High Temporal-Resolution Dynamic PET Image Reconstruction Using a New Spatiotemporal Kernel Method

Guobao Wang

Department of Radiology, School of Medicine, University of California at Davis, Sacramento, CA 95817, USA. gbwang@ucdavis.edu

Abstract

Current clinical dynamic PET has an effective temporal resolution of 5–10 seconds, which can be adequate for traditional compartmental modeling but is inadequate for exploiting the benefit of more advanced tracer kinetic modeling for characterization of diseases (e.g., cancer and heart disease). There is a need to improve dynamic PET to allow fine temporal sampling of 1–2 seconds. However, reconstruction of these short-time frames from tomographic data is extremely challenging as the count level of each frame is very low and high noise presents in both spatial and temporal domains. Previously the kernel framework has been developed and demonstrated as a statistically efficient approach to utilizing image prior for low-count PET image reconstruction. Nevertheless, the existing kernel methods mainly explore spatial correlations in the data and only have a limited ability in suppressing temporal noise. In this paper, we propose a new kernel method which extends the previous spatial kernel method to the general spatiotemporal domain. The new kernelized model encodes both spatial and temporal correlations obtained from image prior information and is incorporated into the PET forward projection model to improve the maximum likelihood (ML) image reconstruction. Computer simulations and an application to real patient scan have shown that the proposed approach can achieve effective noise reduction in both spatial and temporal domains and outperform the spatial kernel method and conventional ML reconstruction method for improving high temporal-resolution dynamic PET imaging.

Keywords

High temporal resolution (HTR); dynamic PET; image reconstruction; maximum likelihood; image prior; kernel method; spatiotemporal correlation

I. Introduction

Dynamic positron emission tomography (PET) can monitor spatiotemporal distribution of a radiotracer in human body. With tracer kinetic modeling [1]–[3], dynamic PET is capable of quantifying physiologically or biochemically important parameters in regions of interest or voxelwise and can be used to detect and characterize many diseases such as cancer or heart diseases. Traditionally compartmental models are used for kinetic analysis of dynamic PET data [1], [2]. Other advanced tracer kinetic models such as the distributed-parameter model [4] and the adiabatic approximations [5] are considered closer to the physiological process than compartmental models. However, those models have not been well explored in dynamic PET because the effective temporal resolution of clinical dynamic PET has been limited to

5–10 seconds (see [6] for an example). This poor temporal resolution is insufficient to justify the use of advanced kinetic modeling in dynamic PET [4].

We aim to improve the effective temporal resolution of clinical dynamic PET imaging to 1–2 seconds. To achieve the high temporal resolution (HTR), short scan duration must be used, which results in very low counting statistics in the dynamic frames. Image reconstruction from the low-count projection data is extremely challenging because the tomography reconstruction is ill-posed and high noise exists in tomographic measurement of short time frames.

In PET, incorporation of image prior information into image reconstruction has become a popular means to improve the quality of reconstructed images [9], [10]. Prior information can be either local smoothness of neighboring pixels or obtained from a co-registered anatomical MRI image (e.g., [11]–[15]) or CT image (e.g., [16], [17]). Most of existing PET image reconstruction methods employ an explicit regularization form (e.g., the Bowshwer's approach [18]) to incorporate image priors and can be complex for practical implementation. Regularization-based methods also commonly require a convergent solution to achieve optimal performance, which is computationally costly and can be inefficient for dynamic PET in which many frames need to be reconstructed. Direct reconstruction is another framework which incorporates kinetic modeling into the reconstruction formula [10], [19]–[21]. The method is statistically efficient when the kinetic model type is well known and all voxels in the field of view can be well described using the same kinetic model. However, this assumption is challenging to meet especially for new radiotracers and new clinical applications where the underlying kinetic model can be different from existing models. Any mismatch of the kinetic model can induce significant bias propagation in the kinetic images [22].

The recent kernel method [23], [28], [30] encodes image prior information in the forward model of PET image reconstruction and requires no explicit regularization. The kernel method is easier to implement and has been demonstrated to be more efficient and to better improve PET image reconstruction than regularization-based methods [23], [28]. Existing kernel methods mainly explore spatial correlations of image pixels to improve image quality in the spatial domain [23], [27], [28]. It has been applied to dynamic PET imaging [23], [27], MRI-guided static PET image reconstruction [28], [29], MRI-guided direct PET parametric image reconstruction [30], [31], and optical tomography [32]. These spatial kernels, however, have a limited ability for suppressing noise in the temporal domain. In HTR dynamic PET imaging, noise variation in the temporal domain can be very severe because many short-time frames are used. It is therefore desirable to include temporal prior knowledge in the kernel method to suppress temporal noise.

In this paper, we extend the spatial kernel method to a spatiotemporal kernel method that allows both spatial and temporal correlations to be encoded in the kernel matrix. We propose a separable spatiotemporal kernel to make the method more computationally tractable and easier to use. The new spatiotemporal kernel method is expected to achieve substantial noise reduction in the temporal domain in addition to the enhancement on image quality by existing spatial kernels. Note that this new method is different from other ongoing

developments that combine the spatial kernel method with direct reconstruction for well-known kinetic models such as the spectral model [30] or the Patlak graphical model [31]. Our focus here is on the reconstruction of HTR time activity curves, allowing new, unknown kinetic modeling to be explored in clinical translation of dynamic PET.

Part of this work was previously presented in the 2017 International Meeting on Fully 3D Image Reconstruction in Radiology and Nuclear Medicine [33]. This work has been substantially extended by an extensive simulation study and an evaluation using real patient data. The rest of this paper is organized as follows. We introduce the generalized theory of spatiotemporal kernel method for dynamic PET reconstruction and describe specific kernels in Section II. We then present a computer simulation study in Section III to validate the improvement of the kernel method over existing methods. The application of new method to real patient data of HTR dynamic PET imaging is presented in Section IV. We further discuss the method and results in Section V. Finally conclusions are drawn in Section VI.

II. Theory

A. Dynamic PET Image Reconstruction

For a time frame m, we denote the PET image intensity value at pixel j by $x_{j,m}$ and the measurement in detector pair i by $y_{i,m}$. The expectation of the dynamic projection data $\bar{y} = \left\{ \bar{y}_{i,m} \right\}$ is related to the unknown dynamic image $x = \{x_{j,m}\}$ through

$$\bar{y} = Px + r$$

(1)

where P is the detection probability matrix for dynamic PET and includes normalization factors for scanner sensitivity, scan duration, deadtime correction and attenuation correction. r is the expectation of dynamic random and scattered events [9].

Dynamic PET projection measurement $y = \{y_{i,m}\}$ can be well modeled as independent Poisson random variables with the log-likelihood function [9],

$$L(y|x) = \sum_{i=1}^{N_i} \sum_{m=1}^{M} y_{i,m}\log\bar{y}_{i,m} - \bar{y}_{i,m} - \log y_{i,m}!,$$

(2)

where N_i is the total number of detector pairs and M is the total number of time frames. The maximum likelihood (ML) estimate of the dynamic image x is found by maximizing the Poisson log-likelihood,

$$\hat{x} = \arg\max_{x \geq 0} L(y \,|\, x).$$

(3)

The expectation-maximization (EM) algorithm [34] with the following iterative update

$$x^{n+1} = \frac{x^n}{P^T \mathbf{1}_N} \cdot (P^T \frac{y}{Px^n + r}),$$

(4)

is often the choice to find the solution, where $\mathbf{1}_N$ is a vector of length $N = N_i \times M$ with all elements being 1, n denotes iteration number and the superscript "T" denotes matrix transpose. The vector multiplication and division are element-wise operations.

B. The Spatiotemporal Kernel Method

The kernel method for tomographic image reconstruction [23] was inspired by the kernel methods for classification and regression in machine learning. Unlike the kernel methods in machine learning [24], the kernel method for image reconstruction has unknown "label" values and the available data for kernel coefficient estimation is the tomographic projection data. Previously the kernel method [23] was derived for frame-by-frame spatial image reconstruction, here we adapt the expressions for spatiotemporal reconstruction.

