A New Sparse Representation Framework for Reconstruction of an Isotropic High Spatial Resolution MR Volume from Orthogonal Anisotropic Resolution Scans

-Supplementary Materials

Yuanyuan Jia*, Ali Gholipour, *Senior Member, IEEE,* Zhongshi He, Simon K. Warfield, *Senior Member, IEEE*

I. The Dictionary Training Algorithm

We utilized Zeyde's algorithm [1] to train the HR and LR dictionary pairs in this paper. To this end, we first trained the LR dictionary using K-SVD (K-Singular Value Decomposition) [2] algorithm based on LR training examples, and obtained the sparse representation. Then the HR dictionary was learned based on the sparse representation and HR training examples.

1) Low resolution dictionary training

LR dictionary training is formulated as:

$$\min_{D_Y} \|Y_T - D_Y\alpha^Y\|^2 \text{ s.t. } \left\|\alpha_p^Y\right\|_0 \le s, \forall\, p = 1,2,\dots,P \quad \text{(S1)}$$

where $Y_T = [y_1, y_2, \dots, y_P]$ is the training LR feature vectors of small patches; $\alpha^Y = [\alpha_1^Y, \alpha_2^Y, \dots, \alpha_P^Y]$, $\left\|\alpha_p^Y\right\|_0 \le s$ means the number of non-zeros in vector α_p^Y is smaller than s, and P is the total number of training patch vectors.

We solve this optimization problem by iteratively updating the dictionary D_Y and sparse coefficient α^Y. In the dictionary update stage, we use the K-SVD algorithm, which generalized from K-means clustering; and use the OMP (orthogonal matching pursuit) algorithm in the sparse representation stage [3], [4]. The dictionary training algorithm is performed by the following pseudo-code:

Initialize the dictionary:
$D_Y^0 \leftarrow$ Randomly select W feature vectors from training set;
$t \leftarrow 1$;
T← The maximum iterations;
WHILE $t < T$
- Sparse coding stage:

FOR each patch y_p
$\alpha_p^Y \leftarrow$Compute the sparse representation by OMP algorithm based on dictionary D_Y^{t-1}.
ENDFOR
- Dictionary updating stage:

FOR each atom $\hat{d}_w$of $D_Y^{t-1}, w = 1,2,\dots,W$;
$E_w \leftarrow Y - \sum_{j\neq w} d_j\alpha_T^j$, α_T^j is the jth row of α^Y;
$\omega_m \leftarrow$The group of indices pointing to examples $\{y_p\}$ that use the atom $\hat{d}_m$.
$\alpha_m^R \leftarrow$Restrict α_T^m by choosing only the columns corresponding toω_m;
$E_m^R \leftarrow$Restrict E_m by choosing only the columns corresponding toω_m;
$U, \Delta, V \leftarrow$ Appling SVD (Singular Value Decomposition) to E_m^R;
$\hat{d}_m \leftarrow$ Choose the first column of U;
$\alpha_m^R \leftarrow$ Multiply the first column of V by$\Delta(1,1)$;
ENDFOR
$t \leftarrow t + 1$;
ENDWHILE

2) High resolution dictionary training

After LR dictionary training stage, a corresponding HR dictionary is constructed, such that HR and LR dictionaries share the same sparse coefficient, and connect the HR and LR patches. Recall that our intention is to recover the HR patches by approximating it as being $x_p = D_X\alpha_p^X$, $\alpha_p^X \approx \alpha_p^Y$. Thus, the corresponding HR dictionary is defined to be the one that minimizes the mean approximation error:

$$D_X = arg\min_{D_X} \sum_p \left\|x_p - D_X\alpha_p^X\right\|_2^2 = arg\min_{D_X} \|X_T - D_X\alpha^X\|_2^2 \quad \text{(S2)}$$

where $X_T = [x_1, x_2, \dots., x_P]$ is the training HR feature vector set corresponding to the LR training examples Y_T. The solution of (S2) is given by the following Pseudo-Inverse expression [1] (given that α^X has full row rank):

$$D_X = X_T(\alpha^X)^+ = X_T(\alpha^X)^T[\alpha^X(\alpha^X)^T]^{-1} \quad \text{(S3)}$$

So based on the sparse representation vector α^X and the corresponding HR training examples, the HR dictionary is obtained.

