds which communicate a message with a melody sound. Typical examples are the ring-tones used throughout the world in mobile phones, and the train departure melodies which are played on station platforms in Japan. Melody sounds incorporate various pitches, so they are easier and more pleasant to listen to. They are sounds which communicate information, and at the same time, they can be used as sounds for dramatizing and creating the "image" of a space [8].

The purpose of this research was integrated design of sign sounds in a space, and research focused on melody-based sign sounds which take listening comfort into account. In contrast with the character of ordinary tunes, it is crucial for sign sounds to accurately communicate various kinds of information in short phrases.

III. INTERACTIVE GENETIC ALGORITHM

An "Interactive Genetic Algorithm" (IGA) [7] is a Genetic Algorithm (GA) [9] which simulates evolution of organisms, where the evaluation part of the GA is handled subjectively by a human being. In problems which cannot be numerically quantified because they involve the impressions and tastes of human beings, optimization is done based on evaluation according to human sensibility. Paying attention to this feature, IGA has been applied in various fields such as art, engineering, and entertainments. For the music field, there have been a lot of researches with IGA too [11]. From these researches, IGA is found to be useful method to reflect human sensibility to create some music better than GA [12], [13].

In contrast with the character of ordinary tunes, it is crucial for sign sounds to accurately communicate various kinds of information in short phrases. In this research, the aim was to develop a technique based on IGA which allows even ordinary users, who cannot create melodies using musical instruments, to simply create good and purposeful sign sounds simply by evaluating based on their own subjectivity as an individual.

IV. A SIGN SOUND DESIGN SYSTEM EMPLOYING AN INTERACTIVE GENETIC ALGORITHM

A. Method of representing sound

The method of representing sound in the proposed system was determined as follows by performing preliminary experiments from various standpoints, and taking those results into consideration.

In the proposed system, note length was determined by taking one 8th note to be the basis for one note. Each gene stores a flag indicating connection of tones, the tone pitch, and velocity information indicating the strength (loudness) of the tone. Note numbers defined using a Standard MIDI File (SMF) were used to represent tone pitch.

In the note numbers in the SMF, 60 is taken to be middle C (Do) on the piano, and the numbers change by 1 for each

semi-tone. The lowest tone is defined to be 0, and the highest tone is defined to be 127. The note numbers used in the research were selected from the C major scale, from which it is easy to generate melodies with a bright sound, and note numbers corresponding to black keys were not used.

Also, for simplicity, the range of notes was set to note numbers 60 to 79 only. Note number 0 (not used for notes) was used for rests, and note number 80 (also not used for notes) was used for the flag indicating note connection. Velocity was expressed with values 0 to 127, with larger values indicating louder tones. However, if the velocity is too small, the tones are hard to hear, and thus the range 50 to 127 was used in this research. Fig. 1 shows the correspondence between tone pitches and note numbers, and Fig. 2 shows the correspondence between chromosome structure and the generated melody.
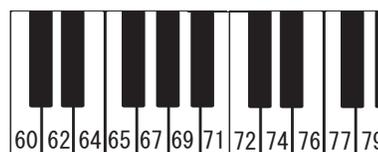

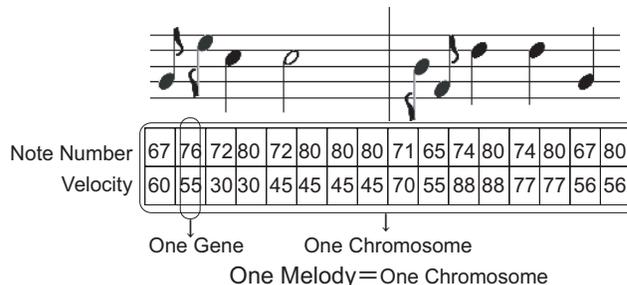Fig. 1. Correspondence between tone pitches and note numbers.

Fig. 2. Correspondence between chromosome structure and the generated melody.

B. Genetic operation

The process of melody generation using IGA is shown in Fig. 3.

1) Generation of initial individual
   The pitches of sounds in the initial individual are generated randomly from the note number values 60 to 79 and the rest 0, which are defined in the SMF. For notes other than the first note, the number 80 (the flag indicating note connection) also becomes a candidate for selection. Velocity is generated randomly, in the same way, from values 50 to 127.

