3rd International Conference on
Spoken Language Processing (ICSLP 94)
Yokohama, Japan
September 18-22, 1994

A DSP-BASED AMPLITUDE COMPRESSOR FOR DIGITAL HEARING AIDS

Yuichi UEDA, Takayuki AGAWA, and Akira WATANABE

Faculty of Engineering , Kumamoto University
2-39-1 Kurokami, Kumamoto 860, Japan

ABSTRACT

Amplitude compression techniques have been used in the hearing aids for the sensory-neural impaired. In this paper, we present a DSP-based amplitude compressor which is based on the conventional filter-bank method, but can realize the relatively smoothed compression characteristics without the FFT method. The system has two kinds of filter-banks. The speech spectrum is estimated using the first bank and a desired compressor is generated at the second one which is similar to FIR filter used in the wide-band compression. Moreover, by estimating the spectrum using the output of the first filter-bank, it is possible to compress adaptively the signal according to the temporal variations of the speech features. We describe the principles of such compression system and its DSP implementation, and show that it is the simplified compression for the DSP realization, in this paper.

1. INTRODUCTION

In recent years, cosiderable attentions have been paied to the digital hearing aids[1], because real-time signal processing techniques have been progressed by the appearance of digital signal processors(DSP). In such digital processings for the hearing aids, the representatives may be the compression systems which intend to pack the speech spectra into the residual auditory range of the hearing impaired.

As such amplitude compression methods, the multi-band compressions using filter-bank are well known, which are suitable for analog processing and intend to compress the range of speech level variation of each band[2],[3]. In these systems, it is reported that audibility of weak speech such as plosives shows improvement, while vowel intelligibility is degraded due to the loss of spectral details[4]. Since the compression is implemented every each band independently, it may be inevitable that the compression discontinuities between the neighboring bands occur frequently. Therefore, in order to reduce the degradation of speech quality, it needs to increase the number of bands as sacrifice of execution time.

On the other hand, as those using the short-term speech spectra, the following methods are known.

i) The short-term speech spectra compressed in spectral region are transformed into speech waveforms by the IFFT and overlap-added those successively[5],[6].

ii) The FIR filter with the desired compression characteristics is generated by the IFFT and then its response is convoluted with the original signal[7].

In this paper, we propose an modified amplitude compressor for the DSP-based hearing aids, which takes advantages of both the filter-bank method and the FIR filter in the FFT method ii). The system uses two kinds of filter banks. One (BANK.A) is for dividing the speech signal into multi-bands and the other (BANK.B) is for generating a compressor as FIR filter without the FFT/IFFT operations. By use of such BANKs, it is possible to vary the lengths of frames adaptively to estimate the short-term spectrum and to execute compressing at shorter execution time. In this paper, we describe the underlying principles of the compressor, the implementation under newly developped DSP system, and several processing results.

2. DESIGN CONSIDERATIONS

Fig.1 shows a block diagram of the proposed compressor which consists of a spectrum estimation using BANK.A, a compressor generation in BANK.B, and a signal convolution by the compressor.

2.1 SPECTRUM ESTIMATION

BANK.A is a conventional filter-bank which consists of a low-pass and four bandpass filters. As shown in Fig.2, those filters are arranged every octave-band and divide the speech signal into five bands with center frequencies of f_1, f_3, f_5, f_7, and f_9, respectively. The output of each filter is squared, averaged, and then converted into log-scale(P_1, P_3, P_5, P_7, P_9).

In order to estimate the levels ($P_{2,4,6,8}$) at intermediate frequencies ($f_{2,4,6,8}$), the five log-powers ($P_{1,3,5,7,9}$) are interpolated by the spline function where terminal points, P_0 and P_9, equal to P_1 and $0.5P_9$ respectively. Those power spectral estimates of P_0 to P_9 are utilized in the following BANK.B.

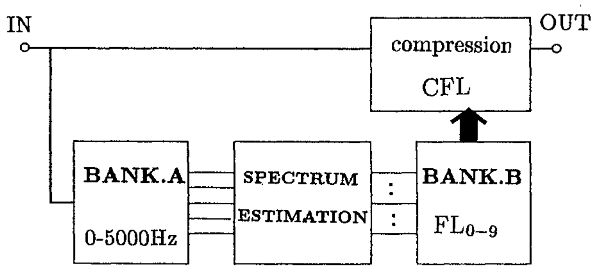


Fig.1 Block diagram of the compression system.

