Audio Super-Resolution Using Analysis Dictionary Learning

Jing Dong, Wenwu Wang
Centre for Vision, Speech and Signal Processing
University of Surrey, Guildford GU 7XH
United Kingdom
Email: {j.dong, w.wang}@surrey.ac.uk

Jonathon Chambers
School of Electrical and Electronic Engineering
Newcastle University, Newcastle upon Tyne NE1 7RU
United Kingdom
Email: Jonathon.Chambers@newcastle.ac.uk

***Abstract*—Super-resolution is an important problem in signal processing. It aims to reconstruct a high-resolution (HR) signal from a low-resolution (LR) input. We consider the super-resolution problem for audio signals in the time-frequency domain and propose a method using analysis dictionary learning. The input to our proposed method is the LR spectrogram matrix of an audio signal, where some rows corresponding to high-frequency information are lost. First, an analysis dictionary is learned from the spectrogram of some related audio signals. The learned dictionary is then applied in an ℓ_1-norm regularization term for the reconstruction of the HR spectrogram. Experimental results with piano signals demonstrate the advantage of the learned dictionaries in reconstructing HR spectrograms.**

***Index Terms*—Sparse represnetation; analysis dictionary learning; super-resolution**

I. Introduction

The goal of super-resolution techniques in signal processing is to enhance the resolution of signals. The super-resolution for images is one of the most active research areas [1], [2] in image processing. However, little work has been done for the super-resolution for audio signals. In the field of audio signal processing, the super-resolution problem can be cast to the problem of reconstructing high-frequency portions of audio signals [3], leading to higher quality audio signals for an improved listening experience. In the time-frequency domain, the LR spectrogram $\mathbf{Y}_L \in \mathbb{R}^{k\times s}$ can be regarded as the remaining part of the HR spectrogram $\mathbf{Y}_H \in \mathbb{R}^{m\times s}$ after the removal of the top $m-k$ high-frequency bins of $\mathbf{Y}_H$, i.e.

$$\mathbf{Y}_L = \mathbf{A}\mathbf{Y}_H, \tag{1}$$

where $\mathbf{A} \in \mathbb{R}^{k\times m}$ is the mapping matrix for removing the first $m-k$ rows of $\mathbf{Y}_H$. The super-resolution method proposed in [3] is based on the well-known sparse synthesis model [4]. This sparse model assumes that a signal can be represented as a linear combination of several atoms (columns) of a synthesis dictionary. In [3], the dictionaries that can represent the HR and the LR spectrograms are assumed to be known and the HR spectrogram is reconstructed with its synthesis dictionary and the sparse coefficients obtained by approximating a pursuit algorithm based on the LR dictionary.

The analysis model for sparse representation, as a counterpart of the synthesis model, has drawn much attention recently [5], [6]. For a signal $\mathbf{y} \in \mathbb{R}^m$, this model assumes that the product of $\mathbf{\Omega} \in \mathbb{R}^{p\times m}$ and $\mathbf{y}$ is sparse, i.e. $\mathbf{x} = \mathbf{\Omega}\mathbf{y}$ with $\|\mathbf{x}\|_0 = p - l$, where the ℓ_0-norm $\|\cdot\|_0$ counts the number of non-zero elements of its argument and $0 \le l \le p$ is the co-sparsity of $\mathbf{y}$. The matrix $\mathbf{\Omega}$ is usually referred to as an analysis dictionary [7], with each row of $\mathbf{\Omega}$ being an atom. The vector $\mathbf{x} \in \mathbb{R}^p$ is the analysis representation of the signal $\mathbf{y}$ with respect to $\mathbf{\Omega}$. In this model, the analysis dictionary $\mathbf{\Omega}$ plays an important role in the analysis representation of the signal $\mathbf{y}$, and the dictionaries learned from a set of training signals show some advantages compared with pre-defined dictionaries [7]. Thus, some algorithms for learning an analysis dictionary have been proposed [7], [8], [9].

In this work, we propose a new method for the audio super-resolution problem using analysis dictionary learning. This method is different from the approach proposed in [3] since it is based on the analysis model. Besides, the dictionary used is learned from training data, rather than pre-defined as in [3]. Our proposed method is introduced in Section II. Simulation results for the super-resolution of piano signals are presented is Section III. Section IV concludes the paper.

