Volume Reconstruction for MRI

Meiqing Zhang, Huirao Nie, Yang Pei, and Linmi Tao

Tsinghua National Laboratory for Information Science and Technology
Department of Computer Science and Technology
Tsinghua University, Beijing 100084, P.R.China
Email: ime.ting@hotmail.com, sangoblin@yeah.net, pyemma1991@gmail.com, linmi@tsinghua.edu.cn

Abstract—One of the challenges in medical imaging is to increase the resolution of 3D MRI (Magnetic Resonance Imaging) signals. This is the problem of 3D signal reconstruction under the condition of very low sampling rate. Based on compressive sensing theory, the Direct Volume Reconstruction (DVR) method is proposed to reconstruct the 3D signal volume-by-volume based on a learned dictionary. DVR is a general method and applicable to any 3D signal as long as it can be sparsely represented. To exploit the nature of the 3D MRI system, the Progressive Volume Reconstruction (PVR) method is further proposed to improve the DVR reconstruction. In PVR, local reconstruction is used to reconstruct in-plane slices, and the output is then forwarded to global reconstruction, in which both the initially sampled and locally reconstructed signals are used together to reconstruct the whole 3D signal. Two separate dictionaries, rather than one, are trained in PVR. In this way, more prior knowledge from the training data is exploited. Experiments on a head MRI dataset demonstrate that DVR achieves much better performance than conventional tricubic interpolation and that PVR considerably improves DVR performance with regard to both PSNR and visibility quality.

Keywords—Magnetic Resonance Imaging (MRI), volume reconstruction, compressive sensing, dictionary learning

I. INTRODUCTION

MRI (Magnetic Resonance Imaging) is a medical imaging technique to visualize tissues and organs inside the body with clear soft-tissue contrast and high spatial resolution. It has become an important and widely used technique to facilitate diagnosis [1]. Although most current MRI systems can already provide resolutions around $1.5 \times 1.5 \times 4.0 \text{mm}^3$ or $1.0 \times 1.0 \times 6.0 \text{mm}^3$ with reasonable Signal-to-Noise-Ratio (SNR) using $1.5T$ or $3T$ machines for static scans, the need is still urgent to increase the resolution to, for instance, $1.0 \times 1.0 \times 1.0 \text{mm}^3$ per volume, or even smaller. Such resolution improvement is very challenging, yet important in clinical diagnosis to visualize certain small pathological changes, such as nasopharyngeal carcinoma, which are subtle even for MRI [1].

Super Resolution (SR) based posterior image processing techniques have flourished in the past decade. In general, SR methods for MRI can be categorized as either Acquisition Model Based (AMB) [1] or Learning Based (LB). The AMB methods reconstruct high-resolution images from several low-resolution images of the same patient, which are obtained by introducing a known rotation or sub-pixel shift. This reconstruction is based on solving the ill-posed inverse problem by using back-projection approaches [2] such as in [3], or regularization based approaches as in [4], [5], [6], [7]. AMB methods are efficient for in-plane detail enhancement if the registration between modalities is accurate. However, these methods require several low-resolution frames of the same patient to reconstruct the high-resolution image, which not only increases scanning time but is also nearly impossible in practice because some of the data to be recovered are from the past.

LB methods incorporate prior knowledge learned from training examples into dictionaries of natural images, and have been extended to 3D MRI for resolution enhancement. For example, based on [8] for color-image super resolution, the first and second order gradient features of the in-plane patches were used to train the dictionary to enhance the through-plane resolution in [9]. Similar work can also be found in [10], which reconstructed intermediate slices using a patch-based sparse representation approach.

Most posterior image processing techniques are used for either through-plane or in-plane resolution enhancement. It is very challenging to accomplish thin 3D volume resolution (e.g., from $1.5 \times 1.5 \times 3 \text{mm}^3$ to $1 \times 1 \times 1 \text{mm}^3$ per volume), or in other words, to increase resolution in all three dimensions simultaneously. The problem stems partly from the fact that low-resolution MRI in a 3D situation is hyper-sparingly sampled, meaning that, the sample rate is sparser by an order than in 2D (e.g., if we want to double each dimension resolution, then the sample rate for a 3D volume is only 12.5%, instead of 25% for the 2D patch). Conventionally, reconstructions in this situation are implemented by 3D image interpolation methods [11] such as tricubic interpolation. However, image interpolation approaches for increasing medical image resolution have side effects such as blurred edges [12].

