COUPLED DICTIONARY TRAINING FOR EXEMPLAR-BASED SPEECH ENHANCEMENT

Deepak Baby[⋆] *Tuomas Virtanen*[†] *Tom Barker*[†] *Hugo Van hamme*[⋆]

[⋆]Department ESAT, KU Leuven, Belgium
[†]Department of Signal Processing, Tampere University of Technology, Finland
{Deepak.Baby, Hugo.Vanhamme}@esat.kuleuven.be, {Tuomas.Virtanen, Thomas.Barker}@tut.fi

ABSTRACT

In exemplar-based speech enhancement systems, lower dimensional features are preferred over the full-scale DFT features for their reduced computational complexity and the ability to better generalize for the unseen cases. But in order to obtain the Wiener-like filter for noisy DFT enhancement, the speech and noise estimates obtained in the feature space need to be mapped to the DFT space, which yield a low-rank approximation of the estimates resulting in a sub-optimal filter. This paper proposes a novel method using coupled dictionaries where the exemplars for the required feature space and the DFT space are jointly extracted and the estimates are directly obtained in the DFT space following the decomposition in the chosen feature space. Simulation experiments revealed that the proposed approach, where the activations of exemplars calculated using the Mel resolution are directly used to obtain the Wiener filter in the DFT space, results in improved signal-to-distortion ratio (SDR) when compared to the system without coupled dictionaries. To further motivate the use of coupled dictionaries, the paper also investigates the use of modulation envelope features for the exemplar-based speech enhancement.

Index Terms— Non-negative matrix factorisation, coupled dictionary training, speech enhancement, modulation envelope

1. INTRODUCTION

Speech recordings taken from realistic environments may contain added degradations along with the required speech signal which reduces its intelligibility as well as results in poor performance in speech processing tasks like automatic speech recognition (ASR), speaker recognition, hearing aids etc. The degradation can be introduced by additive background noise, reverberation, etc., and current state-of-the-art systems employ some mechanism to suppress these artefacts to enhance the speech signal for better performance and/or intelligibility.

Approaches to enhance the speech content in a noisy recording can broadly be classified as supervised and unsupervised techniques. Unsupervised techniques are based on spectral subtraction [1], Kalman filtering [2], make use of the periodic structure in speech [3] etc. Most of these approaches make stationarity assumptions on the noise, which are often invalid on practical data. For the supervised case, the speech and noise model parameters are known a-priori and some of the approaches include codebook-based algorithm [4], models based on hidden Markov models [5] etc. These approaches yield better performance when compared to the unsupervised methods as the noise model is known a-priori.

The author has done this work during his stay at Tampere University of Technology, Finland which was funded with support from the European Commission under Contract FP7-PEOPLE-2011-290000.

In this work, we investigate speech enhancement on a single-channel noisy recording in the presence of additive noise using non-negative matrix factorization (NMF). NMF based models have been successfully deployed in unsupervised [6], semi-supervised [7] and supervised [8] speech enhancement methods. This paper concentrates on a supervised speech enhancement system, where the models for speech and noise estimated from the training data are stored as *exemplars*, and the corresponding model for the noisy speech is decomposed as a sparse linear combination of speech and noise exemplars using NMF. To enhance the DFT (refers to the magnitude of the discrete-Fourier transform throughout this paper) of the noisy speech, the Wiener filter needs to be found in the DFT domain. But for features other than the DFT representation, the obtained speech and noise estimates are to be converted from the feature space to the DFT domain. This makes the DFT representation a trivial choice for the exemplars in this framework.

However, the DFT features suffer from increased computational complexity, poor separation of speech and noise especially in presence of multi-talker babble noise [9], inability to generalize well to unseen cases as it retains the speaker-dependent content for eg. the pitch, etc. which make lower dimensional features a better choice. But in such a system, the resulting speech and noise DFT estimates, which are obtained after extrapolating the estimates from the feature space to the DFT space, will have a degree-of-freedom defined by the dimensionality of the chosen feature space which is typically much less than that of the DFT space. Such a low-rank approximation results in a sub-optimal Wiener filter which cannot account for all the added noise content and yields reduced SDR.