II. The Proposed Method

Based on the sparse analysis model, we assume that analysis representation of the HR spectrogram to be reconstructed is sparse with respect to an analysis dictionary. Since the analysis dictionary learned from some related data usually has the potential to adapt to a signal better as compared with a pre-defined dictionary [7], the first stage of our proposed method is to learn an analysis dictionary from the spectrogram of some HR audio signals. In this stage, the Analysis SimCO algorithm [9] is applied. After that, the HR spectrogram can be reconstructed based on the assumption that it is sparse with respect to the learned dictionary. This is referred to as the spectrogram reconstruction stage. The block diagram of our proposed method is presented in Fig. 1, where $\mathbf{Y} \in \mathbb{R}^{m\times n}$ denotes the training data for the analysis dictionary learning. The details of the the analysis dictionary learning and the spectrogram reconstruction stages will be presented in detail in the following subsections.

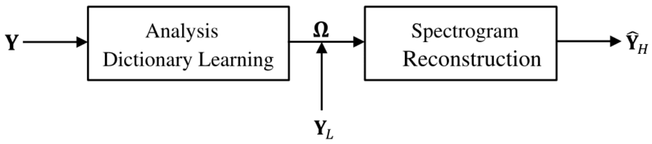

Fig. 1. Block diagram of our proposed method.

A. Analysis Dictionary Learning Stage

Given a set of training data $\mathbf{Y}$, the analysis dictionary learning problem can be written as [10]

$$\{\mathbf{\Omega}^*, \mathbf{X}^*\} = \arg\min_{\{\mathbf{\Omega},\mathbf{X}\}} \|\mathbf{X} - \mathbf{\Omega}\mathbf{Y}\|_F^2 \quad \text{s.t.} \quad \|\mathbf{X}_{:,i}\|_0 = p - l, \ \forall i. \tag{2}$$

This is a general formulation without any additional constraint on $\mathbf{\Omega}$ apart from the co-sparsity constraints $\|\mathbf{X}_{:,i}\|_0 = p-l, \ \forall i$. However, this formulation has ambiguities caused by scaling. In one case, when the training data $\mathbf{Y}$ admits exact sparse representations, there exists a dictionary $\mathbf{\Omega}$ with which the analysis representations of $\mathbf{Y}$, i.e. $\mathbf{X} = \mathbf{\Omega}\mathbf{Y}$, satisfy the co-sparsity constraints. If the dictionary $\mathbf{\Omega}$ is scaled by multiplying a scalar $c \in \mathbb{R}$, the corresponding representations $c \cdot \mathbf{X} = c \cdot \mathbf{\Omega}\mathbf{Y}$ will also satisfy the constraints. Thus, the problem (2) has infinite optimal solutions $c \cdot \mathbf{\Omega}$ and $c \cdot \mathbf{X}$. This may introduce difficulty in optimization. In the other case, if the data $\mathbf{Y}$ admits approximation representations and $\|\mathbf{X} - \mathbf{\Omega}\mathbf{Y}\|_F^2 = \delta$, the value of the cost function with scaled $\mathbf{X}$ and $\mathbf{\Omega}$, i.e. $\|\mathbf{X} - \mathbf{\Omega}\mathbf{Y}\|_F^2 = c^2 \cdot \delta$, can be arbitrarily small. In other words, the cost function is unbounded from below, which makes it impossible to find an optimal solution. In addition, (2) has trivial solutions where $\mathbf{\Omega}$ contains all-zero rows.

In order to avoid these problems, the unit ℓ_2-norm constraints on the rows of $\mathbf{\Omega}$ are applied, leading to the following formulation of the Analysis SimCO algorithm [9]

$$\begin{aligned} \{\mathbf{\Omega}^*, \mathbf{X}^*\} = \arg\min_{\{\mathbf{\Omega},\mathbf{X}\}} & \|\mathbf{X} - \mathbf{\Omega}\mathbf{Y}\|_F^2 \\ \text{s.t.} \quad & \|\mathbf{X}_{:,i}\|_0 = p - l, \ \forall i \\ & \|\mathbf{\Omega}_{j,:}\|_2 = 1, \ \forall j, \end{aligned} \tag{3}$$

