NOISE SUPPRESSION SYSTEM FOR A CAR

Petr Pollák, Pavel Sovka, and Jan Uhlíř

Department of Circuits Theory, Czech Technical University in Prague, Prague, Czech Republic
E-mail: UHLIR@FELD.CVUT.CZ

ABSTRACT

The whole system for noise suppression in speech recorded in a running car was designed. One channel spectral subtraction method with full-wave rectification was chosen because of its robustness, simplicity, and non-musical tone output. The improvement of noise suppression was gained by the repetition of this method. Directional microphones for the signal picking up were chosen to improve the input signal-to-noise ratio (SNR) of corrupted speech signal. Segment speech/pause detector based on energy tracking was used with some prefiltration of corrupted speech to improve detector function.

Keywords: *noise suppression, speech enhancement, spectral subtraction, musical tones, speech/pause detection*

1. INTRODUCTION

Speech recognition and transmission in a noisy environment requires a noise suppression preprocessor of the noisy speech. This task becomes very difficult especially if the noise is highly nonstationary and diffuse. This type of problem arises when a hand-free mobile telephone is used in a car. If, moreover, the minimal cost and simple arrangement inside a car is required then the one or two microphone methods are allowed only.

2. CHOSEN METHODS

Our task was to design a system for noise suppression in a car as a preprocessing system for a hand-free car telephone [1]. After the extensive study, comparisons, experiments, and with respect to the required simplicity of the whole system we decided to use one - maximally two - microphones system and to apply appropriate methods for the car noise suppression. One-channel spectral subtraction method is used for the suppression of the noise caused by the engine and by the vibrations of the car body. Two-channel adaptive noise canceller can be used for a radio noise because it is easy to get a noise reference signal without crosstalks in this case.

The entry conditions for the first steps of our work were the following:

- speech distorted by additive noise from the cab of car running on different surfaces, with closed windows, and without radio noise.
- no cocktail party effects
- fan noise present

To design system for noise suppression system as preprocessing system for hand-free telephone we followed next steps:

- the creation of the database from records of noise and noisy speech signals,
- the analysis of power and power spectral density of car noises and their time dependency, classification with respect to typical characteristics,
- the choice of proper algorithms, their setting and cascading into the whole system, i.e. the speech/pause detection, the noise spectrum estimation in the speech pauses, and the non-musical tone algorithm for noise suppression,
- simulation and experimental verification of the whole system

3. SPEECH/PAUSE DETECTOR

Noise suppression methods require a speech activity detector. We solved the problem of speech and pause detection in noisy speech signal. We use the detector which originates from Harrison's algorithm published in [2]. We suggest an equivalent algorithm which works on signal segments with target to save the relative long processing time of algorithm described above because it works for each sample of signal. Another advantage of this algorithm is better implementation into spectral subtraction method for noise suppression especially when it works in the frequency domain. This algorithm is based on energy tracking and can be described by next equations

$$E = \sum_{n=0}^{N-1} s^2[n] \quad \text{or} \quad E = \sum_{k=0}^{N-1} S^2[k]\,, \tag{1}$$

$$E_p = 1.5 \cdot E_d\,, \tag{2}$$

$$E_d^{\text{new}} = (1-p) \cdot E_d^{\text{old}} + p \cdot E\,, \tag{3}$$

where E is the segment energy computed either from signal samples $s[n]$ or from DFT coefficients $S[k]$ of a signal