To effectively utilise the advantages of the lower dimensional features and to overcome the issue with the low-rank approximation of the resulting estimates, we propose to use coupled dictionaries: a front-end dictionary containing the chosen features to obtain the decomposition, and a back-end dictionary containing the DFT features, the *DFT dictionary*, to reconstruct the estimates directly in the DFT domain. For a reliable reconstruction, the mapping between the corresponding exemplars in both the dictionaries should be one-to-one which is realised by extracting the corresponding exemplars of the coupled dictionaries jointly from the same piece of training data. Since in this framework, the DFT dictionary is over-complete and is coupled to the front-end dictionary, we can enforce a full-rank reconstruction of the Wiener filter in the DFT domain.

For evaluation, we chose two traditional exemplar-based systems as baselines; the first one which uses full-scale DFT as features [10], and the second which uses the Mel-integrated magnitude spectra [11], called the *Mel features*, which results in a Wiener filter with reduced degree-of-freedom. Coupled dictionaries with non-negative representation have been used to increase the spectro-temporal resolution [12]. Here it is used to map low-dimensional spectro-temporal representations to spectral representations with sufficient frequency

resolution. The simulation results obtained on the AURORA-2 database revealed that the proposed system with the Mel features as front-end results in better SDRs when compared to both the baseline systems. The paper also investigates the use of coupled dictionaries for modulation spectrogram (MS) [13] features which has recently been successfully used for blind source separation [14]. The proposed system with MS features also yields improved SDRs over the baseline systems.

2. METHOD

2.1. Compositional model for noisy speech using NMF

In the supervised setting, the exemplars for speech and noise are stored as columns in the dictionary matrices A_s and A_n, respectively. The exemplars may span multiple frames, T to capture temporal dynamics and are reshaped to a vector. The representation for the noisy utterance in the exemplar space, Ψ, is also obtained in the same manner by reshaping overlapping windows of length T [15], which is then decomposed using NMF to get the activations, X, as:

$$\Psi \approx \begin{bmatrix} A_s & A_n \end{bmatrix} \begin{bmatrix} X_s \\ X_n \end{bmatrix} = AX \quad \text{s.t.} \quad X \geq 0. \tag{1}$$

The approximation is done to minimize the Kullback-Leibler divergence between Ψ and AX with additional sparsity constraint on X [16]. The frame-wise speech and noise estimates, $\hat{s}$ and $\hat{n}$ are obtained after removing the windowing effect by adding the components belonging to overlapping windows from the estimates $A_s X_s$ and $A_n X_n$ respectively. The frame-level Wiener filter in the exemplar domain is then obtained as, $W = \hat{s} \oslash (\hat{s} + \hat{n})$, where $\oslash$ denotes the element-wise division.

2.2. Proposed method using coupled dictionaries

In the proposed system, the NMF-based decomposition is obtained in *any* additive and non-negative feature space of choice which serves as the front-end of the speech enhancement system. For simplicity, the front-end features are referred to as the *input features* and the dictionary used to obtain the NMF compositional model is denoted as the input dictionary, $A^{\text{in}} = \begin{bmatrix} A_s^{\text{in}} & A_n^{\text{in}} \end{bmatrix}$. The coupled DFT dictionary, A^{dft} serves as the output dictionary with which the speech and noise estimates are directly obtained in the DFT space using the activations obtained from the front-end, X^{in} as $A_s^{\text{dft}} X_s^{\text{in}}$ and $A_n^{\text{dft}} X_n^{\text{in}}$ respectively.

The proposed system using coupled dictionaries is summarised in Fig. 1. To obtain the dictionaries, each of the coupled exemplars for the input and the DFT dictionaries are extracted from the same piece of training data which span multiple frames of length T, followed by reshaping to form a vector. This will result in speech and noise dictionaries each for the input and DFT exemplar representations which are denoted as A_s^{in}, A_n^{in}, A_s^{dft} and A_n^{dft} respectively. The notations used to explain the test phase are: Ψ_{in} for the noisy speech represented in the input exemplar domain and $[Y]^*$ denotes the matrix obtained after removing the effect of overlapping windows in the windowed observation Y. All matrix divisions should be considered element-wise.

