Evaluation of Wavelet Kernel-Based PET Image Reconstruction

Zahra Ashouri, Guobao Wang, *Senior Member, IEEE*, Richard M. Dansereau, *Senior Member, IEEE*, and Robert A. deKemp, *Member, IEEE*

***Abstract*—Positron emission tomography (PET) imaging is used to track biochemical processes in the human body. PET image quality is limited by noise and several methods have been implemented to improve the quality. Kernel-based image reconstruction is among the methods implemented to increase PET image quality and commonly uses a Gaussian kernel to include spatial correlations from image priors into the forward projection model of PET. Unfortunately, the Gaussian kernel tends to smooth details in the reconstructed image. To reduce noise without losing contrast details, a different kernel is needed. Wavelet kernels can be more efficient than the Gaussian kernel in reducing noise while keeping contrast details by better separating signal from noise and thus, it does not over smooth peak values in the final reconstructed images. In this work, we evaluate a wavelet kernel for kernel-based PET image reconstruction. For this goal, a wavelet kernel approach has been tested using simulated brain data, physical phantom data, and patient data. Reconstruction results are presented and discussed in detail comparing the wavelet kernel method with the Gaussian kernel method. Our results suggest that a wavelet kernel performs better in contrast recovery for phantoms and also results in higher signal-to-noise ratio (SNR) for real patient data.**

***Index Terms*—Kernel-based expectation maximization (KEM), positron emission tomography (PET), image reconstruction, wavelet kernel.**

I. Introduction

POSITRON emission tomography (PET) imaging monitors a patient's *in vivo* radiotracer distribution. PET scanners acquire functional information of physiological processes in the body by measuring the coincident annihilation photons ejected from the radiotracer in the patient's body. These measurements are then reconstructed into cross-sectional images by iterative reconstruction methods.

Maximum-likelihood (ML) expectation maximization (MLEM) is an iterative reconstruction method widely used for PET image reconstruction. This method includes a Poisson model in the reconstruction and results in a closed-form iterative update for reconstruction [2]. Different variations of ML algorithms were developed afterward that tried to overcome the shortcomings of MLEM. MLEM is slow to converge and therefore, a faster variation was proposed by Hudson and Larkin called ordered subset expectation maximization (OSEM) [3]. Another problem with ML-based algorithms is their ill-posed nature where low count statistics will present itself as high noise in the reconstructed image. To improve PET image quality, several methods have been proposed from filtering the reconstructed image to modeling the noise in the reconstruction algorithm and adding it as a smoothing penalty to the likelihood function [4]–[7]. These methods mostly include prior information to improve PET image quality [8].

Prior information can come from other image modalities, mostly the anatomical image modalities, such as computed tomography (CT) or magnetic resonance imaging (MRI). Several methods have used anatomical information as a regularizer in the likelihood function. Bowsher *et al.* [4] proposed a model that imposes greater smoothing among nearby voxels in the PET image when they have higher similarity in MRI signals. A number of papers use the concept of nonlocal means (NLMs) for defining their regularizer term. Nguyen and Lee [9] developed a nonlocal regularizer where the weight is defined based on similarity between two patches in the PET image and also the similarity between two patches centered in the anatomical image. Compressed sensing was also proposed to explore the sparsity in CT image sequences to reconstruct dynamic CT images from highly under sampled projection data sets [5]. This compressed sensing technique can be applied to PET image reconstruction as well [10].

An alternative approach to defining a regularizer is to add a denoising step at the end of image reconstruction. An NLM filter has been applied for denoising the reconstructed image in [6] and [7]. Newer methods have applied deep learning to PET image denoising [11] where a deep neural network is trained for PET image denoising.

The noise in PET images can be tackled by incorporating image prior information directly into the forward projection

Manuscript received June 19, 2021; revised July 28, 2021; accepted August 2, 2021. Date of publication August 6, 2021; date of current version May 3, 2022. This work was supported by the Natural Sciences and Engineering Research Council of Canada. *(Corresponding author: Zahra Ashouri.)*

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Ottawa Health Science Network Research Ethic Boards under protocol No. 20140417-01H.

Zahra Ashouri is with the Cardiac Imaging, Ottawa Heart Institute, Ottawa, ON K1Y 4W7, Canada, and also with the Systems and Computer Engineering, Carleton University, Ottawa, ON K1S 5B6, Canada (e-mail: zashouri@ottawaheart.ca).

Guobao Wang is with the Department of Radiology, University of California at Davis, Davis, CA 95616 USA (e-mail: gbwang@ucdavis.edu).

Richard M. Dansereau is with the Department of Systems and Computer Engineering, Carleton University, Ottawa, ON K1S 5B6, Canada (e-mail: rdanse@sce.carleton.ca).

Robert A. deKemp is with the Cardiac Imaging, Ottawa Heart Institute, Ottawa, ON K1Y 4W7, Canada, and also with the Faculty of Medicine, Ottawa University, Ottawa, ON K1Y 4W7, Canada (e-mail: radekemp@ottawaheart.ca).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TRPMS.2021.3103104.

Digital Object Identifier 10.1109/TRPMS.2021.3103104

model in the form of a kernel [12]–[21]. These papers use either dynamic prior information [12], [13] or MR anatomical information [14]–[18]. The kernel-based expectation maximization (KEM) method proposed by Wang and Qi forms a feature vector from prior information [12]. A set of features from the prior information is extracted for each pixel and the PET image intensity at the specific pixel is a linear function of those features. Then, the kernel method is used to calculate the PET image intensity from the features. Other forms of kernel methods have been proposed in the literature as well. Wang [19] has extended the spatial kernel method into a more generalized spatiotemporal kernel method. Huang [20] has combined denoising with kernel-based reconstruction by applying NLMs denoising after each iteration of KEM update. In that article, the author defined the kernel as a multiplication of the spatial kernel with a temporal kernel that is obtained by comparing feature vectors of the different frames. Gong *et al.* [21] used a kernel method to incorporate MRI information into a Patlak reconstruction model.

Most of the existing work uses a Gaussian kernel [12], but the Gaussian kernel has a smoothing effect on the image, which might cause small details in the image to be smoothed out.

In order to reduce noise while preserving small details, in this article, similar to our initial work [1], we propose using a wavelet kernel for kernel-based image reconstruction with dynamic images as our source of prior information and we have evaluated the method with different datasets.

A part of this work was previously presented in a conference [1]. This work has been substantially extended by an extensive simulation study and an evaluation using physical phantom data and real patient data. The remainder of this article is organized as follows. In Section II, the iterative reconstruction method and kernel-based reconstruction with conventional Gaussian kernel are described briefly. In Section III, the wavelet kernel used is explained. Then, in Section IV, the simulation study is discussed, and the real data study is presented in Section V. Finally, the article is then concluded in Section VI.