II. Evaluation Metrics

To quantitatively and qualitatively evaluate the performance of the proposed method over different MR data sets, we introduce 4 different methods in this section for two scenarios:

1) Image with ground truth

In the experiments, if we have an original HR image, considered as the ground truth, comparing the reconstruction with the original image is a good way to evaluate the results. The following two performance metrics are calculated when the ground truth is available:

Peak Signal-to-Noise Ratio (PSNR) is defined as:

$$\mathrm{PSNR}(X_o, X_h) = 10 \cdot \log_{10}\left(\frac{d}{\mathrm{MSE}(X_o, X_h)}\right) \quad \text{(S4)}$$

where $MSE(X_o, X_h)$ stands for means square error, quantifies the pixel intensity difference between the original HR image X_o and the corresponding SR reconstruction X_h, using $MSE(X_o, X_h) = \frac{1}{|\Omega|}\sum_{k\in\Omega}\left|x_o^k - x_h^k\right|^2$, x_o^k and x_h^k are the image intensity at location k, d is the dynamic range of the intensity value, i.e. $d = max(X_o) - min(X_o)$. Typically, the PSNR values are between 20 dB and 50 dB. A higher value of PSNR indicates a better performance of the reconstruction method.

Structural Similarity Image Metric (SSIM) [5]: It measures the similarity between two images, with a definition that is more consistent with the human visual perception of image quality. Under the assumption that human visual perception is highly adapted to extracting structural information from a scene, SSIM is formulated as:

$$\mathrm{SSIM}(X_o, X_h) = \frac{(2\mu_o\mu_h + C_1)(2\sigma_{oh} + C_2)}{(\mu_o^2 + \mu_h^2 + C_1)(\sigma_o^2 + \sigma_h^2 + C_2)} \quad \text{(S5)}$$

where μ_o and μ_h are the mean intensity of images X_o and X_h, respectively; σ_o and σ_h are the standard deviation of images X_o and X_h, which are estimates of the signal contrast; σ_{oh} is the covariance of X_o and X_h, $C_1 = (K_1L)^2$ and $C_2 = (K_2L)^2$, $K_1 \ll 1$ and $K_2 \ll 1$ are small constants and L is the dynamic range of the intensity values. In this paper, we use $K_1 = 0.01$ and $K_2 = 0.03$. SSIM values are between 0 and 1, where a higher value indicates the better reconstruction results.

2) Images without ground truth

In reality, for example in clinical data, no original HR reference image is available, so no ground truth image is available for the evaluation of results. Alternative methods to evaluate the results are as follows:

Visual inspection: visual assessment of images is also a precious method to compare and judge the benefit of proposed methods; however, it is obviously a subjective method, and also may not be easy when large datasets should be evaluated and compared. In this paper, we display several slices selected from the reconstructed 3D MR image and evaluate the slices by viewing the image details.

Intensity profile: the intensity profile of an image is the set of intensity values taken from regularly spaced points along a line segment or multiline path in an image. The fundamental problem of SR reconstruction can be stated as restoring some high-frequency information (like edges) that has been lost during the acquisition process. An effective SR reconstruction technique should be able to recover these high-frequencies. Intensity profile can show intensity value changes at the interfaces between different tissues, thus may be used as a surrogate measure of how edge features appear and are distinguished in the image. We also evaluate the reconstruction results of our clinical MR experiments based on image intensity profiles in this paper.

III. Experiments on knee MR scans

1) Influence of slice thickness

To study the effect of slice thickness on the proposed method, we respectively produced three orthogonal LR down-sampled MR images with slice thickness of 2-7mm from C001 clinical T2w knee MR image. Then we reconstructed T2w knee HR volume with voxel size 0.625mm×0.625mm×0.625mm from three simulated LR orthogonal knee MR images. Table SI shows the PSNR and SSIM values of the reconstruction results. The proposed algorithm generated the best results in terms of both PSNR and SSIM values. For example, the PSNR/SSIM values obtained from the proposed method were 47.26dB/0.996 in 2 mm, while the results of the other two algorithms are 37.09dB/0.974 and 43.56dB/0.994 respectively. The PSNR/SSIM values dropped as the slice thickness increased. For example, the PSNR/SSIM value of the reconstructed C001 knee MR image was 47.26dB/0.996 when the slice thickness was 2mm, while the PSNR/SSIM value dropped to 40.78dB/0.986 when slice thickness was 3mm.