The proposed method thus can exploit the ability of various feature representations to separate speech from noise and can generate a Wiener filter which has full degree-of-freedom in the DFT space. In this paper, we investigate the use of the proposed approach for various features which will be discussed next.

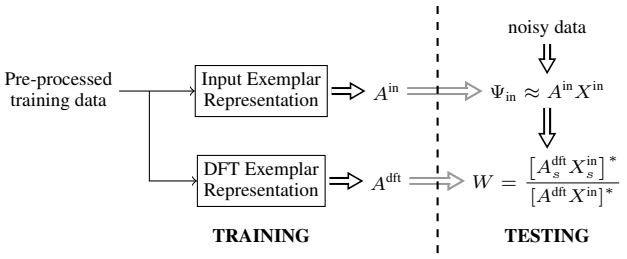


Fig. 1. *Block digram overview of the proposed system using coupled dictionaries.*

3. SYTEM DESCRIPTION

3.1. Mel and DFT baselines

For a fair evaluation, we used two baseline systems for speech enhancement; one where the exemplars are represented in the DFT domain and the second which uses the Mel-integrated spectra for which a conversion is needed to obtain the Wiener filter on the DFT resolution. The DFT exemplars consist of full-resolution magnitude spectra with K bins per frame and segments of T frames are reshaped to get $(K \cdot T)$ dimensional exemplars. The Mel exemplars are obtained by multiplying the DFT segments of size $K \times T$ using the FFT-to-Mel matrix, M, which contains the magnitude response of B Mel bands along its rows, followed by reshaping to vectors of size $(B \cdot T)$. The speech and noise exemplars thus generated are stored as A_s^{dft}, A_s^{mel}, A_n^{dft} and A_n^{mel} respectively for DFT and Mel based systems.

The DFT baseline (DFT BL) results are then obtained after finding the compositional model for noisy speech in the full-resolution DFT domain using the DFT dictionary $A^{\text{dft}} = \begin{bmatrix} A_s^{\text{dft}} & A_n^{\text{dft}} \end{bmatrix}$. The Wiener filter is directly obtained in the DFT domain using the procedure explained in Section 2.1 and is then used to enhance the noisy speech [10].

To obtain the Mel baseline (Mel BL), the speech and noise estimates, $\hat{s}^*$ and $\hat{n}^*$, are first found in the Mel domain using the steps described in Section 2.1 with the Mel dictionary, $A^{\text{mel}} = \begin{bmatrix} A_s^{\text{mel}} & A_n^{\text{mel}} \end{bmatrix}$. These estimates are then extrapolated from the B dimensional Mel space to the K dimensional DFT domain using the transpose of the DFT-to-Mel matrix and the corresponding Wiener filter is then obtained, using element-wise division, as [11] :

$$W^* = \frac{M^T \hat{s}^*}{M^T \hat{s}^* + M^T \hat{n}^*}. \tag{2}$$

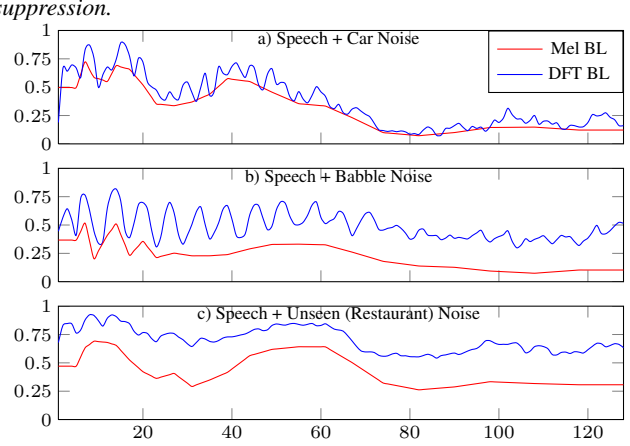
Since M contains triangular shaped filter-banks, this extrapolation is the same as the piece-wise linear interpolation between B points (the Mel filter-bank central frequencies) spread across the 1 to K frequency bins. The resulting filters always fall in the B-dimensional subspace defined by the columns of M^T which cannot account for all the added noise content along the K dimensional DFT space. The enhanced speech obtained after applying this filter on the noisy DFT thus will result in a sub-optimal noise suppression.