II. PET Image Reconstruction

A. Iterative PET Image Reconstruction

The reconstructed image in iterative PET image reconstruction is found by solving an ill-posed inverse problem to find emission image $\mathbf{x}$ from PET projection data $\mathbf{y}$ as

$$\bar{\mathbf{y}} = \mathbf{E}[\mathbf{y}] = \mathbf{P}\mathbf{x} + \mathbf{r} + \mathbf{s} \tag{1}$$

where $\mathbf{P}$ is the projection matrix, $\mathbf{r}$ is random events, $\mathbf{s}$ is scatter events, and E[.] is expectation. To solve (1), the conditional probability or likelihood function for $\mathbf{y}$ given emission image $\mathbf{x}$ is estimated by independent Poisson distributions as given by [8]

$$\mathbf{p}(\mathbf{y}|\mathbf{x}) = \prod_{\mathbf{i}}^{M} \mathbf{e}^{-\bar{\mathbf{y}}_\mathbf{i}} \frac{\bar{\mathbf{y}}_\mathbf{i}^{\mathbf{y}_\mathbf{i}}}{\mathbf{y}_\mathbf{i}!}. \tag{2}$$

The ML estimate of image $\mathbf{x}$ can be found by maximizing (2) or similarly maximizing its log-likelihood equivalent as

$$\mathbf{L}(\mathbf{y}|\mathbf{x}) = \sum_{\mathbf{i}=\mathbf{1}}^{M} \mathbf{y_i log}\left(\bar{\mathbf{y}}_\mathbf{i}\right) - \bar{\mathbf{y}}_\mathbf{i} - \mathbf{log}(\mathbf{y_i}!) \tag{3}$$

where M is the total number of lines of response. The ML estimate of the emission image $\mathbf{x}$ is found by

$$\widehat{\mathbf{x}} = \mathbf{arg}\ \max_{\mathbf{x}\geq\mathbf{0}} \mathbf{L}(\mathbf{y}|\mathbf{x}). \tag{4}$$

The expectation maximization (EM) algorithm for PET image reconstruction, proposed by Shepp and Vardi [2], is commonly used to find the solution to (4) by the iterative update

$$\mathbf{x}^{n+1} = \frac{\mathbf{x}^n}{\mathbf{P}^\mathrm{T}\mathbf{1}_M}\left(\mathbf{P}^\mathrm{T}\frac{\mathbf{y}}{\mathbf{P}\mathbf{x}^n + \mathbf{r} + \mathbf{s}}\right) \tag{5}$$

where $\mathbf{1}_M$ is a vector of length M with elements of 1, n denotes the iteration number, and the superscript "T" denotes a matrix transpose. The vector multiplication and division in (5) are elementwise operations [12].

B. Kernel EM for PET Image Reconstruction

In KEM, the information in the image representation is encoded by using a kernel [12], [22]. The basic idea is to represent the PET image $\mathbf{x}$ by a linear function of transformed features in a high-dimensional space. A feature vector $\mathbf{f_j}$ is picked out for pixel j with pixel intensity x_j. The pixel intensity in the reconstructed image is represented by a linear combination of feature vectors corresponding to the neighboring pixels. From the feature vectors, the kernel is constructed. Using the kernel coefficients, the reconstructed image is formed using

$$x_j = \sum_{\mathbf{l}=\mathbf{1}}^{\mathbf{N}} \alpha_\mathbf{l} \boldsymbol{\kappa}\left(\mathbf{f_j}, \mathbf{f_l}\right) \tag{6}$$

where α_l is the kernel coefficient and $\kappa(\mathbf{f_j}, \mathbf{f_l})$ measures the similarity of features between pixel x_j and pixel x_l in its neighborhood. $\mathbf{N}$ is the total number of pixels in the image. The matrix form of (6) is

$$\mathbf{x} = \mathbf{K}\boldsymbol{\alpha} \tag{7}$$

where the (j, l)th entry of matrix $\mathbf{K}$ is given by $\kappa(\mathbf{f_j}, \mathbf{f_l})$ and $\boldsymbol{\alpha}$ is a vector with elements α_l. Typically, a Gaussian kernel is used, however, other kernels can be chosen, such as the polynomial kernel [12]. The kernel representation of the image $\mathbf{x}$ can be substituted in the PET forward projection model (1) and then the resulting iterative update replacing (5) would be

$$\boldsymbol{\alpha}^{\mathbf{n+1}} = \frac{\boldsymbol{\alpha}^{\mathbf{n}}}{\mathbf{K}^\mathbf{T}\mathbf{P}^\mathbf{T}\mathbf{1_M}} . \left(\mathbf{K}^\mathbf{T}\mathbf{P}^\mathbf{T}\frac{\mathbf{y}}{\mathbf{P}\mathbf{K}\boldsymbol{\alpha}^{\mathbf{n}} + \mathbf{r} + \mathbf{s}}\right) \tag{8}$$

where $\mathbf{n}$ is the iteration number. Using existing OSEM reconstruction with the kernel inserted in the iterative update, the above algorithm can be solved. From it, the PET image is estimated using (7).

C. Dynamic PET Composite Frames and Gaussian Kernel

In order to use the kernel method, we first need to define the source of our prior information. One method of creating the prior information is to form composite image frames. The summation of raw projection data from multiple frames of the dynamic series forms composite frames. This method helps to improve counting statistics and therefore, reduces noise. There is a tradeoff in setting the number of composite frames. A large number of composite frames mean that fewer time frames are used for ensemble averaging to form each composite frame; this will preserve temporal information but noise reduction is hindered in the composite frames. On the other hand, a small number of composite frames, which is the case when more time frames are ensemble averaged to form the composite frames, will result in lower noise but temporal information related to redistribution of the tracer to different tissues may be lost.

After defining composite frames in the projection space, OSEM reconstruction is used to reconstruct composite frames. OSEM reconstruction would be fast with only a few composite frames in comparison to the whole dynamic series. The pixel values in the reconstructed composite frames form the feature vectors as

$$\mathbf{f_j} = \left[x_{j,1}^{comp}, x_{j,2}^{comp}, x_{j,3}^{comp}, \ldots\right] \tag{9}$$

where $x_{j,t}^{\text{comp}}$ is the image intensity value at pixel j in composite frame t. From the feature vectors, the kernel is constructed. The neighborhood can be defined by different methods. One method is to use the k-nearest neighbor (kNN) method [23], which finds the k closest neighbors for each pixel based on the Euclidean distance between $\mathbf{f_j}$ and $\mathbf{f_l}$. Alternatively, a cubic window centered on a pixel of interest can be used to define where its neighboring pixels are located. Then, the chosen kernel is used to calculate the weights based on the neighboring pixels. For the Gaussian kernel, the (j, l)th element of the kernel matrix is calculated as

$$\mathbf{K_{j,l}} = \mathbf{exp}\left(-\frac{\left\|\mathbf{f_j} - \mathbf{f_l}\right\|^2}{\mathbf{2\sigma^2}}\right) \tag{10}$$

where $\mathbf{K_{j,l}}$ is a measure of the similarity of features $\mathbf{f_j}$ and $\mathbf{f_l}$ between pixel x_j and pixel x_l in its neighborhood and σ is the chosen standard deviation (SD). The choice of the neighborhood parameter is based on the voxel size and the tissue being imaged.

III. Proposed Wavelet Kernel Method

Although the Gaussian kernel increases signal-to-noise ratio (SNR) in reconstructed images [12], it tends to over smooth small targets and tissue boundaries. To address this problem, we can propose a kernel that does not have the same smoothing effect of the Gaussian kernel.

A principal component analysis (PCA) can be used to map the data into a new feature space that contains more linearly separable features than those in the original input space. The transformation from an original input space into a transform feature space in which the data are projected onto linear components can be done through the use of kernels. This mapping is referred to as kernel PCA (KPCA) [24].

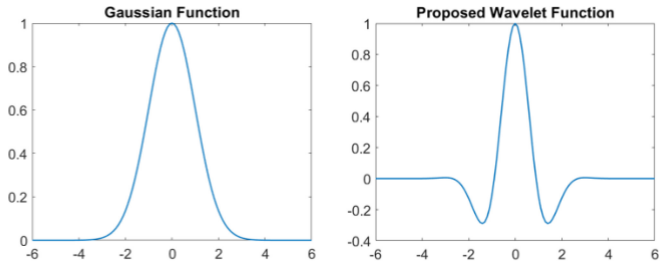


Fig. 1. 1-D Gaussian function versus proposed Morlet wavelet.