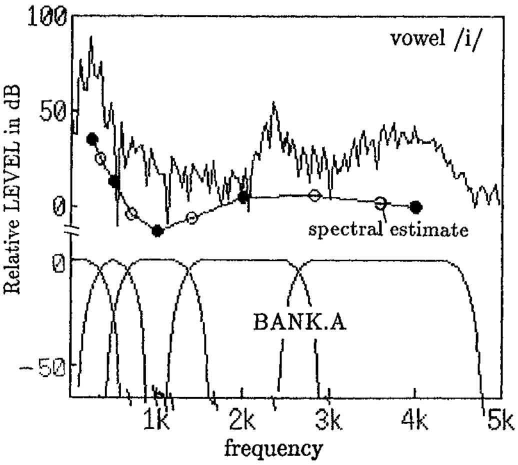


Fig.2 Structure of the filter-bank in BANK.A and spectrum estimation by interpolating those filter outputs

2.2 AMPLITUDE COMPRESSION

A. STRUCTURE of BANK.B

As shown in Fig.3(a) ideally, BANK.B consists of ten fundamental filters (FL_{0-9}), where f', f'_k, and $F'c$ indicate $\log f$, $\log f_k$, and $\log F_c$ respectively. The $k-th$ FL_k is a high-pass one whose gain equals to 0 below f'_k and 1.0 above f'_{k+1}, whereas the 0-th filter (FL_0) is a all-pass one with gain 1.0, as follows:

$$\begin{aligned} FL_k : H_k(f') &= \begin{cases} 0.0 & (0 \le f' < f'_k) \\ \frac{f'_{k+1}-f'_k}{f'-f'_k} & (f'_k \le f' < f'_{k+1}) \\ 1.0 & (f'_{k+1} \le f' \le F'_c) \end{cases} \\ FL_0 : H_0(f') &= \quad 1.0 \qquad (all \;\; f') \end{aligned} \quad (1)$$

where F_c is the maximum frequency of signal. If we add up the impulse responses ($h_{0-9}(n)$) of those filters connected in parallel, a resulting filter becomes to have the following frequency characteristics.

$$H(f)=\sum_{k=0}^{9} H_k(f)=\begin{cases} 1 & (0 \le f < f_1) \\ 10(logf/logF_c) & (f_1 \le f \le F_c) \end{cases} \quad (2)$$

We call this a composite filter, CFL.

Before adding up those h_{0-9}, if $h_k(n)$ is weighted by any gain, g_k, like Eq.(3), a CFL with various characteristics can be generated.

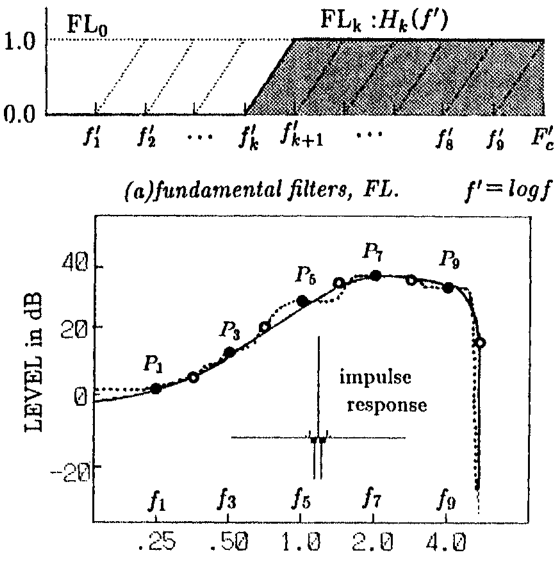
$$h(n) = \sum_{k=0}^{K} g_k h_k(n) \quad (n = -N, -N+1, ..., N) \quad (3)$$

where $2N+1$ equals to the filter length of FL_{0-9} and CFL.

We have designed both the LPF/BPFs in BANK.A and each FL_k in BANK.B by the Kaiser windowing method, whose lengths ($2N+1$) equal to 91 respectively. Fig.3(b) shows an example of CFL characteristics with any g_ks. The dotted curve indicates one of the conventional filter-bank compression using the same g_ks.