where $\mathbf{X}_{:,i}$ is the ith column of $\mathbf{X} \in \mathbb{R}^{p \times n}$. The ℓ_0 quasi-norm $\|\cdot\|_0$ counts the number of non-zeros of a vector and l is the cosparsity. The Analysis SimCO algorithm alternates between two stages: analysis sparse coding and dictionary update. The first stage calculates $\mathbf{X}$ for the given dictionary $\mathbf{\Omega}$. In the dictionary update stage, $\mathbf{\Omega}$ is updated assuming known and fixed $\mathbf{X}$. The optimization framework of Analysis SimCO is presented in Algorithm 1.

1) Analysis Sparse Coding: The purpose of the analysis sparse coding stage is to get the sparse representations $\mathbf{X}$ of the training signals in $\mathbf{Y}$ based on a given dictionary $\mathbf{\Omega}$. The exact representations $\mathbf{X}$ can be calculated directly by simply multiplying the signals in $\mathbf{Y}$ by the dictionary $\mathbf{\Omega}$, that is

$$\mathbf{X} = \mathbf{\Omega}\mathbf{Y}. \tag{4}$$

Algorithm 1 Analysis SimCO

Input: $\mathbf{Y}$, p, l
Output: $\mathbf{\Omega}^\star$
Initialization:
Initialize the iteration counter $t = 1$ and the analysis dictionary $\mathbf{\Omega}^{(t)}$. Perform the following steps.
Main Iterations:
1) Analysis sparse coding: Compute the representations $\mathbf{X}^{(t)}$ with the fixed dictionary $\mathbf{\Omega}^{(t)}$ and the training signals in $\mathbf{Y}$, based on equations (4) and (5).
2) Dictionary update: Update the dictionary $\mathbf{\Omega}^{(t+1)} \leftarrow \mathbf{\Omega}^{(t)}$, using Algorithm 2.
3) If the stopping criterion is satisfied, $\mathbf{\Omega}^\star = \mathbf{\Omega}^{(t+1)}$ and quit the iteration. Otherwise, increase the iteration counter $t = t + 1$ and go back to step 1).

Since the initial dictionary is an arbitrary one, the representations obtained in this way may not satisfy the co-sparsity constraints on $\mathbf{X}$ in (3). A hard thresholding operation is therefore applied to enforce the co-sparsity

$$\hat{\mathbf{X}} = HT_l(\mathbf{X}), \tag{5}$$

where $HT_l(\mathbf{X})$ is the non-linear operator that sets the smallest l elements (in magnitude) of each column of $\mathbf{X}$ to zeros. The representations $\hat{\mathbf{X}}$ obtained via equation (5) are the best approximation of the exact representations $\mathbf{X}$ in terms of the error in Frobenius norm among all the matrices satisfying the co-sparsity constraints.

2) Dictionary Update: The dictionary update stage aims at optimizing the following problem (by fixing $\mathbf{X}$ in (3))

$$\arg\min_{\mathbf{\Omega}} f(\mathbf{\Omega}) = \|\mathbf{X} - \mathbf{\Omega}\mathbf{Y}\|_F^2 \quad \text{s.t.} \quad \|\mathbf{\Omega}_{j,:}\|_2 = 1, \ \forall j. \tag{6}$$

The Stiefel manifold $\mathcal{U}_{m,1}$ is defined as $\mathcal{U}_{m,1} = \{\mathbf{u} \in \mathbb{R}^m : \mathbf{u}^T\mathbf{u} = 1\}$ [11]. Based on this definition, the transpose of each row in $\mathbf{\Omega}$ is one element in $\mathcal{U}_{m,1}$. Thus, the "line" search methods on manifolds can be utilized to deal with problem (6). Here we use the the gradient descent line search method on manifold. We explain below the key points of this method including search direction, line search path and step size respectively. The dictionary update stage is summarized in Algorithm 2.

Algorithm 2 Dictionary Update Stage

Input: $\mathbf{\Omega}^{(t)}$, $\mathbf{X}^{(t)}$, $\mathbf{Y}$
Output: $\mathbf{\Omega}^{(t+1)}$
Main Steps:
1) Calculate the search direction, based on equations (7) and (8).
2) Find a proper step size α using golden section search.
3) Update the dictionary $\mathbf{\Omega}^{(t+1)} \leftarrow \mathbf{\Omega}^{(t)}$, based on equation (9).