2) Presentation and evaluation
   Fig. 4 shows the computer screen presented to the user during evaluation. The number of individuals at this presentation screen and the method of representation were determined as follows based on preliminary experiments.
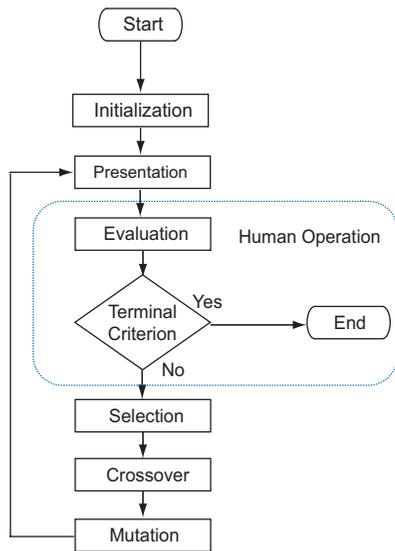
Fig. 3.   Flow Chart of IGA.

Six individuals are presented on the screen as a score (musical notation), and the user can listen with headphones to the melody corresponding to a score by pushing the presented Play button. The system simultaneously displays a melody and score, so it is easier for the user to remember the melody. Using the computer screen shown in Fig. 4, the user listens to the melody as the displayed individual, and evaluates it with a score of 1 to 5 points. With a GA, fitness is calculated using an evaluation function, but with an IGA, the operator subjectively evaluates the individual to determine its fitness. A five level evaluation method was adopted because it is easier for the operator to do the evaluation if it is "rough" to a certain degree [10]. Also, the operator chooses one individual as the "elite individual" he or she wishes to remain in the next generation, and the character of this individual is retained perfectly in one individual of the next generation. The candidate selected as the elite individual in the final generation becomes the final generated sign sound.
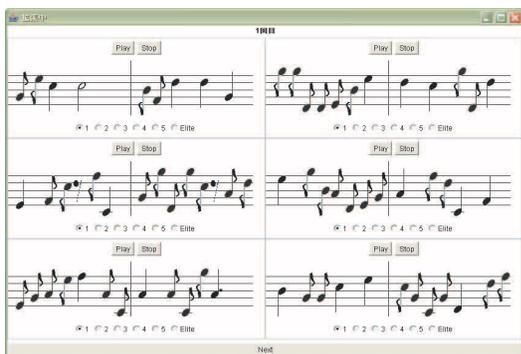

Fig. 4.   Computer screen presented to the user during evaluation.

3) Selection
System designated selection (roulette selection or tournament selection) is performed based on evaluation conducted by the user.

4) Crossover
In the melodies to be optimized, short phrases comprised of a sequence of few notes become important. Therefore, crossover must be done so as to not destroy phrases if at all possible. Experiments were conducted with phrase sizes of 1/4, 1/2 and 1 bar, and it was felt that the characteristics of the parent generation were inherited best by the child generation when phrases were 1 bar. Therefore, phrase size was set to 1 bar, and one-point crossover was performed in phrase units. Fig. 5 shows an example of crossover.
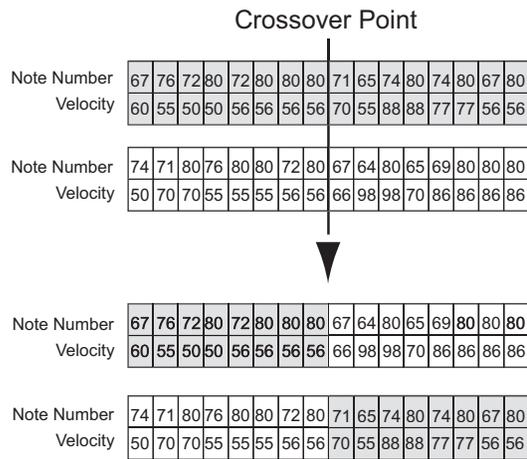

Fig. 5.   Crossover.

5) Mutation
If variation is done uniformly without designating a range for tone pitch, the number of lethal genes increases, and thus variation is done randomly in a range of 2 steps above or below the height of the original tone. Velocity is varied randomly within the defined range.

C. Evolution where part of the melody is not varied

In this research, we considered a method, based on the sign sound design system described in Sections IV-A and IV-B, where part of the melody is fixed, and the other remaining part is evolved. When evaluating sign sounds, there are cases where one part is highly evaluated, even when the evaluation for the entire melody is low. Conversely, there are cases where one part has a low evaluation, even though the overall evaluation is high. For that reason, we felt a good method would be to generate specific phrases beforehand, incorporate those specific phrases when generating an overall melody, and then evaluate the entire melody in accordance with that. Fig. 6 gives an overview of a system incorporating this technique. Here, for example, the end part of the sign sound was regarded as a key part during generation, and a phrase for the end part was generated beforehand in accordance with

