Signal Restoration of Broad Band Speech Using Nonlinear Processing

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 a new system that can enhance the quality of speech signals that have been severely band limited during transmission. We have already proposed a spectrum widening method that utilizes aliasing in sampling rate conversion with digital filtering for spectrum shaping. This paper proposes a quite simple method by adding spectrum in the higher band using nonlinear processing. Implementation procedures are clarified, and its performance is discussed. It is shown that the proposed method offers good performance in terms of spectrum distortion characteristics.

1. INTRODUCTION

Quality enhancement techniques for speech and audio signals are being widely studied, for example, telephone band signals have been converted to wide band signals [2],-[5]. In telephone communication, if the quality of conventional telephone speech is enhanced for traditional telephone connections, the effectiveness of all telephone communication services will be improved.

According to ITU-TS (old CCITT) recommendation G.722, ISDN offers a wideband speech signal with a frequency band from 0.05 to 7 kHz [1]. However, ISDN customers can not now obtain high quality speech when they call a conventional telephone (0.3 to 3.4 kHz frequency band) user. This is one reason why quality enhancement is necessary. There are many approaches to enhance speech quality [2],- [7]. We have proposed a very simple method of improving the sound quality of band limited (telephone) speech signals [5], [6]. It is a spectrum widening method that utilizes aliasing effects in sampling rate conversion with digital filtering for spectrum shaping.

This paper proposes a new method that offers reduced circuit volume [5]. A practical implementation method is described. Performance for actual speech signals is shown.

2. IMPLEMENTATION

First, we introduce the conventional spectrum broadening method with multirate processing, then describe the proposed method.

2.1 Multirate Processing Scheme

A block diagram of the spectrum broadening method in a multirate processing system is shown in Fig.1. 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 sampling by 2. The signal is then expanded by an interpolator with 14 kHz sampling. The signal waveform is filtered by a filter bank, a low pass filter, and a 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 widened signal.

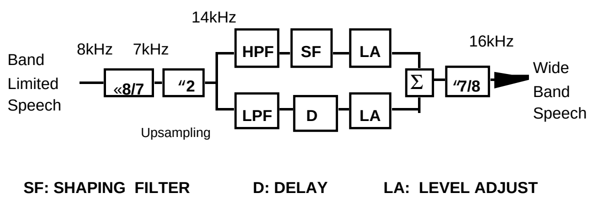


Fig. 1. Block diagram of band limited signal enhancement using spectrum broadening and shaping with multirate processing.

2.2 Proposed Scheme

A block diagram of the proposed system is shown in Fig.2. The sampling rate of the input regular telephone band signal with 8 kHz sampling is expanded by a (1-to-2) interpolator with 16 kHz sampling. The output signal is then filtered into the telephone band by a low pass filter (LPF). After that, a broadened signal is generated by nonlinear processing, that is, in this case the absolute value of the input waveform is generated. The output signal is then filtered by a high pass filter (HPF). The passband of the HPF corresponds to the stopband of the LPF. The high pass filtered signal is shaped by a shaping filter (SF). The high pass filtered signal and low pass filtered signal are summed after level adjustment. This method is a very simple way of generating a widened signal.

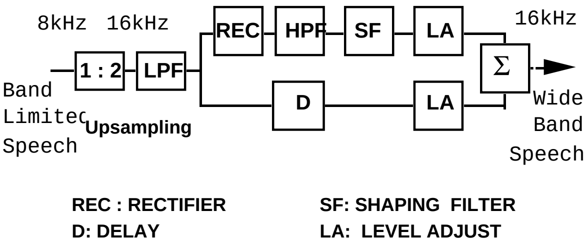


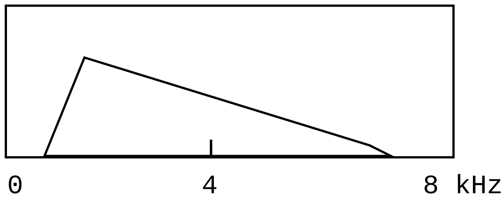
Fig. 2. Block diagram of band limited signal enhancement using spectrum broadening and shaping with nonlinear processing.

The filters utilized are linear phase FIR filters. The delay adjuster matches the phases of the higher and lower band signal paths. For level adjustment, it is important to know the difference between both signals. The level difference ΔL is defined by $\Delta L = L_u - L_L$, where L_u and L_L is the level reference of the upper and lower band, respectively. The level difference is given by level control learned in advance. We describe the optimum ΔL later.

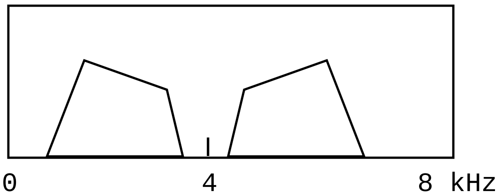
Fig. 3 shows the spectrum of broad band signal generated in the proposed nonlinear processing scheme. The spectrum of higher band speech signal shown in (d) is composed of second order harmonic (and higher) signals of the bandlimited signals.

The frequency responses of the LPF, HPF and the shaping filter are shown in Fig. 4. Cut-off frequencies of the LPF and HPF are around 3.4 kHz.