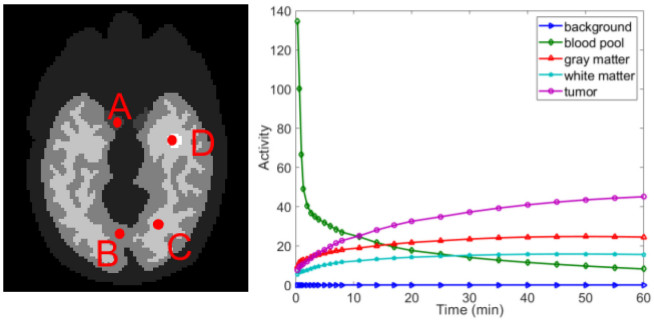
Wavelet KPCA can be used to better capture data similarity measures in the kernel matrix [24]. Wavelet kernels are constructed from a given mother wavelet function to improve the performance of KPCA as the feature extraction method. Many options exist for the mother wavelet, which is just a basis function satisfying the admissibility condition that is shifted and scaled in a wavelet transform. Because of their high flexibility, wavelet kernels have been successfully used in support vector machine (SVM) learning for classifying data [24]. Besides classification, wavelet kernels can be used in the process of feature extraction [24]. One form of single-scale translation invariant (SSTI) wavelet kernel is given as

$$\mathbf{g(v, w)} = \prod_{\mathbf{i=1}}^{\mathbf{N}} \mathbf{h}\left(\frac{\mathbf{v_i} - \mathbf{w_i}}{\mathbf{a}}\right) \tag{11}$$

where $\mathbf{v}, \mathbf{w} \in R^n$, $\mathbf{a}$ is the dilation coefficient, and $\mathbf{h}(\cdot)$ is the mother wavelet. For an SSTI kernel to be used as a kernel in KPCA, the kernel matrix constructed from the SSTI kernel should be positive semidefinite [24]. We used an SSTI wavelet kernel based on the Morlet mother wavelet function given in [25] as

$$\mathbf{h(z)} = \mathbf{cos(1.75z)}\ \mathbf{exp}\left(-\frac{\mathbf{z^2}}{\mathbf{2}}\right). \tag{12}$$

In general, a (real) Morlet mother wavelet consists of a sinusoid $[\cos(\omega z)]$ windowed by a Gaussian function $[\exp(-[z^2/2a^2])]$, so other choices for the parameters can also be made. Other mother wavelets can also be considered, but the Morlet wavelet is often selected given the clear use of a localized frequency. The central frequency ω of the mother wavelet (1.75 in our paper) corresponds to the number of oscillations within the Gaussian window. As stated in the discussion section of [26], which tests different values of ω, low values of ω improve the location detection capabilities of the wavelet. The 1-D form of the Gaussian function and the proposed wavelet function is shown in Fig. 1. Due to the valleys in the wavelet function around peak, the proposed wavelet kernel should be better able to preserve edge information and show higher contrast between the edges and their surrounding compared to the Gaussian function.



(a) (b)

Fig. 2. (a) Digital phantom with A, B, C, and D being regions in the blood pool, gray matter, white matter, and tumor region, respectively. (b) Time activity curve for different regions.

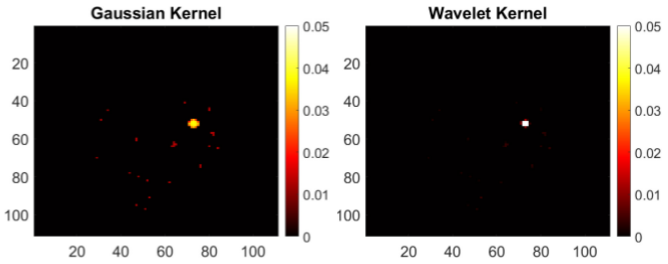


Fig. 3. Basis image reshaped from a column of the two kernel matrices that corresponds to a pixel in the tumor region of the simulated brain.

In (11), $\mathbf{v_i}$ and $\mathbf{w_i}$ are replaced by the features $\mathbf{f^i_j}$ and $\mathbf{f^i_l}$. Therefore, the **(j, l)**th element of the matrix **K** would become

$$\mathbf{K_{j,l}} = \prod_{\mathbf{i=1}}^{\mathbf{Q}} \left[\cos\left(1.75 \left(\frac{\left(\mathbf{f^i_j} - \mathbf{f^i_l}\right)}{\mathbf{a}} \right) \right) \exp\left(- \frac{\left(\mathbf{f^i_j} - \mathbf{f^i_l}\right)^2}{\mathbf{2a^2}} \right) \right] \quad (13)$$

where Q is the number of composite frames. Although the kernel function will have negative values, as explained in [25, Sec. III-B], this kind of wavelet kernel is an admissible kernel since it satisfies Mercer's condition. This wavelet kernel-based image reconstruction is referred to as WKEM in this manuscript.

IV. Simulation Study

A. Simulation Setup

For the simulation study, dynamic PET scans were simulated similar to Wang *et al.*'s paper [12] using a Zubal head phantom [see Fig. 2(a)] [27]. The scanning time consisted of 24 time frames: 4×20, 4×40, 4×60, 4×180, and 8×300 s. Regional time activity curves shown in Fig. 2(b) were assigned to different brain regions. The simulated brain dataset contains dynamic 2-D images of the brain. By forward projecting dynamic activity images and adding scatter and random noise, sinograms were generated. The expected total number of coincidence events over 60 min was 16 million. Ten realizations were generated with the same scatter and noise level.

Fig. 3 shows the basis image reshaped from a column of the Gaussian and wavelet kernel matrices. The chosen column refers to a pixel in the tumor region of the simulated brain data. In this figure, we can see that the chosen pixels in the wavelet kernel are more localized, meaning that the wavelet kernel is able to better identify similar pixels; therefore, it can preserve edge properties better.

B. Implementation Parameters

We chose to have three composite frames, each corresponding to 20 min of the scan time: the first frame consists of first 16 time frames and largely corresponds to the tracer in the blood, the second composite frame consists of the next four time frames and largely refers to the transition of the tracer from blood to tissue, and the final composite frame consists of the last four time frames and largely contains the uptake in important organs and tissues. For the simulation study, a kNN with 48 nearest neighbors was used for both Gaussian kernel and wavelet kernel. Following [12], the kernel parameter σ for the Gaussian kernel was set to 1 and for the wavelet kernel, different values for **a** were tested. For values smaller than 1, the wavelet was too thin and did not produce acceptable reconstruction results. For values larger than 1 mean square error (MSE) increased for most of the frames with the increase in **a** value, especially in early frames in which the WKEM showed greater improvement. Therefore, **a** was chosen to be equal to 1 for the wavelet kernel.

C. Evaluation Method

For each dataset, the evaluation method is defined based on the information available. For the simulation study, the ground truth is available and, therefore, SNR and contrast recovery coefficient (CRC) are defined accordingly and based on using different realizations. For quantitative comparison between different reconstruction methods, CRC of the blood pool and tumor is also calculated by simulating ten noisy realizations and reconstructing them independently. The CRC is calculated by

$$\mathbf{CRC} = \frac{\frac{1}{10}\sum_{i=1}^{10}\left(\frac{R_i - yB_i}{B_i}\right)}{\left(\frac{R_{true} - B_{true}}{B_{true}}\right)} \quad (14)$$

where R refers to region of interest (ROI), B is the background region and i is the ith realization of the simulated data. The white matter was the background region. Background SD was derived pixelwise over ten realizations and was then averaged over all the pixels in the background region and normalized by the true background value.