that concept. The generated phrase was then incorporated into the end part of all individuals when generating overall melodies.
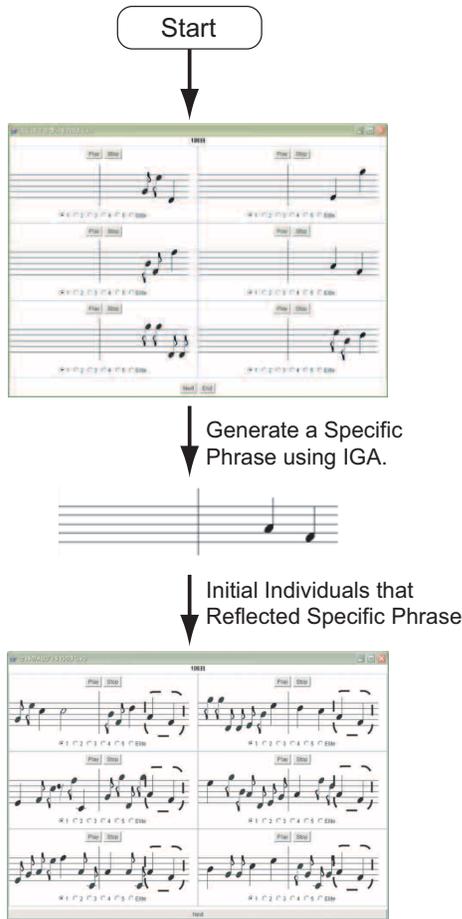

Fig. 6.   Overview of a System of Evolving Part-by-Part.

V. EXPERIMENT

A. Overview of experiment

In order to verify the proposed system, experiments were conducted where sign sounds were generated based on the concept of "a microwave oven alerting the user that warming of food has finished." In this experiment, two cases were compared: the case where melodies were generated using a sign sound design system where the entire melody is evolved (hereafter called a "total evolution system") and the case where melodies were generated using a system where evolution was done while keeping part of the melody fixed (hereafter called a "partial evolution system"). Also, in this experiment sign sounds were generated to alert the user that a process was finished, and thus it was felt that the final part of the melody was the key phrase. Thus a half bar of melody for the end part was generated beforehand, and this was incorporated into the final part when generating all melodies.

The subjects of the experiment were 20 students (male and female in their early 20s) at Doshisha University. The main parameters for each system were as follows: individuals: 6, generations: 5, crossover rate: 1.0, mutation rate: 0.0625, selection: roulette selection. The generated sign sounds were set to be two bars in four-four time. At the end of the experiment, the subjects were asked to fill out a questionnaire with the following items.

1) Evaluation item 1
   Were you able to generate a melody matching the concept using the total evolution system?
2) Evaluation item 2
   Comparing the melody generated using the total evolution system and the melody generated using the partial evolution system, which best matched the concept?
3) Evaluation item 3
   Comparing melody generation using the total evolution system and melody generation using the partial evolution system, which type of melody generation was simpler?

B. Experimental results and analysis

1) Evaluation item 1
   Fig. 7 shows the subject evaluation percentages for evaluation item 1. The figure shows that 65% of the subjects replied that the sign sound generated using the total evolution system either matched the concept or almost matched the concept. This shows that a sign sound design system using IGA can generate sign sounds matching a concept. We can also say that the proposed method of representing sounds, and the methods of coding, crossover and presentation, functioned effectively in sign sound generation.
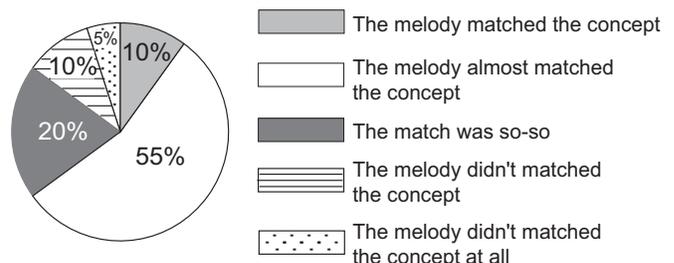

Fig. 7.   Result of Evaluation Item 1.

