Speech Training System for Hearing Impaired Children Using Technology of Voice Recognition

Fumitake SUGANO, Tomoyuki MIZUTANI, Ayano SASAKI,
Takefumi KITAYAMA, Hiroyuki KAMATA and Yoshihisa ISHIDA
School of Science & Technology, Meiji University
1-1-1, Higashi-mita, Tamaku, Kawasaki, 214 Japan
Phone. +81 - 44 - 934 - 7307
Fax. +81 - 44 - 934 - 7909

ABSTRACT - This paper describes a speech training system called " Speech Trainer " using personal computer. For over past twenty years, our laboratory has developed a speech training system for the deaf. This system has been already used by about 80 percent of schools for the deaf in Japan. The system gives characteristics of trainer's voice on CRT display and has some functions such as vowel training, consonant training, pitch accent training using fundamental frequency, articulation training using vocal tract shape and so on.

INTRODUCTION

On the speech training system for the deaf, development of the engineering instrument for teaching the speech information has been trying everywhere by using the sight or the sense of touch. For over past twenty years, our laboratory has developed a speech training system called " Speech Trainer " for the deaf. The speech training system has already been used by about 80 percent of the deaf school in Japan. The training effects with this system have been reported by the deaf school in Kawasaki for a long time. This system using personal computer gives characteristics of voice on CRT display, i.e., vocal tract shape, tongue's position, motion of lips and so on. So deaf children can easily watch the fundamental characteristics of voice and continue the speech training with the sense of a play. Result of voice training can be saved on floppy disk. Therefore deaf children can know the last results of voice training.

This system using DSP is effective to analyze the speech signal in real time, i.e., estimation of the power spectrum, detection of the fundamental frequency, estimation of the vocal tract area function and so on. In this paper, we describe the functions and some applications of speech training system.

ARCHITECTURE OF SPEECH TRAINING SYSTEM

Fig. 1 shows the appearance of speech training system. The speech training system consists of a personal computer and a DSP board as shown in Fig. 2. Reset, bus-hold and interrupt requests to DSP are controlled from I/O port in host computer. In the state of bus-hold, A/D and D/A converters can be directly controlled under the host computer. 2 ports-memory (which enables to transport data between the DSP board and the host computer) of DSP is optionally used. Serial communication between some DSP boards is valid, when these boards are used in one host computer. So speech training for a lot of deaf children can be realized on one host computer. The volume of speech signals is under software control.

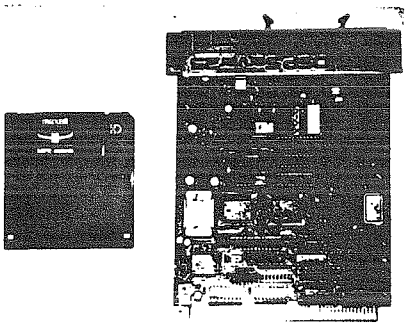

Fig 1. Appearance of speech training system.

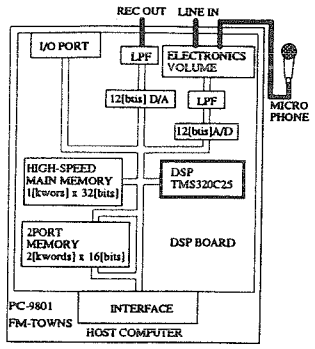

Fig 2. Architecture of speech training system.

FUNCTIONS OF SPEECH TRAINING SYSTEM

We have developed the speech training system with the following functions.

(1) Vowel training.
(2) Consonant training.
(3) Spoken word training.
(4) Pitch accent of speech signal.
(5) Stress accent of speech signal.
(6) Intonation training.
(7) Three-dimensional display of spectra.
(8) A kind of game using the fundamental frequency.
(9) Music training.

Each functions use the speech signal analysis such as estimation of the power spectra, detection of the fundamental frequency, estimation of the vocal tract area function and so on. These analysis methods are calculated by the DSP. Specially, vowel and consonant training applications display the animation of the tongue motion or the lips shape. And they can make the deaf children to understand the fundamental characteristics of speech who have little experience for speech training. Children can enjoy training for a long time because these applications have the sense of a play.

Now, let us show examples of training applications. Firstly, we show vowel training using vocal tract area function. Estimation method of vocal tract area function is shown in Fig. 3 [1]. A_n is vocal tract area and M is analysis order. The estimation method of vocal tract area function is based on Levinson - Durbin algorithm. As pre-processing, adaptive inverse filter is necessity for excluding radiation and glottis characteristics. We use the auto correlation function although the conventional adaptive inverse filter is for time series signal. In this method, auto correlation function $R(i)$ is given from time series signal and is renewed by Eq.(1). Factor ε_n in Eq.(1) is given by Eq.(2). By using DSP, the vocal tract area function is estimated by 50 times/sec (analysis frame: 512, order: 20). So the children can know their vocal tract areas in real time.

$$Rnew(i) = \left(1+\varepsilon_n{}^2\right)Rold(i) - \varepsilon_n\left\{Rold(i+r) + Rold(|i-r|)\right\} \quad (1)$$

$$\varepsilon_n = \frac{Rold(r)}{Rold(0)} \quad (2)$$

Fig. 4 shows the vowel training of /a/. In this training, the deaf children try so as to match to teacher's vocal tract area. A vocal tract area of /a/ is shown on the left of Fig. 4. The painted area on the right

of Fig. 4 is teacher's vocal tract area. On the other hand, the line shows deaf's one. The length of teacher's vocal tract can be changed to coincide with children's one.

