JAIST Repository
https://dspace.jaist.ac.jp/

Title	Multi-channel noise reduction in noisy environments
Author(s)	Li, Junfeng; Akagi, Masato; Suzuki, Yoiti
Citation	Lecture Notes in Computer Science, 4274: 258-269
Issue Date	2006
Type	Journal Article
Text version	author
URL	http://hdl.handle.net/10119/4909
Rights	This is the author-created version of Springer, Junfeng Li, Masato Akagi and Yoiti Suzuki, Lecture Notes in Computer Science, 4274, 2006, 258-269. The original publication is available at www.springerlink.com, http://dx.doi.org/10.1007/11939993_30
Description	

Japan Advanced Institute of Science and Technology

Multi-channel Noise Reduction in Noisy Environments

Junfeng Li[1], Masato Akagi[2], and Yôiti Suzuki[1]

[1] Research Institute of Electrical Communication, Tohoku University,
2-1-1, Katahira, Aoba-ku, Sendai, 980-8577, Japan
[2] School of Information Science, Japan Advanced Institute of Science and Technolgoy,
1-1, Asahidai, Nomi, Ishikawa, 923-1291, Japan
{junfeng@ais.riec.tohoku.ac.jp, akagi@jaist.ac.jp, yoh@ais.riec.tohoku.ac.jp}

Abstract: Multi-channel noise reduction has been widely researched to reduce acoustic noise signals and to improve the performance of many speech applications in noisy environments. In this paper, we first introduce the state-of-the-art multi-channel noise reduction methods, especially beamforming based methods, and discuss their performance limitations. Subsequently, we present a multi-channel noise reduction system we are developing that consists of localized noise suppression by microphone array and non-localized noise suppression by post-filtering. Experimental results are also presented to show the benefits of our developed noise reduction system with respect to the traditional algorithms in terms of speech recognition rate. Some suggestions are finally presented for the further research.

Keywords: Multi-channel noise reduction, beamforming technique, localized noise, non-localized noise, speech recognition.

1. Introduction

Acoustic noise signals dramatically degrade the performance of many speech applications, such as speech communication system and automatic speech recognition system, in noisy environments [1]. For example, for speech communication system, acoustic noises degrade the quality and intelligibility of the received signals. For automatic speech recognition system, acoustic noises cause the mismatch between the training and testing conditions, further decreasing the recognition accuracy in real-world adverse conditions. Therefore, noise reduction must be very useful to improve the performance and robustness of these applications in noisy environments [1].

Though the problem of dealing with acoustic noises has been researched for several decades, it is currently still a challenging research topic. The challenges are mainly caused by the complex and time-varying characteristics of the signals (speech and noise signals) and acoustic environments [1], [2]. Desired speech signals have a broad-band and highly time-varying spectral components. In practical environments, noise signals have very complex and time-varying properties. Take the noise condition in a car environment as an example. Noises generated by winds around the car come from all directions and have slowly time-varying spectral components including coherent and incoherent noise components that are generally modeled as diffuse noise. Noises generated by engine come from certain directions and have slowly time-varying spectral components. Undesired interfering noises (e.g., passenger’s voice and radio), however, have some determinable directions and highly non-stationary speech-like spectral components. Noises with different characteristics from various kinds of sources make it difficult to construct an effective noise reduction system. Furthermore, the characteristics of noises do vary with time and environments in an unpredictable fashion, further increasing the difficulty of designing a noise reduction system. Additionally, considering the practical implementation, real-time processing is generally a “must” for noise reduction systems in real conditions [1], [2].

To suppress various kinds of noises, many noise reduction algorithms have been published in the literature [1], [2]. Generally speaking, all of these noise reduction algorithms can be classified into two categories: single-channel technique and multi-channel technique, according to the number of microphones which are needed in the implementation.