The steepest descent direction is used as the search direction, i.e. the negative gradient of the objective function with respect to $\mathbf{\Omega}$ as follows

$$\begin{aligned}\mathbf{H} &= -\nabla f(\mathbf{\Omega}) \\ &= -\frac{\partial \|\mathbf{X} - \mathbf{\Omega}\mathbf{Y}\|_F^2}{\partial \mathbf{\Omega}} \\ &= 2\mathbf{X}\mathbf{Y}^T - 2\mathbf{\Omega}\mathbf{Y}\mathbf{Y}^T.\end{aligned} \tag{7}$$

The search direction of the jth row of $\mathbf{\Omega}$, i.e. the projection of each row of $\mathbf{H}$ onto the tangent space of $\mathcal{S}$, is [11, pp. 49]

$$\bar{\mathbf{h}}_j = \mathbf{H}_{j,:}(\mathbf{I} - \mathbf{\Omega}_{j,:}^T \mathbf{\Omega}_{j,:}). \tag{8}$$

The line search path for the jth row of $\mathbf{\Omega}$ can be written as

$$\mathbf{\Omega}_{j,:}(\alpha) = \begin{cases} \mathbf{\Omega}_{j,:} & \text{if } \|\bar{\mathbf{h}}_j\|_2 = 0, \\ \mathbf{\Omega}_{j,:}\cos(\alpha\|\bar{\mathbf{h}}_j\|_2) + (\bar{\mathbf{h}}_j / \|\bar{\mathbf{h}}_j\|_2)\sin(\alpha\|\bar{\mathbf{h}}_j\|_2) & \\ & \text{otherwise,} \end{cases} \tag{9}$$

where α is the step size.

In order to find a proper step size α, we apply the golden section search method [12]. This method consists of two stages. In the first stage, it finds a range which contains a local minimum and within which the objective function is unimodal. In the second stage, the golden section ratio is used to successively narrow the range until the minimizer is located and thus α is determined.

B. Spectrogram Reconstruction Stage

Based on the assumption that the HR spectrogram to be reconstructed is sparse with respect to the learned analysis dictionary $\mathbf{\Omega}$, the audio spectrogram reconstruction problem can be cast as the following optimization problem

$$\arg\min_{\hat{\mathbf{Y}}_H} \|\mathbf{\Omega}\hat{\mathbf{Y}}_H\|_1 + \frac{\lambda}{2}\|\mathbf{Y}_L - \mathbf{A}\hat{\mathbf{Y}}_H\|_F^2, \tag{10}$$

where λ is the Lagrangian multiplier to balance the data fidelity term $\|\mathbf{Y}_L - \mathbf{A}\hat{\mathbf{Y}}_H\|_F^2$ and the regularization term $\|\mathbf{\Omega}\hat{\mathbf{Y}}_H\|_1$. The alternating direction method of multipliers (ADMM) [13], [6] is applied to tackle this problem.

The optimization problem (10) is equivalent to the following equality-constrained convex optimization problem

$$\begin{aligned}&\arg\min_{\hat{\mathbf{Y}}_H, \mathbf{Z}} \|\mathbf{Z}\|_1 + \frac{\lambda}{2}\|\mathbf{Y}_L - \mathbf{A}\hat{\mathbf{Y}}_H\|_F^2 \\ &\quad \text{s.t. } \mathbf{Z} = \mathbf{\Omega}\hat{\mathbf{Y}}_H,\end{aligned} \tag{11}$$

by introducing the variable $\mathbf{Z} = \mathbf{\Omega}\hat{\mathbf{Y}}_H$. The augmented Lagrangian function for (11) is