Fig. 5 shows the vowel training of /a/ using power spectrum. In this application, the deaf children must try to match their spectra with the right spectrum. We have a lot of spectral patterns for various lengths of vocal tract. As a method of evaluation in this training, the picture is displayed in reward and its shape is different by the evaluation score.

Fig. 6 shows a kind of game using the fundamental frequency. This is a game for training of intonation which is important for Japanese language. And the airplane must pass through between a star and a ladder. In this training, the range of frequency and the wide of a slit can be freely changed.

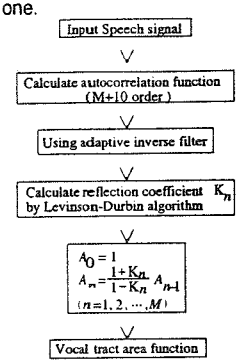

Fig 3. Estimation method of vocal tract area function.

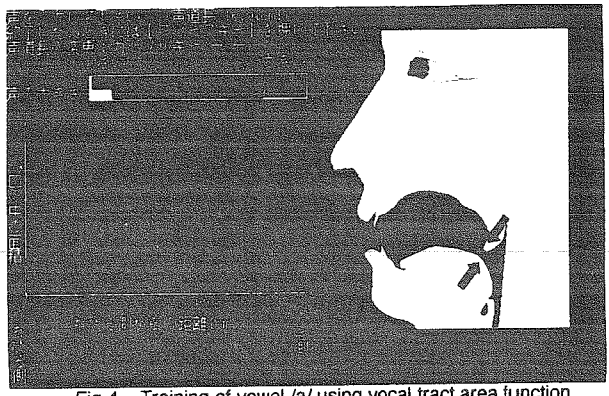

Fig 4. Training of vowel /a/ using vocal tract area function.

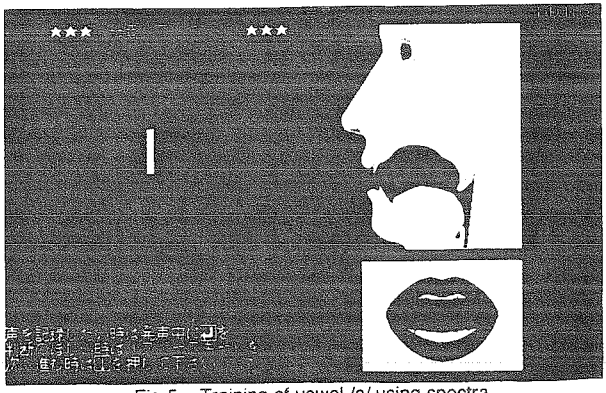

Fig 5. Training of vowel /a/ using spectra.

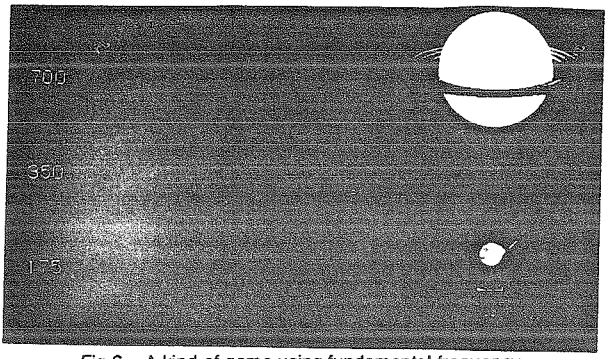

Fig 6. A kind of game using fundamental frequency.

AUTOMATIC EVALUATION AND TEACHING

Training applications mentioned above have been used in the deaf schools. Deaf children are under teacher's guidance in the school. However some adults have little time for speech training under teacher's guidance. Thus they need the automatic evaluation and teaching. As one of the evaluation methods for training result for them, we have used pattern-matching by vocal tract area function as shown in Fig. 7 [2]. Using this method, the deaf can understand the right articulation method. In this system, the length of the right vocal tract is automatically normalized to the their one.

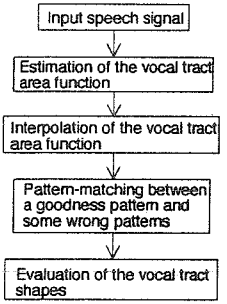

Fig 7. The evaluation method used pattern-matching.

CONCLUSIONS

This paper has presented the speech training system using personal computer. This system is the support system for teachers in the school of the deaf and has been used by about 80 percent in Japan, and has gained highly efficiency. This system is never the substitution of the teachers in the school because teaching must be under teacher. Such a system must have the function by which children can see the training results. On automatic evaluation and teaching, this paper describes an analysis method using vocal tract area function. As the future work, we will developed the training system for dairy conversation using both the fundamental frequency and the power spectrum.

REFERENCES

[1] H. Kamata, Y. Ishida and Y. Ogawa "The Fast Estimation of the Vocal Tract Area Function Using DSP." IEEJ, Vol. 110-7, pp. 773-780 (1990)

[2] F. Sugano, H. Kamata and Y. Ishida "Practice of Monosyllable / Word Utterances Using Technology of Voice Recognition and Its Application to Speech Trainer." Technical Report of IEICE, ET93-122 (1994-01)