A variety of single-channel noise reduction techniques [3], [4], [5], which exploit spectral and temporal differences between the speech and noise signals to suppress acoustical noises, have been proposed for speech enhancement and speech recognition. In real conditions, however, the speech and noise signals are considerably overlapped in the time-frequency domain, which makes it extremely difficult for single-channel techniques to substantially eliminate most of noise components without introducing speech distortion and artifacts (e.g., musical noise). As a result, single-channel techniques achieve very limited improvements in suppressing noise and in enhancing the speech enhancement and recognition performance [2]

In addition to the temporal and spectral characteristics, multi-channel techniques allow to exploit the spatial diversity of the speech and noise signals, resulting in the highly improved noise reduction performance [3]. In most scenarios, desired speech source and interfering noise source are physically located at different positions in the space. Exploiting the spatial diversity of the signals, multi-channel techniques can steer a main beam towards the desired speech source and/or nulls towards the interfering noise sources. Thus, compared to single-channel noise reduction techniques, multi-channel noise reduction techniques are substantially superior in suppressing the interfering signals arriving from the directions other than the specified “look” directions [2]. Additionally, among multi-channel noise reduction algorithms, post-filtering is normally needed to improve the entire performance in practical noisy environments. Therefore, multi-channel noise reduction systems with post-filtering have attracted increasing research interests [2].

In this paper, we first give a review of the state-of-the-art multi-channel (i.e., beamforming based) noise reduction systems ranging from the simple delay-and-sum beamformer to the advanced adaptive beamformers, as well as post-filtering. We then introduce the multi-channel noise reduction system we are developing consisting of localized noise suppression by microphone array and non-localized noise suppression by post-filtering. Experimental results are also presented to illustrate the benefits of our proposed system in terms of speech recognition accuracy in realistic environments where interfering signal and ambient noise are present. We finally provide some suggestions for the further research.

2. State-of-the-art multi-channel noise reduction

In comparison of single-channel noise reduction algorithms, multi-channel noise reduction algorithms have demonstrated a substantial superiority in reducing noise due to their spatial filtering capability. So far, many beamforming based algorithms have been reported in the literature [6], [7], [8], [9], [10], [11], [12], [13], [14]. The beamforming algorithms include fixed beamformer and adaptive beamformer, which are briefly discussed in the following sub-sections. Additionally, the widely used post-filtering algorithms are also discussed. Special attention is paid to the disadvantages of these existing algorithms.

2.1 Fixed beamforming

The first class of beamforming techniques is fixed beamforming. In fixed beamforming techniques, the filter coefficients are normally optimized so that a beam is steered to the direction of the desired signal while suppressing the background noise coming from other directions as much as possible. These optimized filters are fixed, independent of the input signals, and then applied to the multi-channel microphone inputs [1], [2].

The simplest beamformer, referred to as delay-and-sum (DS) beamformer [2], [6], enhances the desired speech signal by summing the in-phase microphone signals after compensating for the arrival time differences of the desired sound signal to each microphone by inserting delays after each microphone, that is, the array is first electronically steered to the look-direction. In other words, in the DS beamforer, the weights of filters are fixed to for all frequencies and all frames. The advantages of the DS beamformer are that it is very simple to implement and that it minimizes the noise sensitivity and hence provides a high robustness against errors in the assumed signal model. However, a large number of microphones are normally needed to obtain an acceptable performance in real-world environments [2]. The superdirective beamformer is another widely studied fixed beamformer [7]. The superdirective beamformer maximizes the directivity index in the direction of the speech source for a diffuse noise field. Actually, the superdirective beamformer minimize the noise power of the beamformer output subject to distortionless response for the “look” direction, hence, it is also a *minimum variance distortionless response* (MVDR) beamformer. The implementation simplicity of the superdirective beamformer leads to its widely use in some known noise field. However, its data-independent property results in that only limited noise reduction performance can be obtained in practical time-varying environments [2].

Fixed beamforming techniques are widely used in the conditions where the acoustical characteristics do not change with time. However, using the fixed beamforming techniques, it is generally not possible to design arbitrary spatial directivity patterns for arbitrary microphone array configuration and design spatial directivity patterns which can be optimized to the time-varying acoustic environments [2], [7].

2.2 Adaptive beamforming

The second class of beamforming techniques is adaptive beamforming. In contrast to fixed beamforming techniques, adaptive beamforming techniques make use of data-dapendent filter coefficients that are adapted to respond to time-varying environments, yielding a better noise reduction performance than fixed beamforming techniques, particularly if the number of interference is small (i.e., smaller than the number of microphones) and in the acoustic environments with low reverberation [1], [2].

