# Joint PET+MRI Patch-Based Dictionary for Bayesian Random Field PET Reconstruction

Viswanath P. Sudarshan[1,2,3](✉), Zhaolin Chen[3,4], and Suyash P. Awate[1]

[1] Computer Science and Engineering Department, Indian Institute of Technology (IIT) Bombay, Mumbai, India
psvish@cse.iitb.ac.in
[2] IITB-Monash Research Academy, Indian Institute of Technology (IIT) Bombay, Mumbai, India
[3] Department of Electrical and Computer Systems Engineering, Monash University, Melbourne, Australia
[4] Monash Biomedical Imaging, Monash University, Melbourne, Australia

**Abstract.** Multimodal imaging combining positron emission tomography (PET) and magnetic resonance imaging (MRI) provides complementary information about metabolism and anatomy. While the appearances of MRI and PET images are distinctive, there are fundamental inter-image dependencies relating structure and function. In PET-MRI imaging, typical PET reconstruction methods use priors to enforce PET-MRI dependencies at the very fine scale of image gradients and, so, cannot capture larger-scale inter-image correlations and intra-image texture patterns. Some recent methods enforce statistical models of MRI-image patches on PET-image patches, risking infusing anatomical features into PET images. In contrast, we propose a novel *patch-based joint dictionary model* for PET and MRI, learning regularity in individual patches and correlations in spatially-corresponding patches, for *Bayesian* PET reconstruction using expectation maximization. Reconstructions on simulated and *in vivo* PET-MRI data show that our method gives better-regularized images with smaller errors, compared to the state of the art.

**Keywords:** PET-MRI · Reconstruction · Joint generative model
Joint dictionary model · Patches · Sparsity
Bayesian Markov random field · EM

## 1 Introduction and Related Work

Multimodal imaging systems [14,19] incorporating positron emission tomography (PET) and magnetic resonance imaging (MRI) acquire both functional and anatomical information to improve clinical diagnosis, therapy, and scientific studies in the human body (e.g., neurology and oncology). PET and MRI images

have *distinctive appearances*. The intrinsic spatial resolution of radioactivity in PET, typically 4–6 mm [1], is far lower than the anatomical resolution in MRI, typically 1–1.5 mm. Spatial resolution in PET is limited by positron range, non-collinearity of the annihilation photons, scatter inside the scintillation crystals, finite crystal dimension, interaction depth, etc. In PET-MRI, PET reconstruction is challenged by the stochasticity in gamma-ray emission and the reliability of attenuation-map estimation without the computed tomography (CT) images; thus, it typically relies on *statistical prior* models on the MRI and PET images.

After expectation maximization (EM) based maximum-likelihood (ML) estimation for PET reconstruction [15], edge-preserving smoothness priors and other gradient based penalties within the EM framework were exploited [7]. The method in [4] relies on hierarchical image segmentation, while that in [11] uses voxel-reweighted fidelity based on Fisher information. Some methods use *edge* information in the associated *anatomical* image through edge locations [5,6,8,9] or mutual information of voxel intensities [16]. However, all these priors [5–9,11,16] use *very local neighborhoods* to enforce piecewise-smooth images and cannot capture larger-scale inter-image (structure-function) correlations and intra-image texture patterns in PET images or associated MRI images. In contrast, we propose a framework to model *regularity within* and *statistical dependencies across* spatially-corresponding *patches* in PET and MRI.

For dynamic PET, recent methods [12,13] use hidden Markov random field (MRF) label priors with a Gaussian mixture model for intensities. Some methods [20] use spatially-varying smoothing using kernel similarity on time curves and others [10] use a Bowsher prior to capture edge information from the anatomical image. However, these methods [10,12,13,20] do *not* capture spatial dependencies across PET and MRI. So, we propose a *joint patch-based dictionary* as a *MRF prior* on the PET-MRI image pair for Bayesian PET reconstruction. Early works on patch-based (multimodal) medical image denoising are in [2,3].

