Proposal of Genetic Operations Reducing the Evaluator Workload to the Voice Quality Conversion Using Interactive GA

Takenori Nishizono Satoshi Noami Yuji Sato
Graduate School of Computer and Information Sciences, Hosei University
3-7-2 Kajino-cho, Koganei-shi, Tokyo, Japan
i04t0014@k.hosei.ac.jp

Abstract

We have already proposed the voice quality conversion system using interactive evolution. Users can convert our voice to various voice qualities because it does not require a codebook. However the practical application of the system requires reducing the evaluator workload. This paper proposes two kinds of genetic operations, that is, “individual complement” and “adjustment after genetic operations” that aim to reduce the evaluator workload by improving the efficiency of tournament selection. The experimental result using the expediential function shows the proposed operations have effect of reduction of the workload.

Keywords: Interactive Evolution; Evolutionary Computation; Voice Quality Conversion;

1. Introduction

The growth of communication technology increases variety of the online communication methods. In these days, as typified by online meeting services, we can communicate by our real voice on line. The development of voice processing technology is important for the growth of application area of these.

The usual voice modification system realizes a specified voice quality by modifying another specified voice according to a codebook. It requires getting the voice data of conversion object and creating the codebook corresponding to the conversion. Therefore, it is only used to improve synthetic voice quality and to convert between the specified speakers. On the other hand, we have already proposed the voice modification system using interactive evolution [1]. Users can convert a voice selected freely to various voice qualities with it because it does not require the codebook. However a system using interactive evolution requires reducing the evaluator workload to put into practical use [2].

We applied tournament selection for evaluator workload reduction to the system. For further improvement, this paper proposes two kinds of genetic manipulation that aims to improve the efficiency of tournament selection, and reports the evaluation result.

2. Voice modification system using interactive GA

Fig. 1 shows the voice modification system using interactive evolution. The system is trisected by a voice conversion part which converts input voice with conversion coefficients, a learning part which search the conversion coefficients with genetic algorithms (GA), and a GUI part. This section explains the system with the focus on the voice conversion part and the learning part.

The voice conversion part treats three prosodic elements – voice frequency, voice amplitude, and speech tempo – as control information. It creates new voice quality by compressing or expanding the above elements according to conversion coefficients that are real number. This paper describes the conversion coefficient of frequency as PIT, of amplitude as POW, and of speech tempo as TMP. The learning part searches PIT, POW, and TMP which realize the target voice quality.

The learning part repeats the search process explained below until the system detects the target. This section explains it by assuming the population size is 2N if it does not apply the operation proposed in this paper. The individuals in population consist of three kinds of chromosome. The each chromosome keeps PIT, POW, and TMP as bit string. First the process performs tournament selection that a user chooses one candidate among three. The selection is repeated N times. Therefore, it has N individuals after finishing the selection. Then it creates new N individuals by performing one-point crossover and mutation to the selected N individuals. At last it creates the population in net generation by adding the N individuals before and

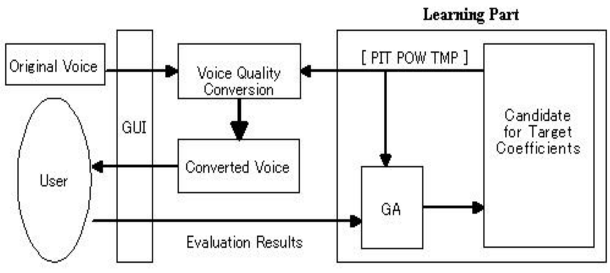

Fig. 1: The system configuration outline of voice quality conversion system.

after performing the crossover and mutation.

The selection presents individual in population at least one time, and forbids selecting a candidate twice. Although the selection reduces the evaluator workload required in one generation, it may increase the generation number until convergence. For further reduction, the next section proposes two kinds of genetic manipulation to reduce the search time.