Adaptive beamforming techniques (e.g., the Frost beamformer) typically solve a *linearly constrained minimum variance* (LCMV) optimization problem [9], keeping the signals arriving from the desired look-direction (i.e., ideally the direction of the desired speech source) distortionless while suppressing the signals from other directions by minimizing the output power. The MVDR beamformer was proven as a special case of the LCMV beamformer under the assumption of zero correlations between the speech signal and the noise signal [2]. A *generalized sidelobe canceller* (GSC) beamformer, first presented by Griffiths and Jim as an alternative implementation structure of the LCMV beamformer, has also been widely researched [9]. The GSC beamformer consists of: a fixed beamformer which electronically steers the microphone array to the direction of interest (i.e., the speech source) and generates the so-called speech reference signal; a block matrix which steers the spatial nulls to the direction of speech source and generates the so-called noise reference signals; and a multi-channel noise canceller which suppress the residual noise components in the speech reference signal by using a multi-channel adaptive filter [9]. In addition, a wide variety of noise reduction algorithms that are based on the GSC beamformer have so far been suggested, which are of interest to be mentioned. Bitzer *et al.* presented an alternative implementation algorithm with a GSC structure of the superdirective beamforme and its performance was also analyzed in a diffuse noise field [10]. Fischer *et al.* proposed to apply a Wiener filter in the upper path of the GSC beamformer to suppress the uncorrelated noise components and then the correlated noise components are then reduced by the adaptive noise canceller in the lower path [11]. Recently, the GSC beamformer was extended to a *transfer function generalized sidelobe canceller* (TF-GSC) beamformer by considering the transfer functions which relate the speech source and the microphones, which was shown to yield high noise reduction performance in real-world environments [12]. Moreover, the theoretical performance of the GSC and TF-GSC beamformers was examined in the diffuse noise field [16], [17].

In all variants of the LCMV and GSC beamformers, adaptive signal processing (e.g., LMS) is normally used to avoid cancellation of the desired speech signal, which introduces low convergence rate in practical conditions and low ability in reducing non-stationary noise (e.g., sudden noise). Moreover, the adaptive beamformers only perform well and provide acceptable performance when the number of interfering noise sources is less than that of the microphones. Their performance will be greatly degraded by the reverberation effect and in the scenario where more noise sources exist (e.g., larger than the number of sensors) [2].

2.3 Post-filtering

Multi-channel beamforming based algorithms provide high noise reduction performance especially for localized noise, however, only limited noise reduction performance is achieved in a diffuse noise field [2], [13], [14]. To further suppress residual noise at the beamformer output, post-filtering is normally needed to improve the noise reduction performance of the entire system in practical environments.

A variety of post-filtering techniques have been presented in the literature [13], [14] [15]. One commonly used multi-channel post-filter, which is based on Wiener filter, was first introduced by Zelinski [15]. The basic assumption behind this post-filter is that noises on different microphones are mutually uncorrelated, corresponding to a perfectly incoherent noise field. This assumption is, however, seldom satisfied in practical environments, especially for closely-spaced microphones and low frequencies, which are characteristics by the high-correlated noise [15].

To suppress the high-correlated noise, Fischer *et al*. proposed a noise reduction system which is based on the GSC beamformer [11]. The GSC beamformer suppresses the spatially coherent noise components, whereas a Wiener filter in the look direction is designed to suppress the spatially incoherent noise components [11]. However, Bitzer *et al*. pointed out that neither the GSC nor the standard Wiener post-filter performs well at low frequencies in a diffuse noise field [16], [17]. Therefore, they proposed to add a second post-filter at the output of a GSC beamformer with standard Wiener post-filter to reduce the spatially correlated noise components [18]. Recently, McCowan *et al*. developed a general expression of the Zelinski post-filter based on the a priori coherence function of the noise field [19]. Although this post-filter was shown to achieve improved speech quality and speech recognition accuracy compared to the Zelinski post-filter using the office room recordings, its performance is expected to be significantly degraded when difference between the “actual” and assumed coherence function exists.

3. Proposed multi-channel noise reduction

3.1 Theoretical principle of proposed system --- multi-channel Wiener filter [2]

The underlying theoretical principle of our proposed multi-channel noise reduction system is the multi-channel Wiener filter, which provides an optimal solution to the problem of multi-channel noise reduction for broadband inputs in *minimum mean square error* (MMSE) sense [2]. With the assumption that the desired signal and noise signals are mutually uncorrelated, Simmer *et al*. showed that the multi-channel Wiener filter can further be decomposed into a MVDR beamformer followed by a single-channel Wiener post-filter [2]. As an extension of this algorithm, we propose a multi-channel noise reduction system consisting of localized noise suppression by a microphone array and non-localized noise suppression by post-filtering. The detailed description is given in the following subsections.

