A SIMPLIFIED SUBBAND NOISE REDUCTION METHOD FOR AUDIO SIGNALS IN TELECOMMUNICATION

Mohona Jahan
Computer Science and Engineering
Brac University
Dhaka, Bangladesh
mohona.jahan@g.bracu.ac.bd

***Abstract*—A method for lowering static background noise in voice time series—which might also be changing—is given. A full-band time sequence is divided into regular grid subbands using a flawless restoration bandpass filter. Every initial code undergoes a detection check, and the outcomes are utilized to dynamically modify the amplitude given to each subband. When the full-band time sequence is rebuilt, the background noise level is lower than the speech frequency. Real-world situations can yield noise reductions of 12 to 18 dJ3 for several purposes. Instances of both telecommunications-level (narrowband) and teleconferencing-level (wideband) speech will be used to illustrate the process. The technique has been shown to significantly improve the quality of speech recorded, as seen in tests employing conventional speech coders.**

***Index Terms*—subband noise reduction, narrowband, wideband, telecommunication**

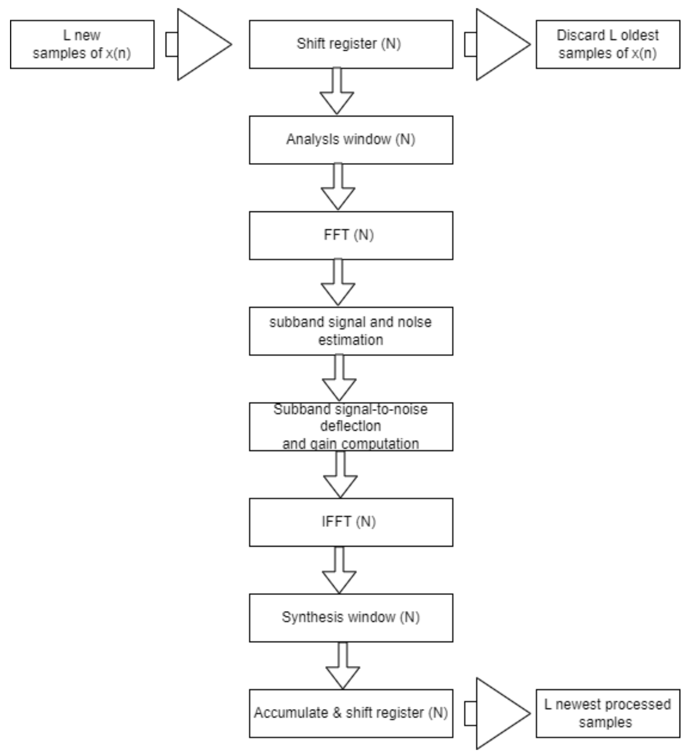
I. Introduction

The understanding that Wiener-filter computing, among other activities, could be achieved by utilizing subband infrastructures emerged with the convergence of electronic multirate perceptual theory in the 1970s and 1980s [I]. The spectrum subtraction method, as proposed by Boll [2], among others, established electronic noise reduction technology for voice improvement. Subsequently, spectrum reduction and a subband gain adjustment approach they called "spectral magnitude expansion" had similarities, according to Etter and Moschytz [3]. Subband noise reduction techniques have lately been the subject of patents [4]. All of these methods are "blind" in the sense that the algorithm only knows the noise-corrupted speech. Therefore, the algorithm must create bootstrap predictions of the signal as well as noise in order to improve the voice signal-to-noise ratio. A subband noise reduction technique for telephone and teleconference-level voice is shown here. The approach presented is comparable to that in [3], but it varies from that approach in terms of subband filter design, how signal and noise frequencies are estimated, and the heuristic processes employed to adjust the subband signal. The method is good in reliability and minimal in computational complexity, fulfilling architectural objectives.

[?]

