



HHS Public Access

Author manuscript

Neuroimage. Author manuscript; available in PMC 2022 June 19.

Published in final edited form as:

Neuroimage. 2022 May 15; 252: 119031. doi:10.1016/j.neuroimage.2022.119031.

Adaptive data-driven motion detection and optimized correction for brain PET

Enette Mae Revilla^a,

Jean-Dominique Gallezot^a,

Mika Naganawa^a,

Takuya Toyonaga^a,

Kathryn Fontaine^a,

Tim Mulnix^a,

John A. Onofrey^{a,b,c},

Richard E. Carson^{a,c},

Yihuan Lu^{a,*}

^aDepartment of Radiology and Biomedical Imaging, Yale University, PO Box 208048, New Haven, CT 06520-8048, USA

^bDepartment of Urology, Yale University, New Haven, CT, USA

^cDepartment of Biomedical Engineering, Yale University, New Haven, CT, USA

Abstract

This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>)

*Corresponding author. Yihuan.lu@yale.edu (Y. Lu).

Credit authorship contribution statement

Enette Mae Revilla: Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing original draft, Writing – review & editing, Visualization. **Jean-Dominique Gallezot:** Software, Validation, Formal analysis, Resources, Writing – original draft. **Mika Naganawa:** Software, Validation, Formal analysis, Resources, Writing – original draft. **Takuya Toyonaga:** Software, Validation, Formal analysis, Resources, Writing – original draft. **Kathryn Fontaine:** Software, Validation, Resources, Data curation, Writing – original draft. **Tim Mulnix:** Software, Validation, Resources, Data curation. **John A. Onofrey:** Software, Validation, Resources, Writing – original draft. **Richard E. Carson:** Methodology, Validation, Writing – original draft, Supervision, Project administration, Funding acquisition. **Yihuan Lu:** Conceptualization, Methodology, Validation, Data curation, Writing – original draft, Writing – review & editing, Visualization, Supervision, Project administration, Funding acquisition.

Declaration of Competing Interest

The other authors declare that they have no conflict of interest.

Ethical approval

The study protocol involving human participants was approved by the Yale Human Investigation Committee, the Yale-New Haven Hospital Radiation Safety Committee, and the Yale University Radiation Safety Committee. Study procedures were performed in accordance with federal guidelines and regulations of the United States for the protection of human research subjects contained in Title 45 Part 46 of the Code of Federal Regulations (45 CFR 46).

Consent to participate

Informed consent was obtained from all individual participants included in the study.

Code availability

The source code used in this study is openly available in GitHub at https://github.com/enetterevilla/COD_MotionDetection.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2022.119031.

Head motion during PET scans causes image quality degradation, decreased concentration in regions with high uptake and incorrect outcome measures from kinetic analysis of dynamic datasets. Previously, we proposed a data-driven method, center of tracer distribution (COD), to detect head motion without an external motion tracking device. There, motion was detected using one dimension of the COD trace with a semiautomatic detection algorithm, requiring multiple user defined parameters and manual intervention. In this study, we developed a new data-driven motion detection algorithm, which is automatic, self-adaptive to noise level, does not require user-defined parameters and uses all three dimensions of the COD trace (3DCOD). 3DCOD was first validated and tested using 30 simulation studies (^{18}F -FDG, $N=15$; ^{11}C -raclopride (RAC), $N=15$) with large motion. The proposed motion correction method was tested on 22 real human datasets, with 20 acquired from a high resolution research tomograph (HRRT) scanner (^{18}F -FDG, $N=10$; ^{11}C -RAC, $N=10$) and 2 acquired from the Siemens Biograph mCT scanner. Real-time hardware-based motion tracking information (Vicra) was available for all real studies and was used as the gold standard. 3DCOD was compared to Vicra, no motion correction (NMC), one-direction COD (our previous method called 1DCOD) and two conventional frame-based image registration (FIR) algorithms, i.e., FIR1 (based on predefined frames reconstructed with attenuation correction) and FIR2 (without attenuation correction) for both simulation and real studies. For the simulation studies, 3DCOD yielded $-2.3 \pm 1.4\%$ (mean \pm standard deviation across all subjects and 11 brain regions) error in region of interest (ROI) uptake for ^{18}F -FDG ($-3.4 \pm 1.7\%$ for ^{11}C -RAC across all subjects and 2 regions) as compared to Vicra (perfect correction) while NMC, FIR1, FIR2 and 1DCOD yielded $-25.4 \pm 11.1\%$ ($-34.5 \pm 16.1\%$ for ^{11}C -RAC), $-13.4 \pm 3.5\%$ ($-16.1 \pm 4.6\%$), $-5.7 \pm 3.6\%$ ($-8.0 \pm 4.5\%$) and $-2.6 \pm 1.5\%$ ($-5.1 \pm 2.7\%$), respectively. For real HRRT studies, 3DCOD yielded $-0.3 \pm 2.8\%$ difference for ^{18}F -FDG ($-0.4 \pm 3.2\%$ for ^{11}C -RAC) as compared to Vicra while NMC, FIR1, FIR2 and 1DCOD yielded $-14.9 \pm 9.0\%$ ($-24.5 \pm 14.6\%$), $-3.6 \pm 4.9\%$ ($-13.4 \pm 14.3\%$), $-0.6 \pm 3.4\%$ ($-6.7 \pm 5.3\%$) and $-1.5 \pm 4.2\%$ ($-2.2 \pm 4.1\%$), respectively. In summary, the proposed motion correction method yielded comparable performance to the hardware-based motion tracking method for multiple tracers, including very challenging cases with large frequent head motion, in studies performed on a non-TOF scanner.

Keywords

PET; Head motion; Data-driven; Motion detection; Motion correction; COD

1. Introduction

Head movement is a major limitation in brain positron emission tomography (PET) imaging, reducing image resolution, lowering apparent concentration in high-uptake regions, introducing attenuation-emission mismatch artifacts, and causing bias in parameter estimates fit by tracer kinetic modeling (Keller et al., 2012). In the past, many methods have been proposed to correct head motion, including frame-based image-registration (FIR) and correction using real-time hardware-based motion tracking (HMT) information (Rahmim et al., 2007; Fulton et al., 2002; Montgomery et al., 2006; Bloomfield et al., 2003; Herzog et al., 2005; Picard and Thompson, 1997; Costes et al., 2009). However, FIR cannot correct for motion within one predefined scan period (intra-frame) while HMT is not routinely used in the clinic, as setup and calibration of the tracking device can be complicated and attaching

markers to each patient increases the logistical burden of the scan. Markerless motion tracking using structured light does not require any physical attachment to the patient, but its performance can be impacted by non-rigid changes in facial expression and different skin colors. Additionally, structured light approaches have not yet been fully validated (Kyme et al., 2018; Olesen et al., 2013).

Recently, several data-driven motion correction approaches (i.e., based on PET raw data itself) have been proposed. Schleyer et al. used a principal component analysis (PCA)-based method, which was first proposed in (Thielemans et al., 2013), to detect head motion with the aid of time-of-flight (TOF) information (Schleyer et al., 2015; Thielemans et al., 2013). Lu et al. proposed another data-driven algorithm, Centroid Of Distribution (COD), to detect patient motion (Lu et al., 2020; 2019b). Both PCA-based and COD-based methods used a one dimensional (1-D) PCA or COD trace, where motion detection was treated as an edge detection problem on the 1-D trace. PCA and COD traces are intrinsically noisy due to the limited count-statistics of the PET raw data. Motion detection also requires tuning multiple user-defined parameters, including the size of a median filter used to smooth the trace and a threshold for detecting motion. Therefore, these parameters control the sensitivity of motion detection, and were empirically set, rather than statistics-based, in both (Schleyer et al., 2015) and (Lu et al., 2020). Thus, these detection algorithms are tracer-distribution and count-level dependent. Note that count-level in the raw data is subject-dependent, and within the same scan, the count-level can vary greatly due to the tracer distribution change as well as the radioactive decay. Optimization of the user-defined parameters is, therefore, non-trivial. Thus, it is of interest to develop a robust data-driven motion detection approach, which does not require subject-dependent or count level-dependent parameter tuning. In (Spangler-Bickell et al., 2021), Spangler-Bickell et al. recently performed an investigation, which focused on the proper frame duration for data-driven rigid motion estimation. Their proposed method utilized ultra-short frames for rigid motion estimation and correction. Comparably, we incorporated both short and long frames for motion estimation and correction. However, in Spangler-Bickell et al. (2021), only static data was used where tracer distribution change over time is insignificant while dynamic data sets were used in this paper. Moreover, the impact of motion correction on the accuracy of absolute quantification was not investigated.

In this study, we propose a new count statistics-based data-driven motion detection algorithm, which does not require user-parameter tuning and uses all three dimensions of the COD trace. We validated the proposed motion detection algorithm using 30 simulated 4-D (3-D + time) dynamic PET studies with large motion for both ^{18}F -FDG and ^{11}C -raclopride (^{11}C -RAC) tracers. For the simulation studies, the proposed motion correction method was compared to two types of FIR algorithms and perfect motion correction, i.e., the same motion used for simulation was also used for correction (Costes et al., 2009). For real studies, the proposed motion correction method was evaluated for human dynamic scans with ^{18}F -FDG ($N=10$), ^{11}C -RAC ($N=10$), ^{11}C -PBR28 ($N=1$), and ^{11}C -MRB ($N=1$); and was compared to HMT with the Polaris Vicra tracking system (NDI Systems, Waterloo, Canada) (referred to as Vicra), which provides continuous head motion monitoring at 20 Hz and was considered as the “gold standard” (Jin et al., 2013).

2. Methods

2.1. Dynamic 4D PET simulation

2.1.1. 4D digital phantom—Fifteen digital phantoms were generated based on real human PET studies, which were previously performed at the Yale PET Center. Each of the 15 subjects underwent separate PET scans with ^{18}F -FDG and ^{11}C -RAC. An individual attenuation map was obtained through a transmission scan before each PET scan. Each subject underwent MR scans and the individual T1-weighted MR images were segmented into 109 regions of interest (ROIs) using FreeSurfer (Fischl et al., 2004; 2002). The ROIs were resliced to the individual PET space based on the MR-PET rigid registration, which was performed using FLIRT with mutual information as the similarity metric (Jenkinson and Smith, 2001). Each individual phantom consists of 109 ROI labels in PET space ($256 \times 256 \times 207$ voxels with $1.219 \times 1.219 \times 1.231$ mm³/voxel).