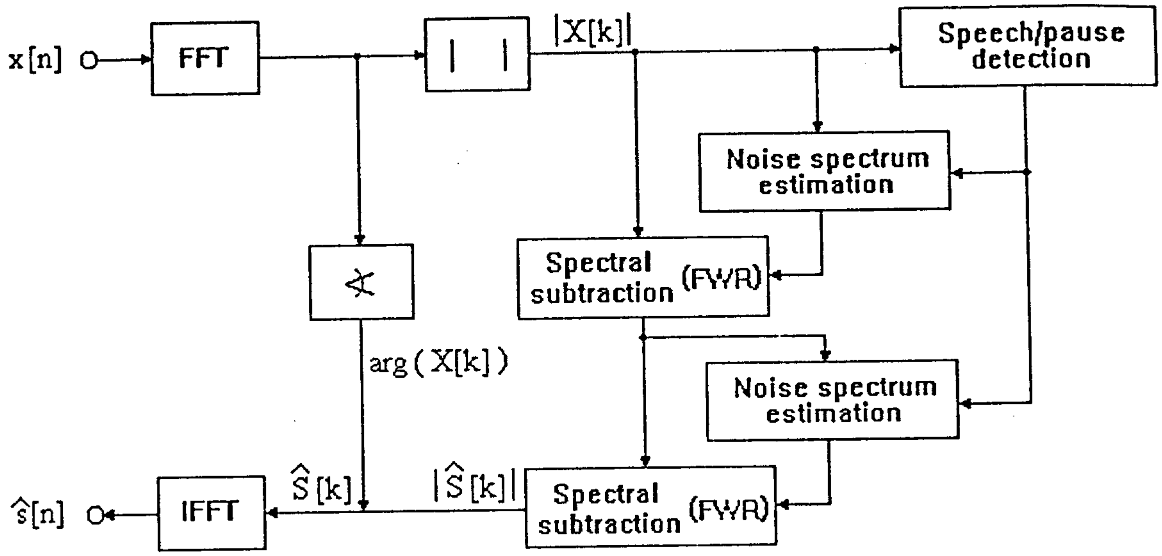

Figure 1: Block scheme of our spectral subtraction method

segment. This energy is compared with the threshold E_p given by background noise energy E_d; i.e. if E is greater than E_p the speech is detected or on the other hand the background noise energy E_d is modified according (3).

This algorithm is very simple and gives good results but it is able to work only for the positive SNR measured as a long-time parameter. Since the SNR in the cab of a car is very often negative we try to improve the input conditions of the detector by some type of prefiltration the input signal. We use filtration by high-pass filter with cut-off frequency $f_c = 350\ Hz$ which can remove dominant strong noise from the car engine. It is possible to use some specified bands for the energy computation too. In the case when the algorithm works in the frequency domain the realization of this filtration is very easy, i.e. by simple choosing of specified spectral components.

Another possibility of increasing the input SNR is in using a directional microphone. The directional microphone was assembled from the pair of small electret microphones. This assembled microphone was placed in the small nick on the front dashboard. Reverberations coming to the rear side of the microphone were attenuated by the piece of velvet. The compromise distance between a speaker and the microphone was found to be *50 cm*. The SNR was greater for the bidirectional microphone than for the omnidirectional microphone no matter of the proximity effect of bidirectional microphone. This SNR improvement was about *3 dB*. This result was confirmed by the critical distance measurement and the measurement of signal powers in the car [1].

4. NOISE SUPPRESSION

If the assumption of the only additive noise presence in the corrupted speech is made, i.e.

$$x[n] = s[n] + d[n], \tag{4}$$

then the use of one channel spectral subtraction method for noise suppression seems to be suitable with respect to its robustness and simplicity. This method is based on the estimation of the magnitude spectrum of enhanced speech signal because of the perceptual aspects of the human audible system [3]. Good estimation of this spectrum is limited by the effect of musical tones. We compared some algorithms of spectral subtraction and musical tone reduction, i.e. spectral smoothing algorithms, median filtering, masking by wide-band noise, or the multiple subtraction of magnitude noise spectrum estimation; presented especially in [4], [5], [6]. We found the best solution for transmission purposes leads to the use of spectral subtraction with pure full-wave rectification which noise suppression was worse but which avoided musical tones generation. The algorithm of this method, see fig.1, can be described by the equation

$$|\widehat{S}[k]| = \begin{cases} \Big| |X[k]| - |\widehat{D}[k]| \Big|, & \text{if } |X[k]| < |\widehat{D}[k]|, \\ |X[k]| - |\widehat{D}[k]|, & \text{otherwise} \end{cases} \tag{5}$$

where $X[k]$ is noisy speech spectrum, $\widehat{D}[k]$ is averaged noise spectrum, and $\widehat{S}[k]$ the estimation of enhanced speech spectrum. The improvement of noise suppression is possible to gain by repetition of this method. The number of operations is equivalent to other methods because FFT and IFFT between particular steps can be omitted. This method which avoids musical tone generation seems to be more effective than other methods, moreover, many musical tone reduction methods distort enhanced speech signal.