3.2. Proposed system with Mel features

The motivation for using Mel features as the input features for the proposed system are:

1. Poor separation capability of DFT Exemplars: In the DFT based system, it has been noticed that many of the speech exemplars are activated for babble noise because of the similarity between the babble noise and speech exemplars [9], which in turn results in a Wiener

Fig. 2. *Filter coefficients obtained for an arbitrary frame containing speech with SNR 0dB as a function of the frequency bins. The color coding is the same for all the figures. a) For car noise which is present in the training set, the DFT baseline filter better captures the formant peaks and valleys when compared to that of the Mel baseline. b) For babble noise, speech exemplars are also activated to model the noise which results in poorer denoising. c) For the unseen restaurant noise, due to poorer modelling, the DFT exemplars result in retaining most of the noise content whereas the filter coefficients of the Mel baseline are quite smaller and result in better noise suppression.*



filter which retains most of the babble noise content (ref. Fig. 2b). Similar situations were observed for unseen noise cases (ref. Fig. 2c) because the DFT exemplars lead to accurate representation of training noise cases which results in poor modelling of unseen noise cases. As a result, NMF will pick speech exemplars also to model the unseen noise content which results in poorer noise suppression. On the other hand, Mel exemplars are found to be better able to differentiate speech from the babble noise and result in better separation (ref. Fig. 2b). The Mel features also have much lower dimensionality when compared to the DFT exemplars and reduces the risk of overfitting to seen noise cases.

2. Piece-wise linear approximation of filter coefficients: As discussed before, even though the Mel exemplars lead to better separation, the low-rank approximation of the coefficients in the DFT domain fails to capture the detailed structure of the underlying speech which can be seen in Fig. 2a. It was observed that both DFT and Mel exemplars yield almost the same separation after NMF, but with the latter resulting in a filter with lesser peaks and valleys, which is essential to capture the formant positions and pitch, yields smaller SDRs.

For evaluation, the Mel and the coupled DFT dictionaries are jointly extracted first. The noisy data is converted to the Mel exemplar representation and is then decomposed using NMF with the Mel dictionary. The speech and noise estimates are then obtained directly in the DFT domain using the coupled DFT dictionary as shown in Fig. 1. This system is referred to as the *Coupled Mel* system.

3.3. Proposed system with MS features

The MS representation was proposed as part of a computational model for human hearing which relies on the low frequency amplitude modulations within various frequency bands [17]. The MS representation for acoustical data is obtained using the procedure explained in [14]. For the NMF based system, this 3D representation of size $b \times T \times B$, is converted to a 2D representation by stacking the truncated spectra belonging to different channels to get a matrix of size $(B \cdot b) \times T$, where b, T and B are the number of truncated bins, number of frames in the MS and number of filter banks used to obtain the MS representation, respectively. Thus for every frame, this representation has $(B \cdot b)$ dimensional features which are referred to as *MS features*.

For evaluation, the Wiener filter is obtained using the procedure depicted in Fig. 1 with MS features as the input features. Since phase information in the MS is disregarded (non-negativity), signal reconstruction is not unique. For instance, any circular temporal shift (modulo the window length) of the DFT will lead to the same MS exemplar. However, this ambiguity can be reduced greatly if the magnitude spectrogram is sampled fast enough using smaller hop sizes [18]. Even though using smaller hop sizes to obtain the MS features lead to temporal oversampling, it is found to be useful for making the mapping nearly one-to-one and make it useful for the proposed setup. This system is referred to as the *Coupled MS* system. To our knowledge, this is the first use of MS features for exemplar-based speech enhancement purpose.

