Adaptive Digital Filtering For Signal Reconstruction Using Spectrum Extrapolation

Hiroshi Yasukawa
NTT Optical Network Systems Labs.
1-2356 Take, Yokosuka, 238-03 Japan
Tel: +81-468-59-3016; Fax: +81-468-55-1283
e-mail: yasukawa@exa.onlab.ntt.jp

Abstract

This paper describes adaptive filtering for signal reconstruction. The speech quality enhancement system by the spectrum extrapolation of the band limited signals is discussed. In telephone communication, the spectrum extrapolation which employs aliasing processing is widely known. In this paper a new implementation using adaptive methods is proposed. This method introduces frequency domain adaptive digital filtering to broaden band limited signals into wide band signals. Implementation of the system and its performance are discussed.

1. INTRODUCTION

Recently, many studies on speech quality enhancement for band limited speech are reported [1] - [13]. In regular telephone communication, ordinary speech signals are band-limited to the 0.3 to 3.4 kHz frequency band. ISDN networks will support enhanced telephone services. For example, telephone speech according to ITU-TS recommendation G.722 has a wide band signal with a frequency band from 0.05 to 7 kHz [14]. Then, only the customers utilizing ISDN with G.722 codecs can be given the high quality speech. However, ISDN customers can not unfortunately obtain high quality speech when they call a conventional telephone (0.3 to 3.4 kHz frequency band) user, and the conventional telephone users as well. This is one reason why quality enhancement is necessary. Furthermore, speech quality in regular telephone communication among conventional telephone users is naturally required to be enhanced. So speech quality enhancement is a frequent study topic.

Many approaches are proposed to solve this problem. One approach proposed is utilizing equalizers to compensate the signal degradation due to the band limitation [1]. Second is the reconstruction of wide band speech from narrow band speech by code book mapping with vector quantization [3]. The former has the advantage of simple configuration, but suffers from the problem of performance limitation due to noise. The second approach offers improved performance, but it introduces additive processing delay into transmission.

This paper focuses on achieving spectrum extrapolation through adaptive digital filtering for speech enhancement. A very simple method has been proposed [4] - [6] that can be easily implemented and has low processing delay. However, circuit parameters must be decided by subjective evaluations. In order to avoid the difficulties inherent in subjective assessment methods, some proposals utilize automatic level control or adaptive digital filters [8]. This paper focuses on the method using adaptive digital filters.

2. CONVENTIONAL SIMPLIFIED SPEECH ENHANCEMENT SYSTEM

2.1 Simplified spectrum broadening method

Let us introduce the conventional system. The input regular telephone band signal with 8 kHz sampling is expanded by interpolation with 16 kHz sampling. The outband signal (including the aliasing signal from the 4 to 7 kHz frequency bands) is generated from the input regular telephone band signal. The signal waveform is filtered by a high pass filter HPF, and is then shaped by a shaping filter. The high pass filtered signal and low pass filtered signals are summed after level adjustment to make the spectrum envelope resemble that of actual wide band signals. This method generates a widen band signal and is very simple.

This is based on the idea that spectrum folding (aliasing) can be utilized and that the empirical knowledge that the higher band signal, even if it resembles noise, can increase the speech quality as experienced by the human auditory sense.

The filters utilized in the system are linear phase FIR filters. The delay adjuster adjusts the difference in delay between the high band signal path and the low band signal path. For level adjustment it is important to know the difference ΔL between both path signal levels. The optimum level difference can be given by subjective assessments. A subjective assessment showed that ΔL values of around 0 or -12 dB were most preferable [4].

The method given in [4] yields an output spectrum with a spectrum gap (split) around the folding frequency of 8 kHz. The spectrum of the output speech waveform is very similar to regular wideband speech signals excluding the band gap. An improvement to this method was proposed in [5], [6]. In this case, the band gap can be sufficiently neglected [5], [6]. Adaptive filtering described in this paper can also be applied.