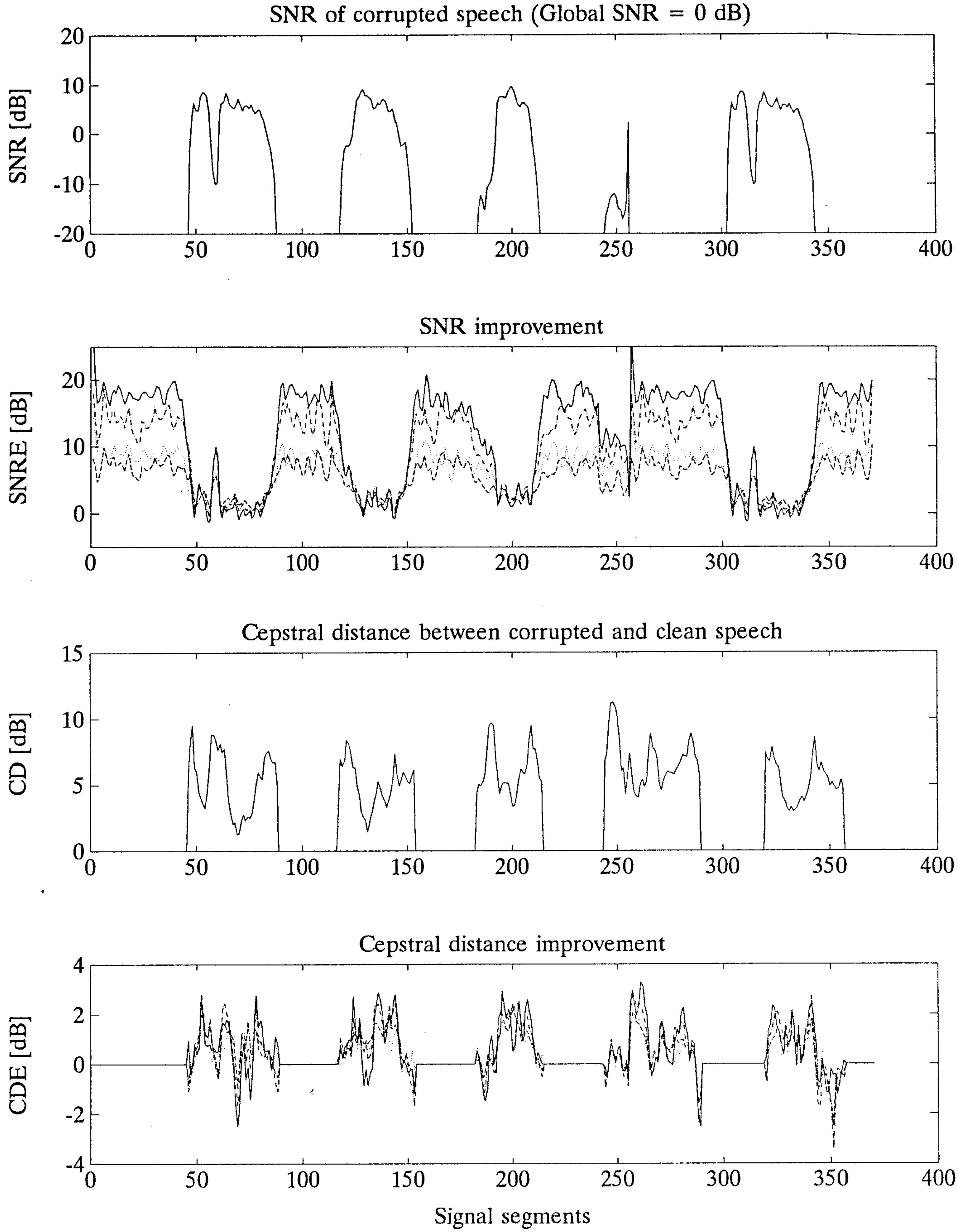

Figure 2: Comparison of local SNR and CD improvement for compared spectral subtraction methods

Another strong limitation of this spectral subtraction method is the described speech/pause detector because the averaging of noise spectrum has to be performed in speech pauses only. The exponential forgetting can be used according formula

$$|\widehat{D}[k]|^{\mathrm{new}} = (1-p) \cdot |\widehat{D}[k]|^{\mathrm{old}} + p \cdot |X[k]| \,. \qquad (6)$$

The enhanced speech signal is reconstructed using the estimation of the magnitude spectrum of enhanced speech and the phase of the input corrupted speech, i.e.

$$\widehat{s}[n] = \mathrm{IFFT}\left\{ |\widehat{S}[k]| \, e^{j \arg X[k]} \right\} . \qquad (7)$$

5. EXPERIMENTS

For our experiments we used signals sampled by frequency of *8 kHz* and quantized to *12 bits*. We realized the experiments with artificially mixed signals in the computer, i.e. clean speech recorded in the car and clean noise recorded in the running car under different conditions with the specified SNR. In this case it is possible to compute some objective criteria of the degree of speech enhancement resp. the level of noise suppression. We used two criteria: *SNR improvement (SNRE)* and *cepstral distance improvement (CDE)*.