In machine learning terminology, the image intensity $x_{j,m}$ at pixel j in time frame m is the "label" value. For each spatiotemporal location, a set of features are identified to form the feature vector $f_{j,m}$, which is also called a "data point" in machine learning. Conceptually we can use a mapping function $\phi(f_{j,m})$, which can be of finite or infinite dimension, to transform the data points $\{f_{j,m}\}$ into a feature space of very-high dimension $\{\phi(f_{j,m})\}$ [24]. By doing this, the "label" value $x_{j,m}$ can be better described as a linear function in the high-dimensional feature space,

$$x_{j,m} = w^T \phi\!\left(f_{j,m}\right)$$

(5)

where w is a weight vector which also sits in the transformed feature space:

$$w = \sum_{j'=1}^{N_j} \sum_{m'=1}^{M} \alpha_{j',m'} \phi\left(f_{j',m'}\right)$$

(6)

with α being the coefficient vector. N_j is the number of pixels in image. By substituting (6) into (5), the kernel representation for the image intensity at pixel j and in time frame m is written as

$$x_{j,m} = \sum_{j'=1}^{N_j} \sum_{m'=1}^{M} \alpha_{j',m'} \phi\left(f_{j',m'}\right)^T \phi\left(f_{j,m}\right)$$

(7)

$$= \sum_{j'=1}^{N_j} \sum_{m'=1}^{M} \alpha_{j',m'} \kappa\left(f_{j,m}, f_{j',m'}\right),$$

(8)

where

$$\kappa\left(f_{j,m}, f_{j',m'}\right) \triangleq \phi\left(f_{j,m}\right)^T \phi\left(f_{j',m'}\right)$$

(9)

is a kernel defined by the kernel function $\kappa(\cdot,\cdot)$ (e.g. radial Gaussian function). The mapping function ϕ is now implicitly defined by the kernel and is not required to be known with an explicit expression. For instance, the second-order polynomial kernel $\kappa(f, f') = (f^T f' + 1)^2$ for a 2-dimensional input data space $f = [f_1, f_2]^T$ implies a mapping from f into a 6-dimensional feature space $\phi(f) = \left[1, \sqrt{2}f_1, \sqrt{2}f_2, f_1^2, f_2^2, \sqrt{2}f_1 f_2\right]^T$ (p 171,[25]). The Gaussian kernel $\kappa(\mathbf{f}, \mathbf{f'}) = \exp(-\|f - f'\|^2)$ corresponds to transforming f into a feature space of infinite dimension, $\phi(f) = \{\phi_k\}_{k=0}^{\infty}$ [26]. For example, a simple case is $\phi_k = \sqrt{2^k/k!}\exp\left(-f^2\right)(f)^k$

when f is a one-dimensional vector $f = [f]^T$. While ϕ may be not or less computationally tractable, the kernel is commonly easy to compute [24].

The image intensity $x_{j,m}$ at pixel j in time frame m is thus described as a linear function in the kernel space but is nonlinear in the original space of the data points $\{f_{j,m}\}$. With x denoting the dynamic image and K the spatiotemporal kernel matrix, the equivalent matrix-vector form of (8) is

$$x = K\alpha,$$

(10)

where $\alpha \triangleq \left\{\alpha_{j,m}\right\}$ denotes the kernel coefficient vector. The dimension of K is $N_\kappa \times N_\kappa$ with $N_\kappa = N_j \times M$. The matrix is square because of its nature defined by the number of "data points".

Substituting the kernelized image model (10) into the standard PET forward projection model (1), we obtain the following kernelized forward projection model for dynamic PET image reconstruction:

$$\bar{y} = PK\alpha + r.$$

(11)

The advantage of using this kernelized model (11) is that image prior knowledge can be incorporated in the forward projection to improve the reconstruction of low-count scans.

A kernelized EM algorithm can be easily derived [23]. The EM update of α at iteration ($n + 1$) is

$$\alpha^{n+1} = \frac{\alpha^n}{K^T P^T \mathbf{1}_N} \cdot \left(K^T P^T \frac{y}{PK\alpha^n + r}\right).$$

(12)

Once the coefficient image α is estimated, the reconstructed dynamic PET image is calculated by

$$\hat{x} = K\hat{\alpha}.$$

$$(13)$$

Note that the EM solution converges to the ML estimate when n is large, which however can result in noisy reconstruction. In practice, early stopping with a reduced number of iterations is commonly used in EM reconstruction to control noise. This mechanism is also used in the kernel-based EM reconstruction.

C. Separable Spatiotemporal Kernel

The kernel matrix K encodes image prior information based on the feature vectors $\{f_{j,m}\}$. For each pixel j in time frame m, we identify a set of features to form $f_{j,m}$,

$$f_{j,m} = [(f_j^s)^T, (f_m^t)^T]^T,$$

$$(14)$$

where the vector consists of two components. f_j^s is the vector for exploring spatial correlations between pixels and f_m^t is for exploring temporal correlations between frames.

We further define the spatiotemporal kernel function $\kappa(f_{j,m}, f_{j',m'})$ to be spatially and temporally separable, i.e.

$$\kappa(f_{j,m}, f_{j',m'}) = \kappa_s(f_j^s, f_{j'}^s)\kappa_t(f_m^t, f_{m'}^t)$$

$$(15)$$

where $\kappa_s(\cdot,\cdot)$ denotes the kernel function for calculating spatial correlations and $\kappa_t(\cdot,\cdot)$ is for calculating temporal correlations.

Thus the overall spatiotemporal kernel matrix K is decoupled into a spatial kernel matrix $K_s \in \mathbb{R}^{N_j \times N_j}$ and a temporal kernel matrix $K_t \in \mathbb{R}^{M \times M}$,

$$K = K_t \otimes K_s,$$

(16)

where $\otimes$ represents the Kronecker product.

Derivation of the spatial kernel matrix K_s has been developed in our previous work [23]. K_s is often formed as a sparse matrix based on image prior data $\{f_j^s\}$ using the method of k-nearest neighbors (kNN):

$$\kappa_s\left(f_j^s, f_{j'}^s\right) = \begin{cases} \exp\left(-\dfrac{\left\|f_j^s - f_{j'}^s\right\|^2}{2\sigma_s^2}\right), & f_{j'}^s \in \text{kNN of } f_j^s \\ 0, & \text{otherwise.} \end{cases}$$

(17)

where σ_s is the Gaussian kernel parameter. In dynamic PET, an effective and efficient means for obtaining $\{f_j^s\}$ is to use composite frames [23]. For example, a one-hour dynamic FDG-PET scan can be first rebinned into three composite frames, each of 20 minutes. From the reconstructed composite images, three time activity points are obtained at each pixel j and used as the feature vector f_j^s to construct the spatial kernel matrix. Readers are referred to [23] for more details.

The spatial kernel method [23] is a special example of the spatiotemporal kernel method with the temporal kernel matrix set to the identity matrix,

$$K_t = I_M.$$

(18)

In this paper, we explore the role of new K_t in the context of HTR dynamic PET imaging.

D. Choice of Temporal Kernels

The (m, m')th element of K_t is obtained by comparing the feature vectors of the frames m and m':

$$\kappa_t\big(f_m^t, f_{m'}^t\big) = \begin{cases} \exp\left(-\dfrac{\big\|f_m^t - f_{m'}^t\big\|^2}{2\sigma_t^2}\right), & |m - m'| < d_t/2, \\ 0, & \text{otherwise.} \end{cases}$$

(19)

where d_t is the width of the time window for neighborhood construction between frame m and frame m'. σ_t is a parameter to adjust the weight calculation.

The simplest form of the temporal feature f_m^t is probably

$$f_m^t \triangleq m$$

(20)

by which K_t becomes a shift-invariant Gaussian smoothing kernel. The parameter σ_t can be determined by setting the window size as twice the full width at half maximum of the Gaussian function,

$$\sigma_t = \frac{d_t}{4\sqrt{2\log 2}}.$$

(21)

This type of kernel can smooth out noise but may also over-smooth sharp signals in the temporal domain.

To make the temporal kernel more adaptive to time varying data, we propose to use the whole sinogram of each frame as the feature vector to capture temporal correlations between frames, i.e.

$$f_m^t \triangleq \tilde{y}_m$$

(22)

where $\tilde{y}_m$ denotes a smoothed version of the raw sinogram y_m of frame m. Using the standard deviation of all the elements $\left\{ d_{m,m'} \right\}_{m=1,\, m'=1}^{M,M}$,

$$d_{m,m'} \triangleq \left\| \tilde{y}_m - \tilde{y}_{m'} \right\|,$$

(23)

the parameter σ_t can be set to:

$$\sigma_t = \frac{1}{M^2} \sqrt{\sum_{m,m'} \left(d_{m,m'} - \bar{d} \right)^2},$$

(24)

where $\bar{d}$ is the mean of $\{ d_{m,m'} \}$.

As both K_s and K_t are very sparse, inclusion of the kernel matrix K in the projection model does not add a significant computational cost in the reconstruction.

III. Validation Using Computer Simulation

A. Simulation Setup

Dynamic ^{18}F-FDG PET scans were simulated for a GE DST whole-body PET scanner in two-dimensional mode using a Zubal head phantom [Figure 1(a)]. The phantom contains several brain regions including brain background, blood region, gray matter, white matter and a tumor (15 mm in diameter). An attenuation map was simulated with a constant linear attenuation coefficient assigned in the whole brain. The scanning schedule consisted of 63 time frames over 20 minutes: 30×2 s, 12×5 s, 6×30 s, 15×60 s.

The blood input function was extracted from a real patient ^{18}F-FDG PET scan and fitted using an analytic form to reduce noise [Figure 1(b)]. Regional time activity curves (TACs) [Figure 1(c-d)] were assigned to different brain regions to generate noise-free dynamic activity images. The early phase of these TACs has fast temporal dynamics and is challenging to reconstruct. The resulting noise-free activity images were first forward projected to generate noise-free sinograms. A 20% uniform background was included to simulate random and scattered events. Poisson noise was then generated with 20 million expected events over 20 minutes.

A total of 20 realizations were simulated and each was reconstructed independently for statistical comparison. As the main goal is to compare the reconstruction methods for HTR dynamic PET imaging, the study focused on the comparisons for the first one minute (thirty 2-second frames) of the dynamic scan.