D. Results

Fig. 4 shows images of the reconstructed brain phantom for two sample time frames using 40 iterations. The WKEM achieved a higher SNR than KEM for the early Frame 2 in which the blood pool has a high activity. For the late Frame 24, the two approaches had the same SNR. Nonetheless, the shape of the tumor using WKEM is more consistent with the ground truth. In Fig. 5, CRC versus background SD percentage is shown for the blood pool and the tumor region by varying the iteration number from 20 to 60 with 20 iterations having

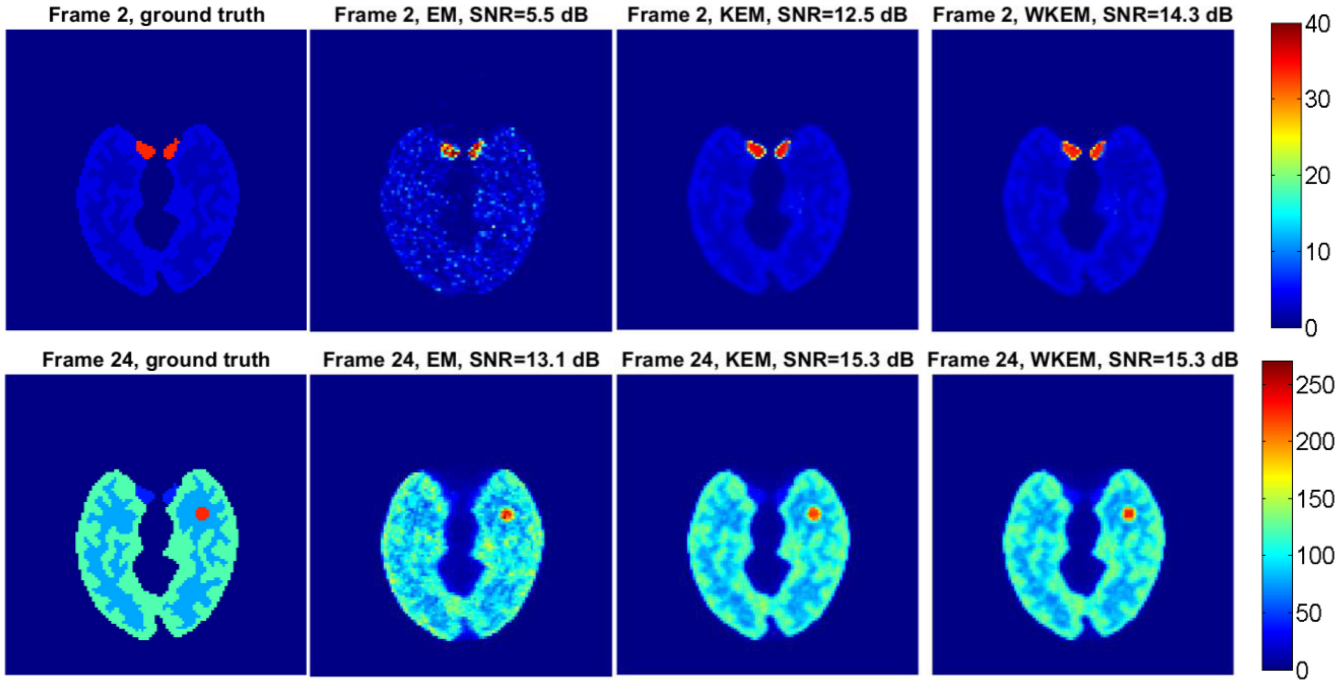


Fig. 4. Ground truth and reconstructed images of three different reconstruction methods for two different time frames of the brain phantom.

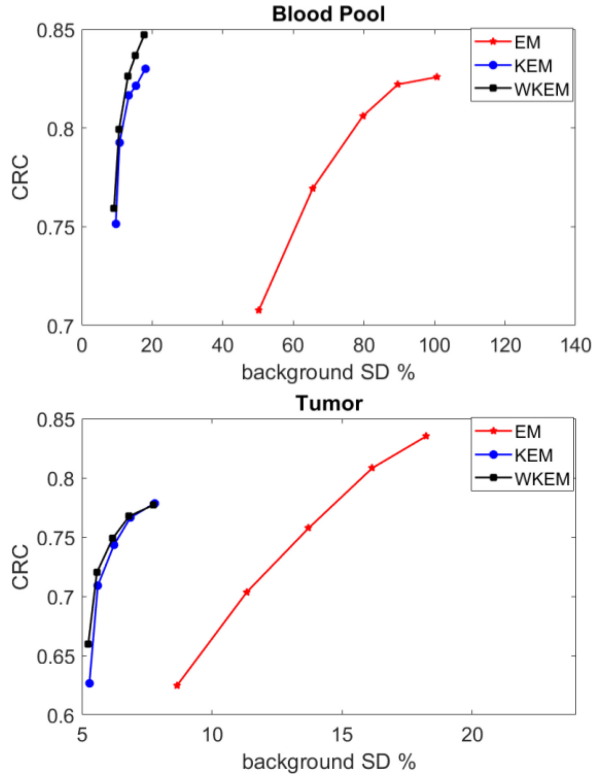


Fig. 5. Contrast recovery versus background standard deviation for Zubal head phantom. Blood pool is calculated in frame 2 and the tumor is calculated in frame 24.

the smallest CRC value and 60 iterations having the largest CRC value in each method. WKEM had a higher CRC than KEM while with a similar background SD. The improvement for the blood pool is higher than for the tumor. This result is also seen in Fig. 4 where SNR and the visual quality of

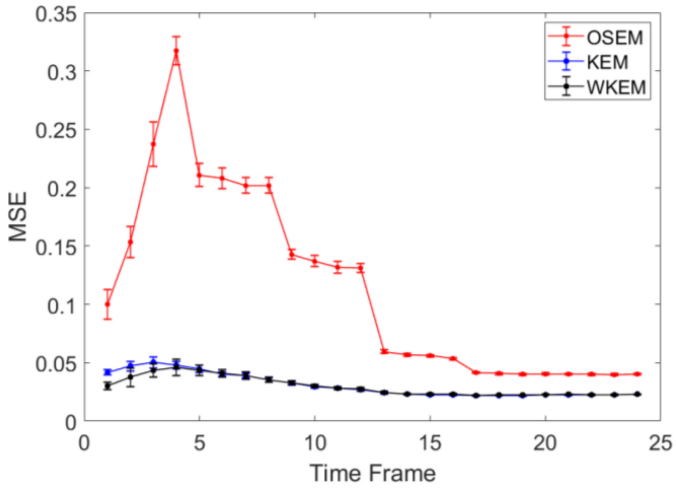


Fig. 6. Mean square error (MSE) for different reconstruction methods shown for each frame with the standard deviation error bars for ten realizations.

the reconstructed images are higher with WKEM for the low count frames, and for the higher count frames (later frames), WKEM is acting similar to KEM. MSE with error bar for all three reconstruction methods is shown in Fig. 6. In this figure, the iteration number is set to 40 and the error bars show the variability between ten realizations. The variability is the SD of the MSE over ten realizations. This figure shows that WKEM has a smaller MSE than KEM for earlier frames, which is consistent with the SNR trends from Fig. 4. For later frames, KEM and WKEM have similar performance in terms of MSE.

V. Real Data

We conducted a physical phantom study and real patient study to further evaluate the WKEM reconstruction method.