Motion was tracked by the Vicra system. Individual motion information, recorded at 20 Hz by the Vicra system during each real scan, was used in the simulation for the same individual phantom. The 15 subjects who underwent the largest motion magnitude were selected out of 57 examined cases for ^{18}F -FDG and 143 for ^{11}C -RAC. The head motion magnitude of any frame within the field-of-view (FOV) was determined from the Vicra data as twice the standard deviation of the location of eight points that were selected as the vertices of a 10-cm side-length cube centered in the scanner FOV. The final motion magnitude was the average of the values from the eight points (Jin et al., 2013).

We simulated ^{18}F -FDG and ^{11}C -RAC studies using the estimated kinetic parameters from two real datasets. Specifically, paired MR scans of the two real studies were segmented using FreeSurfer. For each FreeSurfer ROI, a compartmental model (the two-tissue compartment model for ^{18}F -FDG, and the simplified reference tissue model for ^{11}C -RAC) was used to fit the time-activity curve to generate the tracer kinetic parameters (Gallezot et al., 2020). The kinetic parameters were then used to generate noise-free time-activity curves for each ROI. The same time-activity curves were finally used for each ROI of all 15 subjects for each tracer, i.e., different brains with different motions but with the same tracer dynamics were simulated.

2.1.2. Data simulation—For every simulated study, a 4-D 0–90 min dynamic simulation was performed for ^{18}F -FDG ($N = 15$) and 0–60 min dynamic simulation for ^{11}C -RAC ($N = 15$), in the presence of motion, to generate list-mode data. Simulations were performed for both the non-TOF Siemens HRRT and the TOF Siemens Biograph mCT scanners (Schmand et al., 1999; Jakob et al., 2011). The list-mode TOF forward-projector model is part of the MOLAR (Motion compensation OSEM List-mode Algorithm for Resolution-Recovery Reconstruction) platform (Germino et al., 2017):

$$E(Y_{i,t,\tau}) = D_t \left(\sum_j c_{i,t,j} \zeta_{i,t,\tau,j} L_{i,t} A_{i,t} N_i \lambda_{j,t} \right), \quad (1)$$

where the system matrix element $c_{i,t,j}$ represents the contribution from voxel j to the line-of-response (LOR) i at time t and accounts for scanner geometry, resolution, solid angle, and

motion effects. $\zeta_{i,t,\tau,j}$ is the TOF kernel, which defines the contribution of pixel j to TOF bin τ of LOR i at time t . $L_{i,t}$ is the dimensionless product of decay factor at time t , live time at time t , and the positron branching fraction. The sensitivity normalization factor N_p in units of (counts/second) / (Bq/mL \times mm), converts the forward projection of image $\lambda_{j,t}$ (Bq/mL) to units of counts/second. $A_{i,t}$ is the dimensionless attenuation factor. $E(Y_{i,t,\tau})$ is the expected number of counts in the TOF bin τ of LOR i in time bin t . The total scan time frame T (second) is divided into equal sub-bins of duration D_t (second) indexed by t . For non-TOF HRRT simulations, $\zeta_{i,t,\tau,j}$ is set to 1 and the τ index is eliminated. Data were simulated with and without motion. Compton scatter and randoms were not simulated. Spatial resolution of 2.5 mm Full-Width-Half-Maximum (FWHM) was simulated for HRRT and 4.0 mm was simulated for mCT. TOF resolution of 580 ps in FWHM was used in the mCT simulations.

2.2. Real patient studies

Ten previously acquired human dynamic HRRT ^{18}F -FDG (injected activity: 184 ± 4 MBq) and ten HRRT ^{11}C -RAC (699 ± 66 MBq) studies were analyzed. A transmission scan was used for attenuation correction. Individual T1-weighted MR images were segmented into 109 ROIs using FreeSurfer, which were registered and resliced to the individual PET space based on the MR-PET rigid registration using mutual information (Fischl et al., 2004; 2002; Jenkinson and Smith, 2001).

Two dynamic mCT scans using ^{11}C -PBR28 (0–90 min post injection) and ^{11}C -MRB (0–120 min) were analyzed. ^{11}C -PBR28, a PET radioligand that has high specificity for the 18-kDa translocator protein (TSPO) is used for *in vivo* measurement of neuroinflammation (Hillmer et al., 2020; Owen et al., 2014). ^{11}C -MRB is used to measure brain norepinephrine transporters (NET) (Ding et al., 2010). Vicra HMT was used for motion monitoring for all the real scans for both scanners.

2.3. COD generation and previous detection algorithm

Head motion information was extracted from the PET list-mode data based on the COD trace (Lu et al., 2020; 2019b). To generate COD for TOF raw data, for every event i , the central spatial coordinate of the TOF bin, (X_j, Y_j, Z_j) , was calculated and recorded in mm from the center of the scanner FOV. The (X_j, Y_j, Z_j) values for each event were averaged over a short time interval, e.g., 1 s, to generate raw COD traces in three directions: C_x for lateral, C_y for anterior-posterior (AP), and C_z for superior-inferior (SI) directions. For non-TOF COD, (X_j, Y_j, Z_j) was the central coordinate of the entire LOR (Lu et al., 2020). Note that the random component of COD in each direction follows a normal distribution based on the central limit theorem.

Previously, we developed a semiautomatic algorithm to detect motion based on the COD trace, which required multiple user-defined parameters that are tracer- and noise-level dependent (Lu et al., 2020; 2019b) such as: (1) length of the median filter applied to the COD trace, which was adjusted based on empirical observation; (2) the minimum duration of the motion-free frames (MFF), which was set to 30 s (the length of the shortest dynamic frame); (3) the 5-min maximum length of an MFF as a tradeoff among the sensitivity

to slow motion, the computational cost and registration accuracy and (4) the smoothing kernel for MFF reconstructions. Note that these user-defined parameters are related to motion-thresholding only. The dependency on tracer- and noise-level is further explained under the limitations in the Discussion section. After motion detection, MFF separated by detected motion time points (MTP) were reconstructed without attenuation correction (AC) and rigidly registered to a reference frame to estimate motion between each MFF and the reference frame.

2.4. New adaptive data-driven motion detection algorithm

In this study, we propose a new motion detection algorithm, which automatically detects motion based on the COD trace and is adaptive to different tracers and noise levels without user-defined parameters. First, we provide an overview of the new algorithm (Fig. 1). All the symbols used in the detection algorithm are summarized in Table 1.

Let n be a possible number of MTPs on the COD trace. For a given n , (e.g., $n = 10$) in Fig. 1A, the proposed motion detection algorithm finds the positions of all n MTPs that minimize an estimate of total variability within-MFF, i.e., error Fig. 1.B and C show the scenarios of $n = 20$ and 30 for the same study, respectively. The algorithm will finally choose the n by directly estimating the error level from the COD itself, and the corresponding MTPs are the motion detection results. We will now describe the new algorithm in detail.

2.4.1. RSS error within an MFF—We will evaluate a range of n , and the MTPs/MFFs as follows: for a given detection result, i.e., a set of MTPs and corresponding MFFs for a *given* n , we calculate the mean of the COD trace in one direction, i.e., C , *within* each MFF, and compute the residual sum of squares (RSS) of the fitting. We use C_m to denote the COD segment within the m^{th} MFF. Note that the notation applies to any of the three COD directions. RSS of the m^{th} MFF is computed as:

$$\text{RSS}_m = \sum_k^{N_m} (C_{m(k)} - \bar{C}_m)^2, \text{ and } \bar{C}_m = \frac{1}{N_m} \sum_k^{N_m} C_{m(k)}, \quad (2)$$

where k indexes the sampling interval, e.g., 1 s, within C_m . m indexes the MFF. N_m is the total number of sampling intervals within the m^{th} MFF Eq. (2). is repeated for all m . By taking a sum of RSS over all the MFFs, we obtain the total residual error, i.e., E , of *one* detection scenario:

$$E = \sum_m \text{RSS}_m. \quad (3)$$

2.4.2. Choosing the MTPs for a given n using the pruned exact linear time algorithm—Here, we describe how to use Pruned Exact Linear Time (PELT), a changepoint detection algorithm, to find the n MTPs locations in the COD trace that yields the *lowest* total residual error, i.e., $E_{\min}(n)$ (Killick et al., 2012). Note that PELT guarantees finding $E_{\min}(n)$ on the condition that the noise in the COD trace is normally distributed. PELT was established based on an optimal partitioning (OP) approach, where OP belongs

to one research branch in the changepoint detection literature (Jackson et al., 2005). By adding a pruning phase to the original OP algorithm, Killick et al. (2012), Jackson et al. (2005) reduced the computational cost from $\mathcal{O}(T^2)$ to $\mathcal{O}(T)$, i.e., the computational cost linearly increases with the total data points T . As shown in (Killick et al., 2012), $E_{\min}(n)$ is a monotonically decreasing function of n (see Fig. 2A for example). Naturally, $n(\bullet)$ is a monotonically decreasing function of E_{\min} as well. This characteristic of PELT allows us to convert the motion detection problem, i.e., finding the proper choice of n (and its corresponding MTPs/MFFs), to a problem of finding the proper E_{\min} , since each E_{\min} uniquely corresponds to a set of n MTPs yielded by the PELT. The locations of the n MTPs are the motion detection results for a given E_{\min} . Note that under an extreme scenario, i.e., n equals the number of sampling time points in the COD trace, E_{\min} reduces to zero. This is, of course, a meaningless detection result.

So far, we have introduced PELT, which chooses the optimal MTPs for a *given* n . Next, we will describe how to choose the desired n .

