Quick Individual Fitting Methods of Simplified Hearing Compensation for Elderly People

Kengo Fujita, Tsuneo Kato and Hisashi Kawai

KDDI R&D Laboratories Inc.
2-1-15 Ohara, Fujimino, Saitama, 356-8502 Japan
fujita@kddilabs.jp

ABSTRACT

For simplified hearing compensation, which combines a high-frequency boost, pitch shift and speech rate conversion, two quick individual fitting procedures, based on subjective assessments, are proposed. Since conventional individual fitting procedures impose enormous hearing comparisons on a user for decision of an appropriate parameter set, the proposed procedures split the fitting process into two steps, namely a parameter fitting step for each hearing compensation method and a comprehensive combination step, to reduce the number of comparisons. Fitting experiments with twenty elderly subjects over sixty-five showed that the proposed procedures had effect on hearing with an average of only around 20 comparisons, while the conventional procedure requires hundreds of comparisons.

Index Terms: hearing compensation, elderly people, individual fitting

1. INTRODUCTION

There are a number of elderly people with hearing difficulty. Since the causes and degrees of hearing difficulty vary between individuals, individual fitting of hearing compensation is necessary. We study the individual fitting of hearing compensation to make cellular phone speech more intelligible for elderly users. Unlike a hearing aid, which is fitted to the user's auditory property completely, our fitting system treats a simplified hearing compensation, which combines three hearing compensation methods with discrete levels: a high-frequency boost over 1.2 kHz with five gain levels, a pitch shift with five shift levels, and a speech rate conversion with three conversion levels. An appropriate parameter set of the three hearing compensation methods is selected, based on subjective user assessment.

Several individual fitting methods based on repetitive subjective assessments [1-4] have been proposed. The conventional methods are categorized into two types: a type which presents more than three test speech at a time and requests a user to rank them [1-3], and a type which presents paired test speech and requests a user to choose the better quality one [4]. All the conventional methods are based on GA algorithm. The former type requires very long fitting time because the user has to listen to the test speech a number of times to rank them. As a long fitting time must be tedious to elderly people, a fitting system based on paired comparisons is discussed here. However the use of a paired comparison does not directly reduce the number of comparisons. For example, the number of combinations of five levels of a high-frequency boost, five levels of a pitch shift and three levels of a speech rate conversion is 75 (= 5*5*3). If the 75 parameter sets are compared in round-robin manner, the number of paired comparisons comes to 2775 (= ${}_{75}C_2$), which is too large and not practical.

Therefore two fitting procedures which reduce the number of paired comparisons are proposed. Instead of the paired comparisons in round-robin manner or in GA algorithm, the proposed procedures split the fitting process into two steps: a parameter fitting step for each hearing compensation method and a comprehensive combination step. The proposed procedures are then evaluated in regard to the number of comparisons and fitting effects in fitting experiments for twenty elderly subjects over sixty-five.

This paper is organized as follows. In Section 2, two fitting procedures reducing the paired comparisons are proposed. In Section 3, fitting experiments are conducted based on the proposed fitting procedures with twenty elderly subjects. The number of comparisons and fitting effects are evaluated in the subject's assessment.

2. FITTING METHODS OF SIMPLIFIED HEARING COMPENSATION

Table 1 shows the three hearing compensation methods and their discrete levels that the proposed procedures use. A combination is selected as an appropriate parameter set for a user by repetitive paired comparisons. To reduce the number of comparisons for selecting an appropriate parameter set, the proposed procedures split the fitting process into two steps. In the first step, one of the discrete levels for each hearing compensation method is selected by paired comparisons respectively. In the second step, combinations of "ON" (i.e. the selected level in the first step) and "OFF" of the three hearing compensation methods are compared and an appropriate combination is selected by paired comparisons. The following two procedures A and B are variants in decision algorithm in the first step.

Table 1: Hearing compensation methods and their discrete levels.

Compensation methods	Discrete levels
High-frequency boost	+0, +3, +6, +9, +12 [dB]
Pitch shift	-1/6, -1/12, ±0, +1/12, +1/6 [octave]
Speech rate conversion	1.0, 1.1, 1.2 [times]

Table 2: Combinations of “ON” and “OFF” of the three hearing compensation methods in the second step.

High-frequency boost	○	○	○	×	○	×	×
Pitch shift	○	○	×	○	×	○	×
Speech rate conversion	○	×	○	○	×	×	○

2.1 Procedure A

In the first step of Procedure A (Step A-1 in short), the paired comparisons are conducted among the discrete levels in round-robin manner, and the most frequently chosen level is selected as the appropriate one for each hearing compensation method. The numbers of paired comparisons are 10 (= ${}_5C_2$) for a high-frequency boost and pitch shift, and 3 (= ${}_3C_2$) for a speech rate conversion.

