MULTI-CHANNEL NOISE REDUCTION USING WAVELET FILTER BANK

SIKA Jiri - DAVIDEK Vratislav

Faculty of Electrical Engineering
Czech Technical University
Prague, Czech Republic.
Phone: +420 2 24352291, Fax: +420 2 24310784 , E-mail: sika@feld.cvut.cz

ABSTRACT

This paper deals with the problem of estimation of a speech signal corrupted by an additive noise when observations from two microphones are available. The basic method for noise reduction using the coherence function is modified using Wavelets. The both observations are splitted by filter bank in five narrow bands through the whole used bandwidth (0...4kHz). The coherence functions are then computed for each band and the output speech estimation is reconstructed.

1. INTRODUCTION

The noise reduction in speech signals degraded by an additive noise is an important problem in hand-free mobile car communications. The application of our interest is the enhancement of speech in mobile car telephony to transmit a speech intelligible or more pleasant to listen. The methods developed for solving of noise reduction problem have different performance.

One channel noise cancellation methods based on adaptive processing [2], [3] are popular due their robustness and relative simple implementation. This methods are very powerful in some speech/noise relations. Unfortunately, one channel methods as for example spectral subtraction have limitations in achievement of a marked significant improvement of the noise reduction for a nonstationary and quick changed noises and in the necessity to need silent sequences to estimate the noise spectra.

The multi-channel methods often use the value of coherence function as the criterion to determine whether a speech signal exists or not [1], [4]. The observations picked up by N microphones in running car are each one composed of speech signal and noise. The speech signals are strongly correlated while the correlation between noises is rather weak and depends on the distances between microphones and frequency.

In the paper is introduced a coherence method modified by using of Wavelet filter banks to splitting the whole frequency bandwidth in the five narrower bands. Finally, there are presented practical results obtained on noise speech signals recorded in running car.

2. METHOD DESCRIPTION

The method is described for two microphone observations

$$x_1[n] = s_1[n] + n_1[n] \tag{1}$$

$$x_2[n] = s_2[n] + n_2[n], \tag{2}$$

where $s_i[n]$ is a speech signal and $n_i[n]$ an additive noise from the $i-th$ microphone. The basic idea is to estimate the correlation between the both observations. The speech signals are strongly correlated in the whole transmitted bandwidth, whereas the correlation between noises depends on the microphone distance and the frequency.

The general assumptions of the method are:

- the coherence function between speech signals is close to one
- the coherence function between additive noises is weak.

This second assumption is verified for sufficient distance between microphones. In our experiments the distance of microphones was about 40 cm.

The algorithm was developed in frequency domain. From input signals $x_1[n]$ and $x_2[n]$ are computed the short-time spectrum $X_1[k]$ and $X_2[k]$ using FFT. The data are segmented with a segment length of 256 samples and weighted by Hamming window. The coherence function in $m-th$ segment between channels x_1 and x_2 is computed using power spectral densities of input signals as follows

$$\gamma_{x_1x_2} = \frac{G_{x_1x_2}}{\sqrt{G_{x_1}G_{x_2}}} \tag{3}$$

where G_{x_1} and G_{x_2} are spectral densities of signals $x_1[n]$ and $x_2[n]$ respectively. $G_{x_1x_2}$ is the intersignal spectral density between $x_1[n]$ and $x_2[n]$.

Assuming that the signal to be estimated is speech signal $s_1[n]$, so that the input signal $x_1[n]$ will be processed to give an estimate $\hat{s}_1(t)$. The basic idea of a method is that the coherence function γ is to pass correlated signals and to stop uncorrelated signals.

For the input signal having a coherence function between 0 and 1 the decision about signal character is proposed according criteria:

1. if $|\gamma_{x1x2}|$ is close to 1, the input signal is a speech signal and must be passed without distortion

2. if $|\gamma_{x1x2}|$ is close to 0, the input signal is generally the noise, no speech components of speech are assumed
3. for the intermediate value in the input signal are speech and noise contained and $X_1[k]$ will be weighted by the value of coherence function $|\gamma_{x1x2}|$.