B. AMPLITUDE COMPRESSOR

Using the log-power (P_k), HTL_k(hearing threshold level), and UCL_k(uncomfortable level) at each frequency f_k, a desired gain (GN_k) is calculated as follows:



(a)fundamental filters, FL. $f' = \log f$

(b)an example of CFL

Fig.3 Structure of the fundamental filters, FL_{0-9}, in BANK.B (a) and an example of a CFL characteristics with impulse response of 91-samples (b) [The dotted curve shows the conventional filter-bank compression.]

$$GN_k = P_k(UCL_k - HTL_k - P_{max})/P_{max} + HTL_k \quad (4)$$

where k is 1 to 9 and P_{max} is the dynamic range of input/output signal. Furthermore, the difference between the neighboring gains is converted as follows:

$$g_k = LOG^{-1}(GN_k) - LOG^{-1}(GN_{k-1}) \quad (5)$$

The corresponding FL_k's impulse response is weighted by g_k. Only FL_0's one is directly weighted by $LOG^{-1}P_0$, g_0, so that it may provide the overall gain of the CFL. Repeating such operations from FL_0 to FL_9, a desired CFL's impulse response is obtained, and then the amplitude compression is realized by convolution those with the original signal.

2.3 COMRESSION DYNAMICS

Since the CFL's characteristics is determined by the five spectral levels ($P_{1,3,5,7,9}$), it is important how to select the lengths of averaging durations. For example, if we select a shorter length, CFL may be more sensitive to the instantaneous waveform. In such case, the spectral features of shorter segments with lower energy such as plosives are preserved, but those of longer ones with higher energy such as vowels may be distorted.

From the point of view, we have varied the lengths of averaging duration adaptively, as shown in Fig.4. The following processing was implemented every each band k in BANK.A.

A. DECISION of FRAME LENGTH

At each sampling point, a digitized sample is stored into the signal buffer(SBUF) and the instantaneous energy squared the output of the k-th filter is stored into the energy buffer ($EBUF_k$) successively. When it is the last point in the m-th frame whose length of L_m is determined in the previous $(m-1)$-th one, the past L_m samples in $EBUF_k$ are averaged over, and the short-term energy ($SENG_k$) of the k-th bank is obtained.

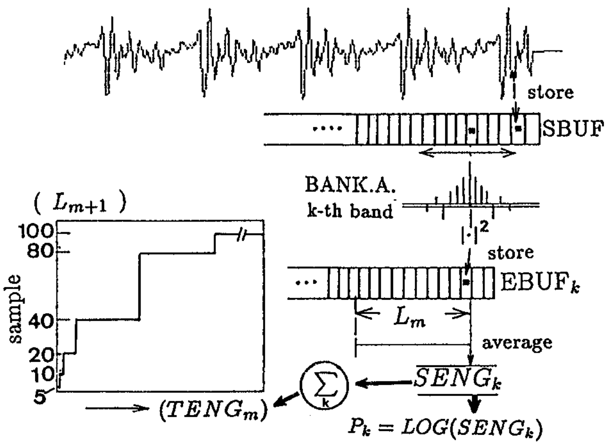


Fig.4 Illustration of the frame length decision.

Since the total energy ($TENG_m$) of the m-th frame added up all $SENG_k$ can be considered as the short-term energy of input signal approximately, the length of next frame (L_{m+1}) is determined by its energy, according to the function depicted in Fig.4. Consequently, in the segments with lower energy, the spectra are estimated during a shorter duration, and vice versa in the segments with higher one.

Fig.5 shows the examples of spectral estimates of /pa/ using the fixed length ($L = 10, 100$) and the adaptive one. It can be seen that spectral features of both the weaker plosive /p/ and the stronger vowel /a/ are estimated by the shorter segments and the longer one respectively, due to the adaptive decision of the frame length.

B. RENEWAL of CFL

Each $SENG_k$ is converted into log-scale, P_k, at the end point of each frame. At the same time, a new CFL is generated using those P_{0-9} interpolated, so compression characteristics becomes to vary according to the variation of signal energies.

Fig.6 shows the original and compressed speech specta of plosive /p/ and vowel /a/ in the same /pa/ in Fig.5.

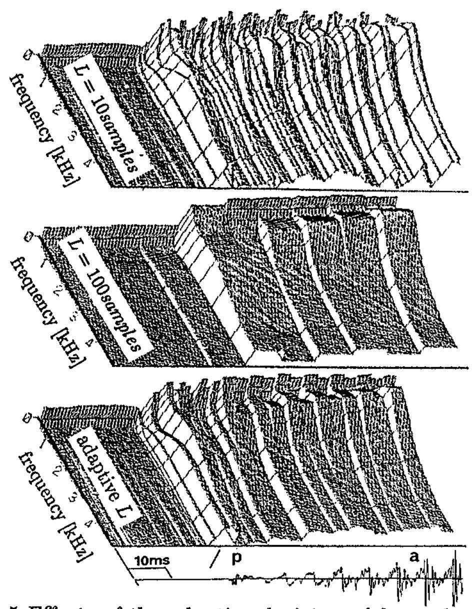


Fig.5 Effects of the adaptive decision of frame lengths.