3.2 Signal model in the proposed system [14]

Let us consider an array of M microphones in a noisy environment. In our research, the observed signal on each microphone consists of three components. The first one is the desired speech signal $s(t)$ arriving from the direction such that the direction in arrival time between two microphones is ξ, The second is localized noise signals $n_p^c(t)$ arriving from the directions such that the time differences are $\delta_{m,p}$ $(p=1,2,\ldots,P)$ and the third is non-localized signal $n^{uc}(t)$ which propagates in all directions simultaneously and is normally modeled as diffuse noise. Thus, the observed signal can further be represented as

$$x_m(t) = s(t-\xi_m) + \sum_{p=1}^{P} n_p^c(t-\delta_{m,p}) + n_m^{uc}(t) \tag{1}$$

Note that the localized noise signals $n_p^c(t)$ are generated by some point noise sources (e.g., fan, radio and competing speakers), which are fixed or movable in the space. Some localized noise sources are spectrally stationary or have slowly time-varying spectral properties (e.g., fan), while others are spectrally highly non-stationary (e.g., competing speech and sudden noise). The non-localized noise signals $n_m^{uc}(t)$ are generally modeled as diffuse noise (e.g., wind noise in car environments) arriving from all directions in the space. In most situations, these kinds of noise sources are spectrally stationary or have slowly time-varying spectral properties.

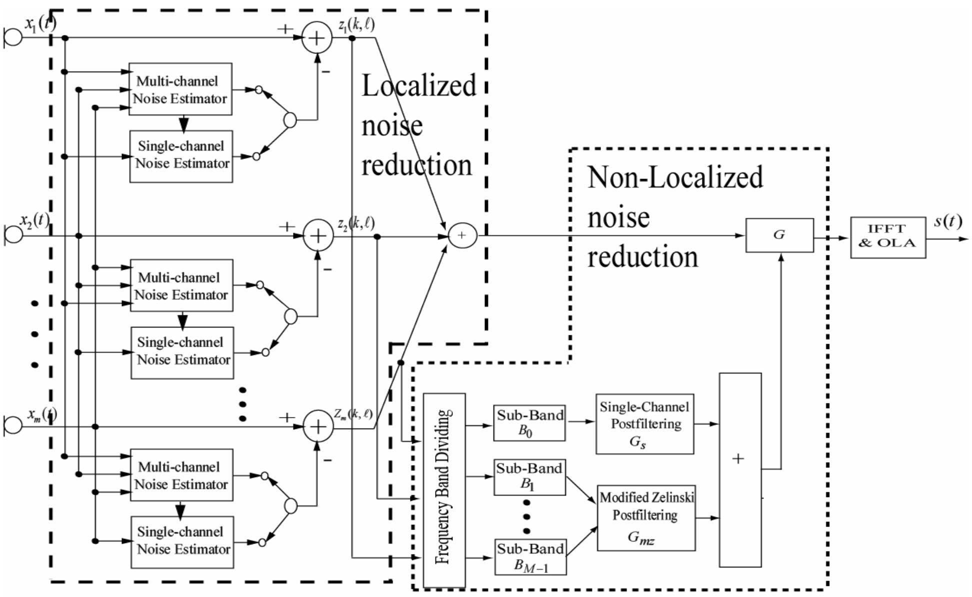

Fig. 1. Block diagram of the proposed noise reduction system.

3.3 Proposed multi-channel noise reduction system

The objective of this research is to reduce both localized and non-localized noises while keeping the desired signal distortionless. In the proposed system, spectra of localized noises are first estimated using a hybrid noise estimation technique which combines a multi-channel approach and a single-channel approach and then subtracted from the spectra of noisy signals in each channel; non-localized noise is then reduced using a hybrid post-filter which is a Wiener filter in theory. The block diagram of the proposed noise reduction algorithm is shown in Fig. 1, including localized noise reduction and non-localized noise reduction.

3.3.1 Localized noise reduction [14], [20], [21]

To deal with localized noise components, we presented a microphone-array noise reduction algorithm based on a beamforming technique. The basic idea of our algorithm is that the spectra of localized noises are first estimated and then subtracted from those of the observed noise signals.