2. 1. Power spectral density computation

The coherence function depends on power spectral densities G_{x_1}, G_{x_2} and $G_{x_1x_2}$. Before description of the used method we show how to compute the noticed power spectral densities required in the approach. They are estimated using short-time spectra $X_1[k,m]$ and $X_2[k,m]$ in a time segment m. The spectral densities are estimated using time averaging by the recursive exponential forgetting formulas

$$G_{x1}(k,m) = (1-\alpha)G_{x1}(k,m-1) + \alpha\,|X_1(k,m)|^2 \ . \quad (4)$$

Similar relation is used for $G_{x2}(k,m)$. For an intersignal power spectral density follows

$$G_{x1x2}(k,m) = (1-\alpha)G_{x1x2}(k,m-1) + \alpha\, X_1(k,m)\, X_2^*(k,m) \quad (5)$$

where forgetting coefficient α was experimentally chosen in the range of 0.2 to 0.3.

2.2 Output spectrum estimation

The approach for the output spectrum $\hat{S}(k)$ estimation was derived by Allen [1] as the method for reverberation reduction. This method was modified for noise reduction [4]. From this method follows the basic idea of our noise reduction advance.

The magnitude of coherence function is computed by

$$|\gamma_{x_1x_2}| = \frac{|G_{x_1x_2}|}{\sqrt{G_{x_1}G_{x_2}}} \ . \quad (6)$$

The value $A(j\varphi)$ is defined by short-time spectra of input signals $x_1[n]$ and $x_2[n]$ as

$$A(j\varphi) = \frac{X_1\, X_2^*}{|X_1\, X_2|} \ . \quad (7)$$

This value is a function of phase difference $e^{j(\varphi_{x1}-\varphi_{x2})}$ between complex spectra $X_1[k]$ and $X_2[k]$.

The output short-time spectra of the estimate $\hat{S}(k)$ is computed by equation

$$\hat{S}[k] = \frac{X_1 + AX_2}{2}|\gamma_{x1x2}| \ . \quad (8)$$

The complex product AX_2 in eq. (8) represents the magnitude of $X_1[k]$ and the phase of $X_2[k]$.

The average of $X_1 + AX_2$ is weighted by module of the coherence function $|\gamma_{x1x2}|$. If input signals $x_1[n]$ and $x_2[n]$ are uncorrelated, it means the absence of speech, the coherence function is close to zero and the average of $X_1 + AX_2$ is switched off. For correlated signals is the estimate of $\hat{S}(k)$ close to average function $X_1 + AX_2$.

3. MODIFIED COHERENCE METHOD

3.1 General Description

The basic idea of our modification lies in the splitting of the input signal in to more frequency bands and processing of the inputs separated for each band. The Fig. 1 shows the basic block diagram of the method.

The input signals $x_1[n]$ and $x_2[n]$ are segmented with segment length of 256 samples and are weighted by Hamming window. The segments overlapping is chosen 50%. The input signals was recorded as stereo signal in real conditions in a car. The segmented signals $x_1[n,m]$ and $x_2[n,m]$ for the *m-th* segment entry to the one of two filter banks. Every filter bank splits the signal to the five frequency bands (0 Hz ... 250 Hz, 250 Hz ... 500 Hz, 500 Hz ... 1 kHz, 1 kHz ... 2 kHz and 2 kHz ... 4 kHz). Due to the used downsampling is the length of the output signals shorter then input. Every segment from the outputs of filter banks are processed by FFT. The short-time spectra $X_{1i}[k]$ and $X_{2i}[k]$ are used to calculate of coherence function and weighting function. The segments of the output spectra estimate $\hat{S}_i(k)$ are transformed to the time-domain using IFFT. The output signal $\hat{s}_1(n)$ is reconstructed using inverse Wavelet filter bank.

3.2. Wavelets filter bank

Processing of the inputs separated for each band shows a way to improve total noise cancellation [6]. The splitting of the input in five bands [5] is used in our modification. The frequency bands was realised as a four-level dyadic tree.