2.4.3. Determining n —First, we apply PELT for a range of n , which generates an E_{\min} vs. n plot (Fig. 2A), up to a predefined maximum n_{\max} , which should be higher than the most times of motion one would like to detect. In this study, we set the default n_{\max} to 300 for a 90-min study and 200 for a 60-min study. We refer to this step as E_{\min} scouting. Next, we find the n that corresponds to a target detection (n_{tar}), which matches a unique E value, E_{tar} . Ideally, n_{tar} and its corresponding MTPs represent the correctly detected motions based on C . Here, we hypothesized that the value of E_{\min} at which head motion is accurately detected (E_{tar}) should be approximately equal to the error for a comparable *hypothetical* scan where no motion occurred (E_{NM}), i.e., same subject with the same injection but without motion. Note that, if no motion occurred, the variation in a COD trace would only be caused by the data noise and tracer distribution change. Of course, for a real study, it is impossible to directly obtain the E_{NM} , since head motion will almost always occur.

Next, we show how to estimate E_{NM} based on the COD trace (with motion) itself, i.e., C . Essentially, we used partial data (without or with minimal motion) of the COD trace to estimate the total RSS contribution from data noise as well as tracer distribution change. We first divide the entire scan into several equal-duration partitions, e.g., 5 min per partition, indexed by s . The use of partitions makes the algorithm adaptive to the COD noise change due to isotope decay and tracer clearance. Based on the PELT detection results using n_{\max} , we use the RSS (RSS_{LONG}) from the *longest* PMFFs (indexed by p) within partition s to estimate the no motion RSS for partition s :

$$\text{RSS}_{\text{NM},s} \approx N_{\text{PART},s} \cdot \frac{\sum_p \text{RSS}_{\text{LONG},p}}{\sum_p N_{\text{LONG},p}}, \quad (4)$$

where $N_{\text{LONG},p}$ and $N_{\text{PART},s}$ are the total numbers of sampling intervals of the p^{th} MFF among the longest PMFFs and the total numbers of sampling intervals of the s^{th} partition, respectively (Fig. 2B). The same process is repeated for every partition. $\text{RSS}_{\text{NM},s}$ from all the partitions are summed and used as the E_{NM} estimate for the entire scan:

$$E_{NM} \approx \sum_s \text{RSS}_{NM,s}. \quad (5)$$

Note that if there are less than P MFFs in one partition, then all the MFFs inside the partition are used to estimate E_{NM} . If an MFF m extends beyond one partition (s), the portion of the MFF m within the partition is used to compute $\text{RSS}_{NM,s}$, i.e., $\text{RSS}_{NM,s} = N_{\text{PART},s} \cdot \frac{\text{RSS}_m}{N_m}$. In this study, we set P equal to 2 (see Discussion).

We use the E_{\min} vs. n plot to find the n which yields the error level E_{tar} , and the results of PELT with n_{tar} shall be the final motion detection MTP set. We repeat the above detection process for all three COD directions, and a union of the MTPs from all the three directions is taken and is used as the final detection results. Note that motion detection in each direction was performed independently, so different n_{tar} will be different for each direction. In the validation studies (Supplementary Material), we found that there is approximately a multiplicative constant between the measured E_{NM} based on simulation studies and the estimated E_{NM} using Eq. (5). Since the measured E_{NM} , which is referred as E_{tar} , is what we are interested in, we introduce a constant $\alpha \geq 1$ to calculate the final E_{tar} :

$$E_{\text{tar}} \approx \alpha E_{NM} = \alpha \sum_s \text{RSS}_{NM,s}, \alpha \geq 1 \quad (6)$$

To determine α , we performed comparable simulations as in the Data simulation section but *without* motion, i.e., data were simulated using the same subjects but without motion. Based on the COD trace without motion, for each study, we measured the E_{NM} , i.e., E_{tar} , based on the MFFs, which were determined during the detection using the COD *with* motion in the motion detection process. In other words, we applied the motion detection results, i.e., the MFFs, from COD with motion to the COD without motion, to obtain E_{tar} . We examined different values of α and found that the difference (average over three directions) between estimated E_{NM} (Eq. (5)) and the E_{tar} was the smallest when $\alpha = 1.0$ and 1.6 for ^{18}F -FDG and ^{11}C -RAC, respectively. Supplemental Table 1 shows the results of E_{tar} and E_{NM} for different scanners and tracers at different directions across 15 subjects. A detailed comparison of E_{tar} and E_{NM} for 5 representative studies is shown in Supplemental Table 2 for ^{18}F -FDG and Supplemental Table 3 for ^{11}C -RAC. Simulation studies (see Fig. 5 and Supplemental Tables 1–3) showed that α equals 1.0 for ^{18}F -FDG, and α equals 1.6 for ^{11}C -RAC; these values were used as the default values in the rest of the paper.

Fig. 2B shows an example of motion detection in one direction for a simulated ^{18}F -FDG scan. Supplemental Fig. 1 shows the pseudo code of the entire adaptive motion detection process.

2.5. Final MFF determination

We observed that, after motion detection, the COD of some MFFs still contain large variations, which represent intra-frame motions and may cause inaccurate downstream motion estimation. Here, we discarded the MFFs with large intra-frame motions. Specifically, we consider an MFF to contain substantial intra-frame motions if the standard

deviation of the COD trace within the MFF is larger than two times the predicted standard deviation of the partition, i.e., $2\sqrt{\text{RSS}_{\text{NM},s}/N_{\text{PART},s}}$, in which the MFF resides Fig. 3.A shows an example of raw COD traces in three directions: C_x for lateral, C_y for anterior-posterior and C_z for superior-inferior directions Fig. 3.B shows the detection result on C_x . The green vertical lines indicate the detected MTPs. The horizontal line segments at the top of the graph indicate a preserved motion-free frame and the short line segments at the bottom of the graph indicate discarded frames due to intra-frame motions. Here, we only discard frames with large intra-frame motions while in Sec. III.B, we will show that frames with too few counts will also be discarded, since motion usually cannot be robustly estimated for those frames.

2.6. Optimized motion estimation and event-by-event corrected reconstruction

Motion between MFFs was estimated and corrected as follows: (1) each MFF was reconstructed using OSEM without attenuation correction; (2) the MFF image was smoothed by a w -mm FWHM Gaussian filter (see below for w), followed by a rigid registration to a 10-min post-injection reference frame; (3) a motion file consisting of all the registration transformation matrices, i.e., motion information, was built and events within one MFF were corrected using the same motion information; (4) MOLAR was used to perform event-by-event motion compensated OSEM reconstruction (2 iterations \times 30 subsets for HRRT and 3 iterations \times 21 subsets for mCT) with attenuation correction. Note that the final MOLAR reconstruction can be performed for any frame duration, which does not depend on the MFF timing, given its event-by-event correction nature. The discarded frames were excluded through the gating function in MOLAR. MFF registration was performed using the BioImage Suite image analysis software (Joshi et al., 2011). Details about the registration parameters can be found in the Supplemental Materials.

To quantitatively evaluate the motion estimation accuracy, we used the mean distance error (MDE, mm) for simulation studies, which is defined as follows: we first computed the center-of-mass (COM) of each FreeSurfer segmented ROI (binary map in PET space, see Supplemental Fig. 2). Ground-truth COM, i.e., COM^{gt} , of each ROI was computed based on the Vicra motion information. Since Vicra motion information was used in the simulation, it provides the ground-truth motion information. The estimated location of each COM, i.e., COM^{est} , based on the registration between each MFF and the reference frame for the entire scan was also computed. MDE is defined as the mean of the Euclidian distance between COM^{gt} and COM^{est} across all ROIs and over all time as:

$$\text{MDE} = \frac{\sum_i^{N_{\text{ROI}}} \left(\frac{\sum_i^{N_{\text{ROI}}} \sqrt{(\text{COM}_{i,t,x}^{\text{gt}} - \text{COM}_{i,t,x}^{\text{est}})^2 + (\text{COM}_{i,t,y}^{\text{gt}} - \text{COM}_{i,t,y}^{\text{est}})^2 + (\text{COM}_{i,t,z}^{\text{gt}} - \text{COM}_{i,t,z}^{\text{est}})^2}}{N_{\text{ROI}}} \right)}{N_{\text{ALL}}} \quad (7)$$

N_{ROI} represents the number of ROI, i indexes the ROI; N_{ALL} represents the number of sampling time points of the entire scan and t indexes the sampling time point. Here, the

COM of each ROI is moved with the ground-truth motion information as well as the estimated motion information, and MDE measures the mean distance between the two.

2.7. Motion correction methods for comparison

Reconstruction results based on the proposed COD-based adaptive motion detection framework (referred as “3DCOD”) were compared with two conventional FIR methods, i.e., FIR1 and FIR2 (see below), and the Vicra-based event-by-event correction. Details regarding Vicra attachment and set-up can be found in Section 2.B in (Jin et al., 2013).

In FIR1, predefined 5-min dynamic frames are first reconstructed using OSEM (2 iterations \times 30 subsets) *with* AC. The same transmission-scan attenuation map is used for attenuation correction for every frame. Each frame reconstruction is smoothed by a 5-mm FWHM Gaussian filter and registered to the reference frame (10 min post injection). As a result, FIR1 suffers from both AC mismatch artifacts and intra-frame motion within each dynamic frame.

FIR2 method is similar to the 3DCOD method, but with the motion information estimated based on the predefined 5-min frames which are reconstructed *without* attenuation correction (NAC). In other words, FIR2 is the same as the 3DCOD method but with MFFs predefined as the consecutive 5-min frames. After the motion file is obtained from registration, MOLAR was used to perform motion compensated OSEM reconstruction (same MOLAR parameters as Section 2.6.) with attenuation correction. As compared to FIR1, FIR2 suffers less from the AC mismatch artifacts but still suffers from intra-frame motions. Furthermore, there would still be some misalignment from missing motion.

2.8. Quantitative analysis

For the simulation studies, motion detection performance was evaluated using *detectability*, which is defined as the ratio between the total number of correctly detected MTPs in all the simulated studies and the total number of ground-truth MTPs (based on the Vicra motion information) in all the simulated studies. The ground-truth MTPs were generated as follows: for each study, FreeSurfer was used to segment the paired MR into 109 ROIs (see Supplemental Fig. 2 for example), which were resliced into the individual PET space. At every time t (20 Hz), the brain location coordinate vector I_b , which is defined as the average coordinate of the COMs of 8 ROIs, i.e., (left and right) frontal, occipital, hippocampus and cerebellum, was computed. I_t was resampled at 1 Hz by averaging over each 20 samples. Euclidian distance ($|I_t - I_{t+1}|$) between I_t and I_{t+1} was computed at all resampled t . Any t that yielded $|I_t - I_{t+1}|$ above a threshold (in mm) was registered as a ground-truth MTP at a given threshold. In this study, nine thresholds were evaluated: 0.5, 1, 2, 3, 4, 5, 6, 7, and 8 mm. If the detection algorithm found a MTP within 1 s before or after a ground-truth MTP, it was registered as a correctly detected MTP. Since 1 s was selected as the minimal interval for generating COD trace, we empirically allow 1 s as the threshold for accepting an MTP match to be selected. Detectability for all three COD directions, as well as using individual C_x , C_y , and C_z , were reported.