To accurately estimate the spectra of localized noise, we proposed a hybrid noise estimation technique in a parallel structure which combines a multi-channel estimation approach and a single-channel approach. The multi-channel estimation approach was implemented using the subtractive beamformer based method since it yields much more accurate spectral estimates for localized noises at most instances. The single-channel estimation approach was implemented using a soft-decision based noise estimation technique due to its ability in estimating the spectrum of non-stationary signal. Thus, the spectrum of localized noise in the *k*-th frequency bin and ℓ-th frame, $\hat{N}^c(k,\ell)$, calculated by this hybrid estimation technique, is given by [14], [20]

$$\hat{N}^c(k,\ell) = \begin{cases} \hat{N}^c_m(k,\ell), & \text{not array nulls}, \\ \hat{N}^c_s(k,\ell), & \text{array nulls}, \end{cases} \tag{2}$$

where $\hat{N}^c_m(k,\ell)$ and $\hat{N}^c_s(k,\ell)$ are the spectral estimates determined by the multi-channel approach [20] and the single-channel approach [22], respectively. The hybrid noise estimation technique is further enhanced by integrating a *robust and accurate speech absence probability* (RA-SAP) estimator [14]. Considering the strong correlation of speech presence uncertainty between adjacent frequency bins and consecutive frames, a RA-SAP estimator is developed. This RA-SAP estimator makes full use of the high estimation accuracy of the multi-channel estimation approach. Therefore, the final estimation accuracy for localized noises is greatly enhanced by the suggested RA-SAP estimator [14], [20]. The estimated spectra of localized noises are subsequently reduced from those of the noisy observations by using the non-linear spectral subtraction. More theoretically important, note that the subtractive beamformer based multi-channel localized noise suppression algorithm is in principle a MVDR beaformer [23].

3.3.2 Non-localized noise reduction [24]

At the output of localized noise reduction, the output signal $Z_m(k,\ell)$ on m-th channel, consisting of desired signal and beamformer-processed non-localized noise $D_m(k,\ell)$, is re-formulated the time-frequency domain as

$$Z_m(k,\ell)=S_m(k,\ell)+D_m(k,\ell). \tag{3}$$

Note that the non-localized noise component $D_m(k,\ell)$ is different from the non-localized component $N_m^{uc}(t)$ at the system input, since the localized noise reduction influences the non-localized components.

To further deal with the residual non-localized noise, we propose a Wiener post-filter with a hybrid structure under the assumption of a diffuse noise field. In the high frequency region, we present a modified Zelinski post-filter which considers and utilizes the correlation of noises on different microphones to improve the noise reduction with minimum speech distortion. The implementation of the modified Zelinski post-filter consists of four steps: determine the transient frequencies (i.e., the first minimum frequency of coherence function of diffuse noise field) according to the microphone array geometry; determine the microphone pairs on which noise is mutually uncorrelated for each frequency; compute the spectral densities of the desired and noisy signals; compute the gain function of the modified Zelinski post-filter. Finally, the gain function of the modified Zelinski post-filter is derived as [24]

$$G_{mz}(k,\ell)=\frac{\dfrac{1}{|\Omega_m(k)|}\sum_{\{i,j\}\in\Omega_m(k)}\Re\left[\hat{\phi}_{Z_iZ_j}(k,\ell)\right]}{\dfrac{1}{|\Omega_m|}\sum_{\{i,j\}\in\Omega_m(k)}\left[\dfrac{1}{2}\left(\phi_{Z_iZ_i}(k,\ell)+\phi_{Z_jZ_j}(k,\ell)\right)\right]}, \tag{4}$$

where Ω_m is the microphone pair sets for *m*-th sub-band on which noises are presumably low correlated, $\Re$ is the real part operation, $\hat{\phi}_{Z_iZ_j}$ and $\hat{\phi}_{Z_iZ_i}$ are the cross- and auto- spectral densities. Note that the first two steps can be done beforehand since they are only dependent on the microphone array geometry and independent of the input signals. Thus, the computational cost will greatly be reduced.

In the low frequency region, a single-channel technique is used to estimate the Wiener filter, given by [24]

$$G_s(k,\ell)=\frac{SNR_{\text{priori}}(k,\ell)}{1+SNR_{\text{priori}}(k,\ell)}, \tag{5}$$

where $SNR_{\text{priori}}(k,\ell)$ is the *a priori* SNR which is updated in a decision-directed scheme, significantly reducing the residual “musical noise” as detailed in [5]. More theoretically important, note that the proposed hybrid post-filter is in principle a Wiener filter [24].