2.2 Spectrum Extrapolation by Multirate Processing

A block diagram of the revised system is shown in Fig. 1 [5], [6]. The sampling rate of the input regular telephone band signal with 8 kHz sampling is converted to 7 kHz with 8-to-7 conversion, following up sampling by 2. The signal is then expanded by an interpolator with 14 kHz sampling. After that, similar processing operations are performed in the subbands as described in 2.1. Namely, the signal waveform is filtered by a filter bank consisting of a low pass and high pass filter. The high pass filtered signal and low pass filtered signal are summed after level adjustment. Finally, the summed signal is converted to 16 kHz with 7-to-8 conversion. This method is a very simple way of generating a continuous widened signal. The filters are FIR type. The high pass filter HPF is a mirror filter of the LPF.

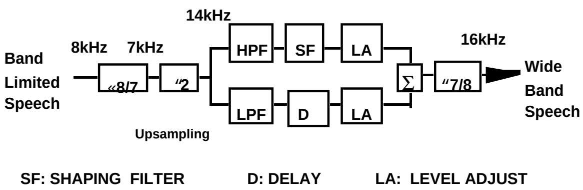

Fig. 1. Block diagram of band limited signal enhancement using spectrum broadening and shaping with multirate processing.

2.3 Spectrum Extrapolation by Nonlinear Processing

A block diagram of the broadening system using nonlinear processing is shown in Fig. 2 [10]. This is a revised version of above mentioned. In this Fig. 2, an absolute circuit is utilized as a nonlinear device. The high pass filter HPF can be implemented without frequency band gaps. The output spectrum envelope of the band limited signal enhancement using nonlinear processing is reconstructed with smooth. The filters are FIR type.

The conventional implementations above mentioned, however, high band spectrum signals are supplied with level adjustment. This is also decided by a subjective assessment.

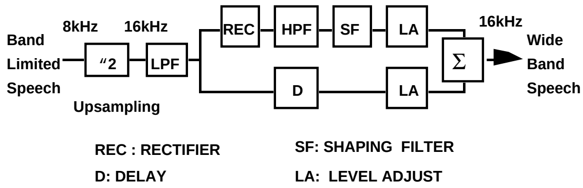

Fig.2. Block diagram of band limited signal enhancement using spectrum broadening and shaping with nonlinear processing.

3. IMPLEMENTATION

3.1 Adaptive Digital Filter

This paper considers a level adjustment scheme that does not depend on subjective assessment. The proposed scheme has two modes. One is the learning (or training) mode for deriving wideband signals from narrow (conventional telephone) band signals utilizing original wideband signals as a reference. The other is the fixed (or regular operation) mode in which the system converts narrow band signals into wideband signals without reference signals, i.e. the system parameters are fixed. A block diagram for the combined representation of both modes is also shown in Fig. 3.

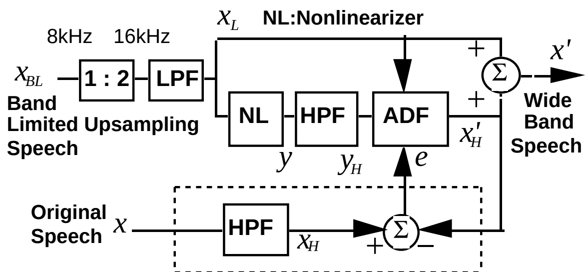

Fig. 3. Block diagram of band limited signal enhancement system using adaptive digital filter.

In this system, the block, indicated by the dotted square line, is operated only in the learning (or training) mode. In Fig. 3, NL, which indicates nonlinearity, is an absolute value circuit in practice, in which high band signals are produced from low band limited signals. ADF in Fig. 3 indicates an adaptive digital filter (ADF). This block generates a high band signal whose spectrum weak approximates the high band signal of the original signal. Finally, a wide band signal is obtained by adding the received low band signal and the newly generated high band signal.

Fig. 3 shows that the ADF is constructed in nonlinear fashion since the system can be assumed as nonlinear mapping between the low band and high band speech signals. From the viewpoint of implementation simplicity, we adopt a frequency domain ADF to achieve spectral estimation. Input signals of the ADF, x_L, y_H, e and etc., are transformed into the frequency domain and we compute the power spectrum. The power spectrum is divided into several sections and averaged.