Very recent works [17,18] on PET reconstruction use a patch dictionary learned from MRI images, but this strategy risks over/underfitting the functional features of PET images by the anatomical features of MRI. In contrast, we propose a *joint generative model*, based on joint patch regularity, for the pair of PET and MRI images and leverage the joint model to reconstruct PET images. Our model subsumes learning PET image patch statistics by jointly learning statistics of PET image patches and spatially-corresponding MRI patches.

We propose a novel *joint generative model* for the pair of images of PET radiotracer activity and MRI magnitude. We rely on a *joint patch-based sparse dictionary* model formulated as a *MRF prior* for Bayesian PET reconstruction using EM. While our model learns the fine-scale and larger-scale regularity in patches in PET and MRI images individually, it also learns correlations between spatially-corresponding patches in MRI and PET images. Reconstructions on simulated and *in vivo* PET-MRI show that our method produces (i) qualitatively better regularity and (ii) quantitatively lower errors and higher structural similarity (SSIM), over the state of the art.

## 2   Methods

We describe our novel generative model on joint PET-MRI, relying on a sparse
joint-dictionary model, and our Bayesian PET image reconstruction using EM.

### 2.1   Generative Model for PET-MRI Using a Joint Sparse Dictionary

We propose a joint MRF-based sparse dictionary model for the pair of MRI
magnitude and PET activity images. Let $X$ model the MRI image and $Y$ model
the PET activity image, in a common spatial coordinate frame. Let $(X, Y)$ be
a MRF with a neighborhood system $\mathcal{N} := \{\mathcal{N}_i\}_{i=1}^{I}$, where $\mathcal{N}_i$ is the neighbors
of voxel $i$. For any voxel $i$ in the MRI or PET images, its neighbors include all
other voxels in the MRI image within a distance $d_X$ and all other voxels in the
PET image within a distance $d_Y$. This paper sets $d_X := 4$ mm, $d_Y := 8$ mm.
This neighborhood system allows us to model the MRF's Gibbs energy in terms
of square-shaped cliques (patches) $\mathcal{P}_i$ of width 5 mm $(d_X + 1)$ in the MRI image
and patches $\mathcal{Q}_i$ of width 9 mm $(d_Y + 1)$ in the PET image. We model the PET-
MRI joint patch $(X_{\mathcal{P}_i}, Y_{\mathcal{Q}_i})$, at each voxel $i$, as a sparse linear combination of
template patches in the joint dictionary $A$ with $J$ atoms, where the $j$-th atom
comprises a MRI template patch $A_j^X$ paired with a PET template patch $A_j^Y$.
Because the PET and MRI images are non-negative, we model the components of
atoms $A_j^X, A_j^Y$ as non-negative and we enforce non-negativity on the coefficients
in the dictionary fit. The prior $P(X, Y|A) := \eta \exp(-G(X, Y, A))$, where the
Gibbs energy $G(X, Y, A) := \sum_{i=1}^{I} \min_{c_i \succeq 0} \beta \|X_{\mathcal{P}_i} - A^X c_i\|_2^2 + (1 - \beta)\|Y_{\mathcal{Q}_i} - A^Y c_i\|_2^2 + \lambda\|c_i\|_1$, $c_i$ is the common coefficient vector used for fitting patches
in both MRI and PET images at voxel $i$, $c_i \succeq 0$ constraints each coefficient
within $c_i$ to be non-negative, and $\eta$ is the normalizing constant. Free parameter
$\beta \in [0, 1]$ balances the quality of fit in the MRI and PET images, adapted to
their relative noise levels. Free parameter $\lambda \in \mathbb{R}^+$ controls the sparsity of the
dictionary coefficients. We tune free parameters $\beta, \lambda$ using cross validation.