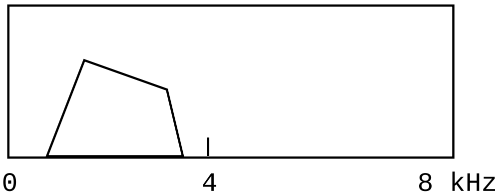
There are two ways of choosing ΔL. One bases ΔL on the circuit parameters that generate the minimum output distortion. The second chooses ΔL using subjective assessment. The latter is the more practical method so we adopted it.



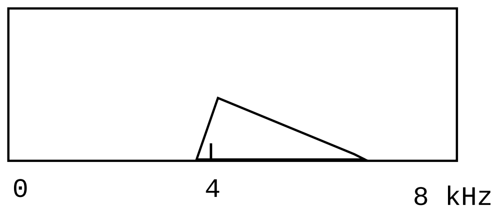
(a) Spectrum of original sending speech signal with 16 kHz sampling



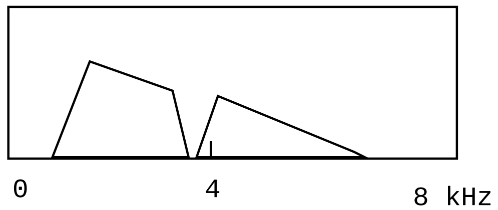
(b) Spectrum of telephone band received speech signal with 8 kHz sampling



(c) Spectrum of speech signal with 16 kHz sampling after upsampling and LPF

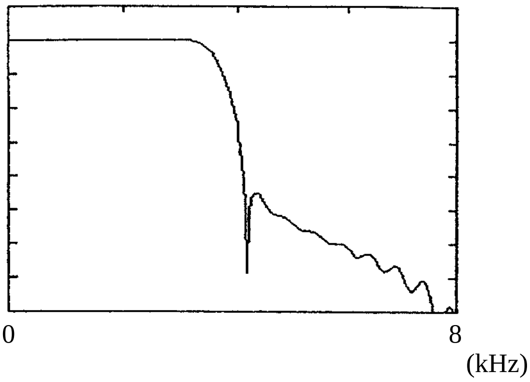


(d) Spectrum of higher band speech signal with 16 kHz sampling after LA

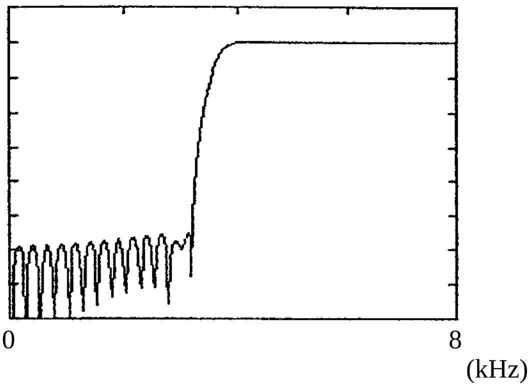


(e) Spectrum of broadening speech signal with 16 kHz sampling by proposed method

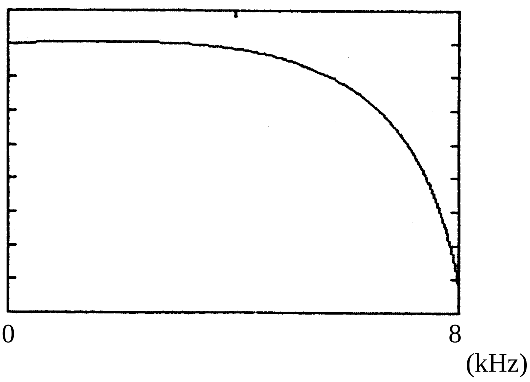
Fig. 3. Spectrum of band limited signal enhancement using spectrum broadening and shaping with nonlinear processing.



(a) Frequency characteristics of LPF.



(b) Frequency characteristics of HPF.



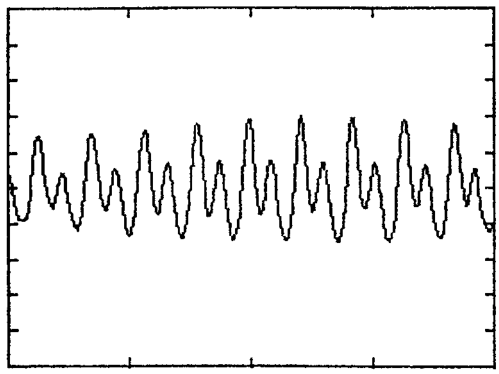
(b) Frequency characteristics of Shaping Filter.

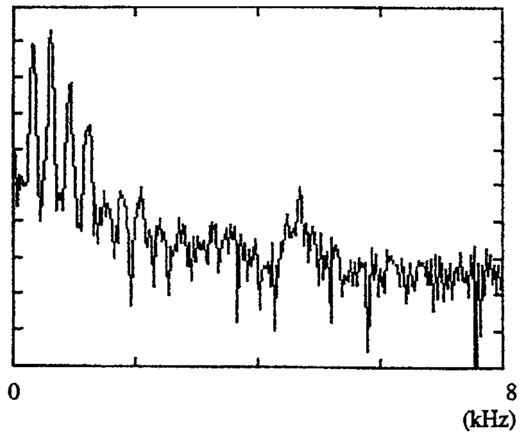
Fig. 4. Amplitude frequency characteristics of Filters. (V: 10 dB/div.)