3.2 Frequency Domain ADF Configuration

The proposed method uses an adaptive system in which level adjustment is made by adaptively utilizing a wide band training signal. A block diagram of the adaptive system is shown in Fig. 4. In this figure, LPF_1 represents a band limitation of the transmission lines. Then we consider telephone band limitation, 0.3-3.4 kHz. LPF_2 is a band limitation for higher band, hence, pass band is less than 7 kHz.

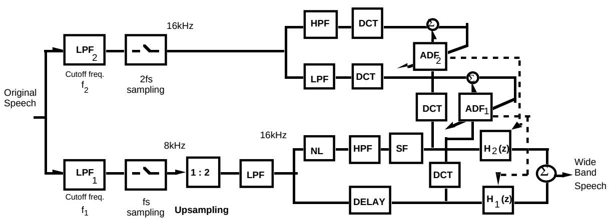

Fig. 4. Block diagram of band limited signal enhancement using adaptive digital filter.

This system adopts a frequency domain adaptive digital filter (FADF). Since this scheme has a subband structure, decimators can be included. Accordingly, the computational burden is strongly reduced according to the ratio of the resampling rate to the input rate. Based on the idea of obtaining quality enhancement by spectrum extrapolation, transformed domain adaptive digital filters are indicated. The FADF includes the DCT (Discrete Cosine Transform) operation because of its easy implementation. The low band ADF provides more precise equalization for LPF_1.

The learning mode can adopt an simple adaptive algorithm. Now, let DCT coefficients vector of the reference path and the processing channel be $c_r(k)$ and $c_p(k)$, respectively. k denotes time index. Then, the difference of DCT coefficients between the reference path and the processing channel is represented by

$$c(k) = c_p(k) - c_r(k). \quad (1)$$

c(k) is added and averaged. So successive adaptation dose not use. Using this averaged c(k), transfer functions are calculated by inverse DCT. Also, let Z transform of original speech signal be X(z).

In the fixed (or usual operation) mode, the ADF parameters are calculated to determine the transfer functions $H_1(z)$ and $H_2(z)$. Finally, overall band limitation, caused by the transfer functions of LPF_1 and the shaping filter, is compensated by $H_1(z)$ and $H_2(z)$.

This system has a disadvantage that the avaraging for all the signals yields uniformed spectrum patern reconstruction. Thus the signal with different spectrum patern from original spectrum is regenerated.

3.3 Selective ADF Configuration

In order to improve the performance, we propose a new structure shown in Fig. 5. In this proposed scheme, Input band limited signals are analyzed in view of their spectrum pattern in the pattern analyzer (PTN). The control signal (output of CNT) for the selective ADF (SADF) is a signal to select the ADF coefficient vector, which is decided by the PTN output. The selective ADF change their filter coeffients according to the input signal spectrum pattern.

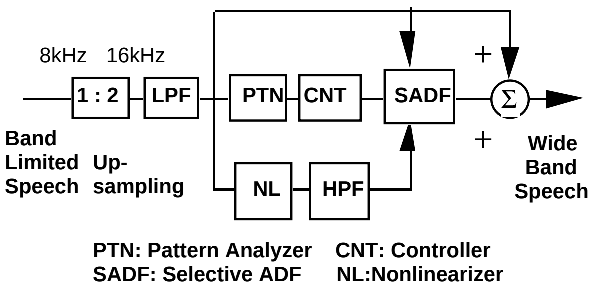

Fig. 5. Block diagram of band limited signal enhancement system using selective adaptive digital filter.

In the pattern analyzer, the spectrum pattern of the input signal can be classified by many ways, for example, a suitable filter bank and power meters. As can be seen in the figure, SADF has both signal components of in-band and out-band. Thus the out-band signals are newly generated and, and the in-band signals are modified, respectively.