B.   Illustration of Temporal Kernels

One way of understanding the temporal kernel matrix K_t is that each column of K_t represents a temporal basis function. For example, the mth column of a kernel matrix K_t represents the temporal basis function centered at the mth time frame. Thus, a time activity curve is a linear combination of these temporal basis functions.

For illustration, a temporal window size of 15 time frames was used to construct temporal kernels. This fixed number of frames used for defining the temporal window leads to varying time windows across the dynamic sequence. The time window size was 30 seconds for early time frames and 15 minutes for late time frames. For calculating the data-driven temporal kernel matrix, the sinogram of each frame was smoothed using Gaussian smoothing with a window size of 7×7 before it was used.

Fig. 2 shows the comparison of temporal basis functions of two types of kernels - Gaussian temporal smoothing kernel (Eq. 20) and data-driven temporal kernel (Eq. 22). These temporal bases correspond to the 5th frame (t=8–10s), 15th frame (t=28–30s) and 55th frame (t=11–12 minute), respectively. For early time frames where fast activity changes occur, the data-driven kernel provides sharper temporal basis functions than the Gaussian smoothing kernel, while the two kernels are similar for late-time frames due to the slower activity change. Thus we expect the data-driven kernel to improve image reconstruction mainly for early-time frames when compared with the Gaussian temporal kernel.

Note that the previously used spatial kernel method has a special temporal kernel matrix - the identity matrix I_M, of which the temporal bases correspond to unit impulse functions for all frames. Because there are no temporal correlations included in this impulse kernel, temporal noise will substantially remain in the reconstruction.

C.   Reconstruction Methods to Compare

Noisy sinograms were reconstructed independently by four different image reconstruction methods: the traditional MLEM method, spatial kernelized EM (KEM-S), and new spatiotemporal kernel method with the Gaussian smoothing temporal kernel (KEM-ST-G) and the spatiotemporal kernel method with the data-driven temporal kernel (KEM-ST-D). Each reconstruction was run for 200 iterations to allow for the investigation of the effect of iteration number.

The spatial kernel matrix K_s was constructed using the same approach as described in [23]. The global kNN search with $k = 48$ was performed over all image pixels. Four composite images (4×5-minute) were obtained by recombining the full 20-minute dynamic data and used for the spatial kernel matrix construction. The temporal kernels were constructed using a window size of 15 frames unless specified otherwise. Different window sizes, ranging from 7 to 33 frames, were also investigated to evaluate the effect of temporal window size on reconstruction quality.

D. Comparison for Overall Image Quality

Image quality of different reconstruction methods were first assessed using the image mean squared error (MSE) which is defined by

$$\text{MSE}(\hat{\boldsymbol{x}}_m) = 10 \log_{10} \frac{\left\| \hat{\boldsymbol{x}}_m - \boldsymbol{x}_m^{\text{true}} \right\|^2}{\left\| \boldsymbol{x}_m^{\text{true}} \right\|^2} \quad (\text{dB}),$$

(25)

where $\hat{\boldsymbol{x}}_m$ is an image estimate of frame m obtained with one of the reconstruction methods and $\boldsymbol{x}_m^{\text{true}}$ denotes the ground truth image of the frame.

Figure 3 shows the true activity image and reconstructed images by the four different methods with 100 iterations for the 5th frame (t=8–10s), 15th frame (t=28–30s) and 55th frame (t=11–12 minute), respectively. As expected, the KEM-S method achieved a significant MSE decrease as compared with the MLEM method. By incorporating temporal correlations, the KEM-ST-G and KEM-ST-D methods further improved the image quality and achieved lower MSE values than KEM-S. The KEM-ST-D method had lower MSE than the KEM-ST-G method for the early-time frames 5 and 15. The two methods had similar MSE for the late-time frame 55.

Figure 4 shows image MSE as a function of iteration number for the three different frames (5th, 15th and 55th). The two spatiotemporal kernel methods KEM-ST-G and KEM-STD had slower convergence rate than the spatial kernel method KEM-S, requiring more iterations to achieve the lowest MSE. This is because of the additional temporal correlation included in the spatiotemporal kernel methods. Note that the MSE curves of the 15th frame ($t = 30$ s) behaved slightly different from that of the 5th and 55th frames and achieved the lowest MSE earlier. This is because the 15th frame had lower activity and count level than the other two frames [see Fig. 1 (c-d) for a comparison].

Figure 5 shows the plots of image MSE of all time frames for different methods. The MSE of each frame is minimized over the number of iteration in each method. The error bars indicate the standard deviation of the MSE over 20 realizations. The three kernel methods outperformed the MLEM reconstruction for all frames. Compared with the spatial kernel method KEM-S, the spatiotemporal kernel method KEM-ST-G improved late-time frames. This is because the TACs of these late-time frames have small temporal changes than early-time frames. Incorporation of temporal correlations thus became beneficial and achieved noise reduction. KEM-ST-D and KEM-ST-G had similar performance especially in the late-time frames. In the early-time frames where activity change is rapid, use of the time invariant temporal kernel in the KEM-ST-G method over-smoothed the temporal signals. Thus KEM-ST-G had even higher MSE than KEM-S. In contrast, KEM-ST-D had a data-adaptive temporal kernel and achieved better results.

E. Comparison for Time Activity Curves

Figure 6(a) and 6(b) respectively show the TACs of a pixel in the blood region and a pixel in the tumor region reconstructed by different methods. The TACs of the first 60 seconds are further shown in Figure 6(c) and 6(d). The reconstructions by MLEM were extremely noisy, especially in the early-time frames because these frames are of only 2-s scan duration. While it achieved a significant noise reduction in the spatial domain, the KEM-S method still resulted in substantial noise in the temporal domain, particularly in the tumor where the tracer uptake is relatively low. The KEMST-G method reduced temporal noise in late-time frames but over-smoothed the early-time peak activities, causing bias in early-time frames. In comparison, the KEM-ST-D reconstruction overcame the limitation of KEM-ST-G and achieved a substantial noise reduction in the temporal domain for both early-time and late-time frames.

To compare different methods quantitatively, we calculated the bias and standard deviation (SD) of the mean uptake in the tumor region by

$$\text{Bias} = \frac{1}{\left\| \boldsymbol{c}^{\text{true}} \right\|} \left\| \bar{\boldsymbol{c}} - \boldsymbol{c}^{\text{true}} \right\|,$$

(26)

$$\text{SD} = \frac{1}{\left\| \boldsymbol{c}^{\text{true}} \right\|} \sqrt{\frac{1}{N_r} \sum_{i=1}^{N_r} \left\| \boldsymbol{c}_i - \bar{\boldsymbol{c}} \right\|^2},$$

(27)

where $\boldsymbol{c}^{\text{true}}$ is the noise-free regional TAC and $\bar{\boldsymbol{c}} = \frac{1}{N_r} \sum_{i=1}^{N_r} \boldsymbol{c}_i$ denotes the mean of N_r realizations. $N_r = 20$ in this study.

Figure 7 shows the trade-off between the bias and SD of different methods in the tumor ROI for two single time points (frame 5 and frame 15) and for the whole TAC of the first 60 seconds. The curves were obtained by varying the iteration number from 10 to 200 iterations with an interval of 10 iterations. Compared with the MLEM and KEM-S methods, the two spatiotemporal kernel methods (KEM-ST-G and KEM-ST-D) had lower noise SD due to the temporal kernels. Being more adaptive to the data, the KEM-ST-D method achieved lower bias than the KEM-ST-G method at any level of noise.

F.   Effect of Temporal Window Size

Figure 8 shows the MSE of regional TAC of the blood region and tumor region as a function of temporal window size. With the window size equal to 1, the spatiotemporal kernel methods KEM-ST-G and KEM-ST-D become the same as the spatial kernel method KEM-ST.

Because the blood region has a high count level (due to high activity), no temporal smoothing was needed. With increasing temporal window size, the Gaussian kernel ST-G reduced the TAC quality due to over-smoothing. The data-driven kernel ST-D was less influenced by the window size and was more stable to reconstruct the blood TAC. In the tumor region, the performance of the ST-G kernel changed sharply, indicating it can be difficult to choose a proper window size. In contrast, the performance of the ST-D kernel was stable for a range of window sizes from 10 to 30 frames. The minimum MSE achieved by the ST-D kernel was also lower than that by the ST-G kernel.

These results indicate that the ST-D kernel is more adaptive to the data and more stable than the ST-G kernel.

IV.   Application to Real Patient Scan

A.   Patient Data Acquisition

A cardiac patient scan was performed on the GE Discovery ST PET/CT scanner at the UC Davis Medical Center in two-dimensional mode. The patient received approximately 20 mCi ^{18}F-FDG with a bolus injection. Data acquisition commenced right after the FDG injection. A low-dose transmission CT scan was then performed at the end of PET scan to provide the CT image for PET attenuation correction. The raw data of the first 90 seconds were binned into a total of 45 dynamic frames with a 2-second duration. The data correction sinograms of each frame, including normalization, attenuation correction, scattered correction and randoms correction, were extracted using the vendor software used in the reconstruction process.

The patient data were reconstructed independently by the traditional MLEM method and three kernel methods: spatial kernel method KEM-S, new spatiotemporal kernel methods KEM-ST-G (Gaussian smoothing temporal kernel) and KEM-ST-D (data-driven temporal kernel) with 100 iterations. The spatial kernel matrix K_s was constructed using five composite frames (3×10 s, 2×30 s). The temporal kernels were constructed using the same parameters as used in the simulation study.