4. EVALUATION EXPERIMENTS

4.1. Experimental setup

The experiments were conducted on the Test sets A and B of the AURORA-2 database which contains utterances of digits from '0-9' and 'oh'. The training set contains 8440 clean speech utterances and 6768 noisy utterances with four different additive noise types (subway, babble, car and exhibition hall). Test set A contains six subsets of noisy utterances corrupted with each of the noise types present in the training data with varying SNRs (-5,0,5,10,15 and 20 dB), along with the corresponding clean utterances. Test set B also has the same number of subsets but for four other noise types (restaurant, train station,street and airport). Thus both test sets contain 24 noisy subsets each along with the corresponding clean utterances. For evaluation, we selected a random subset of 100 utterances from every noise type and the SDR improvements obtained are presented.

The coupled dictionaries were obtained for a temporal context which spans $T = 30$ frames. The noise data required to obtain the noise exemplars were obtained from the noisy utterances using the procedure described in [16]. Both the DFT and Mel exemplars were obtained using a window length of 25 ms and a hop size of 10 ms. $B = 23$ channels were used for the Mel integration of $K = 128$ bins magnitude spectrogram. To obtain the MS dictionary, equivalent rectangular bandwidth filter banks, using Slaney's toolbox [19], with $B = 23$ channels were used to get the band-limited signals. The low-pass filter used had a cut-off frequency of 30 Hz and the magnitude spectra of each of the modulation envelopes were obtained with a window length of 64 ms and hop size of 10 ms. With the sampling frequency of 8000 Hz and FFT with 128 bins within the Nyquist frequency, each of the spectra were truncated to the lowest $b = 5$ bins which were then stacked and reshaped to get the MS exemplar.

The simulation experiments were conducted for coupled dictionaries of size 10000 exemplars each for speech and noise, resulting in dictionary sizes 690×10000, 3840×10000 and 3450×10000 each for Mel, DFT and MS representations respectively. The decomposition was carried out with 700 multiplicative update iterations for the NMF with sparsity constraint. The Mel and the DFT setup used a sparsity penalty of 1.5 for speech and 0.5 for noise exemplars as suggested in [16]. The NMF on the MS exemplars used a sparsity penalty of 1.75 and 0.75 for speech and noise exemplars respectively

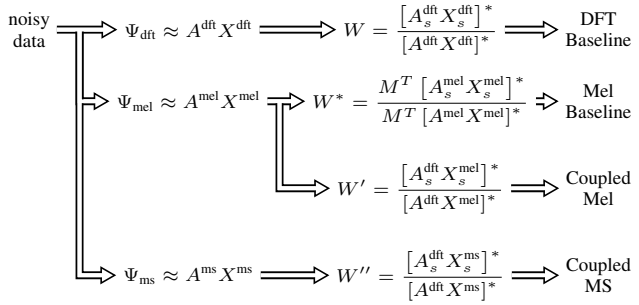


Fig. 3. *Block diagram summary of the processing chains which obtain the Wiener filters for the noisy DFT enhancement for the two baselines and the two proposed coupled systems.*

obtained after a grid search in the range $[0, 3]$ on a subset of 100 utterances in the test set chosen from the complement of the the subset used for evaluation. The iterations were accelerated with GPUs using the MATLAB parallel computing toolbox. The entire processing chain for testing is summarised in Fig. 3 with $\Psi_{\{\cdot\}}$ representing the corresponding exemplar representation in various domains.

From the enhanced magnitude DFT, the complex spectrogram was obtained by using the phase obtained from the noisy speech and the speech signal in the time domain was obtained using the overlap-add method. The resulting enhanced speech data for various systems were compared using the SDR measure in dB using the BSS evaluation toolbox [20].

4.2. Results and discussion

The SDR improvements in dB obtained for the two baseline systems and the proposed algorithm are shown in Fig. 4. The results obtained for test set A, which contain seen noise types are given in Fig. 4a. It can be seen that both the baseline systems yield almost the same performance for all the input SNRs. Notice that, even though the Mel BL involves a low-rank approximation of the estimates, its performance is comparable to that of the DFT BL. This can be attributed to the ability of the Mel features to better separate speech from noise when compared to the DFT features. It is also noticed that, for a system which uses 10000 speech and 4000 noise exemplars each, the performance of the DFT baseline system falls below that of the Mel baseline system, because the higher dimension of the DFT representation demands more exemplars for over-completeness and proper modelling of seen noise.