3. PERFORMANCE EVALUATION

The waveforms and spectra of the original speech signal with 16 kHz sampling before transmission and of the band limited signal in the telephone band are shown in Fig. 5 and Fig. 6.

The waveforms and spectra of the output speech signal of the proposed scheme are shown in Fig. 7.


(a) Waveforms



(b) Spectrum

Fig. 5. Waveforms and spectrum of original speech signal with 16 kHz sampling.

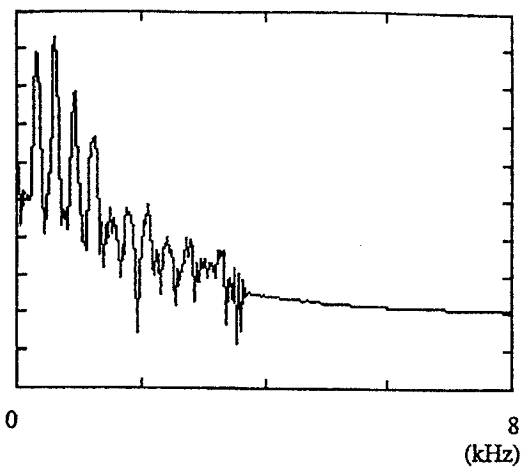
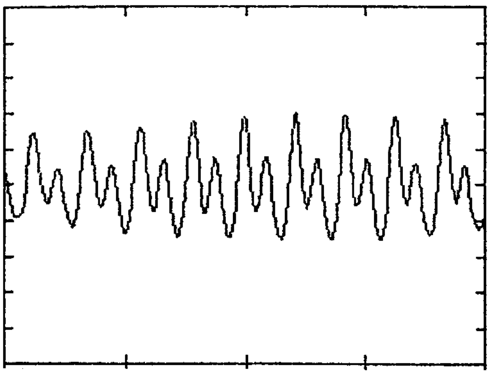
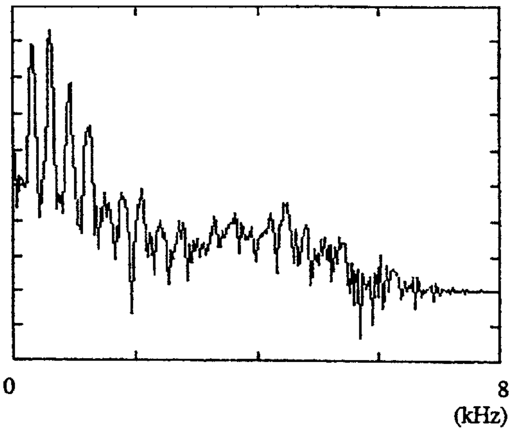


Fig. 6. Spectrum of band-limited speech signal with 16 kHz sampling.

The level difference *D*L is -6 dB. There is no spectrum gap around the folding frequency of 4 kHz.



(a) Waveforms



(b) Spectrum

Fig. 7. Waveforms and spectrum of broadened speech signal with 16 kHz sampling with nonlinear processing.

The spectrum of the output speech waveform is very similar to that of original wideband speech signals. The spectrum distortion level of the output speech signals from the original ones were about -20 dB or less.

The proposed method was also evaluated in a subjective evaluation using a comparison test. The results will be shown in the presentation in detail.

4. DISCUSSION

The proposed quality enhancement method is extremely simple. Although detail results will be reported later, initial subjective evaluations of the proposed method confirm its good performance. It is expected that the characteristics of the proposed method are better than those of other parametric methods, which are tuned into speech properties, because the proposed method does not depend intuitively on speech properties.

The proposed method is also strong against noise and interference periods without speech and so is suitable for actual telephone network conditions.

5. CONCLUSION

A quite simple method for speech quality enhancement by adding spectrum in the higher band using nonlinear processing was proposed for conventional telephone connections and other services. The proposed method utilizes a human auditory characteristic in the frequency domain. This method has many advantages for practical telephone communication. Implementation methods and performance were presented. It was shown that the proposed method offers good performance.

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

[2] D. O. Bowker, et al., "Telephone network speech enhancement", US Patent No. 5195132, Dec. 1990.

[3] Y. M. Chen, et al., "Statistical recovery of wideband speech from narrowband speech", International Conference on Spoken Language Processing, ICSLP ‘92, pp. 1577-1580, 1992.

[4] Y. Yoshida, et al., "An algorithm to reconstruct wideband speech from narrowband speech based on codebook mapping", International Conference on Spoken Language Processing, ICSLP ‘94, 27.3, pp. 1591-1594, Sept. 1994.

[5] H. Yasukawa, "Quality Enhancement of Band Limited Speech by Filtering and Multirate Techniques", International Conference on Spoken Language Processing, ICSLP ‘94, 27.7, pp. 1607-1610, Sept. 1994.

[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.