$$\begin{aligned}L_\gamma(\hat{\mathbf{Y}}_H, \mathbf{Z}, \mathbf{B}) =& \|\mathbf{Z}\|_1 + \frac{\lambda}{2}\|\mathbf{Y}_L - \mathbf{A}\hat{\mathbf{Y}}_H\|_F^2 \\ &+ \gamma\langle \mathbf{B}, \mathbf{\Omega}\hat{\mathbf{Y}}_H - \mathbf{Z}\rangle + \frac{\gamma}{2}\|\mathbf{\Omega}\hat{\mathbf{Y}}_H - \mathbf{Z}\|_F^2 \\ =& \|\mathbf{Z}\|_1 + \frac{\lambda}{2}\|\mathbf{Y}_L - \mathbf{A}\hat{\mathbf{Y}}_H\|_F^2 \\ &+ \frac{\gamma}{2}\|\mathbf{B} + \mathbf{\Omega}\hat{\mathbf{Y}}_H - \mathbf{Z}\|_F^2 - \frac{\gamma}{2}\|\mathbf{B}\|_F^2,\end{aligned} \tag{12}$$

where $\gamma > 0$ is the penalty parameter and $\mathbf{B}$ is called the dual variable. The ADMM algorithm iteratively updates each of the variables $\{\hat{\mathbf{Y}}_H, \mathbf{Z}, \mathbf{B}\}$, while keeping the rest fixed. In the tth iteration, it consists of the following steps

$$\hat{\mathbf{Y}}_H^{(t+1)} = \arg\min L_\gamma(\hat{\mathbf{Y}}_H, \mathbf{Z}^{(t)}, \mathbf{B}^{(t)}) \tag{13}$$

$$\mathbf{Z}^{(t+1)} = \arg\min L_\gamma(\hat{\mathbf{Y}}_H^{(t+1)}, \mathbf{Z}, \mathbf{B}^{(t)}) \tag{14}$$

$$\mathbf{B}^{(t+1)} = \mathbf{B}^{(t)} + (\mathbf{\Omega}\hat{\mathbf{Y}}_H^{(t+1)} - \mathbf{Z}^{(t+1)}). \tag{15}$$

For (13) and (14), there are closed-form solutions [13]

$$\hat{\mathbf{Y}}_H^{(t+1)} = (\lambda\mathbf{A}^T\mathbf{A} + \gamma\mathbf{\Omega}^T\mathbf{\Omega})^{-1}(\lambda\mathbf{A}^T\mathbf{Y}_L + \gamma\mathbf{\Omega}^T(\mathbf{Z}^{(t)} - \mathbf{B}^{(t)})) \tag{16}$$

and

$$\mathbf{Z}^{(t+1)} = ST_{\frac{1}{\gamma}}\{\mathbf{\Omega}\hat{\mathbf{Y}}_H^{(t+1)} + \mathbf{B}^{(t)}\}, \tag{17}$$

where $ST_{\frac{1}{\gamma}}$ is the entrywise soft-thresholding operator defined by

$$ST_{\frac{1}{\gamma}}(\beta) = \begin{cases} \beta - \frac{1}{\gamma}\cdot\mathrm{sgn}(\beta) & \text{if} |\beta| \geq \frac{1}{\gamma}, \\ 0 & \text{otherwise.} \end{cases} \tag{18}$$

It is worth noting that the information of the known LR spectrogram $\mathbf{Y}_L$ should be kept in the estimated HR spectrogram $\hat{\mathbf{Y}}_H$, but the updated $\hat{\mathbf{Y}}_H^{(t+1)}$ obtained by equation (16) cannot guarantee this. Thus, a projection is applied after updating the HR spectrogram via (16), i.e.

$$\hat{\mathbf{Y}}_H^{(t+1)} \leftarrow \mathcal{P}_{\mathbf{Y}_L}(\hat{\mathbf{Y}}_H^{(t+1)}), \tag{19}$$

where $\mathcal{P}_{\mathbf{Y}_L}$ denotes the projection that sets the last k rows of $\hat{\mathbf{Y}}_H^{(t+1)}$ to be identical to $\mathbf{Y}_L$. The spectrogram reconstruction stage is summarized as Algorithm 3.

Algorithm 3 Spectrogram Reconstruction

Input: $\mathbf{\Omega}$, $\mathbf{Y}_L$, $\mathbf{A}$, λ, γ

Output: $\hat{\mathbf{Y}}_H$

Initialization:

Initialize the iteration counter $t = 1$. Perform the following steps.