3. Proposal of genetic operations reducing the evaluator workload

This section proposes two kinds of genetic operations in order to reduce the evaluator workload. First the section 3.1 proposes “individual complement” for reduction of evaluation times required in one generation. The section 3.2 proposes “adjustment after genetic operations” for diversity maintenance of population. At the last of this section, we show the search process which applying the proposed operation.

3.1. Individual complement

The reduction of evaluator workload is realized by curtailment of evaluation times in one generation. However it has adverse effect to search ability because it cut down the population size. The operation “individual complement” realizes the curtailment by complementing the reduction of individuals with individuals created based on the individuals selected in same generation. In addition, this operation aims at improvement of search ability by complementing with individuals created based on bit information of the selected individuals. This operation creates the individuals a reduction number of evaluation times, and uses these to complement.

This section below explains how to create the individual to complement the individual reduction. The individual consists of PIT, POW, and TMP chromosomes that are created by a process explained below. This section explains the process if it creates a PIT chromosome. A similar process creates the others. First the process takes PIT chromosome form the individuals selected in same generation, and creates a group of PIT chromosome. Next it investigates the maximum and the minimum of real-valued PIT kept in the group. Then it calculates an absolute value of difference between the maximum and the minimum of PIT. If the absolute value is over a threshold which is set beforehand, it creates the PIT chromosome by using the upper creation method in Fig. 2. It choices one chromosome from the group, and makes it the PIT chromosome of individual for complement. If the absolute value is under the threshold, it uses the lower method in Fig. 2. It investigates the percentage of bit value in each digit of PIT chromosome in the group of PIT, and creates a probability table of bit value as shown in Fig. 2. Then it creates bit string according to the table, and makes the created bit string the PIT chromosome of the individual for complement.

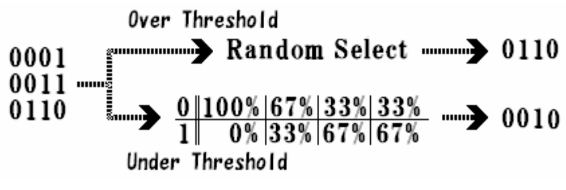

Fig. 2: The creation method of chromosome based on chromosome group

3.2. Adjustment after genetic operations

Next this section proposes the operation “adjustment after genetic operations” in order to maintain diversity of population. The “individual complement” proposed in previous section promotes early detection of the target individual. However it may make the population drop into premature convergence because it reduces diversity. The operation proposed in this section maintains diversity, and tries to improve the effect of the individual complement.

This operation is illustrated by Fig. 3. First it investigates whether the population has two same individuals. If the population has, it performs mutation to chromosome in the one individual with fixed probability. This mutation makes a chromosome mutate within about between -0.256 and 0.256 from the original value if the chromosome has 3 decimal places. In Fig. 3, the lower individual in the squared two individuals in mutated. The convergence speed is easy to become blunt by the target individual. Therefore, this mutation does not take very big mutation amount.

The purpose of this operation is near to the purpose of heterogeneous recombination in CHC (cross generational elitist selection, heterogeneous recombination, and cataclysmic mutation) [3]. However this operation differs from the heterogeneous recombination because this operation maintains the

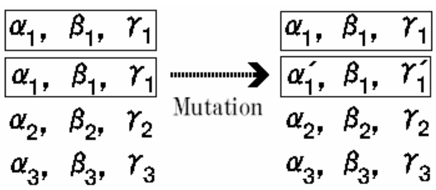

Fig. 3: The image of adjustment after genetic operations

number of individuals. This operation increases diversity with keeping the number of individuals, and maintains the number of candidate's kinds. According to the above, this operation aims at the early detection of target individual.

At last, we explains the search process if the system applies the proposed operations. We assume that the population size is 2N to contrast with the search process explained in section 2. First the process performs tournament selection with threefold choice. The selection is performed 2N/3 times in one generation. The number of evaluation times is decided by the lowest times required to bring up all individuals with the threefold choice evaluation. As the result of this selection, 2N/3 individuals are selected. Next it increases the number of the individuals to N by performing the "individual complement". After that, it creates new N individuals by performing one-point crossover and mutation to the N individuals. It creates 2N individuals by adding the N individuals before and after performing crossover and mutation. Finally, it performs the "adjustment after genetic operations" to the 2N individuals, and makes the individuals the population in next generation. By performing the proposed operations as described above, the operations aims at improvement of efficiency of tournament selection.