B.   Results

Figure 9 shows the comparison of a 10-s low temporal-resolution time frame by standard MLEM with its corresponding 2-s high temporal-resolution time frames reconstructed using the data-driven spatiotemporal kernel method KEM-ST-D. While the 10-s image by MLEM already suffered from noise and showed similar FDG uptakes in the left ventricle and right ventricle, the HTR images demonstrated time uptakes in the two ventricular regions with good image quality even though they have a short 2-second duration. The right ventricle had

decreasing activities and the left ventricle had increasing activities during the same 10 seconds period.

Figure 10 shows the comparison of different image reconstruction methods for reconstructing HTR time frames at t=8–10s, t=28–30s and t=88–90s. The traditional MLEM reconstructions were extremely noisy. The spatial kernel method KEM-S achieved substantial noise reduction. Compared with KEM-S, KEM-ST-G underestimated the FDG uptake in the left and right ventricles due to the Gaussian temporal smoothing. Given that KEM-ST-D is more adaptive to the data than KEM-ST-G, it preserved the improvement by KEM-S in the early frames at t=8–10s and t=18–20s and further reduced the noise in the late frame at t=88–90s.

Figure 11 shows the HTR time activity curves for the pixels in the right ventricle, left ventricle, aorta, and myocardium. The time activity curves of standard temporal resolution by MLEM are also included for comparison. With the standard MLEM reconstruction, the increase of temporal resolution suffered from high noise, thus contaminating any benefit brought by HTR. The kernel-based reconstruction by KEM-S demonstrated low noise for the high-activity time points. However, high noise remained in the low-activity time points. KEM-ST-G smoothed out the noise in late frames but over-smoothed early frames where activity changes fast. In comparison, the KEM-ST-D method achieved satisfactory noise suppression for both early-time and late-time points on the time activity curves.

The results from this patient study have further demonstrated the improvement of the spatiotemporal kernel method for HTR image reconstruction.

V. Discussion

PET scanning is dynamic by its nature and can have an infinitesimal temporal sampling by list-mode data acquisition. The quality of reconstructed images, however, is very much affected by noise and is so far inadequate to support the use of very high temporal resolution, especially for voxel-wise kinetic modeling. To suppress noise, temporal regularization or more generally spatiotemporal regularization may be used by adding a penalty term in the dynamic PET image reconstruction formula, see [7], [8] for examples. This regularization-based method has the disadvantages of being complex and with high computational cost given that there are many time frames (or basis coefficient images) to reconstruct in HTR imaging.

This work continues the development of the kernel method for dynamic PET image reconstruction with the aim to improve HTR imaging. Existing kernel methods mainly exploit spatial correlations and can substantially improve image quality in the spatial domain, as demonstrated in previous works (e.g., [23]) and in this paper (Fig. 3 and Fig. 10). To reduce noise in the temporal domain, the kernel method is generalized in this paper to a spatiotemporal kernel method. The improvement of the new method over the spatial kernel method has been demonstrated using simulation data (Fig. 6) and patient data (Fig. 11).

The comparison studies in this work have been done with dynamic 2D scans for demonstrating the proof of concept. This can be partly justified by the fact that the

Discovery ST scanner is normally performed in 2D mode for cardiac imaging at our institution. The implementation of the method for dynamic fully-3D scans (i.e., fully 4D) is straightforward in principle. Additional attention should be given to handling the increased data size and memory requirement due to the 3D acquisition and large number of time frames in HTR dynamic imaging, which may present some challenges for efficient implementation of the fully 4D spatiotemporal reconstruction.

The data-driven temporal kernel in its current format is derived from global sinogram data. Being adaptive to the overall temporal behavior of radiotracer uptake, such a kernel can provide sharp temporal bases for early-time frames (where fast changes occur) and smooth temporal bases for late-time frames (where activity changes are commonly slow) (Fig. 2). Although it less likely happens in most dynamic PET studies, an atypical region may exist with a very different kinetic behavior, for example fast kinetics at a late time. Then the global temporal kernels would not be able to match well with the local temporal behavior in this atypical region and may result in oversmoothed reconstruction if the temporal window size is large. Hence the method parameters may need to be tailored for specific applications. Our future work will also develop spatially-varying temporal kernels to make the method even more adaptive to spatiotemporal locations.

VI. Conclusion

In this paper, we have developed a spatiotemporal kernel method to incorporate both spatial and temporal prior information into the kernel framework for dynamic PET image reconstruction. The spatiotemporal kernel is separable in the spatial and temporal domains and thus can be easily and efficiently implemented. We conducted a computer simulation to validate the method and tested the method using a patient PET scan. The results from both the simulation study and patient study have shown that the new spatiotemporal kernel method can outperform traditional MLEM and existing spatial kernel methods to achieve a high temporal-resolution of 2-s while maintaining noise at a low level. Future work will include further optimization of the method and a more comprehensive patient study to quantify the improvement.

Acknowledgment

References

[1]. Carson RE, "Tracer Kinetic Modeling in PET," in Bailey DL, Townsend DW, Valk PE, Maisey MN (eds) Positron Emission Tomography, Springer, London, 2005.

[2]. Schmidt KC and Turkheimer FE, "Kinetic modeling in positron emission tomography," Quarterly Journal of Nuclear Medicine, vol. 46, no. 1, pp. 70–85, 2002. [PubMed: 12072847]

[3]. Gunn RN, Gunn SR, and Cunningham VJ, "Positron emission tomography compartmental models," Journal of Cerebral Blood Flow and Metabolism, vol. 21, no. 6, pp. 635–652, 2001. [PubMed: 11488533]

[4]. Muzic RF, Saidel GM, "Distributed versus compartment models for PET receptor studies," IEEE Trans. Med. Imag, vol. 22, no. 1, pp. 11–21, 2003.

[5]. St Lawrence KS, Lee TY, "An adiabatic approximation to the tissue homogeneity model for water exchange in the brain: I. theoretical derivation," Journal of Cerebral Blood Flow & Metabolism, vol. 18, no. 12, pp. 1365–1377, 1998. [PubMed: 9850149]

[6]. Humbert O, Lasserre M, Bertaut A, Fumoleau P, Coutant C, Brunotte F, Cochet A, "Breast cancer blood flow and metabolism on dual-acquisition ^{18}F-FDG PET: correlation with tumor phenotype and neoadjuvant chemotherapy response," Journal of Nuclear Medicine, vol. 59, no. 7, pp. 1035–1041, 2018. [PubMed: 29439010]

[7]. Nichols TE, Qi J, Asma E, Leahy RM, "Spatiotemporal reconstruction of list-mode PET data," IEEE Transactions on Medical Imaging, vol. 21, no. 4, pp. 396–404, 2002. [PubMed: 12022627]

[8]. Li Q, Asma E, Ahn S, Leahy RM, "A fast fully 4-D incremental gradient reconstruction algorithm for list mode PET dData," IEEE Transactions on Medical Imaging, vol. 26, no. 1, pp. 58–67, 2007. [PubMed: 17243584]

[9]. Qi J and Leahy RM, "Iterative reconstruction techniques in emission computed tomography," Physics in Medicine and Biology, vol. 51, no. 15, pp. R541–R578, 2006. [PubMed: 16861768]

[10]. Reader AJ and Verhaeghe J, 4D image reconstruction for emission tomography. Physics in Medicine and Biology, 59(22):R371–R418. [PubMed: 25361380]

[11]. Bai B, Li Q, Leahy RM, "Magnetic resonance-guided positron emission tomography image reconstruction," Seminars in Nuclear Medicine, vol. 43, no. 1, pp. 30–44, 2013. [PubMed: 23178087]

[12]. Chan C, Fulton R, Feng DD, Meikle S, "Regularized image reconstruction with an anatomically adaptive prior for positron emission tomography" Physics in Medicine and Biology, vol. 54, no. 24, pp. 7379–400, 2009. [PubMed: 19934490]

[13]. Tang J, Rahmim A, "Bayesian PET image reconstruction incorporating anato-functional joint entropy," Physics in Medicine and Biology, vol. 54, no. 23, pp. 7063-, 2009. [PubMed: 19904028]

[14]. Somayajula S, Panagiotou C, Rangarajan A, Li Q, Arridge SR, Leahy RM, "PET image reconstruction using information theoretic anatomical priors," IEEE Transactions on Medical Imaging, vol. 30, no. 3, pp. 537–49, 2011. [PubMed: 20851790]

[15]. Vunckx K, Atre A, Baete K, Reilhac A, Deroose CM, Van Laere K, Nuyts J, "Evaluation of three MRI-based anatomical priors for quantitative PET brain imaging," IEEE Transactions on Medical Imaging, vol. 31, no. 3, pp. 599–612, 2012. [PubMed: 22049363]

[16]. Fessler JA, Clinthorne NH, Rogers WL, "Regularized emission image reconstruction using imperfect side information," IEEE Transactions on Nuclear Science, vol. 39, no. 5, pp. 1464–1471, 1992.