Mean and standard deviation of the tracer uptake or standardized uptake value (SUV) percent differences compared with Vicra were computed across all the subjects for eleven

FreeSurfer segmented gray-matter ROIs for ^{18}F -FDG: amygdala, caudate, cerebellum cortex, frontal, hippocampus, insula, occipital, parietal, putamen, temporal and thalamus. For ^{11}C -RAC, the caudate and putamen regions were used. For simulation and real studies, the proposed 3DCOD approach was compared to no motion correction (NMC), FIR1, FIR2 and Vicra. For each correction approach, the MR was registered to the 0–10 min post-injection frame of the individual method.

All the motion detection source code is available on GitHub (https://github.com/enetterevilla/COD_MotionDetection). Detailed information about the software used and computation cost can be found in the Supplemental Materials.

3. Results

3.1. Detection performance

Fig. 4 shows the motion detectability results and false positive rates using HRRT simulation studies (^{18}F -FDG studies (top) and ^{11}C -RAC (bottom)) to compare motion detection based on one direction of the COD (1DCOD) trace vs. 3DCOD. The detectability was above 95% for motion above 2 mm for ^{18}F -FDG and 3 mm for ^{11}C -RAC, using all three directions of COD information (3DCOD, red lines). In comparison, the detectability using only one-COD direction was substantially lower for both scanners. As an example, for motion above 5 mm, the detectability of ^{11}C -RAC in C_y only reached approximately 80% for HRRT. The detectability was lower and the false positive rate was higher for ^{11}C -RAC relative to ^{18}F -FDG, which was expected since ^{11}C -RAC yielded noisier COD and contained more spatially centralized tracer distribution (see Supplemental Fig. 3 for example and see Discussion for details). As expected, the false positive rate decreases as the motion threshold increases. However, our proposed motion detection method showed higher false positive rate than 1DCOD since it incorporates all MTPs from all three COD directions.

Fig. 5 shows the motion detectability and false positive rates at different inflation factor α values (up to 2.0) for simulated HRRT ^{18}F -FDG studies (top row) and ^{11}C -RAC (bottom row). Practically, α controls the tradeoff between the motion detectability and the registration accuracy. Lower α leads to a more sensitive detection but also yields many short MFFs, which may hurt motion estimation accuracy due to higher noise since the motion estimation is performed by the registration between MFFs and a reference frame. In other words, we may be able to detect all the motion, but we may not be able to robustly estimate all of them. Similar to the results shown in Fig. 4, the false positive rate decreases as the α and the motion threshold increase.

Detectability results and false positive rates for the mCT scanner can be found in Supplemental Fig. 4 while the results at different α values can be found in Supplemental Fig. 5. Note that COD with TOF used the center of the TOF bin instead of the center of the LOR, which therefore provided better localization information than non-TOF (see Supplemental Fig. 6 for example). Nevertheless, the motion detection performance and false positive rate of the proposed algorithm was found to be similar between HRRT non-TOF and mCT TOF scanners.

For each tracer, we found that α is consistent across different subjects/scans and is also consistent between mCT and HRRT (see Supplemental Table 1). As empirically validated, α was set to 1.0 for ^{18}F -FDG and 1.6 for ^{11}C -RAC as the default in the rest of the paper. Interestingly, we found that α has very minimal to no effect on the PET quantification. To demonstrate this, we performed additional simulation studies by setting $\alpha = 2.0$ for ^{18}F -FDG and 3.2 for ^{11}C -RAC. Supplemental Fig. 7 and Supplemental Table 4 show the reconstructed results and ROI quantification error, respectively with different α for FDG studies. Supplemental Fig. 8 and Supplemental Table 5 display the results for ^{11}C -RAC. Results showed that there was minimal change across all regions and tracers.

3.2. Motion compensated reconstruction

3.2.1. Simulation—Motion estimation was performed by rigid image registration using the BioImage Suite. Two similarity metrics, i.e., sum of squared differences (SSD) and mutual information (MI), were compared. Detailed results can be found in the Supplemental Materials.

In Fig. 6A, an example of reconstructed HRRT ^{18}F -FDG images are shown. NMC shows very large motion which was largely corrected by 1DCOD while FIR1 and FIR2 failed to yield effective correction. 3DCOD results are slightly better than 1DCOD, with both methods restoring contrast and resolution when compared to ground truth. Quantitatively, with Vicra as the reference, the uptake error mean and standard deviation across the 11 composite brain regions (see Supplemental Table 6 and Supplemental Table 7) for this subject is $-4.9 \pm 3.4\%$ for the 3DCOD method. Averaging regional errors over the 15 subjects, as shown in Table 2, NMC yielded largest mean error -25.4% while FIR1 reduced the error to -13.4% . FIR2 (-5.7%) provided large improvement from FIR1 whereas the 1DCOD method yielded -2.6% mean error. The 3DCOD method produced the smallest mean error at -2.4% , only slightly better than 1DCOD. The inter-subject variation (Average SD in Table 2), averaged over ROIs, was found to be the smallest for 3DCOD (1.4%) and 1DCOD (1.5%) as compared with FIR1 (3.5%), FIR2 (3.6%), and NMC (11.1%). The proposed method outperformed 1DCOD, specifically in the frontal region. The mean absolute error is shown in Supplemental Table 8.

In Fig. 6B, an example of ^{11}C -RAC images are shown, ranking 15/15 ($-6.6 \pm 2.2\%$) in the 3DCOD method in terms of mean percent error compared to Vicra in caudate and putamen. NMC demonstrated the largest error (-34.5%), followed by the FIR1 (-16.1%), FIR2 (-8.0%) and 1DCOD (-5.1%), and the 3DCOD method provided the best result with -3.4% error averaged over 15 subjects and both ROIs. The ROI-level mean of inter-subject deviation was smallest for 3DCOD (1.7%) as compared with the FIR1 (4.6%), FIR2 (4.5%), 1DCOD (2.7%) and NMC (16.1%) Table 3. summarizes these error results. The mean absolute error is shown in Supplemental Table 9.

For the simulated mCT studies, representatives of reconstructed ^{18}F -FDG images are shown in the Supplemental Fig. 10 and quantitative results can be found in Supplemental Table 10. Both 1DCOD and 3DCOD method performed the best with -2.0% mean error as compared with Vicra in mCT studies, followed by FIR2 (-5.8%), FIR1 (-16.7%), and NMC (-27.9%). The mean of inter-subject variation in the ROI-level showed similar trend as compared

to HRRT, with the proposed motion correction method showing the least change (2.0%) compared with other methods (1DCOD: 2.1%, FIR1: 4.2%, FIR2: 3.5%, NMC: 11.5%). Two representatives of reconstructed mCT ^{11}C -RAC images and quantitative results are shown in Supplemental Fig. 11 and Supplemental Table 11, respectively.

3.2.2. Real patient results—In Fig. 7A, an example of 60–90 min ^{18}F -FDG reconstructed images are shown. Mean uptake error of $2.6 \pm 4.0\%$ (ranked 10/10, worst case), as compared to Vicra, was found for 3DCOD. The NMC image shows large head motion in the lateral direction, which was mostly corrected by the 1DCOD method. The proposed 3DCOD method further improved the results, specifically in regions like the frontal cortex and caudate. Even for the worst case scenario, the motion detection algorithm still worked robustly as shown in Table 4, which detailed the quantitative results. The proposed 3DCOD yielded minimal mean error (-0.3%) as compared with FIR1 (-3.6%), FIR2 (-0.6%), 1DCOD (-1.5%) and NMC (-14.9%) across the 10 subjects, using Vicra as the reference. The ROI-level mean of inter-subject deviation was 2.8% for 3DCOD, 4.2% for 1DCOD, 4.9% for FIR1, 3.4% for FIR2 and 9.0% for NMC. For large ROIs which are prone to more motion, e.g., parietal and frontal, the 3DCOD method substantially decreased the percentage error. The mean absolute error is shown in Supplemental Table 12.

Coronal 30–60 min ^{11}C -RAC reconstructed images are shown in Fig. 7B for the worst case (10/10) in the 3DCOD method in terms of mean uptake difference compared to Vicra. Quantitative results are shown in Table 5 for caudate and putamen. Across 10 subjects, NMC showed the largest difference (-24.5%) across all subjects and regions compared with FIR1 (-13.4%), FIR2 (-6.7%), 1DCOD (-2.2%) and 3DCOD (-0.4%). Additionally, the ROI-level mean of inter-subject variation with respect to Vicra was smallest for 3DCOD (3.2%) compared with FIR1 (14.3%), FIR2 (5.3%), 1DCOD (4.1%) and NMC (14.6%). At the caudate, 3DCOD reduced the error from -28.1% (NMC) to -1.8% . The mean absolute error is shown in Supplemental Table 13. The best cases (ranked 1/10) for both tracers can be found in Supplemental Fig. 12.

We also evaluated two studies performed in the mCT scanner to demonstrate that our proposed motion detection and correction method can work in a different scanner. Supplemental Fig. 13A shows coronal 60–90 min SUV images for an mCT ^{11}C -MRB study. Detailed SUV difference results compared to Vicra are shown in Supplemental Table 14. The proposed 3DCOD method yielded smaller mean difference across all regions (2.1%), outperforming NMC (-5.1%). Supplemental Fig. 14 shows the COD traces for this subject, with large motion starting 75 min post injection in the z direction.

Supplemental Fig. 13B (^{11}C -PBR28, mCT) shows coronal 0–90 min reconstructed images. The COD trace for this study, as shown in the Supplemental Fig. 15, indicates the subject underwent large motions in the X and Y directions for the entire scan. 3DCOD restored the image resolution. Visually, 3DCOD and Vicra yielded comparable corrected images. Quantitative results are shown in Supplemental Table 15.