For the proposed setup in test set A, a clear SDR improvement can be seen for both the Coupled Mel and the Coupled MS systems. As discussed above, the Mel features yield a better separation and with the coupled dictionary approach, a "better" Wiener filter is obtained which yields improved SDR. Notice that both the Mel baseline and Coupled Mel systems use the same decomposition in the front-end and results in the same activations. But with the baseline system undergoing a low-rank approximation, it yields lower SDRs when compared to the proposed system where the activations obtained are applied to the coupled DFT dictionary to obtain the estimates. The MS features also perform equally well and as the input SNR increases, impressive SDR improvements are achieved by the Coupled MS system.

For test set B (ref. Fig. 4b), the DFT baseline performs far inferior when compared to the Mel baseline system, unlike test set A. This is because the DFT features yield more accurate representation of the seen noise cases which makes it poor in generalizing to the unseen cases as discussed. The SDRs given by the Mel baseline

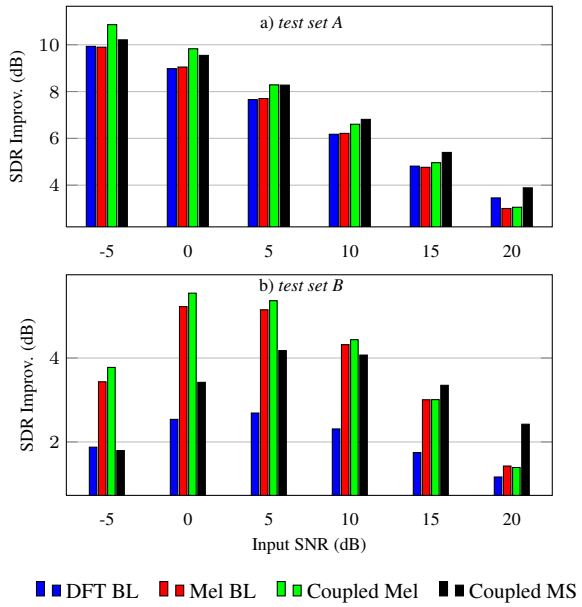


Fig. 4. *SDR Improvements in dB as a function of input SNRs to evaluate and compare the baseline systems with the proposed coupled systems for seen (test set A) and unseen (test set B) noise cases in the AURORA-2 database. Legends are same for both the plots.*

system even in presence of the low-rank approximation suggest that the speech and noise separation achieved by the Mel features is far superior to that of the DFT features.

For the proposed setup in test set B, the Coupled Mel setup result in improved SDRs for all input SNRs except the 20 dB case when compared to both the baseline systems. It can also be seen that, the Coupled MS setup fails to beat the baseline systems for lower input SNRs. This can also be attributed to the increased dimensionality of the MS features when compared to the Mel features which results in overfitting to the seen cases. But as the input SNR increases, MS features yield improvements and especially with input SNR 20 dB, the SDR improvement is more than 1 dB.

5. CONCLUSION AND FUTURE WORK

In this work, we presented a novel method to address the low-rank approximation of the estimates obtained in an exemplar-based speech enhancement system which uses features other than the full-scale DFT features. It has also been shown that the proposed system with coupled dictionaries can be made useful for features where a direct conversion from the feature space to the DFT space is not possible; for e.g. the MS features. The simulation results revealed that the proposed system yields better performance when compared to the baseline systems in terms of SDR. This is the first use of modulation envelope features for the exemplar-based speech enhancement purpose. The paper also presented a comparative study between the speech and noise separation capabilities of various feature representations.

The future work is to further address the dimensionality issues and the overfitting to seen noise cases. Another focus is to investigate the use of coupled dictionaries for other applications and with different feature representations. Further uses of the MS features together with coupled dictionaries for other speech processing related applications also are to be investigated.

6. REFERENCES

[1] S. Boll, "Suppression of acoustic noise in speech using spectral subtraction," *Acoustics, Speech and Signal Processing, IEEE Transactions on*, vol. 27, no. 2, pp. 113–120, 1979.