[17]. Cheng-Liao J, Qi J, "PET image reconstruction with anatomical edge guided level set prior," Physics in Medicine and Biology, vol. 56, no. 21, pp. 6899-, 2011. [PubMed: 21983558]

[18]. Bowsher JE, Johnson VE, Turkington TG, Jaszczak RJ, Floyd CE, Coleman RE, "Bayesian reconstruction and use of anatomical a priori information for emission tomography," IEEE Transactions on Medical Imaging, vol. 15, no. 5, pp. 673–686, 1996. [PubMed: 18215949]

[19]. Tsoumpas C, Turkheimer FE, and Thielemans K, "A survey of approaches for direct parametric image reconstruction in emission tomography," Med. Phys, vol. 35, no. 9, pp. 3963–3971, 2008. [PubMed: 18841847]

[20]. Rahmim A, Tang J, Zaidi H, "Four-dimensional image reconstruction strategies in dynamic PET: Beyond conventional independent frame reconstruction," Medical Physics, vol. 36, no. 8, pp. 3654–3670, 2009. [PubMed: 19746799]

[21]. Wang GB and Qi J, "Direct estimation of kinetic parametric images for dynamic PET," Theranostics, vol. 3, no. 10, 802–815, 2013. [PubMed: 24396500]

[22]. Kotasidis FA, Mehranian A1, Zaidi H, "Impact of time-of-flight on indirect 3D and direct 4D parametric image reconstruction in the presence of inconsistent dynamic PET data," Physics in Medicine and Biology, 61(9):3443–3471, 2016. [PubMed: 27049697]

[23]. Wang GB, Qi J, "PET image reconstruction using kernel method," IEEE Transactions on Medical Imaging, vol. 34, no. 1, pp. 61–71, 2015. [PubMed: 25095249]

[24]. Hofmann T, Scholkopf B, Smola AJ, "Kernel methods in machine learning," The Annals of Statistics, vol. 36, no. 3, pp. 1171–1220, 2008.

[25]. Hastie T, Tibshirani R, Friedman J, "The Elements of Statistical Learning," Springer, second edition, 2008.

[26]. Shashua A, "Introduction to Machine Learning: Class Notes 67577," arXiv preprint, 0904.3664, 2009.

[27]. Spencer B, Qi J, Badawi RD, Wang GB, "Dynamic PET image reconstruction for parametric imaging using the HYPR kernel method," Proceedings of SPIE Medical Imaging Conference, pp. 101324W, 2017.

[28]. Hutchcroft W, Wang GB, Chen K, Catana C, Qi J, "Anatomically-aided PET reconstruction using the kernel method," Physics in Medicine and Biology, 61(18): 6668–6683, 2016. [PubMed: 27541810]

[29]. Bland J, Mehranian A, Belzunce MA, Ellis S, McGinnity CJ, Hammers A, Reader AJ, "MR-guided kernel EM reconstruction for reduced dose PET imaging," IEEE Transactions on Radiation and Plasma Medical Sciences, 2(3): 235–243, 2018. [PubMed: 29978142]

[30]. Novosad P, Reader AJ, "MR-guided dynamic PET reconstruction with the kernel method and spectral temporal basis functions," Physics in Medicine and Biology, 61(12): 46244645, 2016.

[31]. Gong K, Cheng-Liao J, Wang GB, Chen KT, Catana C, Qi J, "Direct Patlak reconstruction from dynamic PET data using kernel method with MRI information based on structural similarity," IEEE Transactions on Medical Imaging, 37(4): 955–965, 2018. [PubMed: 29610074]

[32]. Baikejiang R, Zhao Y, Fite BZ, Ferrara KW, Li CQ, "Anatomical image guided fluorescence molecular tomography reconstruction using kernel method," Journal of Biomedical Optics, 22(5): 055001, 2017.

[33]. Wang GB, "High temporal-resolution dynamic PET image reconstruction using a new spatiotemporal kernel Method," 14th International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine, vol. 14, pp. 805–809, Xi'an, China, 6 2017.

[34]. Shepp LA and Vardi Y, "Maximum likelihood reconstruction in positron emission tomography," IEEE Transactions on Medical Imaging, vol. MI-1, pp. 113–122, 1982.
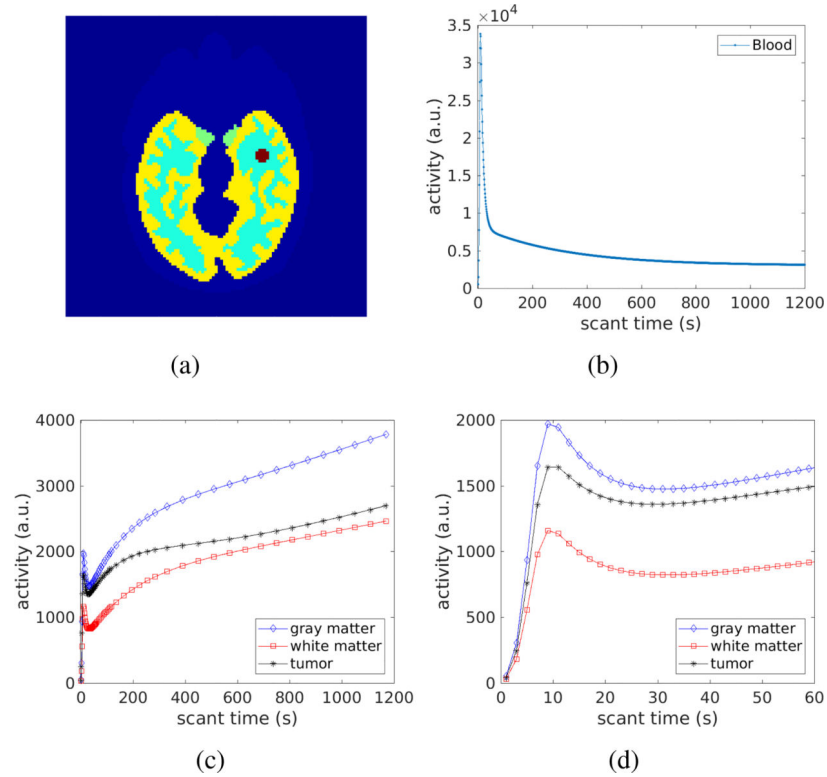
Fig. 1.

The digital phantom and time activity curves used in the simulation study. (a) The Zubal brain phantom; (b) Blood input function; (c) Regional time activity curves of different brain regions; (d) Time activity curves of the first 60 seconds.
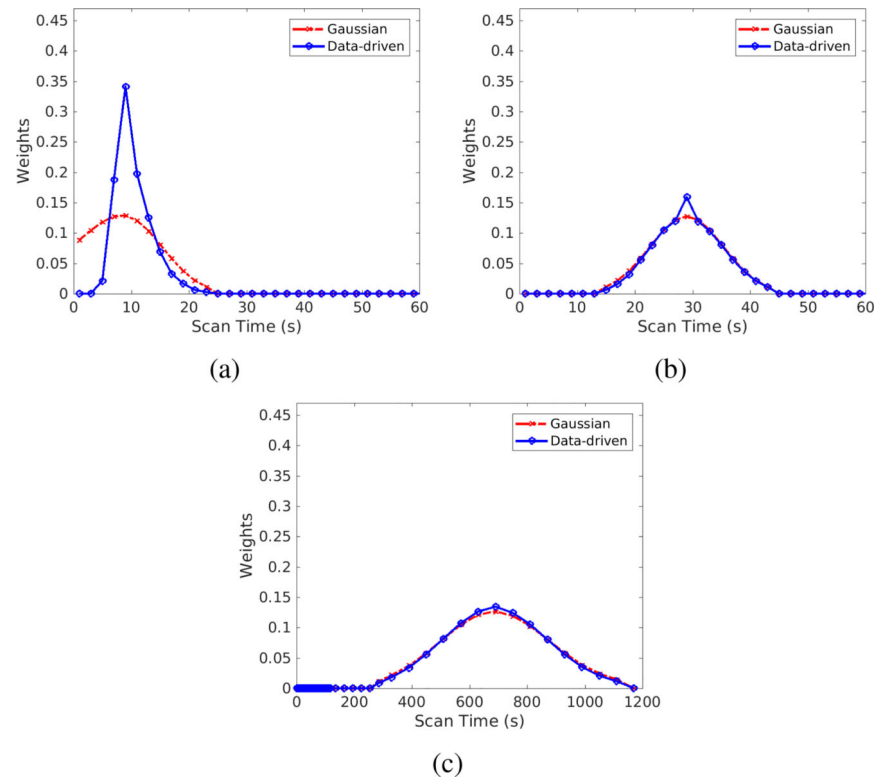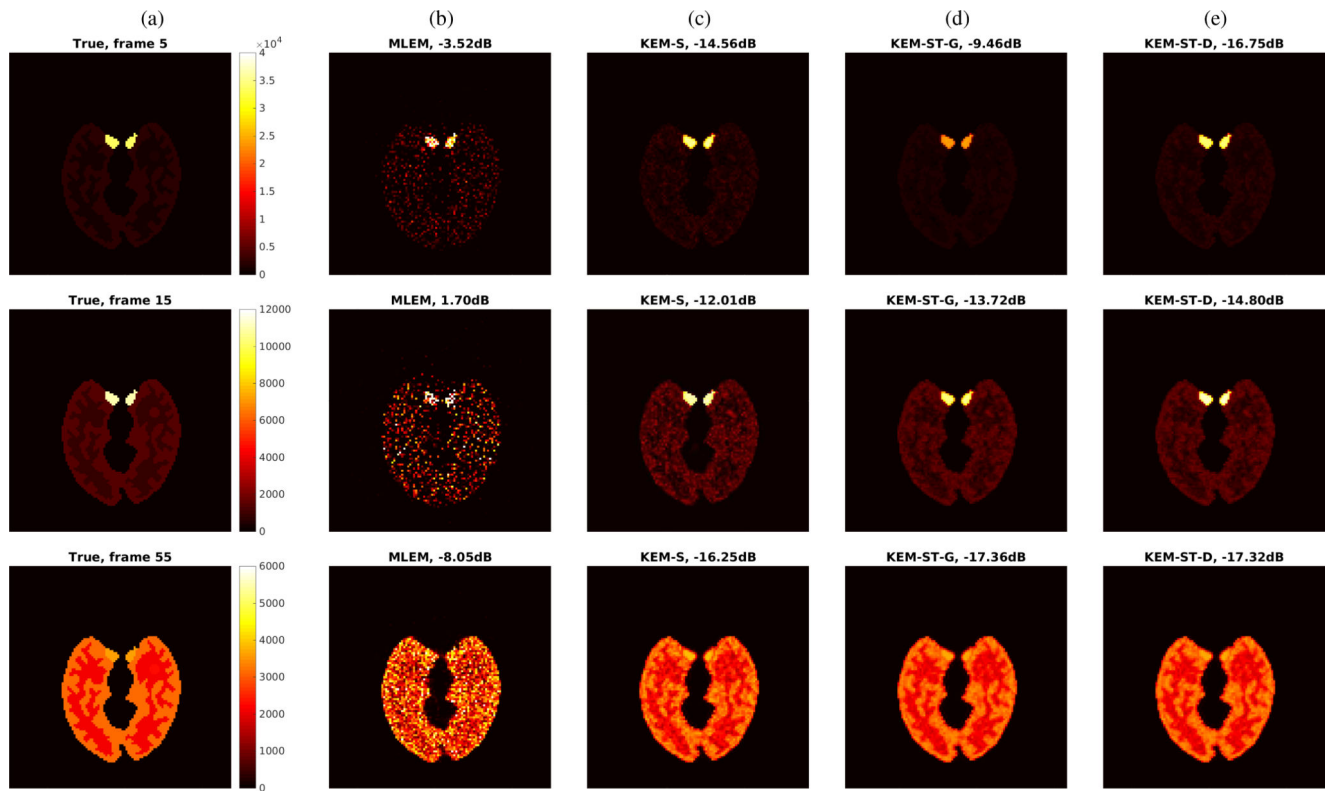
Fig. 2.
Temporal basis functions formed by the Gaussian kernel and data-driven kernel. (a) center located at 5th frame (t=8–10s); (b) centered at 15th frame (t=28–30s); (c) centered at 55th frame (t=11–12 minute).