By combining Compressive Sensing (CS) [13], [14] and the LB approach, this paper first proposes the Direct Volume Reconstruction (DVR) method, which directly reconstructs the 3D MRI volume by volume based on a dictionary trained using Independent Component Analysis (ICA). Experiments indicate that DVR achieves much better performance than tricubic interpolation approaches. DVR is a general method and applicable to any 3D signal as long as it can be sparsely represented. To exploit the special characteristic of 3D MRI that it has higher resolution and correlation in-plane than that through-plane, the Progressive Volume Reconstruction (PVR) method has been devised to improve DVR performance. PVR consists of two steps, local reconstruction and global reconstruction. Multiple dictionaries are used in the two steps of PVR to improve both the adaptivity and diversity of the prior knowledge, which is the 3D nature of MRI. Experiments on a head MRI dataset from a hospital have demonstrated the merits of these approaches.
II. BACKGROUND

Compressive Sensing (CS) is an emerging technique to acquire certain kinds of signals at a sample rate significantly below the well-known Nyquist rate [15]. For a signal vector \( s \in \mathbb{R}^N \) that is of \( K \)-Sparse (i.e., only \( K \) entries are non-zeros), its observation vector \( y \in \mathbb{R}^M (M < N) \) is acquired by the observation matrix \( \Theta \in \mathbb{R}^{M \times N} \) as given in Equation (1).

\[
y = \Theta s
\]  

(1)

In some cases, the sparse signal \( s \) is obtained after applying a certain domain transformation \( \Phi \) (e.g., the discrete cosine transform (DCT)) to the initial signal \( x \) (e.g., a vectorized pixel patch), as shown in Equation (2).

\[
y = \Theta s = \Theta \Phi x
\]  

(2)

Standard CS theory has proved that robust reconstruction of \( s \) (or \( x \)) from observation \( y \) is possible if \( \Theta \) and \( \Phi \) satisfy certain conditions and the number of measurements meets the criterion that

\[
M > cK\log(N/K)
\]  

(3)

where \( c \) is a small constant [16], [17]. It has also been pointed out that CS is a natural fit for MRI because it can be sparsely represented in certain appropriate domains (e.g., wavelet transformation) [18], [19], [20], [21].

In practice, the transformation matrix \( \Phi \) will largely determine the sparsity level of \( s \) and the reconstruction performance. To ensure reconstruction algorithm performance, \( \Phi \) is usually orthogonal and complete, even though sometimes an overcomplete dictionary may be used to improve the sparsity degree of \( s \). Note that if the sparse code \( s \) has been reconstructed, then the initial signal \( x \) can be easily computed as \( x = \Phi^{-1}s \). Defining \( \Phi^{-1} \) as \( D \), Equations (4) and (5) result:

\[
y = \Theta s = \Theta \Phi x = \Theta D^{-1}x,
\]  

(4)

\[
x = Ds.
\]  

(5)

\( D \) may stem from a pre-defined mathematical basis such as a wavelet and Fourier basis, or be trained from diverse samples of one class, in which case it will generally be more efficient and adaptive for signals of the specific class. In this case, it is usually called a dictionary. Each column in the dictionary \( D \) can be called an atom.

Dictionary learning can be seen as a sparse coding problem [22], [23]. To reduce the size of \( D \), training is usually performed in patch space. More explicitly, \( T \) patches of identical size are sampled from multi-dimensional images in the training set, and each patch is reshaped as a vector \( x_i \in \mathbb{R}^n \) (\( i = 1, \ldots, T \)). The principle of training is that \( x_i \) can be sparsely represented using \( D \in \mathbb{R}^{n \times m} (n \leq m) \), as formulated in Equation (6):

\[
x_i \approx De_i \text{ subject to } ||e_i||_0 \ll m, \ i = 1, \ldots, T.
\]  

(6)

All the \( x_i (i = 1, \ldots, T) \) are concatenated along the column direction as \( X = [x_1 \ldots x_T] \in \mathbb{R}^{n \times T} \). Then, by decomposing \( X = DC \) with the sparse constraints on the code matrix \( C \), the dictionary \( D \) can be obtained. There are various algorithms for implementing the factorization, for example, Independent Component Analysis (ICA) [24], the widely used K-SVD [25], [26], and the Non-negative Matrix Factorization (NMF) algorithm [27]. In [23], it was reported that an ICA based dictionary training yields the state-of-the-art inpainting performance for natural images.