Main Steps:

1) Update $\hat{\mathbf{Y}}_H^{(t+1)} \leftarrow \hat{\mathbf{Y}}_H^{(t)}$, based on equations (16) and (19).
2) Update $\mathbf{Z}^{(t+1)} \leftarrow \mathbf{Z}^{(t)}$, based on equations (17) and (18).
3) Update $\mathbf{B}^{(t+1)} \leftarrow \mathbf{B}^{(t)}$, based on equation (15).
4) If the stopping criterion is satisfied, $\hat{\mathbf{Y}}_H = \hat{\mathbf{Y}}_H^{(t+1)}$ and quit the iteration. Otherwise, increase the iteration counter $t = t + 1$ and go back to step 1).

III. Simulation Results

In the analysis dictionary learning stage, two types of training data $\mathbf{Y}$ were tested. The first type is the spectrogram of a combined audio signal consisting of piano note recordings. The database of piano note recordings, found in the website of University of Iowa Electronic Music Studios[1], contains 261 recordings of 44.1 kHz single piano notes. We truncated each recording to 2 seconds by removing the sections with small amplitude in time-domain. All of the truncated recordings are combined as the audio signal to generate the first type of training data. The second type training data is the HR spectrogram of the original audio signals, i.e. $\mathbf{Y}_H$. The dictionaries learned with these two types of training data are referred to as "Type I dictionary" and "Type II dictionary" respectively. The initial dictionaries of Analysis SimCO were generated with the random variables satisfying the i.i.d. Gaussian distribution with zero mean and unit variance and then the rows of the dictionaries were normalized.

We tested our method for the super-resolution of piano signals. Two pieces of 44.1 kHz piano recordings downloaded from the "freesound" website[2] were used as test signals (only the first 262144 points, about 6 seconds, of each signal were used). For each audio signal, Hamming window of length 1024 with 25% overlap was applied to generate the HR spectrogram using short-time Fourier transform (STFT) and thus $\mathbf{Y}_H \in \mathbb{R}^{513\times 431}$. The LR spectrogram $\mathbf{Y}_L$ was obtained by removing several top rows of $\mathbf{Y}_H$. The Analysis SimCO was applied for 100 iterations. The size of $\mathbf{\Omega}$ was 1026×513. The parameters of the spectrogram reconstruction stage are set as $\lambda = \gamma = 1$.

The metric used to evaluate the performance is the normalized reconstruction error in the Frobenius norm, i,e.

$$\delta = \frac{\|\mathbf{Y}_H - \hat{\mathbf{Y}}_H\|_F^2}{\|\mathbf{Y}_H\|_F^2}. \tag{20}$$

The reconstruction errors averaged from ten independent tests are plotted in Fig. 2. The two sub-figures present results of the two test piano signals respectively. The red lines show the results when the random initial dictionaries are used directly for spectrogram reconstrucion, without the involvement of analysis dcitionary learning. In this case, we can see that the reconstruction errors of the two test signals are very similar. Compared with the initial dictionary, the learned dictionaries leads to smaller reconstruction errors, indicating better super-resolution results. For the second test signal, the two types of learned dictionaries obtain similar results. However, for the first test signal, the dictionary learned with the original HR spectrogram (i.e. Type II dictionary) performs better than the dictionary learned with the spectrogram of piano note recordings. This shows that the better selection of training data may have contribution to the super-resolution result.

IV. Conclusion

In this paper we have proposed a method to address the super-resolution problem of audio signals. Our proposed method learns an analysis dictionary from related training data

