Blind Deconvolution for Sparse Acoustic System

Jian Guan[1], Xuan Wang[1], Wenwu Wang[2], and Zongxia Xie[3]

[1] Harbin Institute of Technology,Shenzhen, 518055, China
[j.guan,wangxuan]@cs.hitsz.edu.cn
[2] University of Surrey, Guildford, GU2 7XH, UK
w.wang@surrey.ac.uk
[3] Tianjin University, Tianjin, 30072, China
caddiexie@hotmail.com

Abstract. Blind speech deconvolution aims to estimate both the original speech source and the impulse response of the acoustic channel from the convolutive signal observed by an acoustic sensor. It is an ill-posed and underdetermined problem, since the number of unknown variables is greater than the number of observations. To address this problem, prior information or extra assumption is required to limit the range of possible solutions. In this paper, we consider a sparse acoustic system, where the impulse response of the acoustic channel from the source to the sensor is assumed to be sparse. The blind deconvolution problem is then addressed as an optimisation problem where the cost function to be optimised is defined as a combination of a data fidelity term and a regularisation term. The data fidelity term is used to control the fitting error between the observed signal and the signal reconstructed from the estimated acoustic channel and the estimated speech source. The regularisation term is formed as the L1-norm constraint of the acoustic impulse response which enforces the sparsity of the acoustic channel. The cost function is optimised with a two-stage alternating process iterating between the stage of estimating the acoustic channel with an L1-norm constrained least squares method and the stage of estimating the speech source with a least squares based optimisation method. The performance of the proposed algorithm is demonstrated with an application example of blind dereverberation of early reverberant speech.

Keywords: Blind deconvolution, speech dereverberation, acoustic channel estimation

1 Introduction

Blind deconvolution aims to estimate the unknown source signal and unknown channel given only the information about their convolution. It is an ill-posed problem [1], which often requires prior information, such as sparsity of the signals [2] [6] in order to limit the possible solution set.

In our recent work, we have focused on the blind deconvolution for the single-input and single-output (SISO) acoustic system, where both the speech signal and room impulse response (RIR) are estimated from the reverberant speech. In order to solve the problem, we take into account of the prior informations provided by the RIR of an acoustic system. In some scenarios, RIR may be considered as sparse when the acoustic system has a relatively low reflection level, which is our focus here. By exploiting the sparsity of RIR, a blind deconvolution method is proposed by imposing L1-norm constraint on the RIR of the sparse acoustic system. An alternating minimisation method is proposed to solve the blind deconvolution problem, with two steps: $x - step$ and $h - step$, where the speech signal $\boldsymbol{x}$ and RIR $\boldsymbol{h}$ are estimated in an alternating manner.

2 Proposed Blind Deconvolution Method

2.1 Problem Formulation

The observation $\boldsymbol{y} \in \mathbb{R}^{N+L-1}$ is the reverberant speech obtained by convolving the source speech signal $\boldsymbol{x} \in \mathbb{R}^N$ with RIR $\boldsymbol{h} \in \mathbb{R}^L$ as

$$\boldsymbol{y} = \boldsymbol{x} * \boldsymbol{h}, \tag{1}$$

where * denotes the convolution operation. Here $\boldsymbol{y}$ is given as input for the estimation of unknown $\boldsymbol{h}$ and $\boldsymbol{x}$. However, we usually need some prior information to reduce the solution space for this blind deconvolution problem.

In this work, we assume $\boldsymbol{h}$ is sparse, and impose a sparse constraint on $\boldsymbol{h}$. Hence our proposed blind deconvolution model is formulated as follows

$$F(\boldsymbol{x}, \boldsymbol{h}) = \|\boldsymbol{x} * \boldsymbol{h} - \boldsymbol{y}\|_2^2 + \lambda \|\boldsymbol{h}\|_1 + r(\boldsymbol{x}) \tag{2}$$

where the L2-norm is the data fidelity term, the regularisation $r(\boldsymbol{x})$ is an indicator function accounts for the dynamic range of $\boldsymbol{x}$, and λ is the penalty parameter of the L1-norm based sparse constraint.

In matrix form, the proposed model can be reformulated as

$$\begin{aligned} F(\boldsymbol{x}, \boldsymbol{h}) &= \|\boldsymbol{y} - \boldsymbol{X}\boldsymbol{h}\|_2^2 + \lambda \|\boldsymbol{h}\|_1 + r(\boldsymbol{x}) \\ &= \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}\|_2^2 + \lambda \|\boldsymbol{h}\|_1 + r(\boldsymbol{x}) \end{aligned} \tag{3}$$

where $\boldsymbol{X} \in \mathbb{R}^{(N+L-1)\times L}$ and $\boldsymbol{H} \in \mathbb{R}^{(N+L-1)\times N}$ are the linear convolution matrices constructed from $\boldsymbol{x}$ and $\boldsymbol{h}$ respectively.

2.2 Proposed Method

An alternating minimisation method is proposed to optimise (3), which includes two steps: $h - step$ and $x - step$, where $\boldsymbol{h}$ and $\boldsymbol{x}$ are estimated iteratively, by fixing one, and updating the other.