II. Algorithm Description

A. 3.1 Subband Architecture



figureSubband Noise Reduction Processing

The noise removal technique's signal-flow model is shown in Fig. 1. Subband synthesizing, subband signal-to-noise bending and boost computation, and subband processing (proteolysis) make up the algorithm's three main parts (fullband reconstruction). The discrete Fourier transform (DFT) filter bank approach 121 is used by the subband design to produce a flawless reconstruction filter bank. Although the mathematical framework of this approach is more straightforward, it is the same as a subclass of so-called "polyphase filter banks." Now, for convenience, here is a quick description of this subband filtering technique. A block of new time is added at the beginning of every operating epoch. In an N-sample shift register, a sequence of data is moved. Generally, L = 16 or 32 and N= 64 or 128 for the activity of relevance. The length-N analytical window, which is the prototype FIR filter for the filter bank, is multiplied by the accumulated data before being modified using an N-point DFT. The DFT outputs frequency bins, each of which corresponds to a new structure time-series sample for the associated subband frequency spectrum. The relationship between sample rate and transformation duration determines every subband's wavelength.

During subband evaluation, the subband signal-to-noise bending and gain computing blocks are given the vector of the subband time series. These elements are covered in further detail in Section 3.2. The subband oscillator first performs an inverted DFT transformation on the gain-modified vector of the subband time series in order to recreate the noise-reduced full-band time series. Applying the synthesizing window produces an outcome that is overlaid and stored in the outcome aggregate for N samples. The outcome of the shift-accumulator then generates a block of L-prepared samples. In order to preserve the filter bank's following characteristics, the prototype subband filter, input-output block size L, and transformed block size N are selected:

1. zero subband-level time-domain convolution, 2. zero subband-level frequency analysis convolution, and 3. complete restoration (universal transfer function) when synthesis is done right after inspection, without any further procedure. Additionally, for testing reasons, the analytic and synthesized frames employ a similar prototype FIR filter, which was itself created from a Kaiser data frame. Due to these simplifying assumptions, a modular subband structure is produced, allowing for simple modifications to the filter bank parameters.

B. 3.2 Subband Voice Activity Detection

3.2.1 Signal and Noise Level Estimation: A library of speech activation sensors is depicted receiving subband time series generated by the analytic filter bank in Figure 1. The time-series envelope's long- and short-term standardization are compared for every sensor in the subband. The short-term aggregate determines the transmitted signal while speech is present, whereas the long-term average determines the motionless, or noisy, portion of the subband time series framework. Utilizing nonlinear single-pole recursions, the noise and signal estimations are calculated.

s(i) = as (i-1) +(l-a) lx (i)l (1a)

(14 n(i) = pn(i-1) + (l-p)lx(i)l (1b)

where s(i) and n(i) are indeed the noise and signal estimations, correspondingly, at subband time position I and recursive parameters a and p are provided by

Every subsequent sampling period updates estimations s(i) and n(i) with the size of the subband time-series sample, k(i)l. Based on the connection of Ix(i) I to the present estimation, alpha and beta have various "attack" and "decay" quantities. In a number of speech data processing devices, such as vocoders and speakerphones, the estimation methods (la) and (lb) are commonly employed. These are straightforward and appealing from a mathematical standpoint.

3.2.2 Narrow-Band and Broad-Band Detectors: Immediate speech essentially consists of either bandwidth power (unvoiced communication), narrow-band, multiphase power (vocal communication), or a mixture of both. To enhance sound absorption effectiveness and lessen distortions brought on by failed detection methods, the noise removal algorithm employs a different power sensor for each kind of voice power. The signal-to-noise rate, also known as the bending proportion, is used to determine if narrow-band voice power is present in every subband. d(i) = s(i)/ n(i)........................(4) The findings of two or even more narrow-band deformation proportions are combined in a way that enhances the identification of broad-band power. Take the aggregate bending factor into account.

An arithmetic mean of the K-I-1 subband bending indices, focused on subband j and in relation to the subband j level of noise, is used for subband index j. Rather than a straight (2K+l)-subband average of (4), Eq. (5) is stated in terms of n(j), which enhances the replication of syllable borders and streamlines calculation. Usually K is equal to 2, 3, or 4.