4. Discussion

In this study, we proposed a data-driven head motion detection method that is automated and does not require user-parameter tuning (see Supplemental Table 16 for the recommended default parameters). The new algorithm automatically adapted to the noise changes in the COD for different tracers and was fast to compute (~15 s per study). Registration parameters in the motion estimation process were optimized and the motion correction was performed using event-by-event motion compensated reconstruction. The proposed motion correction method was compared with the hardware-based Vicra method, which was treated as the gold standard for real studies and ground-truth for simulations. The proposed motion correction method was also compared with two conventional frame-based image registration methods.

For both simulation and real studies, ^{18}F -FDG and ^{11}C -RAC were chosen to demonstrate the utility of the algorithm with respect to different tracer distributions and isotopes with different half-lives. Specifically, ^{18}F -FDG is broadly distributed in the brain gray matter with longer decay time while ^{11}C -RAC only contains high uptake in the subcortical region with a much shorter half-life. Image registration for ^{18}F -FDG was found to be more accurate than ^{11}C -RAC, as expected, since ^{11}C -RAC yielded more focal distribution in the central region of the brain which provides fewer image features for image registration, as shown in the registration optimization results (Supplemental Figs. 16, 19). Note that, here, we applied a simple Gaussian filtering, albeit optimized in filter size, to reduce the noise in the MFFs before registration. In the future, we will investigate the use of deep learning-based methods (Lu et al., 2019a) to de-noise the MFFs which may improve the image registration performance.

In the simulation studies for ^{18}F -FDG, parietal and frontal regions yielded relatively larger errors (~5%) for 3DCOD than other regions since both regions usually contain more motion being farther from the rotation axis of the head and are also more prone to attenuation mismatch artifacts for being adjacent to the skull. Overall, in the real studies, the 3DCOD method resulted in slightly better results, i.e., lower mean error, than the simulation studies as compared to Vicra, i.e., ^{18}F -FDG (real: $-0.3 \pm 2.8\%$ error, simulation: $-2.4 \pm 1.4\%$) and ^{11}C -RAC (real: $-0.4 \pm 3.2\%$, simulation: $-3.4 \pm 1.7\%$). We provide three possible reasons for the differences: 1) the studies used for simulation and real studies were different, which exhibited different motion during the scans; 2) simulation is the perfect scenario but our proposed motion correction method is imperfect and still has room for improvement and 3) the real-time hardware-based motion tracking, via Vicra, may be less accurate due to large head motion and fixation issues introducing imperfect motion tracking. Vicra was used as the gold standard motion measure in the real studies, which assumed the Vicra light-reflecting tool to be firmly attached to the subject's head. However, in practice, the tool attachment may fail, e.g., due to hair style, which causes non-rigid motion between the tool and patient head. For example, Fig. 8A shows an example of a Vicra failure case (not used in this study). After applying the Vicra-based motion correction, the mismatch between the early (black and white, 0–10 min) and the late (hot metal, 60–90 min) frames is still evident (Fig. 8A), which suggested that Vicra failed to provide reliable motion information for this study. In contrast, the proposed 3DCOD method, as shown Fig. 8B, outperformed Vicra and

fixed the mismatch between the early and late frames. Note that we observed approximately 2% obvious Vicia failure instances among ~500 studies.

In our previous work (Lu et al., 2020), we used one dimension of the COD trace to detect and correct the motion. Across 10 subjects, the average number of MTPs using the old motion detection method is 29.3 ± 6.7 . After merging the three COD axes in the proposed motion detection algorithm, the average number of MTPs significantly increased to 211.2 ± 66.9 . Fig. 9. shows a representative ^{18}F -FDG HRRT reconstructed images (60–90 min), comparing our proposed 3DCOD approach with the results from the motion detection method with manually selected parameters (referred to as COD*). Detailed quantitative results are shown in Table 6. The mean absolute error is shown in Supplemental Table 17. For regions like amygdala, insula, hippocampus and thalamus, approximately 5% improvement was observed for 3DCOD approach compared with COD*. For caudate, the error was reduced from -12.2% to -2.2% for the representative case. One of the advantages of the new approach is the utilization of three COD directions which considers all possible motions, resulting into lower mean uptake error.

As mentioned earlier, there was a significant increase in the average number of MTPs when using 3DCOD which resulted in higher false positive rates as shown in Fig. 4. False positives potentially lead to short-ened motion-free frames which can affect registration accuracy. However, undetected motion time points (false negatives) are much more likely to lead to motion artifacts in the final reconstructed images than short motion-free frames (false positives) are. Therefore, our approach is more tolerant of false positive MTPs than false negative MTPs.

The motion detection performance of the proposed algorithm was very similar for HRRT non-TOF and mCT TOF scanners. Supplemental Fig. 6 shows a COD trace for a simulated study using the same motion information and tracer but performed on different scanners. Supplemental Fig. 6B displays a somewhat noisier COD for mCT but larger separation when motion occurred. HRRT (Supplemental Fig. 6A) has a smaller separation but the motion is still distinguishable and its COD is less noisy. Note that the noise on each COD trace obeys a normal distribution when motion is absent.

A few limitations of the proposed data-driven motion detection and motion correction approach include: (1) tracer distribution change is rapid immediately after injection. As a result, COD changes quickly in response. Currently, 3DCOD approach is not reliable for motion in the early scan times, i.e., first 2 min post-injection; (2) with the increase of noise, e.g., lower dose, the motion detection accuracy will decrease and (3) our approach does not handle continuous motion perfectly. Therefore, if extremely slow motion occurs, our method may perform suboptimally.

In Spangler-Bickell et al. (2021), Spangler-Bickell et al. recently performed a similar investigation focused on the proper frame duration for data-driven rigid motion estimation. They found that using 440 thousand true and scattered counts per frame will give a mesh error of 1 mm for ^{18}F -FDG whereas we found that MFFs with more than two million counts (true coincidence) were found to provide acceptable registration accuracy, i.e., MDE

< 2 mm for ^{18}F -FDG. They also investigated the FWHM for smoothing and found that 4–6 mm is optimal, which is similar to our findings; 5-mm FWHM was found optimal for ^{18}F -FDG while 7-mm was optimal for ^{11}C -RAC for both scanners. Sprangler-Bickell et al. used two tracers with broad distribution throughout the whole brain (^{18}F -FDG and ^{18}F -florbetaben) while we investigated two tracers (^{18}F -FDG and ^{11}C -RAC) with very different brain distributions. Lastly, Sprangler-Bickell et al. did not account for intra-frame motion, which can have an impact on the motion estimation process.

Different inflation factors (α) were used for ^{18}F -FDG (1.0) and ^{11}C -RAC (1.6) in this paper. However, to further examine the robustness and the sensitivity of motion correction efficacy to the value of α , we tested $\alpha=1.6$ in real ^{18}F -FDG studies (data not shown), and found that the final motion correction performance, i.e., ROI quantification error, was minimally affected. This is understandable since the detection performance between $\alpha=1.0$ and 1.6 was very small for motion larger than 1 mm (Fig. 5). Therefore, practically, for both ^{18}F -FDG and ^{11}C -RAC using HRRT or mCT, the same α choice (1.6) can be used. Other tunable parameters include P (number of MFFs used to predict the $\text{RSS}_{\text{NM},s}$) and the n_{max} , which were not thoroughly tested given the high detectability of the results with these values. Although the paper only examined ^{18}F -FDG and ^{11}C -RAC, we also applied our motion detection method *without* parameter turning, i.e., same set of parameter for all the tracers, to many other tracers. Supplemental Fig. 20 shows examples of motion detection results for six selected tracers with a same set of parameters. In the future, we will test the algorithms for a broader range of PET tracers to further test the robustness of the algorithm.

5. Conclusion

We developed a new COD-based motion detection algorithm that is statistics-based, automatic and uses information from all three COD directions; thus, the algorithm is adaptive to different tracers and acquisition conditions. Preliminary results also indicate that the proposed motion detection and correction method generalize well to other radiotracers acquired with a TOF scanner. The proposed motion correction method yielded $-0.3 \pm 2.8\%$ and $-0.4 \pm 3.2\%$ brain region error for ^{18}F -FDG and ^{11}C -RAC, respectively, across 10 subjects with large head motions for each tracer, using the hardware motion tracking as reference.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

The authors thank Dr. Ansel T. Hillmer for providing the ^{11}C -PBR28 scan data and Dr. David Matuskey for providing the ^{11}C -MRB data. We thank Mr. Yuan Zi (MathWorks®) for the discussion on the PELT algorithm implementation. We would also like to thank the anonymous reviewers for the invaluable comments and suggestions. This work was funded by NIH Grants R03EB027209, R21EB028954 and R01AG052560. Its contents are solely the responsibility of the authors and do not necessarily represent the official view of NIH.

Data availability

The authors confirm that the data supporting the findings of this study are available within the article and/or its supplementary materials.