Reconstructing the length-\( N \) signal \( s \) (or equivalently, \( x \)) from a length-\( M \) signal \( y \) is an ill-posed because \( M < N \). Various algorithms have been proposed to estimate \( s \) by solving the problem formulated in Equation (7):

\[
s = \arg \min \|s\|_p \text{ subject to } \|y - \Theta s\|_2 < \xi,
\]  

(7)

where \( \|s\|_p = \sqrt[p]{\sum_{i=1}^N |s_i|^p} \) and the \( \ell^p \) norm of vector \( s \) is defined as \( \|s\|_p \). In particular, the \( \ell^0 \) norm counts the non-zero entries, i.e., it measures sparsity. However, directly minimizing the \( \ell^0 \) norm is NP-hard. The state-of-the-art for solving Equation (7) when \( p = 0 \) is the recently proposed smoothed \( \ell^0 \) algorithm (SLO) [28] which tries to minimize the \( \ell^0 \) norm directly by using a continuous smooth function to approximate the discrete \( \ell^0 \) norm. SLO is reported to be faster than the previous \( \ell^1 \) norm based approaches such as Basis Pursuit (BP) [29] and Orthogonal Matching Pursuit (OMP) [30] with comparable or better accuracy.

III. VOLUME RECONSTRUCTION METHODS FOR MRI

Although standard CS theory enables robust signal recovery with low sample rate, performance would remain poor if number of measurements \( M \) were insufficient for the \( O(K\log(N/K)) \) requirement. Consequently, directly extending 2D image-reconstruction enhancement methods to 3D situations will be confronted with the signal scarcity problem because the low-resolution image is hyper-sparsely sampled. In other words, the sample rate has been sharply decreased due to the increase in dimension. The consequence is weakened recovery performance. To address this problem, the Direct Volume Reconstruction (DVR) method is first proposed here for MRI volume reconstruction, and the Progressive Volume Reconstruction (PVR) method is then introduced to improve DVR performance. PVR exploits the 3D nature of MRI and the powerful expression capabilities of multiple dictionaries, rather than one, to accomplish a satisfactory reconstruction.

A. Direct Volume Reconstruction (DVR)

Suppose that we want to reconstruct a high-resolution MRI \( H \in \mathbb{R}^{a \times b \times L} \) from the low-resolution image \( L \in \mathbb{R}^{a \times b \times h} \). Note that \( a \times b \) is the in-plane resolution and \( h \) is the length in the slice-selection direction. \( \alpha, \beta, \lambda \) and \( \gamma \) are magnification factors in the three dimensions. In general, it is required that \( \alpha, \beta, \lambda \) all be positive integers.

DVR reconstructs the whole 3D MRI in a volume-by-volume fashion. Suppose that we want to reconstruct a volume \( V_H \in \mathbb{R}^{a \times b \times L} \) from a low-resolution volume \( V_L \in \mathbb{R}^{a \times b \times h} \). Generally, \( V_H \) and \( V_L \) will first be vectorized to \( v_H \in \mathbb{R}^{a \times b \times L} \) and \( v_L \in \mathbb{R}^{a \times b \times h} \) respectively. Note that there are still spatial constraints for pixels in \( v_H \) and \( v_L \); although vectorization weakens them, \( I \) defines how \( v_L \) is obtained from \( v_H \), i.e., \( v_L = Iv_H \).