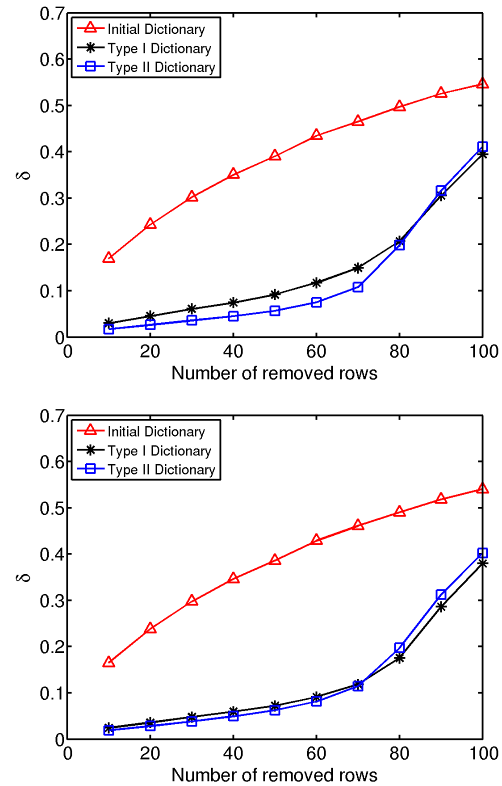

Fig. 2. Experimental results of two audio pieces with different numbers of removed rows in $\mathbf{Y}_H$, using different dictionaries.

and applies this dictionary to reconstruct the HR spectrogram. The analysis dictionary learning is achieved with the Analysis SimCO algorithm which alternates between analysis sparse coding and dictionary learning based on optimization on manifolds. The spectrogram reconstruction is formulated as an optimization with ℓ_1-norm regularization term. The ADMM algorithm is applied to address this optimization problem. Simulation results show that the learned dictionaries have advantage in reconstructing the HR spectrograms as compared with the initial dictionaries. The selection of training data is also important in affecting the results.

V. Acknowledgement

This work was supported by the Engineering and Physical Sciences Research Council (EPSRC) Grant number EP/K014307/1 and the MOD University Defence Research Collaboration in Signal Processing. We thank Prof. Jonathon Chambers for proofreading the manuscript.

[1] http://theremin.music.uiowa.edu/MISpiano.html

[2] The piano recordings were downloaded from {https://freesound.org/people/Lemoncreme/sounds/186942/} and {https://freesound.org/people/Bradovic/sounds/164718/}.

References

[1] S. C. Park, M. K. Park, and M. G. Kang, "Super-resolution image reconstruction: a technical overview," *IEEE Signal Process. Mag.*, vol. 20, no. 3, pp. 21–36, 2003.

[2] J. Yang, J. Wright, T. S. Huang, and Y. Ma, "Image super-resolution via sparse representation," *IEEE Trans. Image Process.*, vol. 19, no. 11, pp. 2861–2873, 2010.

[3] B. P. Keegan, S. K. Tjoa, and K. R. Liu, "Super-resolution of musical signals using approximate matching pursuit," in *2011 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, 2011, pp. 81–84.

[4] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4311–4322, 2006.

[5] S. Nam, M. E. Davies, M. Elad, and R. Gribonval, "The cosparse analysis model and algorithms," *Appl. Comput. Harmon. Anal.*, vol. 34, no. 1, pp. 30–56, 2013.

[6] M. Yaghoobi, S. Nam, R. Gribonval, and M. Davies, "Constrained overcomplete analysis operator learning for cosparse signal modelling," *IEEE Trans. Signal Process.*, vol. 61, no. 9, pp. 2341–2355, 2013.

[7] R. Rubinstein, T. Peleg, and M. Elad, "Analysis K-SVD: A dictionary-learning algorithm for the analysis sparse model," *IEEE Trans. Signal Process.*, vol. 61, no. 3, pp. 661–677, 2013.

[8] S. Hawe, M. Kleinsteuber, and K. Diepold, "Analysis operator learning and its application to image reconstruction," *IEEE Trans. Image Process.*, vol. 22, no. 6, pp. 2138–2150, 2013.

[9] J. Dong, W. Wang, and W. Dai, "Analysis SimCO: A new algorithm for analysis dictionary learning," in *Proc. Int. Conf. Acoust., Speech, and Signal Process.*, 2014, pp. 7193–7197.

[10] S. Ravishankar and Y. Bresler, "Learning sparsifying transforms," *IEEE Trans. Signal Process.*, vol. 61, no. 5, pp. 1072–1086, 2013.

[11] P.-A. Absil, R. Mahony, and R. Sepulchre, *Optimization Algorithms on Matrix Manifolds*. Princeton University Press, 2009.

[12] W. Dai, T. Xu, and W. Wang, "Simultaneous codeword optimization (SimCO) for dictionary update and learning," *IEEE Trans. Signal Process.*, vol. 60, no. 12, pp. 6340–6353, 2012.

[13] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2011.