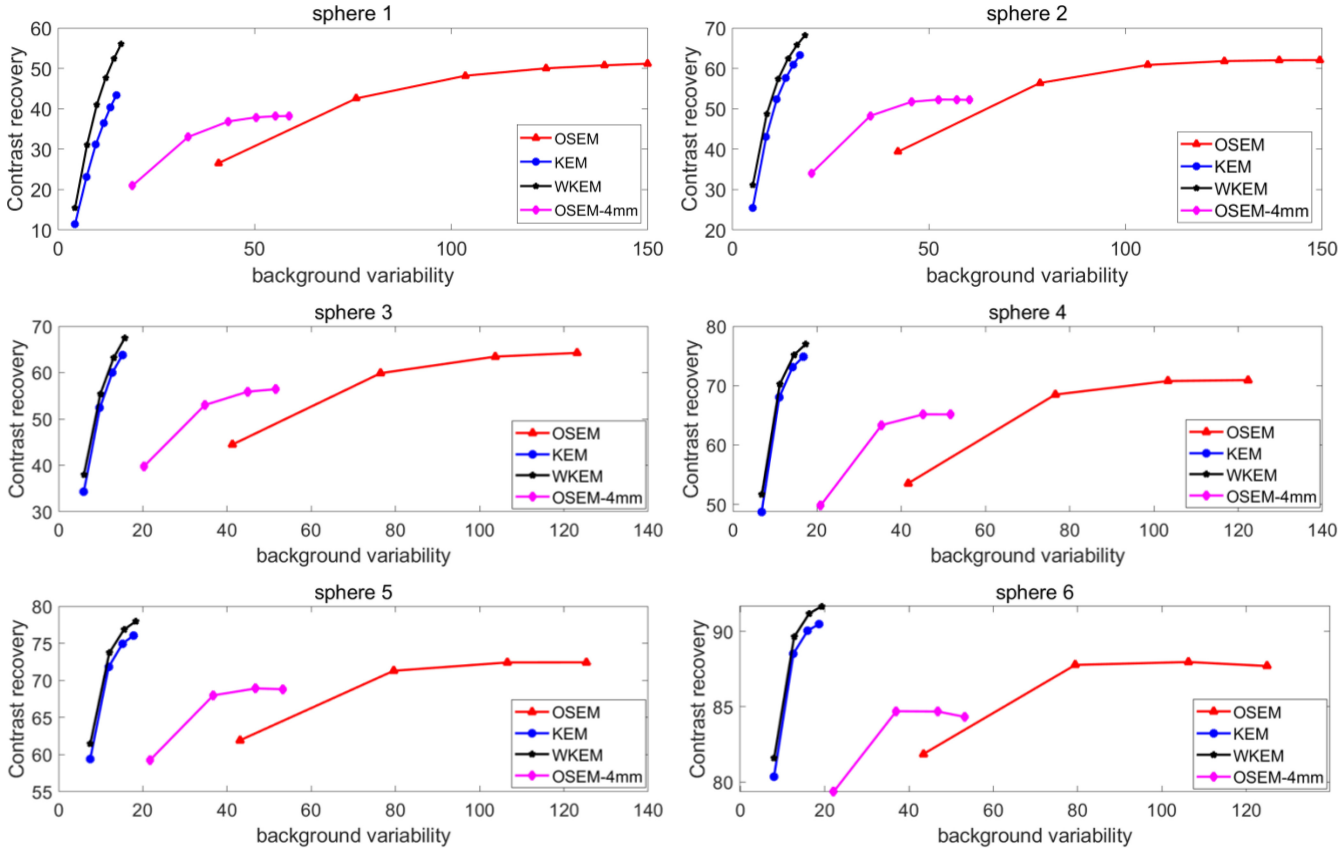


Fig. 7. Contrast recovery versus background variability for all sphere sizes, with sphere 1 being the smallest sphere and sphere 6 being the largest sphere.

For the implementation of the kernel method in each study, a cubic window was chosen as the neighborhood and different sizes of the neighborhood window were tested for each study. A window of $7 \times 7 \times 7$ showed effective in noise reduction while not losing significant features of the tissue being imaged. Note that using this local neighborhood instead of global kNN decreases the computational complexity such that the method is more practical for 3-D reconstruction [12]. To include the spatial distance between the pixel of interest and neighboring pixels in the cubic neighborhood, we have given a Gaussian weight to the pixels in the neighborhood such that the pixel of interest has the largest weight and as the pixels are further away in the neighborhood, they would have smaller weight in the kernel. These weights are then multiplied in the calculation of kernel coefficients in (10) and (13). This spatial weighting was not applied in the simulated 2-D brain data since using a global kNN was feasible. However, for the real data, we have dynamic volumetric data, which means the data in each time frame are 3-D and kNN is not desirable since it is computationally inefficient and therefore, spatial weighting is included in the kernel construction. To make the kernel sparser, the 50 highest weight values in each neighborhood are selected to form the matrix and the rest of the weights are set to zero. This reduction in the number of kernel coefficients will speed up the reconstruction and it does not change the reconstructed image noticeably. For both the wavelet kernel and Gaussian kernel, the parameters σ and $\mathbf{a}$ were set to 1.

A. NEMA Phantom

To evaluate the performance of reconstruction, we used the National Electrical Manufacturers Association (NEMA) standard phantom [28]. A dynamic study was performed using this phantom and the resulting sinograms are reconstructed with OSEM, KEM, and WKEM reconstruction methods. The NEMA phantom consists of six spheres with varying diameters, which are filled with radiotracer. For the measurements, each sphere is filled with 1.5 MBq of F-18 diluted in about 55 mL of water. The background is injected with 80 MBq of F-18 diluted in about 10 L of water. The dynamic frames follow an exponential determined by F-18 decay. The scan time is 15 min and consists of 26 time frames as follows: 15×10, 5×30, 5×60, and 1×300 s. Composite frames were chosen so that each corresponds to 5 min of the scan time: the first composite frame is from the first 20 time frames, the second composite frame contains the next 5 time frames, and the final composite frame corresponds to the last time frame. For the NEMA phantom, we still have a known ground truth. Therefore, for comparison, we looked into the contrast and noise calculations in each of the spheres [29]. For this goal, a ROI with the same size of the sphere is drawn manually

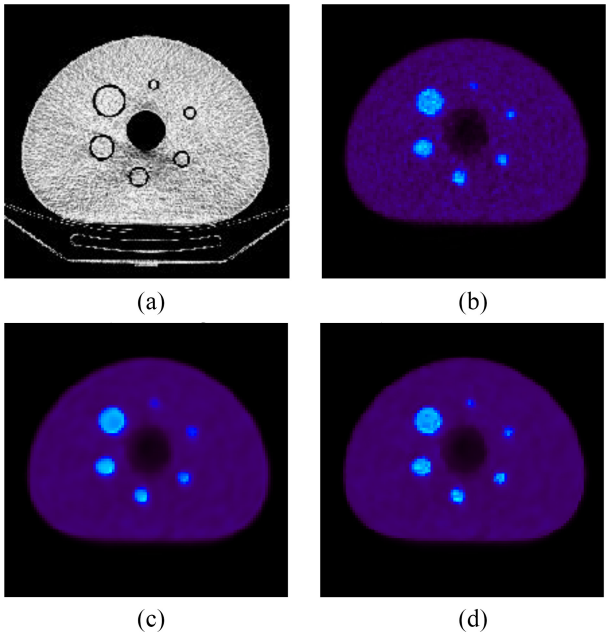


Fig. 8. Reconstructed image of a frame from NEMA phantom. (a) CT image. (b) OSEM 4mm. (c) KEM. (d) WKEM.