In practice, for efficiency, \( I \) is not computed directly; instead, a boolean mask matrix is equivalently computed whose
Algorithm 1 Direct Volume Reconstruction (DVR)

Input: $L$, $p$, $q$, $v$, $\alpha$, $\beta$, $\lambda$, $D$

Output: Reconstructed high-resolution MRI $\hat{H}$

1: Initialize $\hat{H} \in R^{a \times b \times h}$ from $L \in R^{a \times b \times h}$ by zero padding
2: Initialize $H_1 \in R^{a \times b \times h}$ with zeros
3: Initialize a matrix $G \in R^{a \times b \times h}$ with zeros. $G$ counts how many times each position has been recovered
4: for $i = 1$ to $\alpha a$ step $\alpha$ do
5: for $j = 1$ to $\beta b$ step $\beta$ do
6: for $k = 1$ to $\lambda h$ step $\lambda$ do
7: Extract volume $V_{H} \in R^{p \times q \times \lambda v}$ from $\hat{H}$ starting at position $(i, j, k)$
8: Extract corresponding volume $V_L \in R^{p \times q \times v}$ from $L$
9: Vectorize $V_{H}$ to $v_{H}$
10: Vectorize $V_L$ to $v_L$
11: Computing $I$ which transforms $v_{H}$ to $v_{L}$
12: Obtain sparse code $\hat{c}$ by solving Equation (8)
13: Estimate $\hat{v}_{H} \approx D \hat{c}$
14: Obtain $V_{H} \in R^{p \times q \times \lambda v}$ by reshaping $\hat{v}_{H}$
15: Add $V_{H}$ to $H_1$ at corresponding position
16: Increase the values in volume $V_{G} \in R^{p \times q \times \lambda v}$ (starting at position $(i, j, k)$ in $G$) by 1
17: end for
18: end for
19: end for
20: Divide $H_1$ by $G$ element-wise, yielding $\hat{H} = H_1 / G$
21: return $\hat{H}$

The local reconstruction step reconstructs a local area made up of $\Delta h$ (a small number, e.g., 1-3) successive slices each time, for which the volume is $p_1 \times q_1 \times \Delta h$. In the local area, the signal is highly correlated. Moreover, the sample rate is improved to $1/(\alpha \times \beta \times \lambda)$. Consequently, its reconstructions are hypothesized to have low error so that they can be used as part of the known signals as input for the following global reconstruction step. Dictionaries for these two steps can be trained separately; consequently, two dictionaries ($D_1$ and $D_2$ in Algorithm 2), rather than one ($D$ in Algorithm 1), are used in the reconstruction. In this way, more prior knowledge can be exploited. Meanwhile, different volume sizes are beneficial for better adaptivity. For instance, in local reconstruction, correlations are high, so that $p_1 \times q_1$ can be relatively high to improve the dictionary expression capabilities; in global reconstruction, a relatively small volume $p_2 \times q_2 \times v$ can be used to ensure that intensities and structures in volumes of different positions follow traceable patterns.

Algorithm 2 Progressive Volume Reconstruction (PVR)

1: $\hat{L} = \text{LocalReconstruction}(L, p_1, q_1, \Delta h, \alpha, \beta, D_1)$ \hspace{1cm} $\triangleright$ sample rate: $1/(\alpha \times \beta)$
2: $\hat{H} = \text{GlobalReconstruction}(\hat{L}, p_2, q_2, v, \lambda, D_2)$ \hspace{1cm} $\triangleright$ sample rate: $1/\lambda$
3: return $\hat{H}$
4: function \text{LOCALRECONSTRUCTION($L$, $p_1$, $q_1$, $\Delta h$, $\alpha$, $\beta$, $D_1$)}
5: Divide $L$ into $L_i, i = 1, \ldots, h/\Delta h$, where $L_i \in R^{a \times b \times \Delta h}$
6: for $i \leftarrow 1$ to $h/\Delta h$ do
7: $L_i = \text{DVR}(L_i, p_1, q_1, \Delta h, \alpha, \beta, 1, D_1)$
8: end for
9: Concatenate $L_1$ along the third dimension: $\hat{L} = \text{cat}(3, L_1, \ldots, L_{h/\Delta h}) \in R^{a \times b \times h}$
10: return $\hat{L}$ \hspace{1cm} $\triangleright$ as input for the second step
11: end function
12: function \text{GLOBALRECONSTRUCTION($\hat{L}$, $p_2$, $q_2$, $v$, $\lambda$, $D_2$)}
13: $\hat{H} = \text{DVR}(\hat{L}, p_2, q_2, v, 1, 1, \lambda, D_2)$
14: return $\hat{H}$ \hspace{1cm} $\triangleright$ final reconstructed 3D MRI
15: end function

is accurate if both a good learned dictionary and a robust reconstruction algorithm were used. This recovered signal is further to be treated as known for the next step, where the whole image is finally reconstructed. These two steps are called local reconstruction and global reconstruction respectively.
IV. EXPERIMENT