References

- Bloomfield PM, Spinks TJ, Reed J, Schnorr L, Westrip AM, Livieratos L, et al. , 2003. The design and implementation of a motion correction scheme for neurological PET. *Phys. Med. Biol.* 48, 959–978. [PubMed: 12741495]
- Costes N, Dagher A, Larcher K, Evans AC, Collins DL, Reilhac A, 2009. Motion correction of multi-frame PET data in neuroreceptor mapping: simulation based validation. *Neuroimage* 47, 1496–1505. [PubMed: 19481154]
- Ding YS, Singhal T, Planeta-Wilson B, Gallezot J–D, Nabulsi N, Labaree D, et al. , 2010. PET imaging of the effects of age and cocaine on the norepinephrine transporter in the human brain using (S,S)-[(11)C]O-methylreboxetine and HRRT. *Synapse* 64, 30–38. doi:10.1002/syn.20696. [PubMed: 19728366]
- Fischl B, Salat DH, Busa E, Albert M, Dieterich M, Haselgrove C, et al. , 2002. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron* 33, 341–355. doi:10.1016/s0896-6273(02)00569-x. [PubMed: 11832223]
- Fischl B, van der Kouwe A, Destrieux C, Halgren E, Segonne F, Salat DH, et al. , 2004. Automatically parcellating the human cerebral cortex. *Cereb. Cortex* 14, 11–22. doi:10.1093/cercor/bhg087. [PubMed: 14654453]
- Fulton RR, Mickle SR, Eberl S, Constable CJ, 2002. Correction for head movements in positron emission tomography using an optical motion-tracking system. *IEEE Trans. Nucl. Sci.* 49, 116–123.
- Gallezot JD, Lu Y, Naganawa M, Carson RE, 2020. Parametric imaging with PET and SPECT. *IEEE Trans. Radiat. Plasma Med. Sci.* 4, 1–23.
- Germino M, Gallezot JD, Yan J, Carson RE, 2017. Direct reconstruction of parametric images for brain PET with event-by-event motion correction: evaluation in two tracers across count levels. *Phys. Med. Biol.* 62, 5344–5364. doi:10.1088/1361-6560/aa731f. [PubMed: 28504644]
- Herzog H, Tellmann L, Fulton R, Stangier I, Kops ER, Bente K, et al. , 2005. Motion artifact reduction on parametric PET images of neuroreceptor binding. *J. Nucl. Med.* 46, 1059–1065. [PubMed: 15937320]
- Hillmer AT, Matuskey D, Huang Y, Nabulsi N, Ropchan J, Carson RE, 2020. Tobacco smoking in people is not associated with altered 18 kDa-translocator protein levels: a Positron Emission Tomography study. *J. Nucl. Med.* 61, 1200–1204. [PubMed: 32005773]
- Jackson B, Scargle JD, Barnes D, Arabhi S, Alt A, Gioumouisis P, et al. , 2005. An algorithm for optimal partitioning of data on an interval. *IEEE Signal Process. Lett.* 12, 105–108.
- Jakoby BW, Bercier Y, Conti M, Casey ME, Bendriem B, Townsend DW, 2011. Physical and clinical performance of the mCT time-of-flight PET/CT scanner. *Phys. Med. Biol.* 56, 2375–2389. [PubMed: 21427485]
- Jenkinson M, Smith S, 2001. A global optimisation method for robust affine registration of brain images. *Med. Image Anal.* 5, 143–156. [PubMed: 11516708]
- Jin X, Mulnix T, Gallezot JD, Carson RE, 2013. Evaluation of motion correction methods in human brain PET imaging—a simulation study based on human motion data. *Med. Phys.* 40, 102503. [PubMed: 24089924]
- Joshi A, Scheinost D, Okuda H, Belhachemi D, Murphy I, Staib LH, et al. , 2011. Unified framework for development, deployment and robust testing of neuroimaging algorithms. *Neuroinformatics* 9, 69–84. [PubMed: 21249532]
- Keller SH, Sibomana M, Olesen OV, Svarer C, Holm S, Andersen FL, et al. , 2012. Methods for motion correction evaluation using 18F-FDG human brain scans on a high-resolution PET scanner. *J. Nucl. Med.* 53, 495–504. [PubMed: 22331217]

- Killick R, Fearnhead P, Eckley IA, 2012. Optimal detection of changepoints with a linear computational cost. *J. Am. Stat. Assoc.* 107, 1590–1598.
- Kyme AZ, Se S, Meikle SR, Fulton RR, 2018. Markerless motion estimation for motion-compensated clinical brain imaging. *Phys. Med. Biol.* 63, 105018. doi:10.1088/1361-6560/aabd48. [PubMed: 29637899]
- Lu W, Onofrey JA, Lu Y, Shi L, Ma T, Liu Y, et al. , 2019a. An investigation of quantitative accuracy for deep learning based denoising in oncological PET. *Phys. Med. Biol.* 64, 165019. doi:10.1088/1361-6560/ab3242. [PubMed: 31307019]
- Lu Y, Gallezot JD, Naganawa M, Ren S, Fontaine K, Wu J, et al. , 2019b. Data-driven voluntary body motion detection and non-rigid event-by-event correction for static and dynamic PET. *Phys. Med. Biol.* 64, 065002. [PubMed: 30695768]
- Lu Y, Naganawa M, Toyonaga T, Gallezot JD, Fontaine K, Ren S, et al. , 2020. Data-driven motion detection and event-by-event correction for brain PET: comparison with Vicra. *J. Nucl. Med.* 61, 1397–1403. [PubMed: 32005770]
- Montgomery AJ, Thielemans K, Mehta MA, Turkheimer F, Mustafovic S, Grasby PM, 2006. Correction of head movement on PET studies: comparison of methods. *J. Nucl. Med.* 47, 1936–1944. [PubMed: 17138736]
- Olesen OV, Sullivan JM, Mulnix T, Paulsen RR, Hojgaard L, Roed B, et al. , 2013. List-mode PET motion correction using markerless head tracking: proof-of-concept with scans of human subject. *IEEE Trans. Med. Imaging* 32 (2), 200–209. doi:10.1109/TMI.2012.2219693. [PubMed: 23008249]
- Owen DR, Guo Q, Kalk NJ, Colasanti A, Kalogiannopoulou D, Dimber R, et al. , 2014. Determination of [(11)C]PBR28 binding potential *in vivo*: a first human TSPO blocking study. *J. Cereb. Blood Flow Metab.* 34, 989–994. doi:10.1038/jcbfm.2014.46. [PubMed: 24643083]
- Picard Y, Thompson CJ, 1997. Motion correction of PET images using multiple acquisition frames. *IEEE Trans. Med. Imaging* 16, 137–144. [PubMed: 9101323]
- Rahmim A, Dinelle K, Cheng JH, Shilov MA, Segars WP, Lidstone SC, et al. , 2007. Accurate event-driven motion compensation in high-resolution PET incorporating scattered and random events. *IEEE Trans. Med. Imaging* 27, 1018–1033.
- Schleyer PJ, Dunn JT, Reeves S, Brownings S, Marsden PK, Thielemans K, 2015. Detecting and estimating head motion in brain PET acquisitions using raw time-of-flight PET data. *Phys. Med. Biol.* 60, 6441–6458. [PubMed: 26248198]
- Schmand M, Wienhard K, Casey ME, Eriksson L, Jones WF, Reed RH, et al., 1999. Performance evaluation of a new LSO high resolution research tomograph-HRRT. In: *Proceedings of the IEEE Nuclear Science Symposium and Medical Imaging Conference (Nss/Mic)*, 2, pp. 1067–1071.
- Spangler-Bickell MG, Hurley SA, Deller TW, Jansen F, Bettinardi V, Carlson M, et al. , 2021. Optimizing the frame duration for data-driven rigid motion estimation in brain PET imaging. *Med. Phys.* doi:10.1002/mp.14889.
- Thielemans K, Schleyer P, Dunn J, Marsden PK, Manjeshwar RM, 2013. Using PCA to detect head motion from PET list mode data. In: *Proceedings of the IEEE Nuclear Science Symposium and Medical Imaging Conference (Nss/Mic)*.

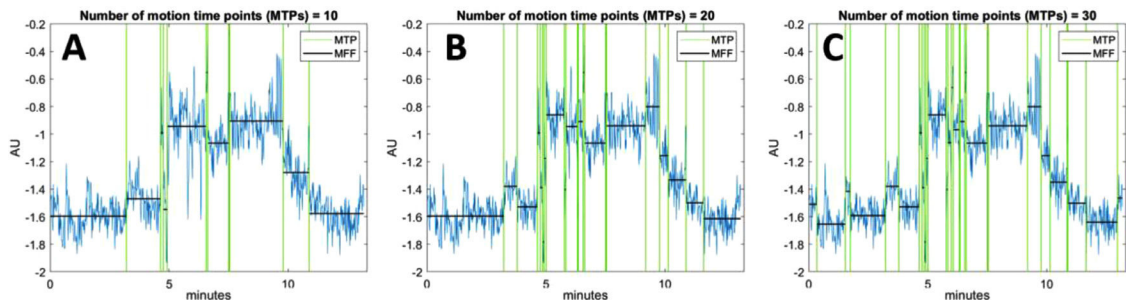


Fig. 1.

COD example (in the lateral direction) with motion detection under (A) 10 detected motion time points (MTPs, B) 20 MTPs and (C) 30 MTPs. Green vertical lines represent the MTPs and the black horizontal lines represent the mean of the COD segment within each motion-free frame (MFF), which is the period between two adjacent MTPs. AU: Arbitrary Unit.

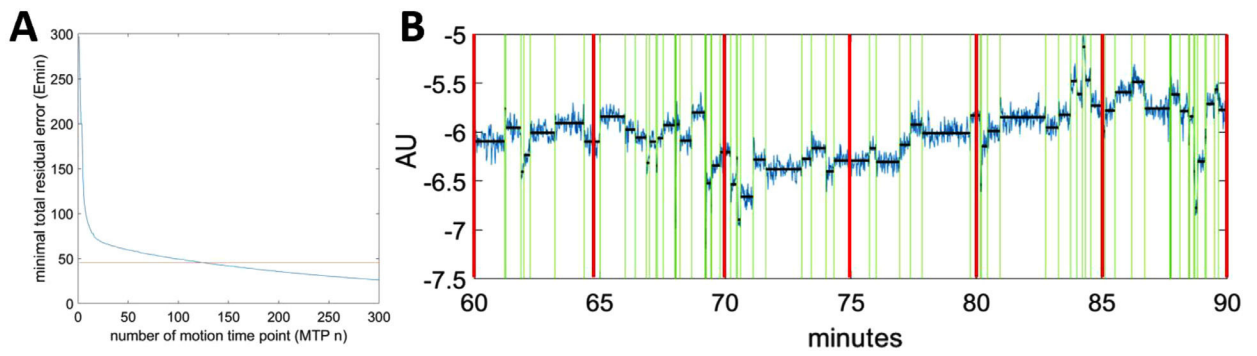


Fig. 2.

(A) E_{\min} vs number of MTPs (n). The intersection between the total residual error (E_{\min}) at different MTPs and E_{tar} corresponds to a set of n MTPs. (B) An example of detection result in one COD direction (lateral) for a simulated ^{18}F -FDG scan. Red vertical lines show the equal-duration partitions. The unit for this measure is arbitrary, which we denote as Arbitrary Unit (AU).