2) Influence of noise power

To evaluate the impact of the noise in the proposed algorithm, three orthogonal LR T2w knee MR scans of C004 with 3mm slice thickness were respectively produced. Different percentage noise (1%, 3%, 5%, 7%, and 9%) levels were used to investigate the noise influence. Table SII presents the reconstruction accuracy values in terms of PSNR and SSIM on C004 T2w knee MR image. The proposed method obtained the best results in most cases. But when the noise power was 9%, the proposed method got worse SSIM value than Cub-Ave algorithm. In Table SII, the PSNR/SSIM values dropped as the noise level increased. For example, the PSNR/SSIM value was 41.52dB/0.968 when the noise power was 1%, while the PSNR/SSIM value was 36.85dB/0.902 when the noise power was 3%.

IV. The computational time of the proposed algorithm based on different number inputs

In fact, the quality of the reconstructed image improves as a larger number of LR MR images are fused. But the computational time also increases proportional with the number of LR MR images. Table SIII shows the computational time of the different scenarios, including the proposed method based on single, two orthogonal and three orthogonal simulated T2w 3D-MR brain images with 2-7mm slice thickness. When the slice thickness was 3mm, the average computational time of single-frame SR reconstruction was 3.59 min, the computational time of SR reconstruction using two scans was 7.28 min, and the computational time of SR reconstruction using three orthogonal scans was 10.79 min. As the slice thickness increased, the computational time decreased.

Table SI
Accuracy of reconstructed images under the influence of slice thickness on C001 T2w knee MR image

Slice thickness (mm)		2	3	4	5	6	7
Cub-Ave	PSNR (dB)	37.09	33.65	31.59	30.19	28.90	28.11
	SSIM	0.974	0.945	0.918	0.892	0.865	0.845
Cub-Wav	PSNR (dB)	43.56	35.03	32.64	30.92	29.48	28.58
	SSIM	0.994	0.962	0.939	0.912	0.886	0.865
Proposed Method	PSNR (dB)	**47.26**	**40.78**	**37.81**	**35.76**	**34.15**	**33.04**
	SSIM	**0.996**	**0.986**	**0.976**	**0.963**	**0.950**	**0.938**

Table SII
Accuracy of reconstructed images under the influence of noise power on clinical C004 T2w knee MR image

Noise power		0%	1%	3%	5%	7%	9%
Cub-Ave	PSNR (dB)	39.76	38.69	35.89	34.41	33.46	32.71
	SSIM	0.953	0.947	0.888	0.848	0.817	**0.788**
Cub-Wav	PSNR (dB)	41.05	39.32	36.08	34.48	33.48	32.70
	SSIM	0.967	0.953	0.891	0.849	0.816	0.785
Proposed Method	PSNR (dB)	**46.87**	**41.52**	**36.85**	**34.84**	**33.67**	**32.79**
	SSIM	**0.993**	**0.968**	**0.902**	**0.855**	**0.819**	0.785

Table SIII
The computational time of the proposed algorithm based on single, two orthogonal and three orthogonal 3D-MR simulated brain images (in minutes).

Slice thickness (mm)		2	3	4	5	6	7
Single MR Image	Training time	7.11	3.10	1.83	1.05	0.78	0.55
	Reconstruction time	0.90	0.49	0.39	0.30	0.26	0.22
	Total time	8.01	3.59	2.22	1.35	1.04	0.77
Two orthogonal MR image	Training time	14.32	6.29	3.81	2.21	1.47	1.06
	Reconstruction time	1.75	0.99	0.74	0.66	0.51	0.47
	Total time	16.07	7.28	4.55	2.87	1.98	1.53
Three orthogonal MR image	Training time	21.18	9.28	5.55	3.49	2.24	1.60
	Reconstruction time	2.15	1.51	1.17	0.93	0.77	0.70
	Total time	23.33	10.79	6.72	4.42	3.01	2.30

REFERENCES

[1] R. Zeyde, M. Elad, M. Protter, "On single image scale-up using sparse-representations," *in Proc. Of the 7th Inter. Conf. Curves and Surfaces*, Berlin, Heidelberg, 2012, pp. 711-730.

[2] M. Aharon, M. Elad, A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Trans. Signal Proces.*, vol. 54, pp. 4311–4322, Nov. 2006.

[3] Y.C. Pati, R. Rezaiifar, P.S. Krishnaprasad, "Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition," in *Conf. Rec. 27th Asilomar Conf. Signals, Syst. Comput.*, Pacific Grove, CA ,1993, vol. 1, pp. 40-44.

[4] J.A. Tropp, "Greed is good: Algorithmic results for sparse approximation," *IEEE Trans. Inf. Theory*, vol. 50, pp. 2231-2242, Oct. 2004.

[5] Z. Wang, A.C. Bovik, H.R. Sheikh, E.P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, pp. 600–612, Apr. 2004.