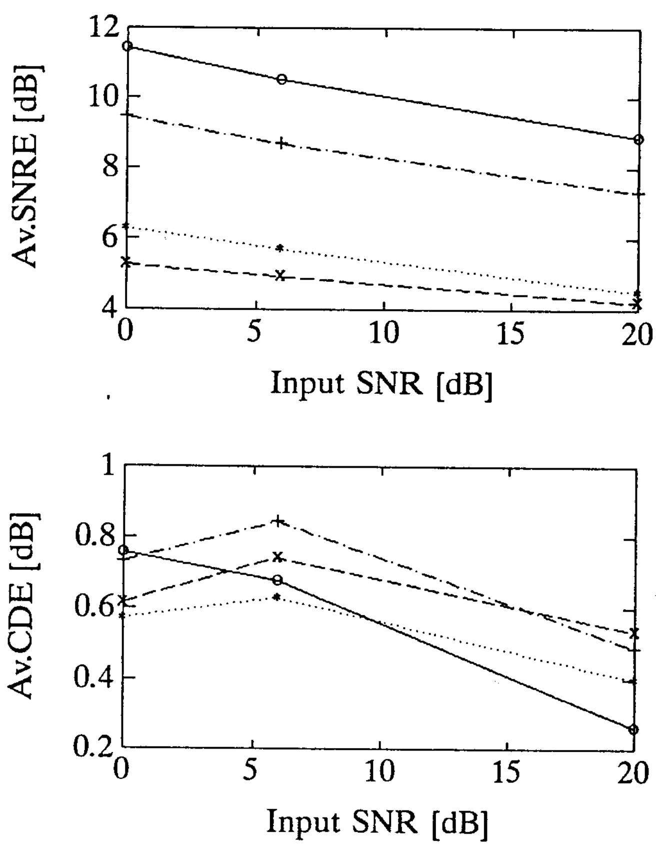

Figure 3: Average SNRE and average CDE dependency on SNR of input signal for compared spectral subtraction methods

The results of the comparison are presented on fig.2 and fig.3 where dashed line is for Kang-Fransen method, dashdotted for method with half-wave rectification, dotted for the results after first step of our method, and solid one for final results of our method.

Experiments was repeated with real corrupted speech signals recorded in the running car and with real speech/pause detection to confirm results with simulated signals.

6. CONCLUSION

The integrated system for noise suppression in the car was designed. The noise suppression of this system is about *5* to *8 dB* depending on the noise and speech characteristics. This system was optimized with respect to the minimum of musical tone presence in enhanced speech.

7. ACKNOWLEDGEMENT

We would like to thank the US WEST company for its support to our research especially to Mr. Hynek Hermansky.

REFERENCES

[1] Petr Pollák, Pavel Sovka, and Jan Uhlíř. The enhancement of noisy speech for transmission and recognition. Research report USWEST agreement #CS140191071, Czech Technical University, Faculty of Electrical Engineering, Prague, May 1993.

[2] William A. Harrison, Jae S. Lim, and Elliot Singer. A new application of adaptive noise canceller. *IEEE Transactions on Acoustics, Speech and Signal Processing*, Vol.ASSP-34(No.1):21–27, February 1986.

[3] Jae S. Lim and Alan V. Oppenheim. Enhancement and bandwidth compression of noisy speech. *Proceedings of the IEEE*, Vol.67(No.12):1586–1604, December 1979.

[4] Steven F. Boll. Suppresion of acoustic noise in speech using spectral subtraction. *IEEE Transaction on Acoustics, Speech and Signal Processing*, Vol.ASSP-27(No.2):113–120, April 1979.

[5] M.Berouti, R.Schwartz, and J.Makhoul. Enhancement of speech corrupted by acoustic noise. In *Proceedings of the IEEE Conference on Acoustics, Speech, and Signal Processing*, pages 208–211, April 1979.

[6] G. S. Kang and L. J. Fransen. Quality improvement of LPC-processed noisy speech by using spectral subtraction. *IEEE Transactions on Acoustics, Speech and Signal Processing*, Vol.ASSP-37(No.6):939–942, June 1989.