**Learning a Joint PET-MRI Dictionary.** We learn the joint dictionary
$A$, comprising atom pairs $\{(A_j^X, A_j^Y)\}_{j=1}^{J}$, from a training set of $T$ high-
quality PET-MRI images $\{(\mathring{X}^t, \mathring{Y}^t)\}_{t=1}^{T}$, as the maximum-a-priori estimate
$\arg\max_A \prod_{t=1}^{T} P(\mathring{X}^t, \mathring{Y}^t|A)$ subject to scale constraints on the atoms $\|A_j^X\|_2^2 +
\|A_j^Y\|_2^2 \leq 1, \forall j$, positivity constraints on the atoms $A_j^X \succeq 0, A_j^Y \succeq 0, \forall j$, and
positivity constraints on the coefficients $c_i \succeq 0$. Our design of the dictionary
learning (and coding) formulation leads to *smooth convex* problems for optimiz-
ing (i) $A$ given $c$ and (ii) $c$ given $A$; both involve quadratic objective functions
and convex constraint sets. We solve by iterative alternating optimization of $A$
and $c$, where each update uses projected gradient descent with an adaptive step
size to ensure improvement in the objective function value. We initialize atoms
through (i) k-means++ on the subset of joint patches $(\mathring{X}_{\mathcal{P}_i}, \mathring{Y}_{\mathcal{P}_i})$ with variance
significantly higher than the noise variance, excluding constant patches, and

(ii) adding two constant atoms each modeling patches in MRI or PET that are constant or have the minimum intensity as positive.

## 2.2   MAP Reconstruction of PET with Joint-Dictionary Prior, EM

**Likelihood Model for PET.** Given $Y_i$ as the rate of gamma-ray emission counts at each voxel $i$, let $\alpha_{i,d}$ be the emission fraction reaching detector $d$ in a ring of $D$ detectors. We model $\alpha_{i,d}$ as voxel strip integrals, as in STIR (github.com/UCL/STIR). At detector $d$, let the emission rate coming from voxel $i$ be $Y_{i,d} := Y_i\alpha_{i,d}$ and let the total emission rate be $\sum_{i=1}^{I} Y_{i,d}$. Let the observed emission at detector $d$ be $Z_d \sim \text{Poisson}(\sum_{i=1}^{I} Y_{i,d})$. Equivalently, let the part of the observed emission at detector $d$ coming from voxel $i$ be $Z_{i,d} \sim \text{Poisson}(Y_{i,d})$. Then, the likelihood $P(Z|Y) := \prod_{d=1}^{D} P(Z_d|\sum_{i=1}^{I} Y_{i,d}) = \prod_{d=1}^{D}\prod_{i=1}^{I} P(Z_{i,d}|Y_{i,d})$.

Given the observed PET data $\{z_d\}_{d=1}^{D}$, the observed MRI image $X$, and the joint dictionary $A$, our reconstructed PET activity image $Y$ is the MAP estimate: $\arg\max_Y P(Y|z,x,A) = \arg\max_Y P(z|Y)P(Y,x|A)$.

**EM for PET Image Reconstruction.** We propose to solve the MAP estimation problem for parameter $Y$, given observations $\{z_d\}_{d=1}^{D}$, using EM. We model $Z_{i,d}$ as the hidden variable. We initialize the iterative EM optimization with the PET image estimate based on filtered back-projection followed by standard EM [15] without any priors. At iteration $m$, let the current estimate of the PET image be $y^m$. The E step designs the function $Q(Y; y^m) := E_{P(\{Z_{i,d}\}_{i=1,d=1}^{I,D}|y^m,\{z_d\}_{d=1}^{D};X,A)}[\log P(z|Y) + \log P(Y,x|A)]$, where the expectand is the sum of the complete-data log likelihood and the log prior. The M step maximizes the $Q(Y; y^m)$ function to produce the updated reconstruction estimate $y^{m+1}$ using the update rule based on the one-step-late modified-EM [7] strategy adapted to our PET-MRI joint-dictionary modeling framework. Thus, the M step updates the reconstruction estimate, for each voxel $i$, as

$$y_i^{m+1} := \left(\sum_{d=1}^{D} \frac{y_i^m \alpha_{i,d} z_d}{\sum_{i=1}^{I} y_i^m \alpha_{i,d}}\right) \bigg/ \left(\left.\frac{\partial G(y,x,A)}{\partial y_i}\right|_{y_i=y_i^m} + \sum_{d=1}^{D} \alpha_{i,d}\right), \forall i, \quad (1)$$