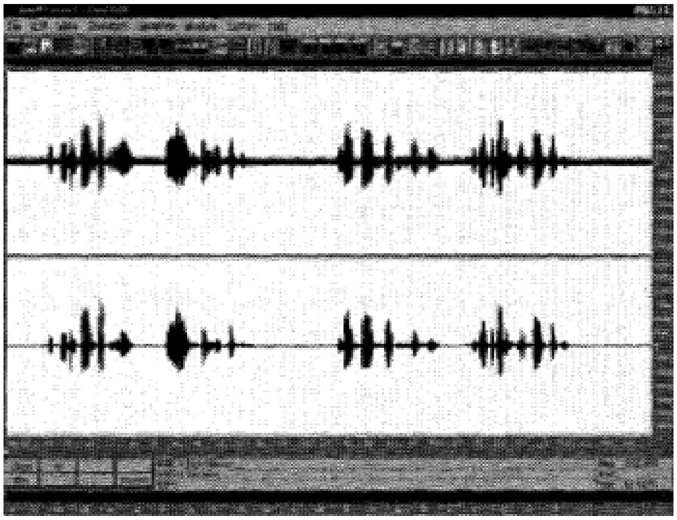
3.3 Subband Gain Computation: The narrow-band and broad-band bending rates are integrated to provide a signals gaining experience for every subband at time index i. The gain is provided by g(i) = 1, if $(i) > 1.0($ i), otherwise........................6(a) Where

and where p is the growth ratio and y and r are the voice identification thresholds. The bending proportions are claimed to suggest a surety of voice power at values known as thresholds y and r. Subband time series with deflections over the threshold are sent with unity gain to the synthesis bank. Subband time series with deflection below the threshold are transferred to the synthesis bank using gain determined by (6b). determines how quickly the gain g(i) decomposes for bending ratios below unity. Since p = I in the present implementation, g(i) degrades linearly further with subband bending ratio (s).

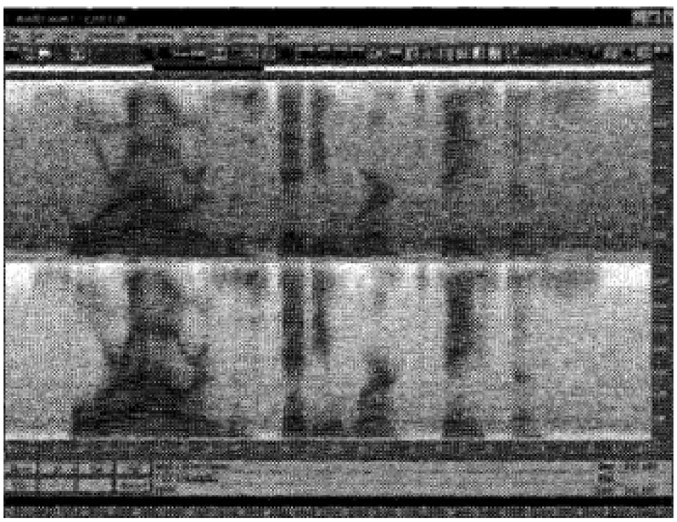
III. Results

The results of analyzing loud speech using the suggested subband noise removal strategy are shown in a sequence of figures in Fig. 2. A section of the original time series for a string of phonetically balanced phrases captured in a car moving at highway speeds is shown in the top trace of Figure 2 (a). The voice was captured using a wireless phone's microphone preamplifier channels, which were afterwards digitalized at an 8 KHz sample rate. The result of the subband noise - reducing method is displayed in the lower trace in (a) for the appropriate noise-reduced section. Spectrograms matching the time series are shown in Fig. 2(b).

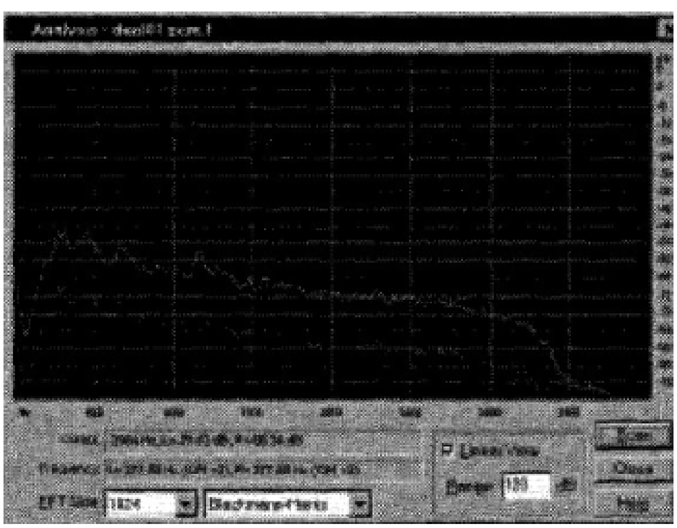
For the time period [3.5,6.0] seconds, h(a) (second sentence). The spectrograms demonstrate, at least visually, that there is no discernible distortion introduced by the noise reduction technique. The aggregated ambient noise power spectra for the unprocessed and noise-reduced time series are shown in Fig. 2(c). As can be observed, the noise level of the treated time series is around 18 dB lower throughout the band than that of the original time series.