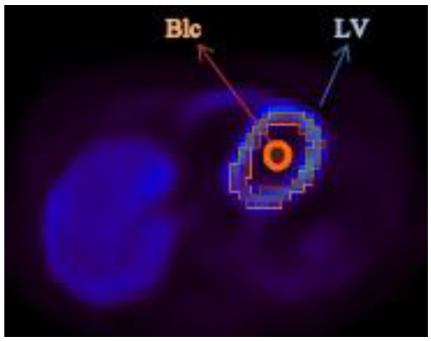


Fig. 9. Example showing left ventricle (LV) wall segment and blood cavity (Blc) area inside the left ventricle.

for each sphere with the help of the corresponding CT image and the mean value in these spheres is the average number of counts. For comparison, contrast recovery is calculated for each reconstruction method using

$$\mathbf{CR}_j = \frac{\frac{C_{s,j}-C_{B,j}}{C_{B,j}}}{\frac{a_s-a_B}{a_B}} \times \mathbf{100}\ [\%] \tag{15}$$

where $C_{s,j}$ is the average counts in the ROI for sphere j and $C_{B,j}$ is the average counts in the background ROI for sphere j. The background ROIs are drawn inside the phantom away from the spheres and are concentric to each other with diameter equal to the referring sphere. a_s and a_B are the activity concentration in the hot spheres and background, respectively. The percentage of background variability N_j is a measurement of image noise for sphere j with SD_j (in an ideal case $= 0\%$) and is calculated using

$$N_j = \frac{SD_j}{C_{B,j}} \times 100\ [\%]. \tag{16}$$

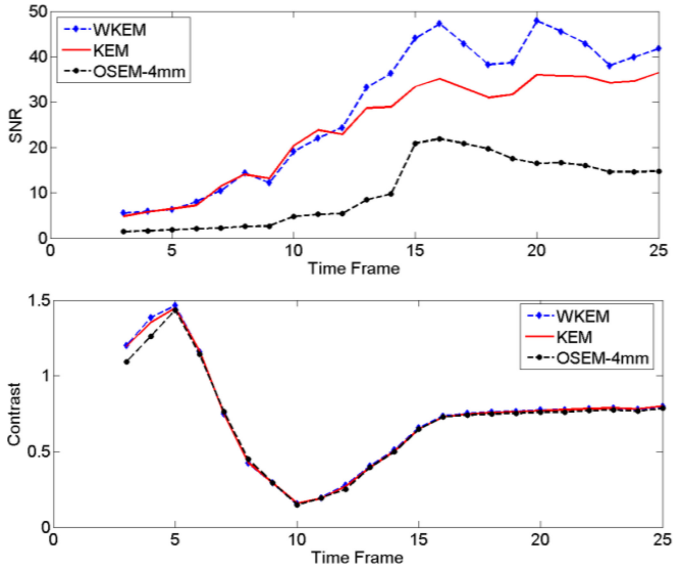


Fig. 10. Average SNR and contrast over all frames for OSEM (with 4mm post filtering), KEM, and WKEM.

SD_j is the SD in the background ROI for sphere j. Fig. 7 shows contrast recovery versus background variability for different sphere sizes by increasing the iteration number. Contrast recovery and background variability are averaged over the time frames. In this study, there are 24 subsets in the reconstruction, and the number of iterations varies from 1 to 4 (for smaller spheres, the iteration number is increased to 6 to have convergence for CRC) with one iteration having the smallest background variability value and 4 (6 for sphere 1 and 2) iterations having the largest background variability value in each method. As can be seen from Fig. 7, both KEM and WKEM have significantly smaller background variability in comparison to OSEM. However, WKEM has the best performance in terms of contrast recovery and background variability tradeoff, especially for smaller spheres. In the smallest sphere (sphere 1), OSEM achieves a much higher contrast recovery value in comparison to WKEM, which is expected as KEM and WKEM have smoothing included in the reconstruction process. But, since the background variability is larger for OSEM, overall, it is not as effective as WKEM. Also, OSEM is clinically always followed by post-filtering to reduce noise and as the figure suggests, OSEM with 4mm postfiltering reduced contrast recovery to a level lower than WKEM. Fig. 7 also demonstrates that for the smallest sphere with the highest iteration number, there is an increase of around 20% in the contrast recovery value between KEM and WKEM, which is significant. There is also an increase of around 10% for sphere 2 between KEM and WKEM. For other larger spheres (spheres 3–6), the increase in the contrast recovery value is around 1%–3%. These results affirm the statement about this wavelet kernel being better than the Gaussian kernel at retrieving small objects. This result can also be seen in Fig. 8, in which a reconstructed frame of the NEMA phantom is shown and the contrast, especially in the smallest sphere, is best preserved in WKEM.