In the second step of Procedure A (Step A-2 in short), the paired comparisons are conducted among all the combination of “ON” and “OFF” of the three hearing compensation methods except for the all “OFF” case. Table 2 shows the combinations, where a circle indicates “ON” and a cross indicates “OFF”. The number of the paired comparisons is 21 (= ${}_7C_2$) between the two of the seven cases. If the selected level of a hearing compensation method in the first step is the default level (+0 dB for a high-frequency boost, ±0 octave for a pitch shift and 1.0 times for a speech rate conversion), the level of the hearing compensation method is fixed at the default level, and the hearing compensation method becomes out of comparison. If the default level is selected in one hearing compensation method in the first step, the number of combination of “ON” and “OFF” except for the all “OFF” case becomes three, and the number of paired comparisons is 3 (= ${}_3C_2$). If the default level is selected in two or three hearing compensation methods in the first step, no paired comparison is required in the second step, because only one combination of three hearing compensation methods remains.

The maximum and minimum total numbers of comparisons are 44 and 23 respectively, as shown in Table 3.

2.2 Procedure B

In Step A-1, all the paired combinations of the discrete levels are compared regardless of user preference shown up to the point. The number of comparisons is thought to be reduced efficiently by restricting the range of the level stepwise. In the first step of Procedure B (Step B-1 in short), the paired test speech is determined by a deterministic algorithm which restricts the range of the level stepwise based on the result of the former comparison. Figure 1 shows the deterministic algorithm of Step B-1 for a high-frequency boost. Firstly, the maximum (+12 dB) and minimum levels (+0 dB) are compared. Next paired levels are determined, based on the result of the former comparison. As shown in Figure 1, the appropriate level is to be decided in three or four comparisons for a high-frequency boost. Similarly, a pitch shift with five levels requires three or four comparisons, and a speech rate conversion with three levels requires two.

The second step of Procedure B (Step B-2 in short) is the same as Step A-2 described in Section 2.1.

The maximum and minimum total numbers of comparisons are 31 and 8 respectively, as shown in Table 3.

Table 3: Numbers of paired comparisons in Procedure A and Procedure B.

		Procedure	
		A	B
Step 1	High-frequency boost	10	3 or 4
	Pitch shift	10	3 or 4
	Speech rate conversion	3	2
Step 2	No default level	21	21
	One default level	3	3
	Two or three defaults	0	0
Total number	Maximum	44	31
	Minimum	23	8

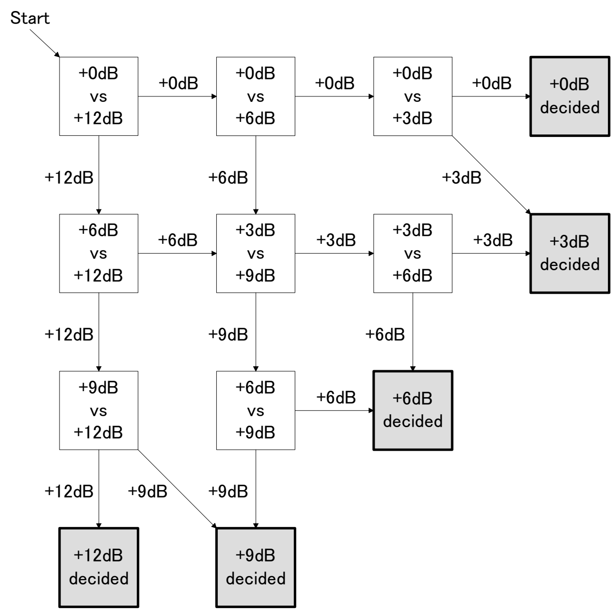

Figure 1: Deterministic algorithm of Step B-1 (High-frequency boost).

3. EXPERIMENTS

3.1 User application

Figure 2 shows the GUI of the fitting system. The system plays paired test speech. The user listened to the paired speech and just chooses the better quality one by clicking one of two buttons. The system and user repeat the operation to decide the appropriate parameter set.