(a) Raw (upper) and noise-reduced (lower) speech



(b) Spectrograms corresponding to (a) for the interval [3.5,6.0] sec. (second sentence). (120 dB gray scale.)



(c) Background noise power spectra for the interval [6.0, 7.5] sec. (center of the series in (a))

A. *Discussion*

A heuristic method for background noise suppression uses the narrow- and broad-band identification data of (4) and (5) and the gain extension algorithm of (6b). Other plans might be developed, and more efficiency benefits can be achieved by combining the processes described here with other, complementary ones. Whenever speech recognition is carried out using inconsistent detection statistics, such as energy performance comparisons, the subband structure significantly improves speech detection. This is certainly relevant for vocal communication, which includes lots of distinct spectral analysis that is melodically connected. If the identification statistics just include the subband of frequencies in which the element is present, the identification of any one of such elements is greatly improved. This is so that the subband filter may function for that speech component as a vaguely defined Wiener filter.

IV. Conclusion

As a solution, a secondary subband filter design with a subband bandwidth greater than the narrow-band filter bank might be used to identify broad-band components. This method is preferred from a theoretical perspective because broad-band elements are more suited to the filter bank's greater subband bandwidth, which is itself a consistent mechanism. An inconsistent mixture of adjacent subband energies is produced by the sum in (5a), which is a suboptimal statistics in the notion described earlier. Nevertheless, the effectiveness disparity between the two techniques is minimal for low to substantial K in (5).

Example Citation: [1]

References

[1] E. J. Diethorn, "Subband noise reduction methods for speech enhancement," in *Audio Signal Processing for Next-Generation Multimedia Communication Systems*. Springer, 2004, pp. 91–115.

References

[1] S. F. Boll, "Suppression of Acoustic Noise in Speech Using Spectral Subtraction," IEEE Trans. Acoust., Speech, and Signal Proc., Vol. ASSP-27, No. 2, April 1979.

[2] W. Etter and G. S. Moschytz, "Noise Reduction by NoiseAdaptive Spectral Magnitude Expansion," J. Audio Eng. Soc., Vol. 42, No. 5, May 1994.

[3] B. M. Helf and P. L. Chu, "Reduction of background noise for speech enhancement," U.S. Patent 5550924, March 13, 1995.

[4] R. E. Crochiere and L. R. Rabiner, Multirate Digital Signal Processing, Prentice-Hall, Inc., 1983.

[5] Gay, S. L., Benesty, J. (Eds.). (2012). Acoustic signal processing for telecommunication (Vol. 551). Springer Science Business Media.

[6] Kirovski, D., Malvar, H. S. (2003). Spread-spectrum watermarking of audio signals. IEEE transactions on signal processing, 51(4), 1020-1033.

[7] Djendi, M., Bendoumia, R. (2013). A new adaptive filtering subband algorithm for two-channel acoustic noise reduction and speech enhancement. Computers Electrical Engineering, 39(8), 2531-2550.

[8] Diethorn, E. J. (2004). Subband noise reduction methods for speech enhancement. In Audio Signal Processing for Next-Generation Multimedia Communication Systems (pp. 91-115). Springer, Boston, MA.

[9] Czyzewski, A., Krolikowski, R. (1999, October). Noise reduction in audio signals based on the perceptual coding approach. In Proceedings of the 1999 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics. WASPAA'99 (Cat. No. 99TH8452) (pp. 147-150). IEEE.

References

[1] E. J. Diethorn, “Subband noise reduction methods for speech enhancement,” in *Audio Signal Processing for Next-Generation Multimedia Communication Systems*. Springer, 2004, pp. 91–115.