4. Evaluation experiment

4.1. Evaluation method

Since the voice modification system adopts interactive GA, repeating the experiment of this system tasks the evaluator. Therefore, as a matter of convenience, this paper experiments with an evaluation function described below.

The evaluation function calculates a point of each candidate, and choices a candidate which has the lowest point. The point of a candidate is decided by a total of three values. These are absolute value of difference between each conversion coefficient of the target and the candidate. The experimental system searches the target with it, and records the generation number when the system detects the target.

This experiment was performed with the parameters below. The population size was 30. Therefore, it performs the selection 10 times in one generation if the system applies the "individual complement", and performs 15 times if it does not apply. PIT and TMP have 3 decimal places, and POW has 2. The bit length of PIT or TMP chromosome is 13, and that of POW is 11. The crossover rate is 1.0. The mutation rate was 0.3. The threshold used in the individual complement was set 0.0025 to PIT or TMP, and set 0.025 to POW. The mutation rate in the adjustment after genetic operations was set 0.4. A coefficient value of the initial individuals was created randomly within the range from 0.4 to 2.5 if it is PIT, within the range from 0.4 to 2.5 if it is POW, and within the range from 0.5 to 2.5 if it is TMP.

4.2. Experimental results

This section reports the evaluation results. We performed experiment to the 4 kinds of the system; the system that does not apply the proposed operations, that applies one of the proposed operations, and that applies both proposed operations. Each experiment is performed 5000 times, and graph-ized the result. The target conversion coefficients of these experiments were set 1.5 to PIT, set 4.0 to POW and set 0.8 to TMP.

First we show the result applying the individual complement by Fig. 4. The horizontal axis of the graph shows the generation number when the system detected the target. The vertical axis shows the number of times which the system could detect on the generation number of the horizontal axis. The result shows the operation decreased the average number of generation required until detection from 50.5426 to 43.314, the shortest number form 19 to 16, and the maximum number from 128 to 110. On the other hand, the application of adjustment after genetic operations decreased the shortest number of generation required until the target detection from 19 to 14, and the maximum number from 128 to 98. However, it had little effect to the average number.

Next, this section shows the result when the system applied both proposed operations. Fig. 5 shows the experimental result. This experiment took the best result in this paper. The average number of generations was 38.3840, the shortest number was 12, and the maximum number was 84. In addition, as shown in Fig. 6, we compare this result with the result in the case the system applied half of the proposed operations. Although the adjustment after genetic operations had little effect to the average number in the previous result, it had effect when it is applied with the individual complement. By applying it to the system that applies the individual complement, it decreased the average number of generations from 43.3140 to 38.3840, the shortest number from 110 to 84, and the maximum number from 16 to 14.

4.3. Discussion

Fig. 4 shows the individual complement is effective to improve the efficiency of tournament selection. It decreases evaluation times required in one generation by 2/3 in addition to the reduction effect of the generation number. The adjustment after genetic operations did not have effect to the average number of generations required until target detection. However, Fig. 7 shows it reduces the average number when applied system applied the individual complement. We consider that this result shows the adjustment after genetic operations complements diversity which is decreased by the individual complement.

The section below estimates how much the evaluation times were reduced by the proposed operations according to the experimental results. If the system did not apply the proposed operations, it performed evaluation 15 times in one generation and required 50.5426 generations until target detection on an average. Therefore, it gave users 758.139 times evaluation on an average. On the other hand, if the system applied, it evaluated 10 times and required 38.3840 generations. Therefore, it required 383.84 times evaluation on an average. Altogether, the proposed operations decreased the evaluation times by about 49%. It means that the proposed operations are effective to the efficiency improvement of tournament selection.