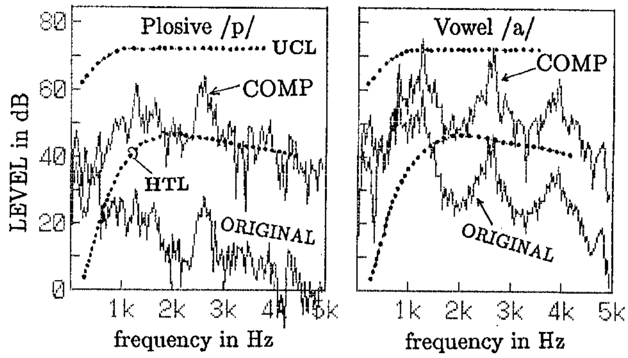


Fig.6 Examples of the compressed speech spectra

3. IMPLEMENTATION

3.1 DSP SYSTEM CONFIGURATION

Fig.7 shows a block diagram of the DSP system which has been developed newly using a 32_*bit* floating-point DSP(TMS320C31). Although the principal configulation is similar to the previous system[3] using a DSP(TMS320C25), the performances improve in both the machine cycle which becomes from 100 to 50ns, and the on-chip RAM in DSP which becomes 256 to 1960 words.

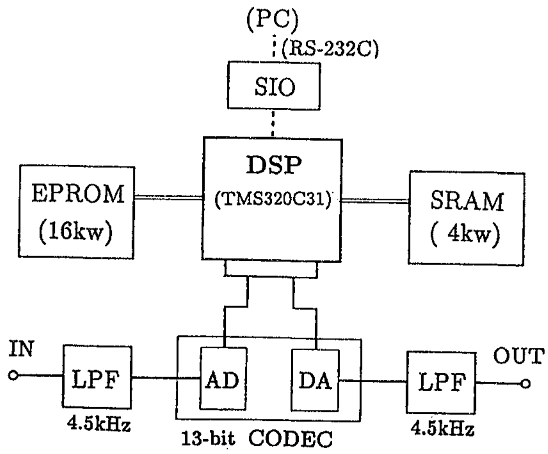


Fig.7 Block diagram of the DSP system

The system has the interface with a personal computer (PC), which is used to generate programs, download those, and so on. A monitor program stored in the EPROM initiates the system and transfers the program/parameters and the tables of LOG/LOG^{-1} from the PC to the SRAM.

The input signal is digitized by a 13_*bit* linear codec with a sampling rate (9.76kHz) and entered into DSP through a serial receiving port. Since the sampling period is 102.4μs, it is possible to execute operations of 2048 machine-cycles in real-time.The processed data through the serial transmitting port in DSP is DA-converted and smoothed by a LPF.

3.2 DSP IMPLEMENTAION

The principal parameters and buffers, those memory size, and execution cycles are listed in Table.1. Those values were not selected under the best condition based on the auditory evaluations, but under the condition when the DSP system could execute in real-time.

Table.1 Lists of parameters used in designing

		mem.size	exec.cycle
BANK.A	91-th*5ch	5*46w	308
BANK.B	91-th*9ch	9*46w	1123
(CFL)	91-th	—	102
SBUF	max.25.6ms	max.256w	–
EBUF	max.10.0ms	max.100w	–
LOG	(table)	4096w	–
LOG^{-1}	(table)	256w	–

A. INTERRUPT ROUTINE

Operations to be executed every each sampling period (102.4μs) are as follows.

1. Shift the signal buffer (SBUF) and store a digitized sample in it. [INPUT]
2. Convolute the LPF/BPF coefficients in BANK.A with the signal sequence in SBUF.
3. Shift the energy buffer ($EBUF_k$) and store the squared output of 2. in it.
4. Convolute the generated composite filter(CFL) with the signal sequence in SBUF and write its result into the serial transmitting port. [OUTPUT]

The convolutions at 2. and 4. are most efficient on the DSP which has a multiply/accumulate with data move instruction with a single cycle. Those regular operations were implemented as the interrupt routine every 102.4μs.

B. MAIN ROUTINE

On the other hand, the following operations occurring at the end point of each varying frame were implemented as main routine.