Fig. 3.
True activity image and reconstructed images at iteration 100 by different reconstruction methods for the 5th frame (t =10 s, top row), 15th frame (t = 30 s, middle row) and 55th frame (t = 12 minutes, bottom row). (a) True images, (b) MLEM reconstructions, (c) reconstructions by the spatial kernel method (KEM-S), (c) by the spatiotemporal kernel method with the time-invariant Gaussian smooth kernel (KEM-ST-G), (e) by the spatiotemporal kernel method with a data-driven temporal kernel (KEM-ST-D).
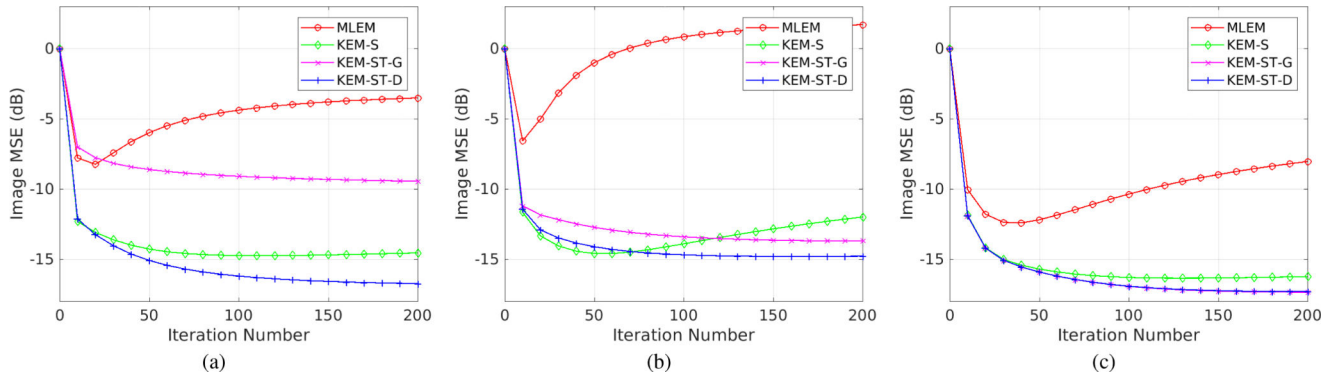
Fig. 4.

Plots of image MSE as a function of iteration number for different reconstruction methods.