4. Experiments and results

We investigated the performance of the proposed noise reduction system using the speech enhancement experiments and the comprehensive speech recognition experiments. The noise reduction system was first performed on the multi-channel noisy signals, enhancing the desired speech signals. For the recognition experiments, these enhanced speech signals were further fed into the speech recognizer for recognizing the utterance. The performance improvements caused by the proposed noise reduction system (PRO-MAPF) were finally compared to those obtained by the traditional *delay-and-sum beamformer followed by Wiener post-filter* (DSWF) [11].

4.1 Speech enhancement experiments

To assess the performance of the proposed noise reduction algorithm, an equally-spaced linear array consisting of three microphones with the inter-element spacing of 10 cm was mounted in a car. The noise recordings were performed across all channels simultaneously at the sampling frequency of 12 kHz. The target signals and the interfering signals were the Chinese province/city names, uttered by one male and one female. The target speaker was placed in the front of the microphone array and the interfering speaker was placed with DOA of 60 degrees to the right. The integrated noise signals were first generated by mixing the car noise signals and the interfering signals at the same energy level. The observed noisy signals were created by adding the integrated noise signals into the target speech signals at 5 dB.

The speech enhancement results are plotted in Fig. 2. Fig. 2 (b) shows that the speech signals (北京，上海，广东，天津，重庆，内蒙古，宁夏，河南) were corrupted by both the interfering signals (广西，海南，四川，贵州，云南，西藏，香港) and the car noises. Fig. 2 (c) illustrates that the output of the DSWF is characterized by the high-level noise components (both the low-frequency car noises and the interfering signals). In contrast, the PRO-MAPF suppress almost all interfering signals and the car noises even in the regions where the speech and interfering signals are overlapped in the time-frequency domain, as shown in Fig. 2 (d). These results show that the PRO-MAPF is powerful in suppressing both localized and non-localized noise components.

4.2 Speech recognition experiments

For speech recognition, the non-localized noises were the car noises same as used in speech enhancement experiments. The speech data were selected from AURORA-2J database for training and testing. For training, 8440 sentences uttered by 55 persons were used. For testing, two sets of noise-corrupted data were generated. The first data set (*Set A*) involved the addition of the car noise recordings and 1001 test sentences at different SNRs from 0 to 20dB with 5dB step. The second data set (*Set B*) involved the addition of the multi-channel car noises and a passenger's voice which was Japanese digit /iti/ with DOA of 60 degrees to the right, across 1001 test sentences at the different SNRs same as in *Set A*. Note that *Set B* corresponds to a realistic context for a typical car condition where a passenger is speaking.

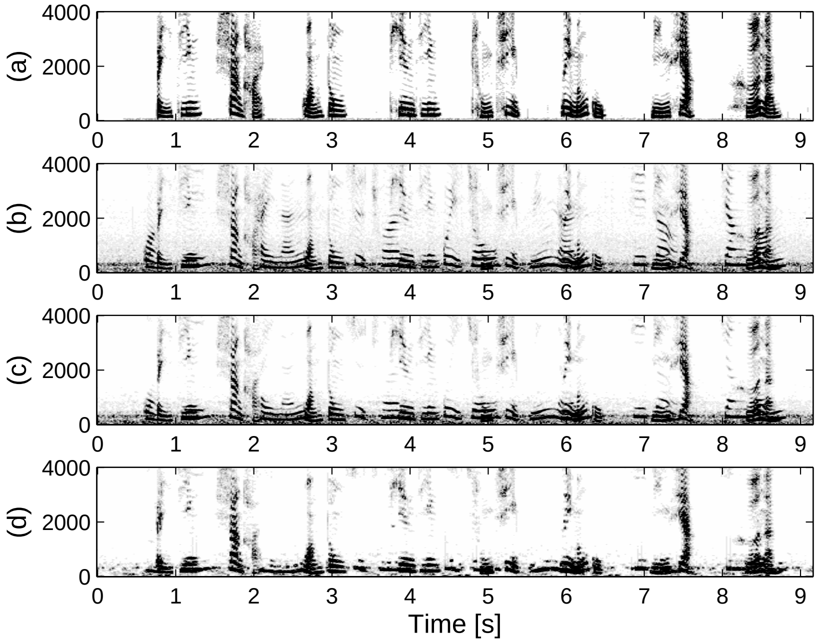

Fig. 2. Speech spectrograms. (a) Clean speech signal (北京，上海，广东，天津，重庆，内蒙古，宁夏，河南); (b) Noisy signal at the first microphone (SNR = 5 dB); (c) DSWF output; (d) PRO-MAPF output.