where we evaluate $\partial G(y,x,A)/\partial y_i$ at $y_i = y_i^m$ by (i) fitting the dictionary $A$ to the image pair $(x, y^m)$, as dictated within $G(y,x,A)$, to produce optimal coefficients $\{c_i^*\}_{i=1}^{I}$ (our formulation makes this is a quadratic programming problem having efficient solvers leading to global optima) and then (ii) taking the partial derivative of $(1-\beta)\sum_{i=1}^{I} \|Y_{\mathcal{Q}_i} - A^Y c_i^*\|_2^2$ with respect to $y_i$. EM iterations stop when the relative change in the estimates $y^m$ and $y^{m+1}$ is small.

## 3   Evaluation, Results, and Discussion

We compare our method with 5 other reconstruction methods: (i) EM without priors (MLEM) [15], (ii) EM with edge-preserving Huber-loss based MRF

prior (HuberMRF), (iii) joint total variation (JTV) prior [6], (iv) parallel level set (PLS) [5,6] prior, (v) MRI-patch dictionary prior for PET reconstruction (MRI-Dict) [18]. For all methods relying on the information in the MRI image, including JTV and PLS, we fix the acquired MRI image to the ground truth. For all methods we tune the underlying free parameters to give the best results qualitatively and quantitatively. We evaluate on 3 datasets: (i) simulated phantom used in [5,6], (ii) simulated BrainWeb used in [5,6], and (iii) *in vivo* PET-MRI. For quantitative evaluation we use SSIM [21] and relative root mean squared error (RRMSE), i.e., ratio of RMSE between $y^{\mathrm{true}}$ and $y^{\mathrm{estimate}}$ to RMS of $y^{\mathrm{true}}$.
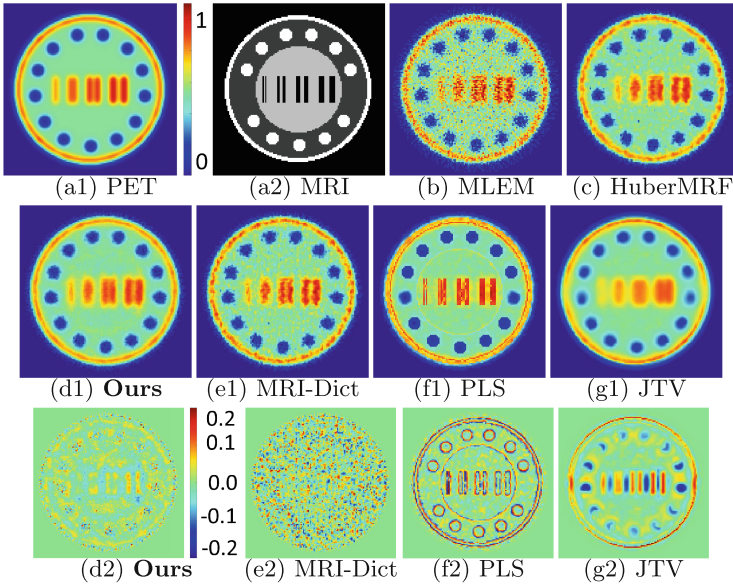


**Fig. 1. Validation: Simulated Phantom. (a1)–(a2)** PET-MRI ground truth. **(b), (c), (d1)–(g1)** PET reconstructions using various methods. **(d2)–(g2)** Residual (reconstructed - truth) images for the results in (d1)–(g1). **RRMSE: Ours 0.06**, MRI-Dict 0.10, PLS 0.09, JTV 0.08, HuberMRF 0.09, MLEM 0.18. **SSIM: Ours 0.92**, MRI-Dict 0.81, PLS 0.86, JTV 0.90, HuberMRF 0.84, MLEM 0.61.