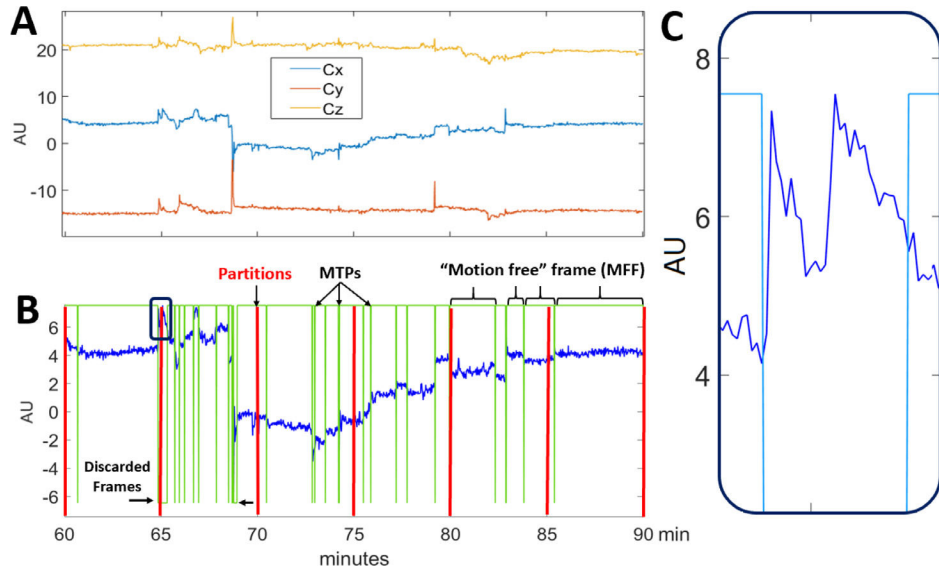


Fig. 3. COD motion detection results for a real ¹⁸F-FDG study. (A) COD in lateral (C_x), anterior-posterior (C_y) and superior-inferior (C_z) directions. (B) Example of motion detection using C_x . Green vertical lines indicate motion time points (MTPs). Top horizontal line segments indicate a preserved motion-free frame (MFF) and short bottom line segments show discarded frames, where 43 s (2.4%) of the entire 30 min were discarded. Red vertical lines show the equal-duration partitions. (C) Example of a discarded frame (magnified version of black box region in (B) at 65 min) with over-frequent motion.

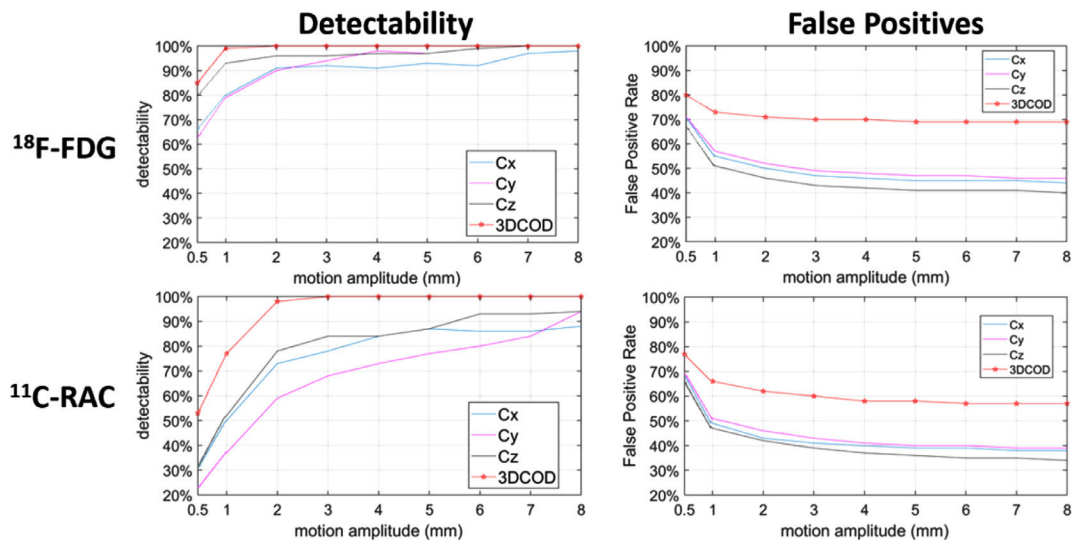


Fig. 4. Comparison of detectability and false positive rate results between 1DCOD and 3DCOD approaches for simulated ^{18}F -FDG (top) and ^{11}C -RAC (bottom) HRRT studies (without TOF). Red line with stars indicates the 3DCOD method. The inflation factor α was set to 1.0 and 1.6 for ^{18}F -FDG and ^{11}C -RAC studies, respectively. Each data point in the detectability figures is the ratio between the total number of correctly detected MTPs and the total number of ground-truth MTPs in all the simulated studies at a given motion amplitude.

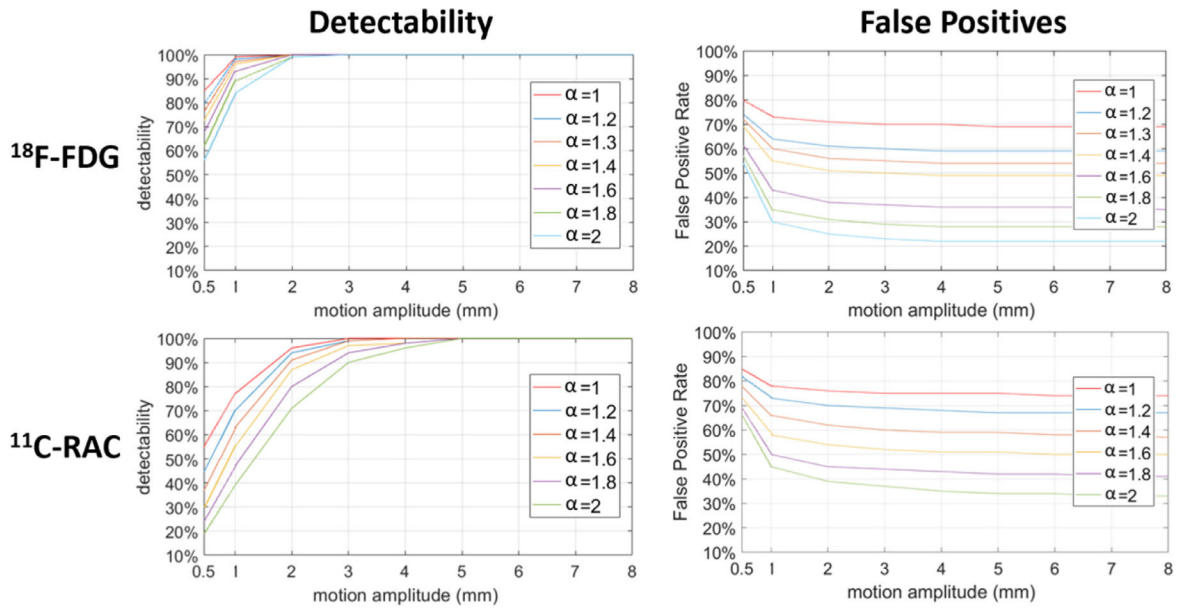


Fig. 5. Detectability and false positive rate results for 3DCOD as a function of α for $^{18}\text{F-FDG}$ (top) and $^{11}\text{C-RAC}$ (bottom) simulated for the HRRT scanner.

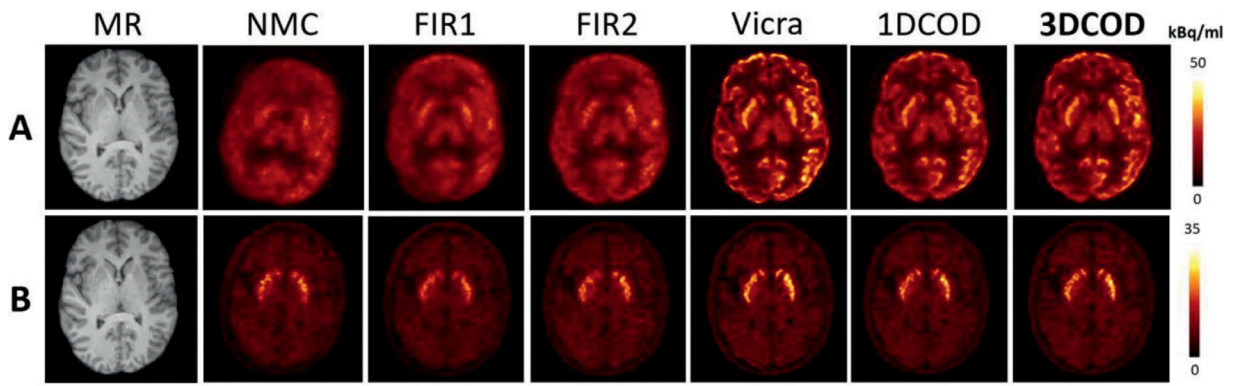


Fig. 6. Sample slices of motion-corrected reconstructions of simulated HRRT (A) ^{18}F -FDG study (60–90 min) and (B) ^{11}C -RAC study (30–60 min). Studies are ranked 15/15 (worst case) based on mean difference for the 3DCOD-based approach. The best cases (1/15) can be found in Supplemental Figure 9.

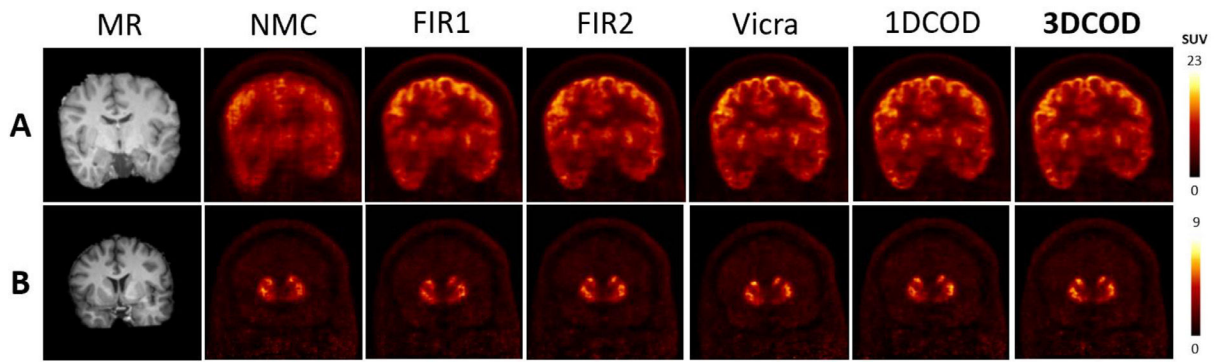


Fig. 7. Sample slices of motion-corrected reconstructions of real (A) ^{18}F -FDG (60–90 min) and (B) ^{11}C -RAC (30–60 min) studies. Studies are ranked 10/10 (worst case) based on mean difference for the 3DCOD-based approach.