A. Data Acquisition and Settings

The dataset consisted of the 3D head MRI data from 9 different subjects in hospital. The high-resolution images were captured using a Siemens® 3T Magnetom Vision™ with a resolution of $1.0 \times 1.0 \times 1.33\, \text{mm}^3$ (TR=253.8ms, TE=3.39ms). The image size was $256 \times 256 \times 128$. The leave-one-out approach was used to augment the dataset. Specifically, head images from eight subjects were used to train the dictionary using FastICA [31], which yielded a complete dictionary $D \in R^{N \times N} (N = p \times q \times v)$ for a volume of size $p \times q \times v$. Then the remaining subject’s head image was used as $H$ (the ground-truth image) in the test. Only the middle 60 slices of $H$ were used for the test, meaning that, $H \in R^{256 \times 256 \times 60}$. The low resolution image $L$ was obtained by evenly and discretely sampling pixels from $H$. Magnification factors in the three dimensions were assumed to $\alpha = \beta = \lambda = 2$ (i.e., the size of the reconstructed MRI is $256 \times 256 \times 60$).

B. DVR

Detailed comparisons of the PSNR between conventional tricubic interpolation and the proposed DVR method with different volume sizes can be found as Table I. The first PSNR column is the result of testing the 1st patient, while the second is from the 7th patient. Dictionaries were trained based on the different volume sizes can be found as Table I. The first PSNR column is the result of testing the 1st patient, while the second is from the 7th patient. Dictionaries were trained based on the corresponding left 8 images.

In DVR, only one dictionary is used. The best reconstruction is achieved with an $8 \times 8 \times 5$ volume, and the corresponding PSNR is on average 0.73 dB higher than that of conventional tricubic interpolation. A visual comparison can be found in Figure 1.

As can be seen from Table I, as the size of the dictionary increases (i.e., the volume size increases), the expression capability of the dictionary will at first increase and the reconstruction will become better. However, when the size reaches a certain point, reconstruction performance becomes negatively related to volume size. The underlying reasons for this will be discussed later.

C. PVR

The PVR method consists of local reconstruction and global reconstruction, as shown in Algorithm 2. In local reconstruction step, local areas $L_1, \ldots, L_h/\Delta_h$ ($L_i \in R^{a \times b \times \Delta_h}$, where $a = b = 128$) are reconstructed to obtain $\hat{L}_i \in R^{a \times b \times \Delta_h}$, using the volume $p_1 \times q_1 \times \Delta_h$. Then $L \in R^{a \times b \times v}$ is constructed by concatenating $\hat{L}_i$ along the third dimension. The output $\hat{L}$ is then chosen (in these experiments, the corresponding volume is $p_1 = p_2 = 10$, $\Delta_h = 3$) as the input for the second step, global reconstruction, which will reconstruct the final $H \in R^{2a \times 2b \times 2h} = R^{256 \times 256 \times 60}$ from $\hat{L}$ using a volume $p_2 \times q_2 \times v$.

1) Local Reconstruction: The results of local reconstruction with respect to the PSNR (dB) and volume $p_1 \times q_1 \times \Delta_h$ are detailed in Table II.

2) Global Reconstruction: Using the result of local reconstruction as input for the global reconstruction step, the final output of the PVR method is detailed in Table III. On average, the best reconstruction is $33.4324\, \text{dB}$, which is about $0.27\, \text{dB}$ higher than the best output from the DVR method and about 1 dB higher than the result of tricubic interpolation. A visual comparison can be found in Figure 2.