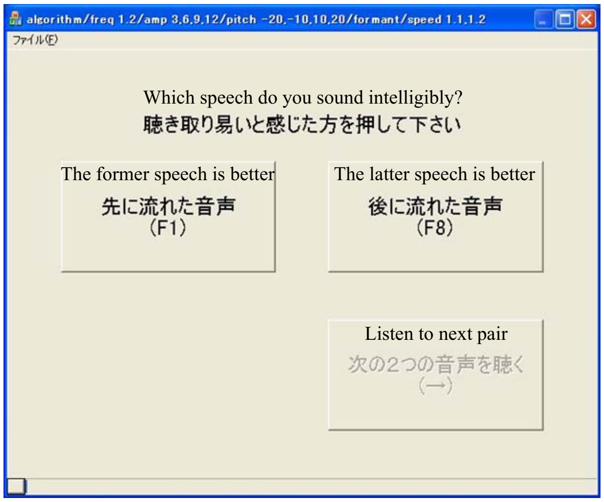

Figure 2: GUI-based user fitting system.

3.2 Procedure of experiments

Table 4 shows the configurations of fitting experiments. The experiments, based on Procedure A and Procedure B were conducted for ten male and female elderly subjects over sixty-five respectively. The test speech set consisted of sixteen utterances, which were uttered by four male and female speakers respectively, selected from a standard Japanese speech database [5]. Each test speech consisted of a pair of two semantically-independent sentences, uttered by a male and female speaker, respectively. The subjects listened to the test speech through headphones in a sound-proof room.

After the fitting experiments were conducted, the fitting effects of Procedure A and Procedure B were assessed. Subjects, unaware of whether they were listening to the original and speech with the selected parameter set, listened to the pair, and then chose one of three alternatives: "The former speech is better", "The latter speech is better" and "Difficult to say which is better". The subjects listened to four paired test speech and assessed each (i.e. four votes for a subject).

3.3 Number of paired comparisons

Table 5 shows the average numbers of paired comparisons and the reduction rates from Procedure A to Procedure B for three groups, R, I and A. Group R consists of twelve subjects whose numbers of comparisons were *reduced* from Procedure A to Procedure B. Conversely, Group I consists of eight subjects whose numbers of comparisons were *increased* from Procedure A to Procedure B. Group A indicates *all* subjects. For Group A, the average number of comparisons in Procedure A is 24.8 and that in Procedure B is 18.8, which represents a 24.4% reduction. For Group R, the reduction rate was 56.0%. On the other hand, the increase rate was 27.8% for Group I.

Focusing on the difference between Group R and Group I, Procedure B shows a significant difference between the two while Procedure A shows little difference. In Procedure B, the number of comparisons for Group I is larger than twice that for Group R. The cause of the difference is investigated in detail in Table 6, which shows the distributions of subjects with respect to the numbers of selected default levels in Step A-1 and Step B-1 for Group R and Group I. For Group R, more than one default level is selected for all subjects except for a subject in Step A-1. When the number of selected default levels is more than one in the first step, the number of comparisons in the second step is 3 or 0 as shown in Table 3. The difference in the numbers of comparisons between Step A-2 and Step B-2 is slight. Therefore the reduction in the number of comparisons from Step A-1 to Step B-1 directly resulted in a reduction in the total number of comparisons from Procedure A to Procedure B for Group R. Table 3 shows that the maximum number of reduction from Step A-1 to Step B-1 for a high-frequency boost is 7 (= 10-3). Similarly, the maximum reduction for a pitch shift is 7 (= 10-3) and that for a speech rate conversion is 1 (= 3-2). These numbers lead the maximum reduction of the number of comparisons from Step A-1 to Step B-1 to 15 (= 7+7+1) in total. On the other hand, all subjects of Group I have more than one selected default levels in Step A-1 whereas the same subjects have zero in Step B-1. Since 21 comparisons are needed for no default level in Step A-2 and Step B-2, as shown in Table 3, the numbers of comparisons in Step B-2 are 18 or 21 larger than those in Step A-2. Therefore, the increased numbers of comparisons from Step A-2 to Step B-2 exceeded the

Table 4: Configurations of fitting experiments.

Subjects	10 males and 10 females over 65 years
Listening instrument	SENNHEISER HD270
Listening level of original speech	78 dBSPL at ERP
Ambient noise	25 dBA
Original speech	16 utterances selected from NTT-AT "SPEECH DATABASE FOR TELEPHONEMETRY 1994"
Speakers	4 males and 4 females

Table 5: Average numbers of paired comparisons and the reduction rates from Procedure A to Procedure B for Group R, Group I and Group A.

	Group		
	R (#12)	I (#8)	A (#20)
Procedure A	25.8	23.4	24.8
Procedure B	11.3	29.9	18.8
Reduction rate	56.0%	-27.8%	24.4%

Table 6: Distributions of subjects with respect to the numbers of selected default levels in Step A-1 and Step B-1 for Group R and Group I.