(a) frame 5, (b) frame 15, (c) frame 55.
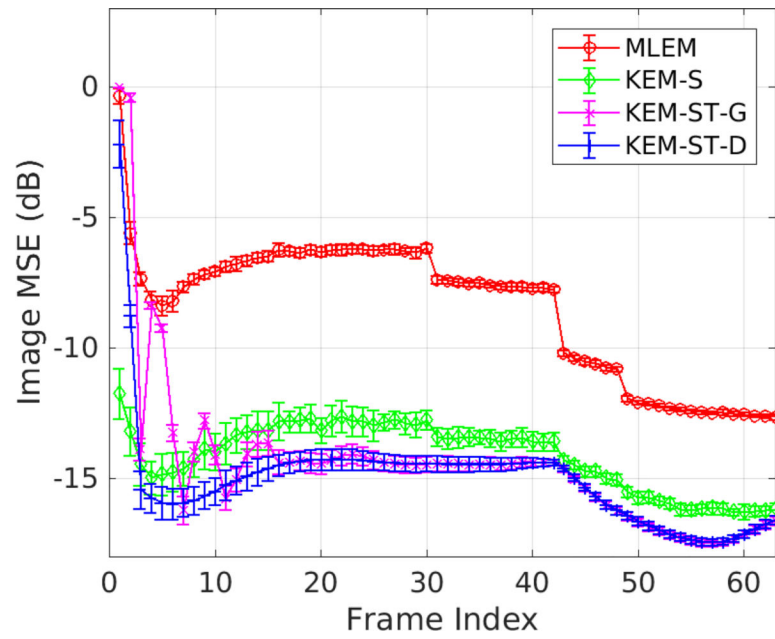
Fig. 5.
Plots of image MSE of all time frames reconstructed by different reconstruction methods.
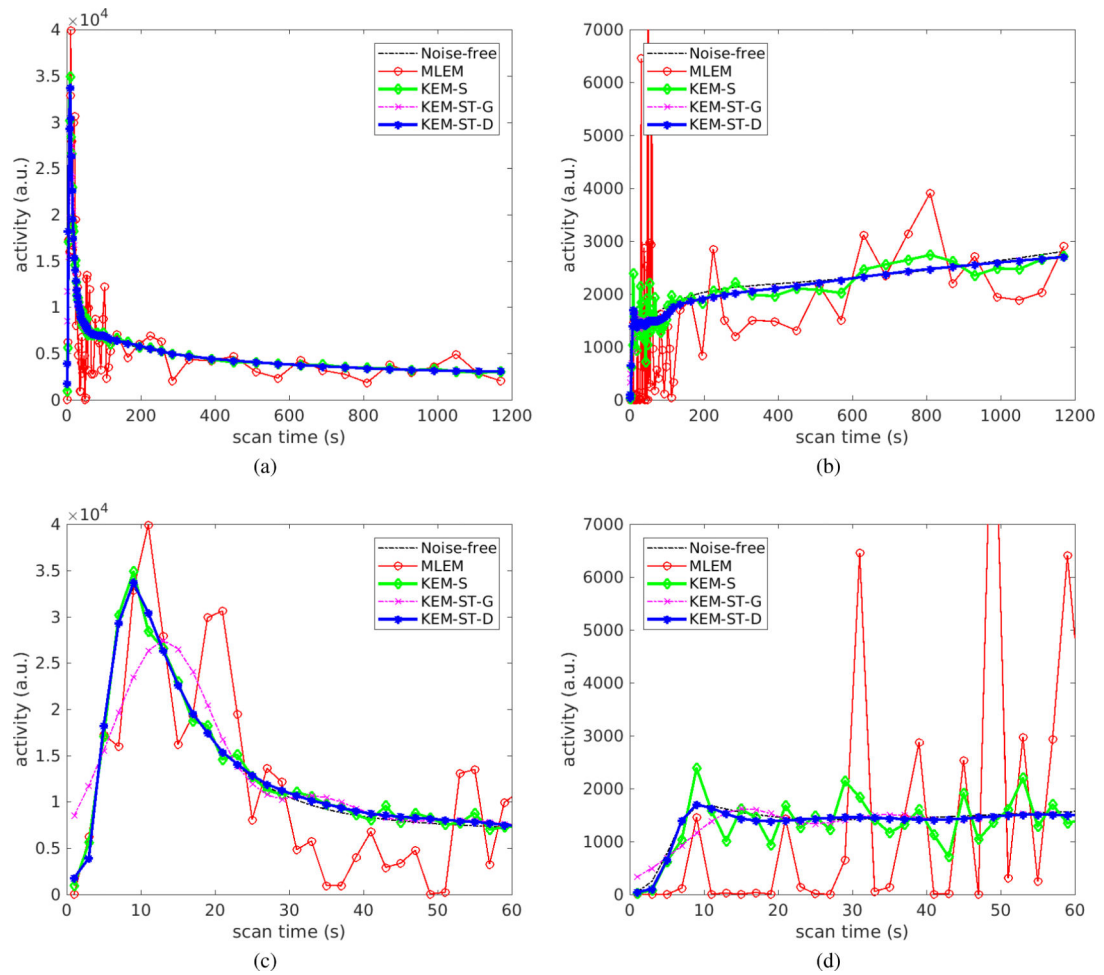
Fig. 6.
Time activity curves reconstructed by different reconstruction methods. (a) Blood TAC, 0–20 minutes, (b) tumor TAC, 0–20 minutes, (c) blood TAC, 0–60 seconds, (d) tumor TAC, 0–60 seconds.
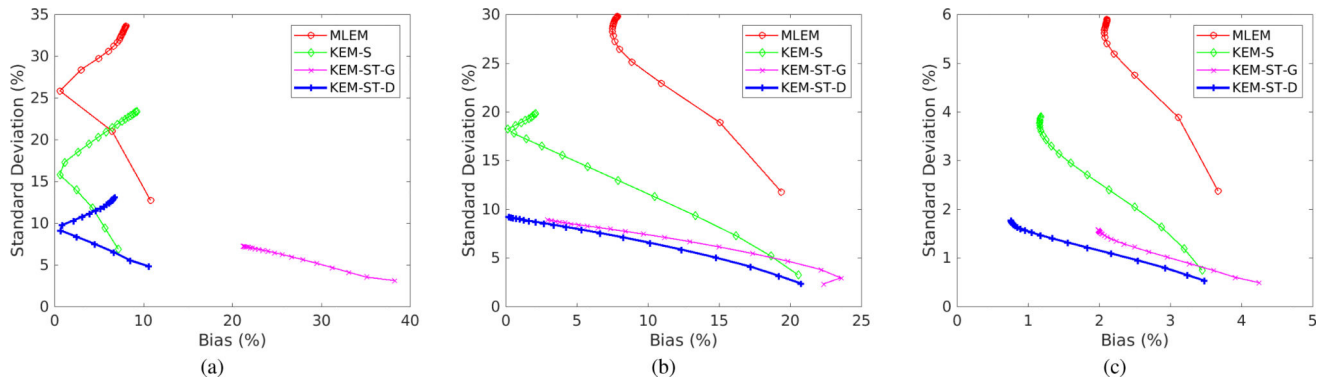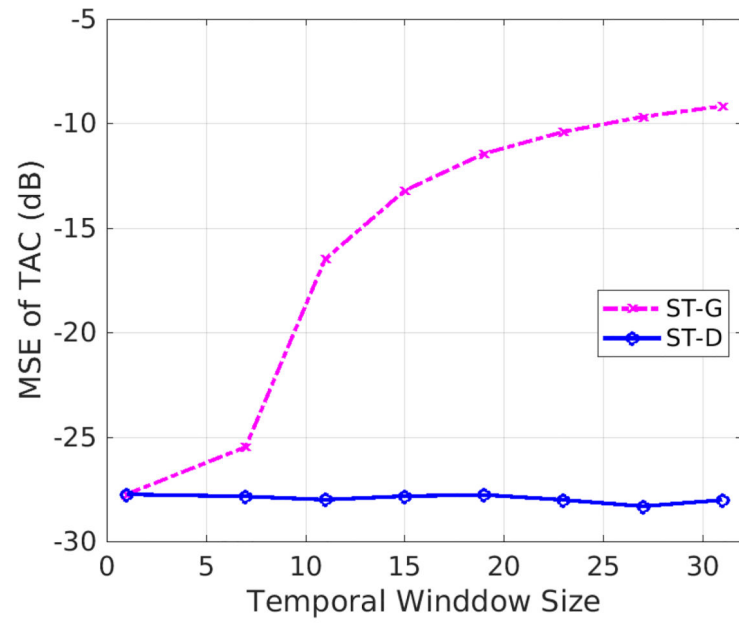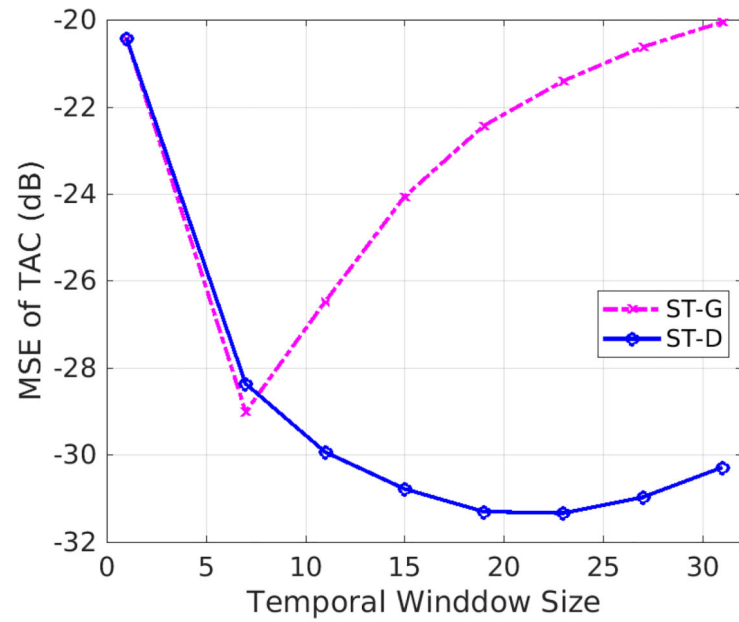
Fig. 7.
Standard deviation versus bias trade-off of tumor ROI quantification by varying iteration numbers from 10 to 200 iterations. (a) frame 5 ($t = 8 - 10$-s), (b) frame 15 ($t = 28 - 30$-s), (c) whole TAC of the first 60 seconds.

(a) Blood

(b) Tumor

Fig. 8.
Effect of temporal window size on quantification of regional TACs in the simulation study.
(a) Blood ROI, (b) Tumor ROI.