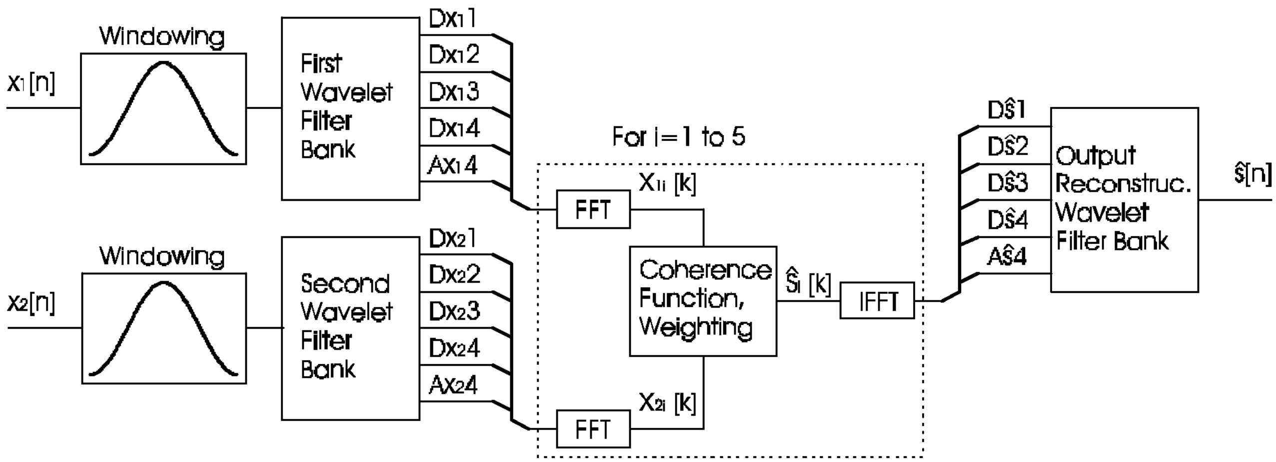

Fig. 1: Block diagram of the modified coherence method using Wavelet filter banks

The input filter bank Fig. 2 was realised by quadruple repetition of a simple block which contain a high-pass decomposition filter, a low-pass decomposition filter and two downsampling blocks. Output from a high-pass decomposition filter after downsampling is described as detailed signal (Dx_i) and output from low-pass filter after downsampling is described as approximated signal (Ax_i)

The output filter bank Fig. 3 is similar to input filters bank but it was realised in reverse order. Downsampling is replaced by upsampling and the decomposition filters are replaced by reconstruction filters.

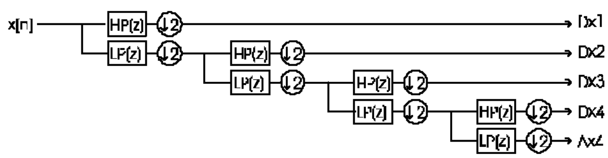

Fig. 2: Block diagram of the five bands input filter bank

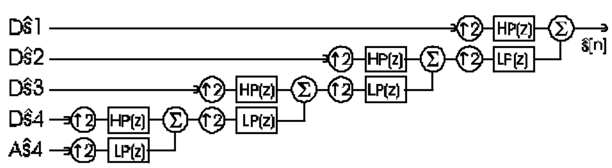

Fig. 3: Block diagram of the five bands output filter bank

The frequency bands of filter bank was elected with consideration for spread out energy of input signal and not complicated realisation. For processing of an acoustic signal in frequency band ($0 \ldots 4\ kHz$) are sufficed five bands with octave bands.

This splitting in five bands rate was elect as a compromise between computation demand and the achieved results. In the developed method is used multiple-level decomposition. The global algorithm for computation of an output of the filter bank lies in repetition of the simple algorithm. This simple algorithm splits one input to two outputs. The approximated signal (Ax_i) contains the low frequency components obtained by LP filter and the detailed signal (Dx_i) contains the high frequency components obtained by the complementary HP filter. As the input for second section is used the approximated signal of the first section.

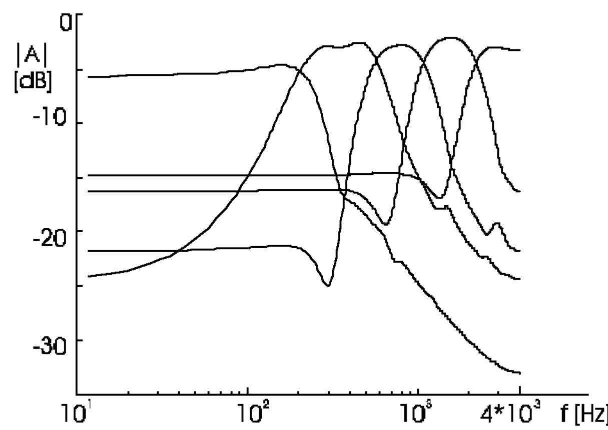

Fig. 4: Frequency characteristic of realised filter bank

The difference between Wavelet filter bank and classic filter bank (FIR and IIR) is primarily at downsampling of the output signal, therefore output signals from Wavelet based filter bank have a shorter number of samples and the computation demands are less. After separation of two input signals to ten signals (five for each one) are performed five coherence algorithms. The system output signal is composed from five output signals of the coherence functions composed by inverse Wavelet transform (IWT). IWT is computed in the same way as WT in reverse order. There exist many types of basis for Wavelet transform. For our application the Wavelet transform on Daubechies basis seems to be convenient.