[2] V. Grancharov, J. Samuelsson, and Bastiaan Kleijn, "On causal algorithms for speech enhancement," *Audio, Speech, and Language Processing, IEEE Transactions on*, vol. 14, no. 3, pp. 764–773, 2006.

[3] J.R. Jensen, J. Benesty, M.G. Christensen, and S.H. Jensen, "Enhancement of single-channel periodic signals in the time-domain," *Audio, Speech, and Language Processing, IEEE Transactions on*, vol. 20, no. 7, pp. 1948–1963, 2012.

[4] T.V. Sreenivas and P. Kirnapure, "Codebook constrained Wiener filtering for speech enhancement," *Speech and Audio Processing, IEEE Transactions on*, vol. 4, no. 5, pp. 383–389, 1996.

[5] Y. Ephraim, "A Bayesian estimation approach for speech enhancement using hidden Markov models," *Signal Processing, IEEE Transactions on*, vol. 40, no. 4, pp. 725–735, 1992.

[6] T. Virtanen, "Monaural sound source separation by nonnegative matrix factorization with temporal continuity and sparseness criteria," *Audio, Speech, and Language Processing, IEEE Transactions on*, vol. 15, no. 3, pp. 1066–1074, 2007.

[7] G.J. Mysore and P. Smaragdis, "A non-negative approach to semi-supervised separation of speech from noise with the use of temporal dynamics," in *Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on*, 2011, pp. 17–20.

[8] J.F. Gemmeke, T. Virtanen, and A. Hurmalainen, "Exemplar-based sparse representations for noise robust automatic speech recognition," *Audio, Speech, and Language Processing, IEEE Transactions on*, vol. 19, no. 7, pp. 2067–2080, 2011.

[9] N. Mohammadiha, P. Smaragdis, and A. Leijon, "Supervised and unsupervised speech enhancement using nonnegative matrix factorization," *Audio, Speech, and Language Processing, IEEE Transactions on*, vol. 21, no. 10, pp. 2140–2151, 2013.

[10] T. Virtanen, R. Singh, and B. Raj, Eds., *Techniques for Noise Robustness in Automatic Speech Recognition*, Wiley, 2012.

[11] Jort F. Gemmeke, Tuomas Virtanen, and Antti Hurmalainen, "Exemplar-based speech enhancement and its application to noise-robust automatic speech recognition," in *International Workshop on Machine Listening in Multisource Environments*, 2011, pp. 1–6.

[12] Juhan Nam, Gautham J. Mysore, Joachim Ganseman, Kyogu Lee, and Jonathan S. Abel, "A super-resolution spectrogram using coupled PLCA," in *INTERSPEECH*, Makuhari, Japan, 2010, pp. 1696–1699.

[13] S. Greenberg and B.E.D. Kingsbury, "The modulation spectrogram: in pursuit of an invariant representation of speech," in *Acoustics, Speech, and Signal Processing, 1997. ICASSP-97., 1997 IEEE International Conference on*, 1997, vol. 3, pp. 1647–1650 vol.3.

[14] T. Barker and T. Virtanen, "Non-negative tensor factorization of modulation spectrograms for monaural sound source separation," in *Proc. INTERSPEECH*, 2013.

[15] J. F. Gemmeke and T. Virtanen, "Noise robust exemplar-based connected digit recognition," in *Proc. ICASSP*, March 2010, pp. 4546 –4549.

[16] J. F. Gemmeke, T. Virtanen, and A. Hurmalainen, "Exemplar-based sparse representations for noise robust automatic speech recognition," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 19, no. 7, pp. 2067 –2080, Sept. 2011.

[17] C. Plack, *The Sense of Hearing*, Lawrence Erlbaum Associates, ch. 10, 2005.

[18] D. Griffin and J.S. Lim, "Signal estimation from modified short-time fourier transform," in *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP '83.*, 1983, vol. 8, pp. 804–807.

[19] Malcolm Slaney, "Auditory toolbox," 1994.

[20] E. Vincent, R. Gribonval, and C. Fevotte, "Performance measurement in blind audio source separation," *Audio, Speech, and Language Processing, IEEE Transactions on*, vol. 14, no. 4, pp. 1462–1469, 2006.