2) Evaluation item 2
   Fig. 8 shows the subject evaluation percentages for evaluation item 2. The figure shows that melodies matching the concept could be generated better when sign sounds were generated using the partial evolution system-i.e. by generating a specific phrase suited to the end part beforehand, and the generating the overall melody by incorporating that generated phrase.
3) Evaluation item 3
   Fig. 9 shows the subject evaluation percentages for evaluation item 3. The figure shows that sign sounds can be more easily generated when the key phrase is generated beforehand using the partial evolution system.
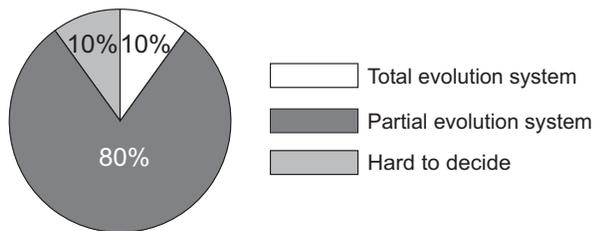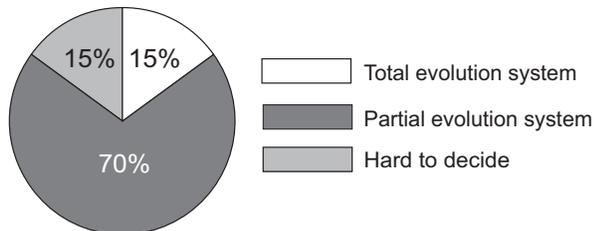
Fig. 8.   Result of Evaluation Item 2.

Fig. 9.   Result of Evaluation Item 3.

Evaluation items 1, 2 and 3 showed that this sign sound design system enables generation of sign sounds matching a concept even by ordinary people who cannot create melodies. It was also found that, when generating sign sounds, the technique of generating the part of the melody thought to be crucial before evaluating the entire melody (taking into consideration factors like the message qualities of melody), and then incorporating that into the generation of all melodies, is an effective technique for generating sign sounds matching a concept.

VI. CONCLUSION AND FUTURE ISSUES

In this research, we developed a system for generating melody-based sign sounds matching a concept, using IGA and subjective evaluation by people, and verified the effectiveness of that system. Evaluation experiments using the system showed that, in the proposed system, the method of representing tones, and the method of coding, crossover and presentation to the user during evaluation, functioned properly, and the system was effective for generating sign sounds. Also, it was possible to generate good sign sounds, even when the entire 2 bars were evolved, but good sign sounds could be generated more simply by using the partial evolution method where evolution is done by generating part of the melody beforehand, and incorporating that into other parts. This is thought to be because the desires of the user can be more effectively reflected in the functionality part of communicating the message of the sign sound when a phrase thought to be crucial is generated beforehand as a specific phrase. Issues for the future include: saving multiple specific melodies which are generated beforehand, generating harmonized melodies, and reducing the fatigue of the user during evaluation.

REFERENCES

[1] W. Gaver   The SonicFinder : An Interface That Uses Auditory Icons   Human-Computer Interaction   Vol.4   No.5, 1989, pp.67-94

[2] M.M. Blattner   D.A. Sumikawa, R.M. Greeberg, Earcons and Icons : Theire Structure and Common Design Principales   Human-Computer Interaction, Vol.4, 1989, pp.11-44

[3] Wake, S. H. and Y. Okada, "Notification Sound Design for Human Interfaces," Journal of the Japanese Society for the Science of Design (Design Science Research), Vol. 49, No. 5, 2003, pp. 41-50 (in Japanese)

[4] Nishida, K. and Y. Shinjo, "Easy-to-Understand, Pleasant Notification Sounds which are Elderly-Friendly," Sharp Technical Journal, Vol. 77, 2000, pp. 48-52 (in Japanese)

[5] Japanese Standards Association, "Guidelines for the elderly and people with disabilities-Auditory signals on consumer products," 2002 (in Japanese)

[6] Sammy   NetWorks,   http://www.sammy-net.jp/news/pdf/mupass04120201.pdf, 2004

[7] Takagi, H., T. Unemi, and T. Terano, "Interactive Evolution Computation," Genetic Algorithms 4, Sangyo Tosho, 2000, pp. 325-361 (in Japanese)

[8] Nakamura, H., "Entertainment and Sound: Creation of Sound Environments and the Psychological Effects of Sound," AMBUSINESS, 2000 (in Japanese)

[9] Goldberg,D, Genetic Algorithms in Search Optimization and Machine Learnig, Addision Wesley, Reading,Mass   1989

[10] K.Oya,M.Osaki,H.Takagi, An input method using discrete fitness values for interactive GA, J.of Inteligent and Fuzzy Systems, Vol.6, No.1, 1998, pp.131-145

[11] H Takagi, Interactive Evolutionary Computation: Fusion of the Capabilities of EC Optimization and HumanEvaluation, Proc. of the IEEE, vol.89, No.9, pp.1275- 1296, 2001

[12] M.Unehara and T.Onisawa, Interactive Music Composition System, Proc. of 2002 IEEE International Conference on Systems, Man, and Cybernetics, Hammamet,Tunisia, 2002

[13] Roads, C, The Computer Music Tutorial, MIT Press, 1996