**Validation: Simulated Phantom.** For the simulated phantom (Fig. 1(a1)–(a2)), we sufficiently blur the PET image to reproduce the *lower resolution in PET* [1], relative to MRI. The EM reconstruction without any prior (Fig. 1(b)) retains a lot of the noise compared to prior-based methods. EM with the HuberMRF prior (Fig. 1(c)) gets rid of most of the random noise. JTV and PLS (Fig. 1(g1), (f1)) leverage the anatomical structure in the MRI, encouraging edges in the PET reconstruction to occur at the same spatial locations as the edges in the MRI image. They improve over HuberMRF, but the gradient-based penalty limits the quality of reconstruction of the (i) blue circular blobs,
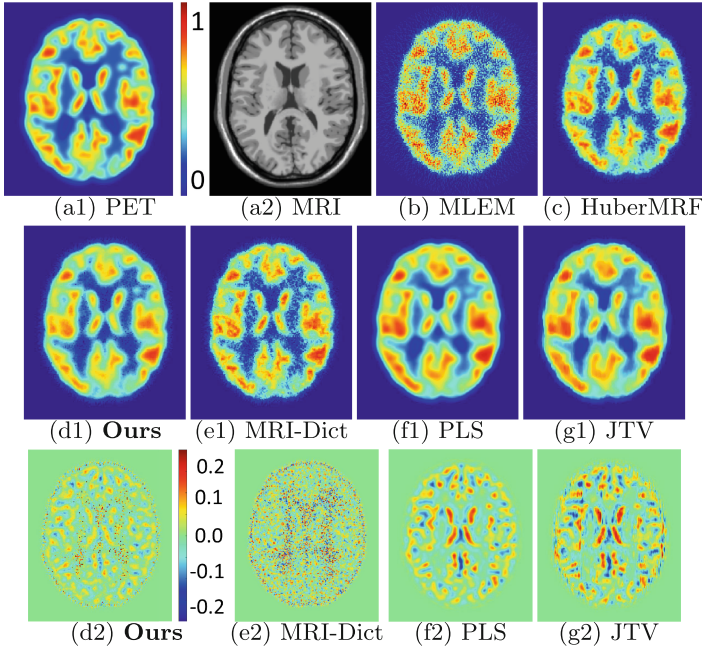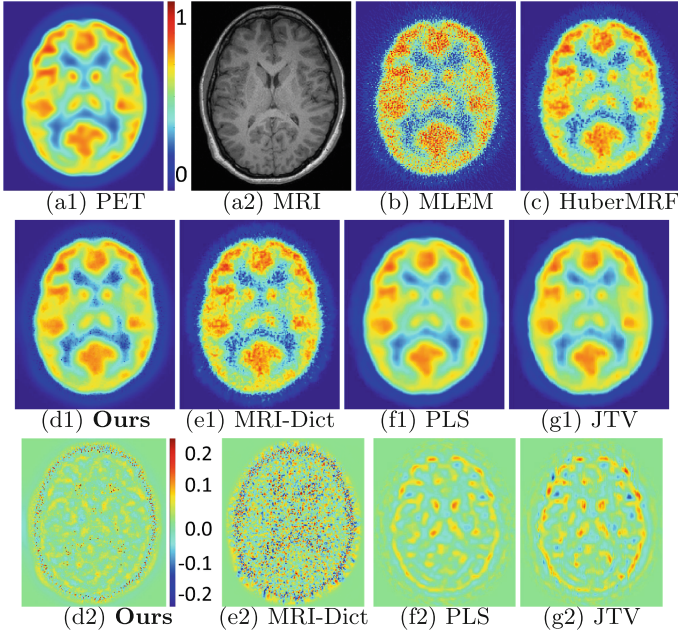
**Fig. 2. Validation: BrainWeb-based Phantom. (a1)–(a2)** PET and MRI ground truth. **(b), (c), (d1)–(g1)** PET image reconstructions using various methods. **(d2)–(g2)** Residual (truth - reconstructed) images for the results shown in (d1)–(g1).

(ii) red circular outside rim, and (iii) red parallel bars in the center. Using a MRI-patch statistical model to reconstruct PET images (PET patch intensities being significantly smoother than MRI patch intensities) results in overfitting of the dictionary to the noise (Fig. 1(e1)). Our reconstruction (Fig. 1(d1)) using a joint patch-based dictionary model maintains both fine-scale regularity, in the form of smoothness, and larger-scale regularity by preservation of structures like the straightness and separability of the red bars, circularity of the blue blobs, and the continuity in the red outer ring. Our reconstruction has much smaller residual magnitudes (Fig. 1(d2)–(g2)) compared to all other methods, and is closest to the ground truth qualitatively and quantitatively (Fig. 1).