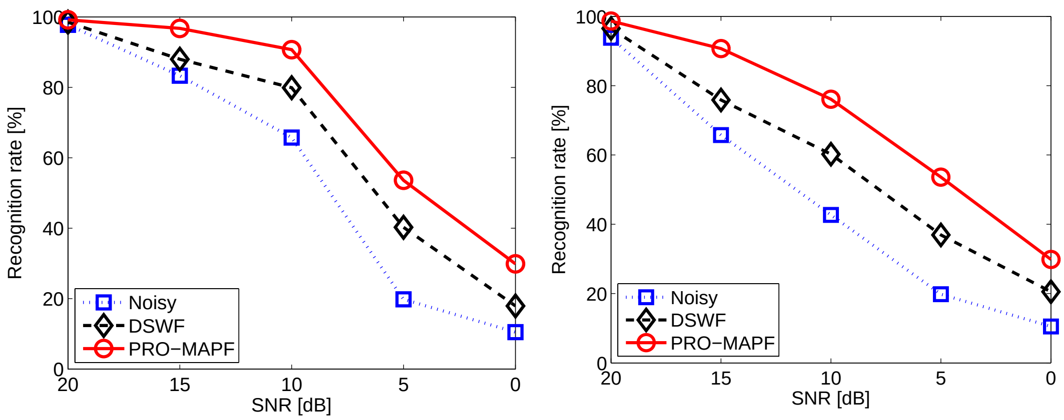

Fig. 3. Speech recognition results for the testing data *Set A* (left) and for the testing data *set B* (right)..

4.2 Experimental results

The recognition results for the noise reduction systems (DSWF and PRO-MAPF) in two noise conditions (*Set A* and *Set B*) are presented in Fig. 3.

As Fig. 3 (left) shows, for data *Set A*, all tested noise reduction algorithms provide some degree of performance improvement in speech recognition rate compared with noisy inputs. The average recognition rate improvement achieved by DSWF algorithm amounts to 6.0% with respect to noisy inputs. Whereas, the highest recognition rate improvement of about 18.6% was achieved by our PRO-MAPF. The recognition rate improvements drastically increase as the noise level increase. Moreover, in very high SNR conditions, all the tested algorithms provide just slight performance improvement compared with the noisy inputs, which is reasonable since the inputs are “clean” enough and a relatively high recognition rate is obtained in these conditions.

Concerning the recognition results for data *Set B* shown in Fig. 3 (right), we can observe that PRO-MAPF also demonstrates highest recognition rate at all SNRs. In this noise condition, the recognition rate goes down greatly for unprocessed noisy testing data. Recognition rate improvements of 11.5% and 23.2% were demonstrated by the DSWF and PRO-MAPF algorithms. The highest recognition rate of PRO-MAPF can be attributed to the fact that it is successful in dealing with both passenger’s interfering speech and diffuse car noise simultaneously with minimum speech distortion, resulting in the higher speech recognition rate.

5. Suggestions for further research

In this research, we have so far developed a noise reduction algorithm that is designed using microphone array and post-filtering in noisy environments. Its performance was evaluated in various car noise conditions and was further shown to outperform many traditional noise reduction algorithms in terms of speech recognition rate. However, the proposed noise reduction algorithm should be further improved in the following ways. (1) So far, the input microphone signals were assumed to be perfectly time-aligned in advance, that is, the desired speech signals were assumed to come from the front of the microphone array. In the practical implementation, it is necessary to take into account of the transfer function between the desired speech source and microphones. (2) Because of the small-size microphone array, improving the robustness of the noise reduction system against imperfections, such as the imperfection of microphone positions, is necessary for the real-world implementation, which is suggested as well for further research. (3) Moreover, in the real-world environments, the performance degradation of hands-free speech recognition systems is caused by not only acoustic background noise, but also reverberation and acoustic echoes. To further improve the performance of many speech applications in practical conditions, it is necessary to further deal with reverberation and acoustic echoes by combining the proposed noise reduction algorithm with other advanced dereverberation and echo cancellation techniques.