![Figure 1](image-url)

(a) Original low-resolution MRI. (b) Reconstruction by tricubic interpolation. (c) Reconstruction by DVR method. (d) Ground-truth image. Figure 1. This figure demonstrates the original low-resolution MRI ($128 \times 128 \times 30$) and a comparison of reconstructions by tricubic interpolation and by the proposed DVR method and the ground-truth image (for convenience, only one slice is presented). Test image is of the 1st patient. Sample rate is 12.5%, and $\alpha = \beta = \lambda = 2$ (i.e., the size of the reconstructed MRI is $256 \times 256 \times 60$).
Figure 2. A comparison of DVR and PVR reconstructions with the ground-truth image. The test image is from the 1st patient. The slice selected here is an intermediate one inserted by DVR and PVR, and therefore the corresponding original low-resolution slice is shown as semitransparent to indicate that it actually does not exist. The sample rate was 12.5%, and $\alpha = \beta = \lambda = 2$ (i.e., the size of the reconstructed MRI was 256 $\times$ 256 $\times$ 60).

Table II. RESULTS OF THE FIRST STEP (LOCAL RECONSTRUCTION) OF PVR WITH RESPECT TO PSNR (dB) AND A VOLUME $p_1 \times q_1 \times \Delta h$, WHERE $p_1 \times q_1$ IS THE IN-PLANE RESOLUTION.

<table>
<thead>
<tr>
<th>$p_1 \times q_1 \times \Delta h$</th>
<th>1st</th>
<th>7th</th>
<th>1st</th>
<th>7th</th>
<th>1st</th>
<th>7th</th>
</tr>
</thead>
<tbody>
<tr>
<td>4 $\times$ 4 $\times$ 1</td>
<td>33.2730</td>
<td>33.3156</td>
<td>33.2743</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 $\times$ 4 $\times$ 5</td>
<td>33.2945</td>
<td>33.3465</td>
<td>33.2955</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 $\times$ 6 $\times$ 3</td>
<td>33.3011</td>
<td>33.3011</td>
<td>33.2867</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 $\times$ 6 $\times$ 5</td>
<td>33.3814</td>
<td>33.3814</td>
<td>33.4324</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 $\times$ 6 $\times$ 7</td>
<td>33.3471</td>
<td>33.3471</td>
<td>33.3471</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 $\times$ 8 $\times$ 3</td>
<td>33.2976</td>
<td>33.2976</td>
<td>33.2922</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 $\times$ 8 $\times$ 5</td>
<td>33.3329</td>
<td>33.3329</td>
<td>33.3329</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 $\times$ 8 $\times$ 7</td>
<td>32.8663</td>
<td>32.8663</td>
<td>32.8663</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 $\times$ 8 $\times$ 9</td>
<td>32.5679</td>
<td>32.5679</td>
<td>32.5679</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table III. PSNR OF THE FINAL PVR OUTPUT. ON AVERAGE, THE BEST RECONSTRUCTION IS 33.4324 dB, WHICH IS ABOUT 0.27 dB HIGHER THAN THE BEST OUTPUT FROM THE DVR METHOD; AND ABOUT 1 dB HIGHER THAN THE RESULT OF TRICUBIC INTERPOLATION.

<table>
<thead>
<tr>
<th>$p_2 \times q_2 \times t$</th>
<th>PSNR, 1st image</th>
<th>PSNR, 7th image</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>4 $\times$ 4 $\times$ 1</td>
<td>33.2730</td>
<td>33.3156</td>
<td>33.2743</td>
</tr>
<tr>
<td>4 $\times$ 4 $\times$ 5</td>
<td>33.3696</td>
<td>33.3011</td>
<td>33.3471</td>
</tr>
<tr>
<td>6 $\times$ 6 $\times$ 3</td>
<td>33.3011</td>
<td>33.3011</td>
<td>33.2867</td>
</tr>
<tr>
<td>6 $\times$ 6 $\times$ 5</td>
<td>33.3814</td>
<td>33.3814</td>
<td>33.4324</td>
</tr>
<tr>
<td>6 $\times$ 6 $\times$ 7</td>
<td>33.3471</td>
<td>33.3471</td>
<td>33.3471</td>
</tr>
<tr>
<td>8 $\times$ 8 $\times$ 3</td>
<td>33.2976</td>
<td>33.2976</td>
<td>33.2922</td>
</tr>
<tr>
<td>8 $\times$ 8 $\times$ 5</td>
<td>33.3329</td>
<td>33.3329</td>
<td>33.3329</td>
</tr>
<tr>
<td>8 $\times$ 8 $\times$ 7</td>
<td>32.8663</td>
<td>32.8663</td>
<td>32.8663</td>
</tr>
<tr>
<td>8 $\times$ 8 $\times$ 9</td>
<td>32.5679</td>
<td>32.5679</td>
<td>32.5679</td>
</tr>
</tbody>
</table>