5. Conclusion

This paper proposed the "individual complement" and the "adjustment after genetic operations" to reduce the evaluator workload in the voice modification system using interactive GA which applied tournament selection. As a result of the experiment using a function which evaluates voices replaced by users, the proposed operations reduced the evaluation times required until target detection about 49%. In future, the system requires further improvement of a system side and interface side to the system easy to use.

References

[1] Sato Y.:"Voice Conversion Using Interactive Evolution of Prosodic Control", Proc. Of the 2002 Genetic and Evolutionary Computation Conference (GECCO-2002), Morgan Kaufmann Publishers, pp.1204-1211 (2002).

[2] H. Takagi: Interactive Evolutionary Computation: Fusion of the Capabilities of EC Optimization and Human Evaluation. In: "Tutorial Book of the 2001 Congress on Evolutionnary Compuation", IEEE Press, 2001.

[3] D.E. Goldberg, "Genetic Algorithms in Search, Optimization, and Machine Learning", Addison-Wesley, 1989.

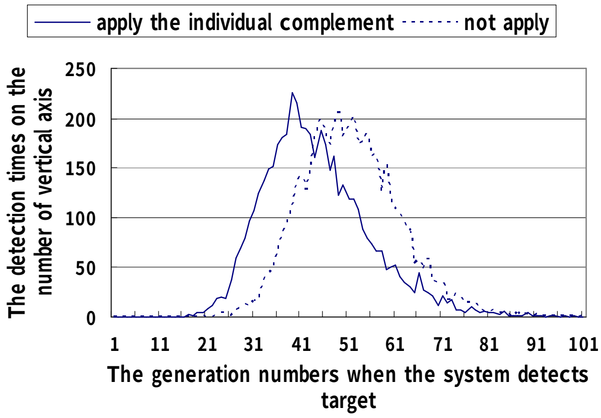

Fig. 4: The result applying the individual complement

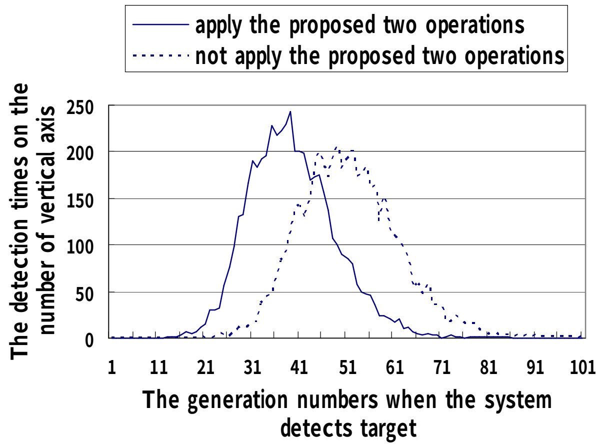

Fig. 5: The result applying the proposed operations

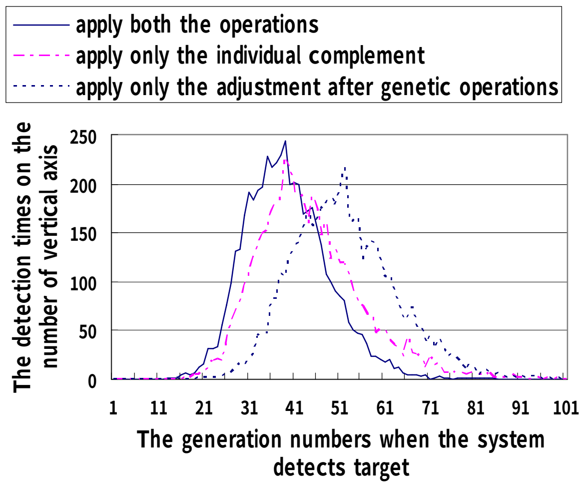

Fig. 6: The comparison between the result applying the proposed operations and the result applying the one-half of operations