	Step	Number of selected default levels at Step 1		
		0	1	2 or 3
Group R (#12)	A-1	1	4	7
	B-1	0	8	4
Group I (#8)	A-1	0	1	7
	B-1	8	0	0

Table 7: Average numbers of votes in the assessments of fitting effects.

	Selected parameter set	Difficult to say which is better	Original Speech
Procedure A	2.35	0.65	1.00
Procedure B	2.00	0.50	1.50

Table 8: Average numbers of votes in the assessments of fitting effects for Group R and Group I.

	Procedure	Average number of votes		
		Selected parameter set	Difficult to say	Original speech
Group R (#12)	A	2.08	0.92	1.00
	B	2.33	0.67	1.00
Group I (#8)	A	2.75	0.25	1.00
	B	1.5	0.25	2.25

reduced numbers from Step A-1 to Step B-1 for Group I. This fact explains why the total numbers of comparisons in Procedure B are larger than that in Procedure A for Group I.

Taking a high-frequency boost as an example, the default level cannot be selected unless the user chooses "+0 dB" for all comparisons in the deterministic algorithm of Step B-1. For a pitch shift and speech rate conversion, there are similar cases. It is thought that non-default level tends to be selected by the deterministic algorithm compared to the decision algorithm based on the most frequently choices.

3.4 Assessments of fitting effects

Table 7 shows the average numbers of votes in the assessments of fitting effects. For both Procedure A and Procedure B, the average numbers of votes of the speech with the selected parameter set are larger than that of the original speech. This fact can indicate that the speech with the selected parameter set based on Procedure A or Procedure B is effective for improvement in hearing. Since the average number of votes for the speech with the selected parameter set on Procedure A is larger than that on Procedure B and the average number of votes for the original speech on Procedure A is smaller than that on Procedure B, the speech with the selected parameter set based on Procedure A is more effective than that on Procedure B.

Table 8 shows the average numbers of votes for Group R and Group I, as described in Section 3.2. For Group R, the numbers of votes for the speech with the selected parameter set on both Procedure A and Procedure B are larger than those for the original speech. However, for Group I, the number of votes for the speech with the selected parameter set on Procedure A is larger than that for the original speech whereas the number of votes for the speech with the selected parameter set on Procedure B is smaller than that for the original speech. As described in Section 3.2, non-default level tends to be selected by the deterministic algorithm. Some of the inappropriate non-default levels are selected in Step B-2 and then could have adversely caused deterioration in hearing.

4. CONCLUSIONS

.

For simplified hearing compensation, which combines a high-frequency boost, pitch shift and speech rate conversion, two quick individual fitting procedures, based on repetition of a paired subjective comparison, were proposed. To reduce the number of comparisons, both the proposed procedures split the fitting process into two steps. The first step is a parameter fitting step for each hearing compensation method, and the second step is a comprehensive combination step. The fitting experiments for twenty elderly subjects over sixty-five showed that the average numbers of comparisons in two procedures were 24.8 and 18.8 respectively, whereas the number of round-robin comparisons for every possible combination is 2775. The assessments of fitting effects by comparing the speech with the selected parameter set and the original speech indicated that the speech with the selected parameter set on both procedures was effective for improvement in hearing. The deterministic algorithm in the second step of Procedure B can reduce the number of comparisons efficiently, while the tendency of selecting non-default level might adversely cause increase of comparisons and deterioration in hearing.

5. REFERENCES

[1] H. Takagi and M. Ohsaki, "IEC-based Hearing Aid Fitting," *Proc. IEEE System, Man, and Cybernetics (SMC '99)*, vol. 3, pp. 657-662, 1999.

[2] H. Takagi, "Interactive Evolutionary Computation: Fusion of the Capacities of EC Optimization and Human Evaluation," *Proceedings of the IEEE*, vol. 89, no. 9, pp. 1275-1296, 2001.

[3] H. Takagi, S. Kamohara and T. Takeda, "Introduction of Soft Computing Techniques to Welfare Devices," *Proc. IEEE Soft Computing Methods in Industrial Applications (SMCia/99)*, pp. 116-121, 1999.

[4] E. A. Durant, G. H. Wakefield, D. J. Van Tasell and M. E. Rickert, "Efficient Perceptual Tuning of Hearing Aids With Genetic Algorithms," *IEEE Trans. Speech and Audio Processing*, vol. 12, pp. 144-155, 2004.

[5] *SPEECH DATABASE FOR TELEPHONOMETRY 1994*, NTT Advanced Technology Company, 1994.