Note that if global reconstruction is not used in the second step, but instead, tricubic interpolation is used to reconstruct the final result from the best output of the first step, the PSNR is only 32.7299 dB, which is about 0.7 dB lower than the result obtained by global reconstruction. It indicates the necessity of global reconstruction as the second step.

D. Discussion

1) The 3D nature of MRI: These experimental results indicate that the 3D nature of MRI plays an important role in good reconstruction. Data correlation exists not only in-plane, but also in the through-plane. For example, as shown in Table I, the results of DVR with volume $4 \times 4 \times 5$ are better than with $4 \times 4 \times 3$, $6 \times 6 \times 5$ is better than $6 \times 6 \times 3$, and so on. However, when the volume reaches a certain point, volumes at different positions in the MRI may not share enough structural similarities, or in other words, they do not follow consistent spatial and intensity patterns, which will significantly weaken the expressive capabilities of the atoms in the dictionary. Locally similar and globally diverse are important features of 3D MRI. Considering these features in algorithm design and engineering practice would improve the results.

2) Adaptivity and Power of Multiple Dictionaries: The MRI data are highly correlated locally and become more and more diverse with distance, especially in the slice-selection direction. The increase in dictionary size is actually a paradox. First, the larger the dictionary, the richer its expression capability theoretically should be. However, in practice, a larger dictionary size means that the volume size should be larger when sampling volumes to train the dictionary. The intensities in a volume would then become highly diverse and follow no pattern, and both their spatial and intensity distribution patterns may have little similarity to those of other volumes, counterproductively making it hard to represent signals using atoms in the dictionary. This dilemma is evidenced by Table I, in which the PSNR is first positively and then negatively correlated with volume size.

The PVR method tries to take advantage of this dilemma to improve DVR performance. In the two steps, two dictionaries of different sizes can be used to accomplish better reconstruction, as well as providing better adaptivity. In the first step, larger values of $p_1$ and $q_1$ can be used, but in the second step, $p_2$, $q_2$, and $t$ must be carefully constrained. From this perspective, PVR provides better adaptivity to the locally similar and globally diverse nature of 3D MRI than other methods. Moreover, two dictionaries actually contain more priori knowledge than only one.

V. Conclusion

This paper has first proposed the DVR method to increase the resolution of all three dimensions of an MRI volume simultaneously, yielding much better performance than the conventional tricubic interpolation. DVR is a general method for the reconstruction of any 3D signal which can be sparsely represented. To exploit the special characteristic of 3D MRI that it has higher resolution and correlation in-plane than in the through-plane, the PVR method has been proposed here to improve DVR performance further. Experiments demonstrate that the 3D nature of MRI and multiple dictionaries considerably benefit reconstruction from hyper-sparse sampled low-resolution MRI. Hyper-sparse sampled means that the sample rate is lower in an order than in 2D. For example, the sample rate is only 12.5% when resolution is doubled in each of the three dimensions, instead of 25% in two dimensions.
PVR uses the 3D nature of MRI to improve reconstruction performance, and moreover, two dictionaries, rather than one, can be used. In this way, more prior knowledge can be exploited. Although the dictionary trained using ICA is already complete, multiple dictionaries can handle the hyper-sparse situation better than a single one, as well as providing better adaptivity. Experiments on a head MRI dataset indicated the merits of the proposed DVR and PVR methods. They achieved satisfactory results with regard to PSNR and visual effects.

ACKNOWLEDGMENT

This work was supported in part by the National Natural Science Foundation of China under Grants 61272232 and MOST under Grants 2012AA011602. The authors also would like to thank anonymous reviewers who gave valuable suggestions that has helped to improve the quality of the paper.

REFERENCES