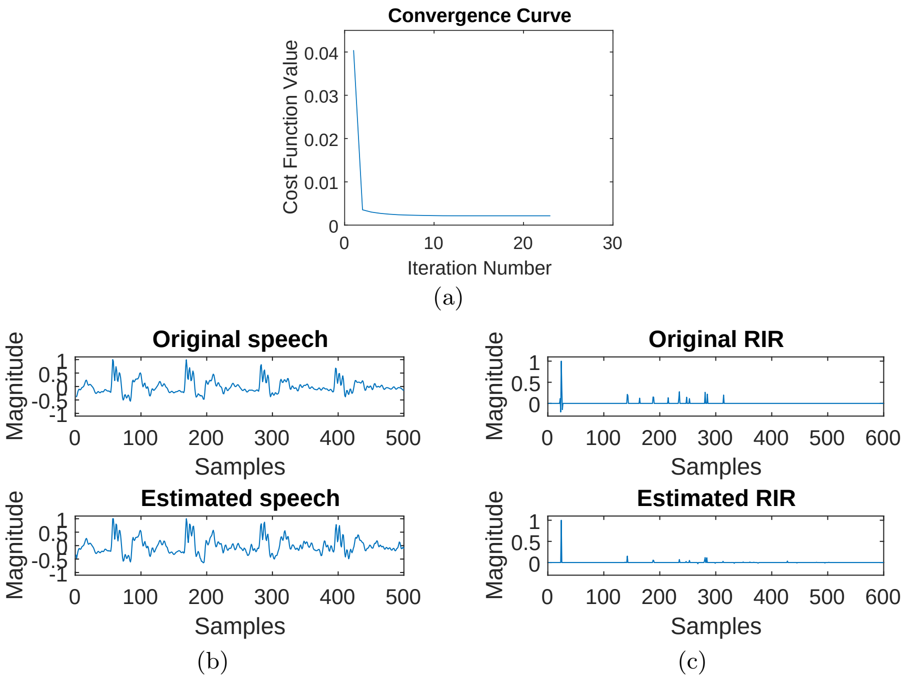

Fig. 1: Results illustration of the proposed method for blind dereverberation.

$\boldsymbol{h} - \boldsymbol{step}$ In this step, we suppose $\boldsymbol{x}$ is known, $\boldsymbol{h}$ can be updated by optimising the following problem

$$\boldsymbol{h}^{(k+1)} = \underset{\boldsymbol{h}}{\operatorname{argmin}} \|\boldsymbol{y} - \boldsymbol{X}^{(k)}\boldsymbol{h}\|_2^2 + \lambda\|\boldsymbol{h}\|_1 \tag{4}$$

Here we use the CVX toolbox [5] to update $\boldsymbol{h}$ at each iteration, once we obtain $\boldsymbol{h}^{(k+1)}$, the convolution matrix $\boldsymbol{H}^{(k+1)}$ can then be constructed, and used to update $\boldsymbol{x}$ at this iteration.

$\boldsymbol{x} - \boldsymbol{step}$ In this step, $\boldsymbol{x}$ can be updated by minimising the following function

$$\boldsymbol{x}^{(k+1)} = \underset{\boldsymbol{x}}{\operatorname{argmin}} \|\boldsymbol{y} - \boldsymbol{H}^{(k+1)}\boldsymbol{x}\|_2^2 + r(\boldsymbol{x}), \tag{5}$$

where a variable metric forward-backward method [3, 4] is applied to optimise Equation (5).

3 Experiments and Results

In this section, we carry out an experiment to illustrate the performance of our proposed method for blind deconvolution of the acoustic system. The observation

$\boldsymbol{y}$ is the convolution of source speech signal $\boldsymbol{x}$ and RIR $\boldsymbol{h}$, where $\boldsymbol{h}$ is generated by room image model, we selected the early reflections and remove the late reverberations of the generated RIR. The length $\boldsymbol{x}$ is 500 samples, and the length of $\boldsymbol{h}$ is 600 samples. Here $\boldsymbol{x}$ is initialised with the source signal added with 0 dB white Gaussian noise.

As can be seen from Figure 1, the proposed method can converge at a small number of iterations, and the estimated signal is similar to the original signal. The estimated RIR resembles most impulses of the original RIR. Our future work will consider the modelling of both the early reflections and late reverberations of RIR.

4 Conclusion

In this paper, we proposed a blind deconvolution method for sparse acoustic system, where the sparsity of RIR is used as the additional prior information. Experiment shows that our method can recover both the signal and RIR.

5 Acknowledgement

The work was conducted when J. Guan was visiting the Centre for Vision, Speech and Signal Processing (CVSSP), the University of Surrey. This work was supported by the Engineering and Physical Sciences Research Council (EPSRC) Grant number EP/K014307/2 and the MOD University Defence Research Collaboration in Signal Processing, and partially supported by International Exchange and Cooperation Foundation of Shenzhen City, China (No.GJHZ20150312 114149569).

References

1. Benichoux, A., Vincent, E., Gribonval, R.: A fundamental pitfall in blind deconvolution with sparse and shift-invariant priors. In: Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP) (2013)
2. Chi, Y.: Guaranteed blind sparse spikes deconvolution via lifting and convex optimization. IEEE Journal of Selected Topics in Signal Processing 10(4), 782–794 (2016)
3. Chouzenoux, E., Pesquet, J.C., Repetti, A.: Variable metric forward–backward algorithm for minimizing the sum of a differentiable function and a convex function. Journal of Optimization Theory and Applications 162(1), 107–132 (2014)
4. Chouzenoux, E., Pesquet, J.C., Repetti, A.: A block coordinate variable metric forward–backward algorithm. Journal of Global Optimization pp. 1–29 (2016)
5. Grant, M., Boyd, S., Grant, M., Boyd, S., Blondel, V., Boyd, S., Kimura, H.: Cvx: Matlab software for disciplined convex programming, version 2.1. Recent Advances in Learning and Control pp. 95–110
6. Repetti, A., Pham, M.Q., Duval, L., Chouzenoux, E., Pesquet, J.C.: Euclid in a taxicab: Sparse blind deconvolution with smoothed regularization. IEEE Signal Processing Letters 22(5), 539–543 (2015)