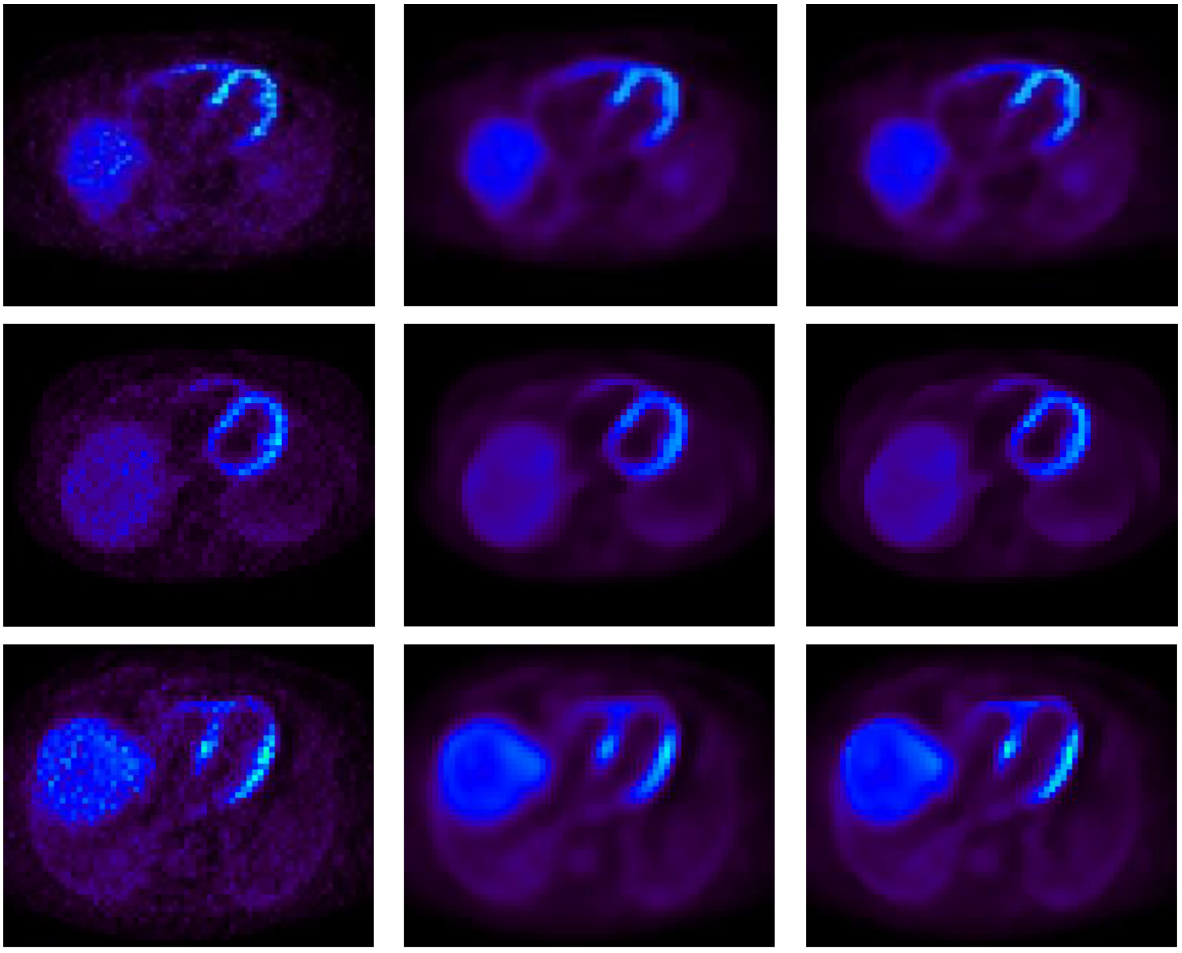


Fig. 11. Transaxial slice of the heart for OSEM, KEM, and WKEM. Each row corresponds to one of the patients and shows the last frame. (a) OSEM-4mm. (b) KEM. (c) WKEM.

B. Patient Data

The patient data are from Ottawa Heart Institute and were acquired as part of a research ethics board approved study. All images were anonymized before analysis. For patient data, there is no ground truth available and, therefore, SNR and contrast are defined based on tissue regions. To evaluate the KEM and WKEM methods on real patient data, i.e., with no ground truth, SNR and contrast are defined as

$$\textbf{SNR} = \frac{\text{mean(LV)}}{\text{std(Blc)}} \tag{17}$$

$$\textbf{Contrast} = \frac{|\text{mean(LV)} - \text{mean(Blc)}|}{\text{mean(LV)}}. \tag{18}$$

where "LV" makes reference to the left ventricle (LV) wall and the blood cavity (Blc) inside the LV is represented as "Blc" [22]. These regions are shown in Fig. 9. The LV wall is segmented using the FlowQuant software developed at the Ottawa Heart Institute [30] and the mean value of the segmented LV wall is calculated. In order to calculate background noise, the SD of an area inside the Blc of the LV is calculated. The mean value of this region is used for contrast calculation.

The study includes dynamic PET scans acquired from ten patients with C-11-hydroxyephedrine (HED) tracer on a GE D600 PET/CT scanner. The scan time was 60 min split into 25 time frames, which is defined as follows: 9×10, 3×30, 2×60, and 11×300 s. Three composite frames were created, each corresponding to 20 min of scan time; therefore, the first composite frames was from first 17 frames, the second composite frame was derived from the next four frames, and the last one is from the last four frames of the dynamic series. For this study, the iteration number was 3 with 32 subsets. The Gaussian kernel parameter σ is set to 1 and for WKEM, **a** is 1 as discussed.

Fig. 10 shows the average SNR and contrast over all ten patients for different time frames and different reconstruction methods. In the first couple of frames of the dynamic series, the tracer has not been detected since there is a very low number of photons. This will cause the mean value and SD to be zero or very close to zero in these time frames; therefore, the resulting SNR and contrast value would be inconsistent with the rest of the frames. Because of this reason they have been removed from this study and are not shown in the SNR and contrast calculations. As can be seen from (18),

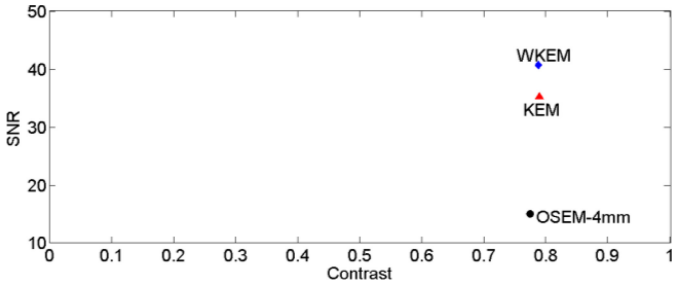


Fig. 12. SNR versus contrast of uptake frames (last four frames) averaged over all patients for OSEM (with 4mm postfiltering), KEM, and WKEM.

the contrast is the normalized difference of the regions in the LV and Blc. In early frames, the tracer is mainly in the blood so the difference is large. Then, at some point in time, the tracer is in both the blood and LV and, therefore, their difference is small. Then, later, the tracer is taken up in the LV and again the difference between these two regions is large. Fig. 11 shows a transaxial slice of the heart for three patients. As explained in Section III, a wavelet kernel is better able to preserve edge information. Fig. 11 affirms this statement by showing that WKEM preserves high intensity values better than KEM, especially in the LV wall.

In Fig. 12, SNR versus contrast is shown for WKEM and is compared to KEM and OSEM (with 4mm postfiltering). For this figure, the uptake frames, which are the last four frames, are included, and the mean value of contrast and SNR over the last four frames of all patients is calculated. From Fig. 12, we see that WKEM presents a higher SNR value compared to KEM and OSEM with postfiltering and it also preserves contrast.

VI. Discussion and Conclusion

Morlet wavelet is used here as a proof of concept. The method is still applicable to other wavelets as well. Morlet wavelet is picked because of its simple form and closed form, which makes it easily fitted in the kernel reconstruction framework. Other Morlet wavelets and in general, other wavelet functions maybe tested to have a more comprehensive overview of the method.

Our patient data are a limited dataset. For a more comprehensive comparison, it would be useful to have more patient data with other tracers as well. This work did not include motion compensation in the reconstruction and may further benefit from a reconstruction method, which included motion compensation. The emphasis in this manuscript is on temporal spatial information from the movement of the radioactive tracer. We did not focus on the effect of the method in diagnosis accuracy. Our future work will be to further improve the kernel method by using the multiresolution property of wavelets with dynamic imaging as prior information and also using other anatomical imaging methods as our prior information.

In this article, a wavelet kernel method is evaluated with computer simulation, physical phantom, and patient scan datasets for PET image reconstruction. Our results from simulated data show that WKEM can achieve higher contrast recovery than conventional KEM for low count frames. The NEMA phantom study suggests WKEM can reconstruct the phantom with higher contrast recovery, especially for smaller spheres. In line with that the study using patient data demonstrates clear improvement of the SNR in WKEM method over KEM and OSEM methods.

References

[1] Z. Ashouri, C. R. Hunter, B. A. Spencer, G. Wang, R. M. Dansereau, and R. A. DeKemp, "Dynamic PET image reconstruction using the wavelet kernel method," in *Proc. IEEE Nucl. Sci. Symp. Med. Imag. Conf. (NSS/MIC)*, Oct. 2019, pp. 1–3, doi: 10.1109/NSS/MIC42101.2019.9059890.

[2] L. A. Shepp and Y. Vardi, "Maximum likelihood reconstruction for emission tomography," *IEEE Trans. Med. Imag.*, vol. 1, no. 2, pp. 113–122, Oct. 1982, doi: 10.1109/TMI.1982.4307558.

[3] H. M. Hudson and R. S. Larkin, "Accelerated image reconstruction using ordered subsets of projection data," *IEEE Trans. Med. Imag.*, vol. 13, no. 4, pp. 601–609, Dec. 1994, doi: 10.1109/42.363108.

[4] J. E. Bowsher *et al.*, "Utilizing MRI information to estimate F18-FDG distributions in rat flank tumors," in *Proc. IEEE Symp. Conf. Record Nucl. Sci.*, vol. 4, Oct. 2004, pp. 2488–2492, doi: 10.1109/NSSMIC.2004.1462760.