4. RESULTS

The developed method was tested on real speech signals degraded by the additive noise from running car. The signals were recorded using two microphones with distance about 40 cm.

The results achieved by modified coherence method are illustrated in figure 5. In fig. 5 a) is shown the clear speech signal. The speech signal degraded by additive noise from running car is in figure 5 b). In figures 5 c) and d) are illustrated both noise cancellation methods, the basic coherence method - fig. 5 c) and the modified coherence method using Wavelet filter bank - fig. 5 d).

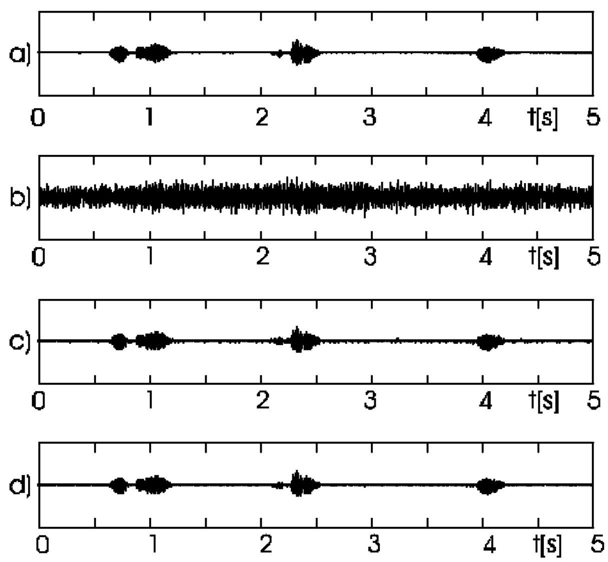

Fig. 5: Comparison of noise cancellation using basic and modified coherence methods.

5. CONCLUSIONS

In this paper we propose the coherence method for noise reduction modified by Wavelet based filter bank. This method was tested on real speech signals recorded by two microphones in a car. The results was compared to those obtained with basic coherence method, derived from Allen's method and spectral subtraction. Under subjective tests we obtained speech signal more pleasant to listen. The advantage of developed coherence method consists in its simply application for more input channels. The application of Wavelet transform in coherence method gives an approved performance in noise reduction.

The modification of basic algorithm using filter bank may be used not only for coherence method, but also for modification of another methods, for example for a spectral subtraction [3] and for the others.

The advantage of Wavelet filter banks is a simply implementation and fast calculation, but unfortunately, in comparison of Wavelet filter banks with classic filter banks using FIR and IIR filters, Wavelets have a wider transient band lower attenuation in stop band for the same order of the filter.

The modified algorithm presented in this paper was implemented in MATLAB. In the present time we are implementing the developed algorithm on a floating-point digital signal processor for testing of noise reduction in real-time in a car.

6. REFERENCES

[1] Allen, J. B. - Berkley, D. A. - Blauert, J.: Multimicrophone Signal Processing Technique to Remove Room Reverberation from Speech Signals. J. Acoust. Soc. Am., 1977, pp 912-915

[2] Boll, S.F.: Suppression of Acoustic Noise in Speech Using Spectral Subtraction. ASSP-27, No.2, April 1979, pp 113-121.

[3] Davidek, V. - Sovka, P. - Sika, J.: Real-Time Implementation of Spectral Subtraction Algorithm for Suppression of Acoustic Noise in Speech. In. proc. Of the 4rd European Conference on Speech Communication and Technology, EUROSPEECH'95, Madrid, Sept. 1995, pp 141-144.

[4] Bouquin, R. L. - Faucon, G.: Using the Coherence Function for Noise Reduction. IEE proceedings - I, Vol. 139, No.3, June 1992, pp 276-280.

[5] Strang, G. - Nguyen, T.: Wavelets and Filter Banks. Wellesley-Cambridge Press, 1996.

[6] McAulay, R. J. - Malpass, M. L.: Speech Enhancement Using a Soft-Decision Noise Suppression Filter. in Lim, J. S.: Speech Enhancement. Prentice-Hall, 1983, pp 74-82.