**Validation: BrainWeb-Based Phantom.** We simulated PET-MRI from BrainWeb MRI and segmentation, akin to the scheme in [5], but our PET image is much smoother than the MRI image, as exhibited in *in vivo* imaging. We learn our joint dictionary $A$ on patches from five slices and reconstruct 50 slices. Our reconstructions (i) better preserve fine-scale and larger-scale structure (Fig. 2) and (ii) have better RRMSE and SSIM (Fig. 4(c)), over all other methods.

**Results: *In vivo* Brain PET-MRI.** We collected data for 5 subjects using a 3T PET-MRI Siemens scanner (PET slice thickness 2 mm and T1 MRI 1 mm$^3$ voxels). We learn the joint patch dictionary from one subject and reconstruct

**Fig. 3. Results on *in vivo* Brain PET-MRI. (a1)–(a2)** PET and MRI ground truth. **(b), (c), (d1)–(g1)** PET image reconstructions using various methods. **(d2)–(g2)** Residual (truth - reconstructed) images for the results shown in (d1)–(g1).
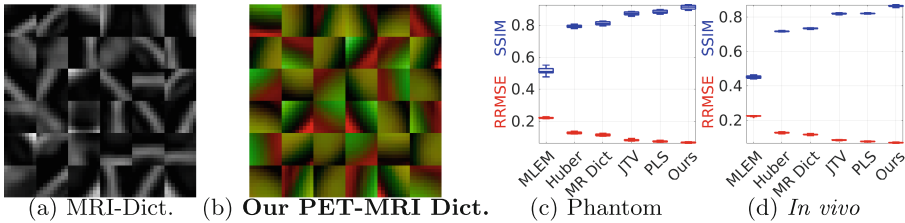


**Fig. 4. Results: Quantitative Evaluation. (a)** Dictionary of MRI patches, used by [18] (MR-Dict). **(b)** Our joint dictionary of PET-MRI patches (MRI patch appears as red channel; corresponding PET patch as green channel). **(c)–(d)** Box plots for RRMSE and SSIM for 50 slices from BrainWeb simulation and 4 subjects *in vivo*.

50 slices from the other subjects. Unlike the dictionary of MR patches in MRI-Dict (Fig. 4(a)), our joint dictionary (Fig. 4(b)) captures regularity in individual patches (which is distinctive for PET and MRI) and PET-MRI correlations in spatially-corresponding patches. Our reconstructions are qualitatively (Fig. 3) and quantitatively (Fig. 4(d)) better than all other methods.

**Conclusion.** We propose a novel *joint generative model* for the PET-MRI image pair, relying on a *joint patch-based dictionary model*, formulated as a *MRF prior*

for Bayesian PET reconstruction using EM. While our model learns the fine-scale and larger-scale regularity in patches in PET and MRI images individually, it also learns structure-function correlations between spatially-corresponding patches in MRI and PET images. Our reconstructions on simulated and *in vivo* PET-MRI improve qualitatively and quantitatively over the state of the art.