(a) Standard, 10-20 s

(b) HTR, 10-12 s

(c) HTR, 12-14 s

(d) HTR, 14-16 s
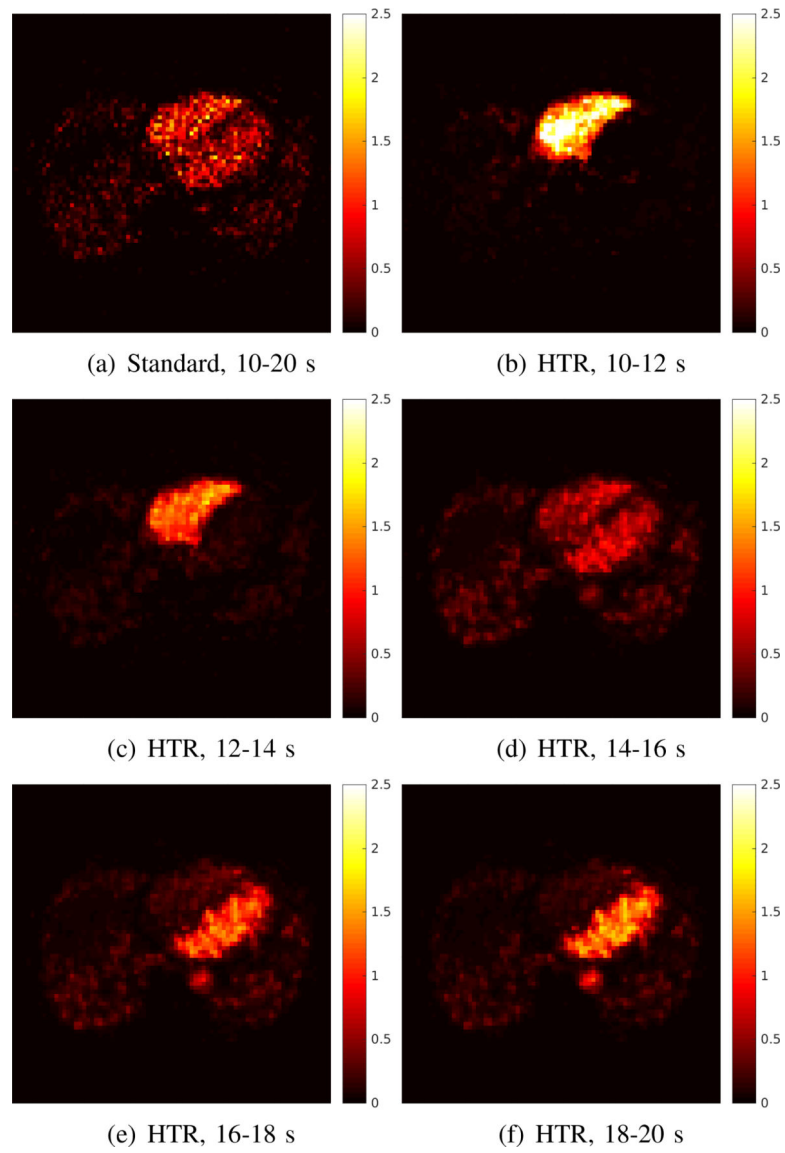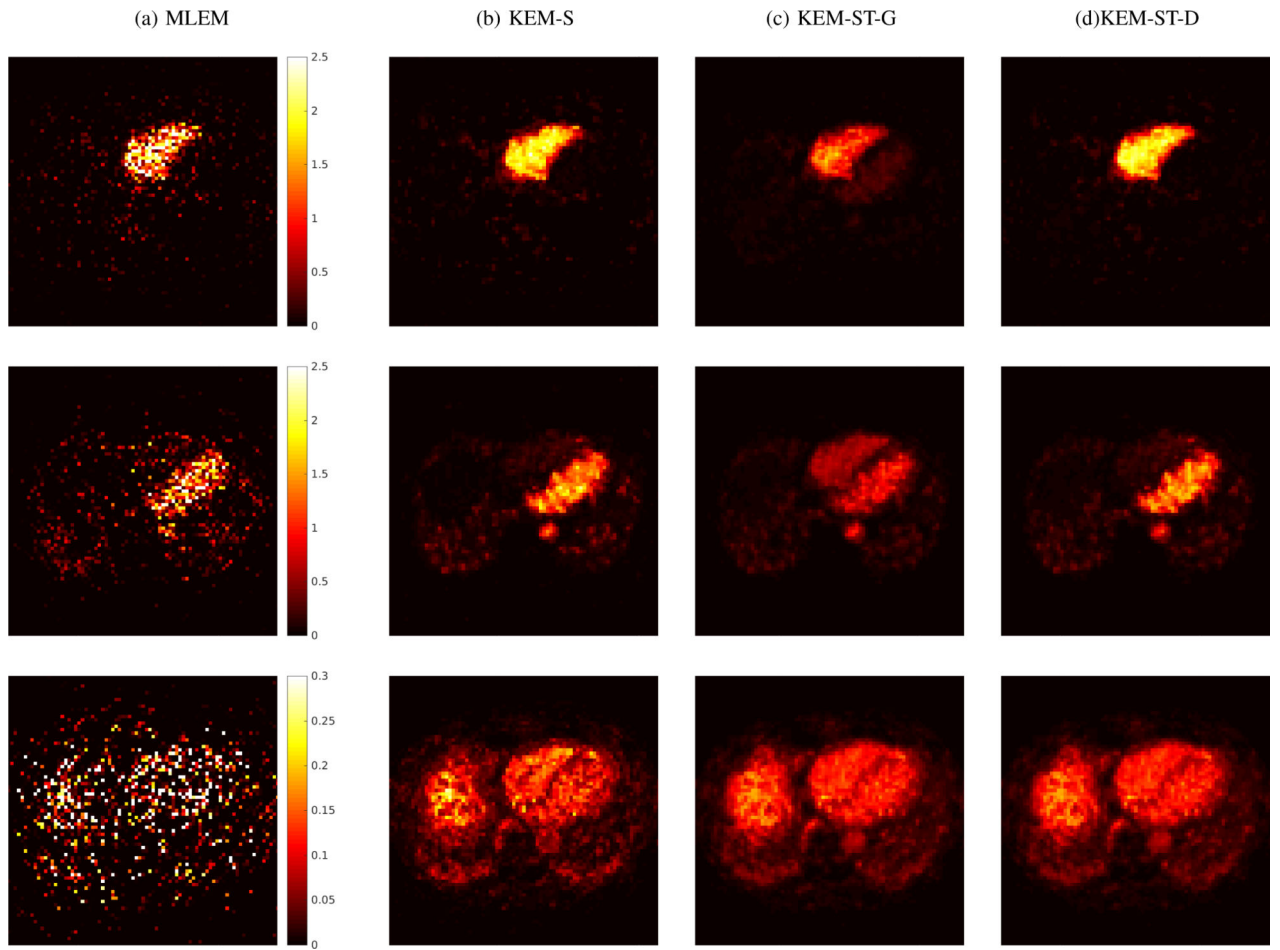
(e) HTR, 16-18 s

(f) HTR, 18-20 s

Fig. 9.
Comparison of a low-resolution time frame starting from 10 to 20s with its corresponding 2-s HTR time frames by the spatiotemporal kernel method for the patient study. (a) 10-s frame at 10–20s by MLEM, (b-f) HTR time frames by KEM-ST-D: (b) 10–12s, (c) 12–14s, (d) 14–16s, (e) 16–18s, (f) 18–20s.

(a) MLEM   (b) KEM-S   (c) KEM-ST-G   (d)KEM-ST-D

Fig. 10.
Comparison of different methods for reconstructing HTR patient data at $t = 8 − 10s$ (top row), $t = 18 − 20s$ (middle row) and $t = 88 − 90s$ (bottom row). (a) MLEM, (b) KEM-S, (c) KEM-ST-G, (d) KEM-ST-D.
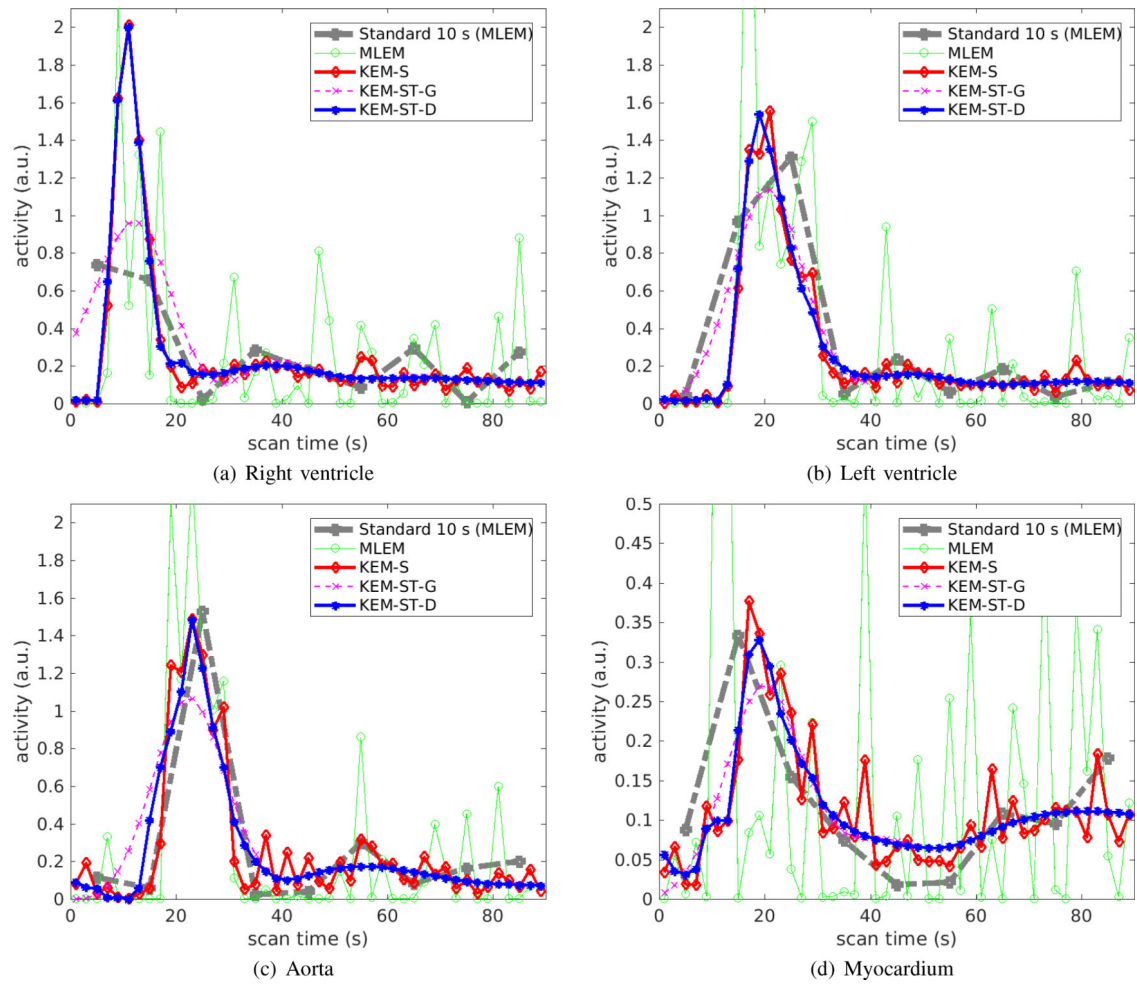
Fig. 11.
Time activity curves reconstructed by different methods for the pixels in different regions.
(a) right ventricle, (b) left ventricle, (c) aorta, (d) myocardium.