6. Reference

1. J. Benesty, S. Makino and J. Chen (eds.), Speech Enhancement, Springer, Verlag, 2005.
2. M. S. Brandstein and D. B. Ward (eds.), Microphone Arrays: Signal Processing Techniques andApplications, Springer-Verlag, 2001.
3. S. F. Boll. “Suppression of acoustic noise in speech using spectral subtraction,” *IEEE Trans. on ASSP,*, vol. ASSP-28, no. 2, pp. 113-120, 1979.
4. Y. Ephraim and D. Malah, “Speech enhancement using a minimum mean-square error short-time spectral amplitude estimator,” *IEEE Trans ASSP*, vol. 32, pp. 1109-1121, 1984.
5. -----, “Speech enhancement using a minimum mean-square error log-spectral amplitude estimator,” *IEEE Trans. on ASSP*, vol. 33, no. 2, pp. 443-445, 1985.
6. K. U. Simmer and A. Wasiljeff, “Adaptive microphone arrays for noise suppression in the frequency domain,” In *Proc. WAAC*, 1992.
7. H. Cox, R. M. Zeskind and M. M. Owen, “ Robust adaptive beamforming,” *IEEE on ASSP*, vol. ASSP-35, no. 10, pp. 1365-1375, 1987.
8. O. L. Frost, “An algorithm for linearly constrained adaptive array processing,” *In Proc. Of the IEEE*, vol. 60, no. 8, pp. 926-935, 1972.
9. L. J. Griffiths and C. W. Jim, “An alternative approach to linear constrained adaptive beamforming,” *IEEE Trans. on Antennas and Propagation*, vol. AP-30, pp. 27-34, 1982.
10. J. Bitzer, K. D. Kammeyer, and K. U. Simmer, “An alternative implementation of the superdirective beamformer,” *In Proc. IEEE WASSPAA1999,* pp. 1-4, 1999.
11. S. Fischer and K. D. Kammeyer, “Beamforming microphone arrays for speech acquisition in noisy environments,” *Speech Communication*, pp.215-227, 1996.
12. S. Gannot, D. Burshtein, and E. Weinstein “Signal enhancement using beamforming and nonstationary with applications to speech,” *IEEE Trans. on Signal Processing,* vol. 49, no. 8, pp. 1614-1626, 2001.
13. I. Cohen, “Multi-channel post-filtering in non-stationary noise environments,” *IEEE Trans. on Signal Processing*, vol. 52, no. 5, pp. 1149-1160, 2004.
14. J. Li and M. Akagi, “A noise reduction system based on hybrid noise estimation technique and post-filtering in arbitrary noise environments,” *Speech Communication*, vol. 48, no. 2, pp. 111-126, 2006.
15. R. Zelinski, “A microphone array with adaptive post-filtering for noise reduction in reverberant rooms,” *In Proc. ICASSP*, pp. 2578-2581, 1988.
16. J. Bitzer, K. U. Simmer and K. D. Kammeyer, “Multichannel noise reduction: Algorithm and Theoretical limits,” *In Proc. ESPC*, pp. 105-108, 1998.
17. ----- , “Theoretical noise reduction limits of the generalized sidelobe canceller for speech enhancement,” *In Proc. ICASSP*, pp. 100-103, 1999.
18. J. Bitzer, K. U. Simmer and K. D. Kammeyer, “Multi-microphone noise reduction by post-filter and superdirective beamformer,” *In Proc. IWAENC*, pp. 100-103, 1999.
19. I. A. McCowan and H. Bourland, “Microphone array post-filter based on noise field coherence,” *IEEE Trans. on Speech and Audio Processing*, vol. 11, pp. 709-716, 2003.
20. J. Li and M. Akagi, “Noise reduction using hybrid noise estimation techniques and post-filtering,” *In Proc. ICSLP*, pp. 2705-2708, 2004.
21. -----, X. Lu and M. Akagi, “A noise reduction system in arbitrary noise environments and its applications to speech enhancement and speech recognition,” *In Proc. ICASSP*, pp. 277-280, 2005.
22. I. Cohen and B. Berdugo, “Speech enhancement for non-stationary noise environments,” *Signal Processing*, vol. 81, pp. 2403-2418, 2001.
23. J. Li and M. Akagi, “Noise reduction based on generalized subtractive beamformer,” *Acoustical Science and Technology*, vol. 27, no. 4, 2006.
24. -----, “A hybrid microphone array post-filter in a diffuse noise field,” *In Proc. Eurospeech*, pp. 2313-2316, 2005.