1. Average over the energy sequence with the frame length in $EBUF_k$ and convert those into log-scale, where the look-up table of *LOG*-function was utilized.
2. Estimate the speech spectrum by interpolating the log-powers of five bands
3. Calculate the desired gain of each band using the spectral level at 2., HTL_k, and UCL_k data, and then differentiate the gains of neiboring bands and convert those into g_k using the look-up table of LOG^{-1}-functions.
4. Weight the coefficients of each fundamental filter (FL_k) by g_k, sum up those, and then renew the composite filter (CFL).
5. Define the length of the next frame by summing up the five short-term energies.

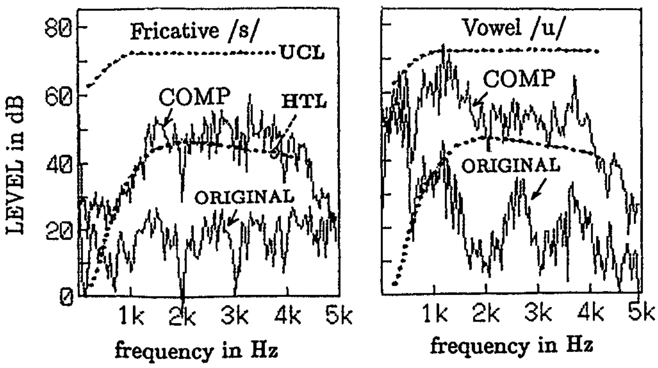


Fig.8 Examples of the compressed speech spectra by DSP implementation in real-time

3.3 COMPRESSION EXAMPLE

Fig.8 shows the examples of DSP implementaion in real time which are the short-term spectra of fricative /s/ and vowel/u/ in utterance /su/ by a male. The original speech and the compressed one are shown respectively. The *UCL*s and *HTL*s, are the values simulating an example with hearing loss in higher frequency area. It notes that the compressed spectra are packed into the audible area and the detail spectral structures of both segments are preserved.

4. CONCLUSIONS

The compression results presented above clearly show that use of a composite filter offers possibilities for the smoothed compression in spectral region. Indeed, a use of FIR filter transformed by IFFT method, whose points increase, can realize more detail compression. However, it may be impossible to implement under the DSP operations in real-time. In the case of proposed compressor, the condition similar to the IFFT method may be realized by increasing the number of fundamental filters, but in oder to decide the number , it needs to employ the auditory evaluation.

On the other hand, in spectrum estimation it may be also able to use the FFT method. Under such condition, spectral features can be obtained in detail and a more desirable compression may be realized. But, as mentioned above, the number of FFT-points decides which a DSP operation is possible or not. Furthermore, the FFT method using the fixed frame length is difficult to estimate spectra according to the temporal features of speech, that is, plosives, vowels, fricatives and so on. The proposed spectral estimation based on the filter-bank can estimate only the gross spectral features, whereas the temporal variation of those can be extracted among the relatively flexible segment length.

The various parameters used in this proposed compressor are set up experimentally as a pilot system, and those desired values should be decided by the auditory evaluations in future.

REFERENCES

[1] H.Levitt, "Digital hearing aids : A tutorial review," J.Rehabili.Res.Dev. **24**(4), pp.7-20, 1987

[2] R.E.Morley,JR, A.M.Engebretson, and J.G.Totta, "A Multiprocessor Digital Signal Processing System for Real-time Audio Applications," IEEE Trans.on ASSP, Vol.**ASSP-34**, No.2, pp.225-231, Apr.1986

[3] Y.Ueda and A.Watanabe,"Computer-assisted system for digital hearing aid and its application to the fitting procedure," Proc. 14th ICA, H3-3, Sep.1992

[4] S.DeGennaro, L.D.Braida, and N.I.Durlach,"Multichannel syllabic compression for severely impaired listeners," J.Rehabili.Res.Decv., **23**, pp.17-24,1986

[5] D.K.Bustamante and L.D.Braida, "Principal-component amplitude compression for the hearing impaired," J.Acoust. Soc.Am., **82**(4), pp.1227-1242, Oct.1987

[6] H.Levitt and A.C.Neuman,"Evaluation of orthogonal polynomial compression," J.Acoust.Soc.Am., **90**(1), pp.241-252, July.1991

[7] F.Asano, Y.Suzuki, and T.Sone,"Amethod of desinging and realizing a digital hearing aid that compensates loudness for impaired listeners," J.A.S.J. **47**(6), pp.373-379, 1991 (in Japanese)