# References

1. Aklan, B., et al.: Impact of point-spread function modeling on PET image quality in integrated PET/MR hybrid imaging. J. Nucl. Med. **57**(1), 78–84 (2016)
2. Awate, S.P., Whitaker, R.T.: Nonparametric neighborhood statistics for MRI denoising. In: Christensen, G.E., Sonka, M. (eds.) IPMI 2005. LNCS, vol. 3565, pp. 677–688. Springer, Heidelberg (2005). https://doi.org/10.1007/11505730_56
3. Awate, S.P., Whitaker, R.T.: Feature-preserving MRI denoising: a nonparametric empirical-Bayes approach. IEEE Trans. Med. Imag. **26**(9), 1242–55 (2007)
4. Bowsher, J., Johnson, V., Turkington, T., Jaszczak, R., Floyd, C., Coleman, R.: Bayesian reconstruction and use of anatomical a priori information for emission tomography. IEEE Trans. Med. Imag. **15**(5), 673–86 (1996)
5. Ehrhardt, M., et al.: PET reconstruction with an anatomical MRI prior using parallel level sets. IEEE Trans. Med. Imag. **35**(9), 2189–99 (2016)
6. Ehrhardt, M., et al.: Joint reconstruction of PET-MRI by exploiting structural similarity. Inverse Probl. **31**(1), 015001 (2014)
7. Green, P.: Bayesian reconstructions from emission tomography data using a modified EM algorithm. IEEE Trans. Med. Imag. **9**(1), 84–93 (1990)
8. Knoll, F., Holler, M., Koesters, T., Otazo, R., Bredies, K., Sodickson, D.: Joint MR-PET reconstruction using a multi-channel image regularizer. IEEE Trans. Med. Imag. **36**(1), 1–16 (2017)
9. Leahy, R., Yan, X.: Incorporation of anatomical MR data for improved functional imaging with PET. In: Colchester, A.C.F., Hawkes, D.J. (eds.) IPMI 1991. LNCS, vol. 511, pp. 105–120. Springer, Heidelberg (1991). https://doi.org/10.1007/BFb0033746
10. Loeb, R., Navab, N., Ziegler, S.: Direct parametric reconstruction using anatomical regularization for simultaneous PET/MRI data. IEEE Trans. Med. Imag. **34**(11), 2233–47 (2015)
11. Nuyts, J., Fessler, J.: A penalized-likelihood image reconstruction method for emission tomography, compared to post-smoothed maximum-likelihood with matched spatial resolution. IEEE Trans. Med. Imag. **22**(9), 1042–52 (2003)
12. Pedemonte, S., Bousse, A., Hutton, B.F., Arridge, S., Ourselin, S.: 4-D generative model for PET/MRI reconstruction. In: Fichtinger, G., Martel, A., Peters, T. (eds.) MICCAI 2011. LNCS, vol. 6891, pp. 581–588. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-23623-5_73
13. Pedemonte, S., Cardoso, M.J., Arridge, S., Hutton, B.F., Ourselin, S.: Steady-state model of the radio-pharmaceutical uptake for MR-PET. In: Ayache, N., Delingette, H., Golland, P., Mori, K. (eds.) MICCAI 2012. LNCS, vol. 7510, pp. 289–297. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-33415-3_36
14. Pichler, B., Wehrl, H., Kolb, A., Judenhofer, M.: PET/MRI: the next generation of multi-modality imaging? Semin. Nucl. Med. **38**(3), 199–208 (2008)
15. Shepp, L., Vardi, Y.: Maximum likelihood reconstruction for emission tomography. IEEE Trans. Med. Imag. **1**(2), 113–22 (1982)

16. Somayajula, S., Panagiotou, C., Rangarajan, A., Li, Q., Arridge, S., Leahy, R.: PET image reconstruction using information theoretic anatomical priors. IEEE Trans. Med. Imag. **30**(3), 537–49 (2011)
17. Tahaei, M., Reader, A.: Patch-based image reconstruction for PET using prior-image derived dictionaries. Phys. Med. Biol. **61**(18), 6833 (2016)
18. Tang, J., Wang, Y., Yao, R., Ying, L.: Sparsity-based PET image reconstruction using MRI learned dictionaries. In: IEEE 11th International Symposium on Biomedical Imaging (ISBI), p. 1087 (2014)
19. Vandenberghe, S., Marsden, P.: PET-MRI: a review of challenges and solutions in the development of integrated multimodality imaging. Phys. Med. Biol. **60**, R115 (2015)
20. Wang, G., Qi, J.: PET image reconstruction using kernel method. IEEE Trans. Med. Imag. **34**(1), 61–71 (2015)
21. Wang, Z., Bovik, A., Sheikh, H., Simoncelli, E.: Image quality assessment: from error visibility to structural similarity. IEEE Trans. Imag. Proc. **13**(4), 600 (2004)