[5] G.-H. Chen, J. Tang, and S. Leng, "Prior image constrained compressed sensing (PICCS): A method to accurately reconstruct dynamic CT images from highly undersampled projection data sets," *Med. Phys.*, vol. 35, no. 2, pp. 660–663, Feb. 2008, doi: 10.1118/1.2836423.

[6] A. Buades, B. Coll, and J. M. Morel, "On image denoising methods," Technical Note, CMLA (Centre de Mathematiques et de Leurs Applications), Cachan, France, 2004.

[7] C. Chan, S. Meikle, R. Fulton, G.-J. Tian, W. Cai, and D. D. Feng, "A non-local post-filtering algorithm for PET incorporating anatomical knowledge," in *Proc. IEEE Nucl. Sci. Symp. Conf. Rec. (NSS/MIC)*, Oct. 2009, pp. 2728–2732, doi: 10.1109/NSSMIC.2009.5401971.

[8] J. Qi and R. M. Leahy, "Iterative reconstruction techniques in emission computed tomography," *Phys. Med. Biol.*, vol. 51, no. 15, pp. R541–R578, Aug. 2006, doi: 10.1088/0031-9155/51/15/R01.

[9] V.-G. Nguyen and S.-J. Lee, "Incorporating anatomical side information into PET reconstruction using nonlocal regularization," *IEEE Trans. Image. Process.*, vol. 22, no. 10, pp. 3961–3973, Oct. 2013, doi: 10.1109/TIP.2013.2265881.

[10] G. Wang, C. Catana, and J. Qi, "Joint reconstruction of attenuation and emission from PET data using MR-derived attenuation prior," presented at the IEEE Nucl. Sci. Symp. Med. Imag. Conf., Anaheim, CA, USA, Nov. 2012.

[11] K. Gong, C. Catana, J. Qi, and Q. Li, "PET image reconstruction using deep image prior," *IEEE Trans. Med. Imag.*, vol. 38, no. 7, pp. 1655–1665, Jul. 2019, doi: 10.1109/TMI.2018.2888491.

[12] G. Wang and J. Qi, "PET image reconstruction using kernel method," *IEEE Trans. Med. Imag.*, vol. 34, no. 1, pp. 61–71, Jan. 2015, doi: 10.1109/TMI.2014.2343916.

[13] B. Spencer, J. Qi, R. D. Badawi, and G. Wang, "Dynamic PET Image reconstruction for parametric imaging using the HYPR kernel method," in *Proc. Phys. Med. Imag.*, vol. 10132, Mar. 2017, Art. no. 101324W, doi: 10.1117/12.2254497.

[14] W. Hutchcroft, G. Wang, K. T. Chen, C. Catana, and J. Qi, "Anatomically-aided PET reconstruction using the kernel method," *Phys. Med. Biol.*, vol. 61, no. 18, pp. 6668–6683, 2016, doi: 10.1088/0031-9155/61/18/6668.

[15] P. Novosad and A. J. Reader, "MR-guided dynamic PET reconstruction with the kernel method and spectral temporal basis functions," *Phys. Med. Biol.*, vol. 61, no. 12, pp. 4624–4644, 2016, doi: 10.1088/0031-9155/61/12/4624.

[16] K. Gong, G. Wang, K. T. Chen, C. Catana, and J. Qi, "Nonlinear PET parametric image reconstruction with MRI information using kernel method," in *Proc. Phys. Med. Imag.*, vol. 10132, Mar. 2017, Art. no. 101321G, doi: 10.1117/12.2254273.

[17] J. Bland *et al.*, "MR-guided kernel EM reconstruction for reduced dose PET imaging," *IEEE Trans. Radiat. Plasma Med. Sci.*, vol. 2, no. 3, pp. 235–243, May 2018, doi: 10.1109/TRPMS.2017.2771490.

[18] J. Bland, M. A. Belzunce, S. Ellis, C. J. McGinnity, A. Hammers, and A. J. Reader, "Spatially compact MR-guided kernel EM for PET image reconstruction," *IEEE Trans. Radiat. Plasma Med. Sci.*, vol. 2, no. 5, pp. 470–482, Sep. 2018, doi: 10.1109/TRPMS.2018.2844559.

[19] G. Wang, "High temporal-resolution dynamic PET image reconstruction using a new spatiotemporal kernel method," *IEEE Trans. Med. Imag.*, vol. 38, no. 3, pp. 664–674, Mar. 2019, doi: 10.1109/TMI.2018.2869868.

[20] H.-M. Huang, "Dynamic PET reconstruction using the kernel method with non-local means denoising," *Biomed. Signal Process. Control*, vol. 68, Jul. 2021, Art. no. 102673, doi: 10.1016/j.bspc.2021.102673.

[21] K. Gong, J. Cheng-Liao, G. Wang, K. T. Chen, C. Catana, and J. Qi, "Direct Patlak reconstruction from dynamic PET data using the kernel method with MRI information based on structural similarity," *IEEE Trans. Med. Imag.*, vol. 37, no. 4, pp. 955–965, Apr. 2018, doi: 10.1109/TMI.2017.2776324.

[22] Z. Ashouri, C. R. Hunter, B. A. Spencer, G. Wang, R. M. Dansereau, and R. A. deKemp, "Kernel-based reconstruction of C-11-hydroxyephedrine cardiac PET images of the sympathetic nervous system," in *Proc. 41st Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2019, pp. 832–835, doi: 10.1109/EMBC.2019.8856752.

[23] J. H. Friedman, J. L. Bentley, and R. A. Finkel, "An algorithm for finding best matches in logarithmic expected time," *ACM Trans. Math. Softw.*, vol. 3, no. 3, pp. 209–226, Sep. 1977, doi: 10.1145/355744.355745.

[24] S. Xie, A. T. Lawniczak, S. Krishnan, and P. Lio, "Wavelet kernel principal component analysis in noisy multiscale data classification," *Int. Scholarly Res. Notices*, vol. 2012, Jul. 2012, Art. no. 197352, doi: 10.5402/2012/197352.

[25] L. Zhang, W. Zhou, and L. Jiao, "Wavelet support vector machine," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 34, no. 1, pp. 34–39, Feb. 2004.

[26] P. S. Addison, J. N. Watson, and T. Feng, "Low-oscillation complex wavelets," *J. Sound Vib.*, vol. 254, no. 4, pp. 733–762, Jul. 2002, doi: 10.1006/jsvi.2001.4119.

[27] I. G. Zubal, C. R. Harrell, E. O. Smith, Z. Rattner, G. Gindi, and P. B. Hoffer, "Computerized three-dimensional segmented human anatomy," *Med. Phys.*, vol. 21, no. 2, pp. 299–302, 1994. [Online]. Available: https://doi.org/10.1118/1.597290

[28] *Performance Measurements of Positron Emission Tomographs*, NEMA Standards NU 2-2007, 2007.

[29] S. Ziegler, B. W. Jakoby, H. Braun, D. H. Paulus, and H. H. Quick, "NEMA image quality phantom measurements and attenuation correction in integrated PET/MR hybrid imaging," *EJNMMI Phys.*, vol. 2, no. 1, p. 18, Dec. 2015, doi: 10.1186/s40658-015-0122-3.

[30] R. Klein *et al.*, "Intra- and inter-operator repeatability of myocardial blood flow and myocardial flow reserve measurements using rubidium-82 pet and a highly automated analysis program," *J. Nucl. Cardiol.*, vol. 17, no. 4, pp. 600–616, Aug. 2010, doi: 10.1007/s12350-010-9225-3.