4. DISCUSSION

The number of pattern size is expected to be a small number according to other preliminary studies. This yields an extremely small computation burden for the hardware implementation. It also considered that the proposed system is robust aginst input speech contaminated by interference signals such as noise or complexed speech.

The proposed method has an advantage that the output spectrum distortion can be reduced less than that in the conventional method. This method yields an output spectrum without a spectrum gap around the folding frequency of 4 kHz.

The new method offers an extremely small computation burden and small training iteration volume. These advantages are important for hardware implementation. We obtained good performance for the proposed method in terms of spectrum distortion.

The proposed method may also pave the way for being available system design without help of subjective assessment.

5. CONCLUSION

This paper proposed a new implementation of the speech enhancement method using a frequency domain adaptive digital filter. The selective digital filtering method is introduced. Implementation procedures and its performance are discussed.

The effect of kind of speech, gender and individuality etc. on this method are for further study.

ACKNOWLEDGMENTS

The author wishes to thank Dr. T. Aoyama, Mr. I. Yamashita and Mr. K. Yoshida of NTT Optical Networks Systems Laboratories for their advice and encouragement.

REFERENCES

[1] D. O. Bowker, et al, “Telephone network speech enhancement”, US Patent No. 5195132, Dec. 1990.

[2] Y. M. Chen, et al., “Statistical recovery of wideband speech from narrowband speech”, the Proc. of International Conference on Spoken Language Processing, ICSLP’92, pp. 1577-1580, 1992.

[3] Y. Yoshida, et al., ”An algorithm to reconstruct wideband speech from narrow-band speech based on codebook mapping”, the Proc. of 1994 International Conference on Spoken Language Processing, ICSLP’94, S27.3, pp. 1591-1594, Sept. 1994.

[4] H. Yasukawa, ”Quality enhancement of band limited speech by filtering and multirate techniques”, the Proc. of 1994 International Conference on Spoken Language Processing, ICSLP’94, S27.7, pp. 1607-1610, Sept. 1994.

[5] H. Yasukawa, “Spectrum broadening of telephone band signals using multirate processing for speech quality enhancement”, IEICE Trans. on Fundamentals, EA, pp. 996-998, Aug. 1995.

[6] H. Yasukawa, "Enhancement of telephone speech quality by simple spectrum extrapolation method”, European Conference on Speech Communication and Technology, EUROSPEECH '95, pp. 1545-1548, Sept. 1995.

[7] C. Avendano, et al., "Beyond Nyquist: Towards the recovery of broad-bandwidth speech from narrow-bandwidth speech", European Conference on Speech Communication and Technology, EUROSPEECH '95, pp. 165-168, Sept. 1995.

[8] H. Yasukawa, “Signal restoration of broad band speech using adaptive digital filter”, the Proc. of 1994 International Workshop on Intelligent Signal Processing and Communication Systems, ISPACS’94, P4.9, pp. 405-408, Oct. 1994.

[9] Y. Nakatoh, et al., “Recovery of wideband speech from narrowband speech based on piecewise linear mapping”, (in Japanese) Technical report of IEICE, US95-76, EA95-70, Jan 1996.

[10] H. Yasukawa, "Signal restoration of broad band speech using nonlinear processing", VIII European Signal Processing Conference, EUSIPCO'96, Sept. 1996.

[11] Y. Tanaka, et al., “Telephone speech enhancement by bandwidth expansion and spectral equalization”, (in Japanese) Acoustical society of Japan Autumn Convention report, 1-P-6, Oct. 1994.

[12] H. Tasaki, et al., “Reconstruction of wideband speech from narrowband CELP codes”, (in Japanese) Acoustical society of Japan Autumn Convention report , 1-5-9, Oct. 1994.

[13] K. Nagami, et al., “A method of generation of wideband speech from band-limited speech by autocorrelation function”, (in Japanese) Acoustical society of Japan Spring Convention report , 2-P-10, March 1996.

[14] CCITT recommendation G.722, Blue Book, Vol. III, Fascicle III.4, Geneva 1989.