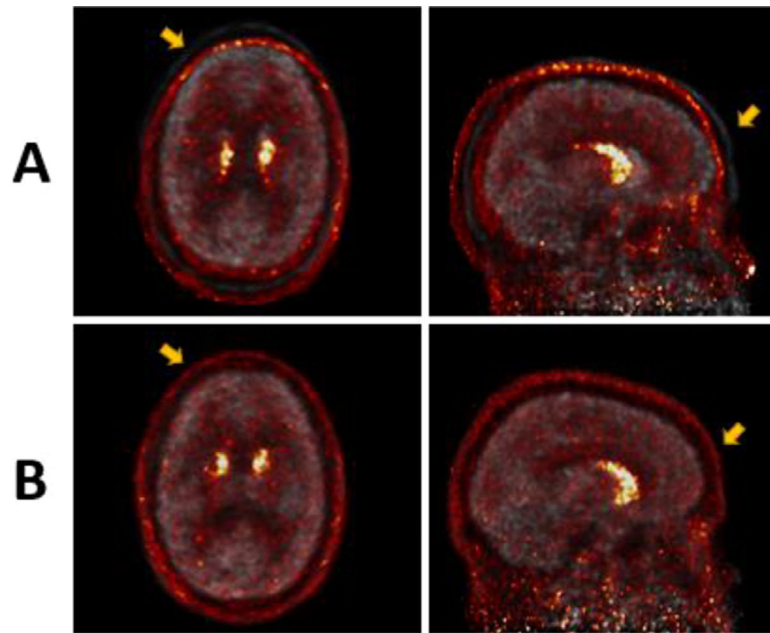


Fig. 8. Sample slices of motion-corrected reconstructions of a HRRT ^{11}C -RAC study (black and white colormap: 0–10 min, hot metal: 60–90 min) using (A) Vicra and (B) 3DCOD methods. Vicra shows a clear misalignment between the early and late frames while 3DCOD shows a good alignment.

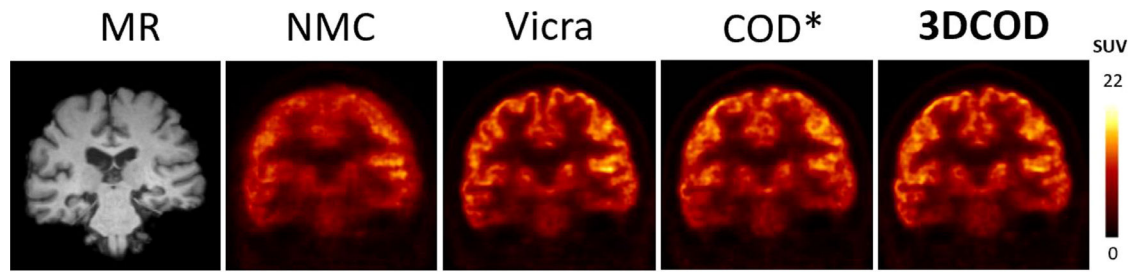


Fig. 9. Sample slices of motion-corrected reconstructions of a real ^{18}F -FDG (60–90 min) HRRT study showing the proposed 3DCOD approach compared with NMC, Vicra and COD* (example result from our previous work).

Table 1

Definition of symbols used in the motion detection algorithm.

Symbols	Meaning
$C; C_m; C_{m(k)}; \bar{C}_m$	COD trace in an arbitrary direction; COD segment within the m^{th} motion free frame (MFF); COD segment of the k^{th} sampling interval within the m^{th} MFF; averaged $C_{m(k)}$ over all k .
RSS m ; $\text{RSS}_{\text{NM},s}$	Fitting residual sum of squares (RSS) of the m^{th} MFF; RSS of the partition s under hypothetical no-motion (NM) scenario.
N_m	Total number of sampling intervals within the m^{th} MFF.
$N_{\text{LONG},p}$	Total numbers of sampling intervals of the p^{th} MFF among the longest MFFs within one partition.
P	Number of MFFs used to predict the $\text{RSS}_{\text{NM},s}$.
$N_{\text{PART},s}$	Total number of sampling intervals within the s^{th} equal-duration partition.
$E_{\text{min}}(n)$	Lowest total residual error for n motion time points (MTPs) in a COD trace.
$E_{\text{max}}; n_{\text{tar}}$	Predefined maximum number of MTPs; the number of MTPs that corresponds to a target detection.
$E_{\text{tar}}; E_{\text{NM}}$	Target error level; a comparable hypothetical scan where no motion occurred E_{NM}
α	Inflation factor
N_{ROI}	Number of brain ROI
N_{ALL}	Number of time sampling points of an entire scan
l_i	Brain location coordinate at time t_i defined as the average coordinate of 8 COMs of (left and right) frontal, occipital, hippocampus and cerebellum.
l_i	Euclidian distance between l_i and l_{i-1} .
MDE	Mean distance error
i	Subscript i , line-of-response (LOR)

Mean \pm Standard deviation (SD) uptake error (%) compared to Vicia for simulated 18 F-FDG (60–90 min) for the HRRRT scanner across 15 subjects.

Table 2

ROI	NMC	FIR1	FIR2	IDCOD	3DCOD
Amygdala	-13.9 \pm 23.8	-3.2 \pm 3.3	-0.9 \pm 1.7	0.6 \pm 1.9	-0.4 \pm 1.7
Caudate	-31.6 \pm 13.6	-9.5 \pm 5.1	-7.7 \pm 5.6	-2.8 \pm 2.0	-2.8 \pm 1.9
Cerebellum	-18.8 \pm 12.1	-4.3 \pm 4.2	-1.7 \pm 2.1	-0.3 \pm 0.9	-0.2 \pm 0.7
Frontal	-42.9 \pm 9.5	-13.6 \pm 5.4	-12.3 \pm 5.6	-6.1 \pm 2.4	-5.4 \pm 2.1
Hippocampus	-13.5 \pm 6.9	-4.6 \pm 3.4	-3.3 \pm 2.7	-1.4 \pm 0.8	-1.0 \pm 0.8
Insula	-14.2 \pm 7.3	-4.8 \pm 2.8	-4.3 \pm 2.5	-1.6 \pm 0.8	-1.3 \pm 1.1
Occipital	-24.0 \pm 7.1	-10.9 \pm 6.7	-6.9 \pm 4.0	-4.0 \pm 2.1	-4.0 \pm 1.9
Parietal	-34.5 \pm 10.7	-13.1 \pm 5.2	-9.5 \pm 4.6	-5.5 \pm 2.2	-5.4 \pm 2.0
Putamen	-34.0 \pm 12.0	-7.4 \pm 4.5	-6.0 \pm 4.3	-2.4 \pm 1.4	-1.9 \pm 1.3
Temporal	-30.2 \pm 9.6	-10.0 \pm 4.9	-7.3 \pm 3.8	-3.2 \pm 1.4	-3.1 \pm 1.3
Thalamus	-22.2 \pm 9.6	-5.0 \pm 2.9	-3.2 \pm 2.6	-1.4 \pm 0.8	-1.3 \pm 0.9
Ave difference (%)	-25.4	-13.4	-5.7	-2.6	-2.3
Ave SD (%)	11.1	3.5	3.6	1.5	1.4

Mean \pm Standard deviation (SD) uptake error (%) compared to Vicia for simulated ^{11}C -RAC (30–60 min) for HRRT scanner across 15 subjects.

Table 3

ROI	NMC	FIR1	FIR2	IDCOD	3DCOD
Caudate	-35.2 \pm 15.1	-17.6 \pm 5.2	-9.1 \pm 5.4	-5.8 \pm 3.3	-3.9 \pm 2.4
Putamen	-33.8 \pm 17.0	-14.6 \pm 4.0	-6.9 \pm 3.6	-4.4 \pm 2.0	-2.9 \pm 1.1
Ave difference (%)	-34.5	-16.1	-8.0	-5.1	-3.4
Ave SD (%)	16.1	4.6	4.5	2.7	1.7

Mean \pm Standard deviation (SD) uptake error (%) compared to Vicia for real 18 F-FDG (60–90 min) study on HRRT scanner across 10 subjects.

Table 4

ROI	NMC	FIR1	FIR2	IDCOD	3DCOD
Amygdala	-6.9 \pm 11.6	-4.0 \pm 2.9	-0.3 \pm 2.4	1.4 \pm 6.8	0.7 \pm 3.3
Caudate	-23.5 \pm 10.6	-4.6 \pm 7.9	-4.1 \pm 5.7	-7.1 \pm 6.1	-2.9 \pm 4.0
Cerebellum	-10.8 \pm 11.7	-4.2 \pm 3.4	0.9 \pm 2.5	0.9 \pm 2.6	0.4 \pm 2.2
Frontal	-22.9 \pm 9.1	-3.1 \pm 6.5	-2.9 \pm 4.0	-4.0 \pm 4.1	-1.6 \pm 3.2
Hippocampus	-9.0 \pm 4.6	-4.0 \pm 2.8	-2.5 \pm 1.6	-3.9 \pm 3.0	-1.9 \pm 3.1
Insula	-7.7 \pm 5.5	-2.8 \pm 3.0	-0.4 \pm 1.6	1.3 \pm 1.7	1.1 \pm 1.5
Occipital	-11.5 \pm 9.3	-1.7 \pm 8.1	2.2 \pm 8.0	2.1 \pm 6.3	1.8 \pm 5.4
Parietal	-16.5 \pm 11.3	-3.2 \pm 6.5	-2.4 \pm 3.7	-3.7 \pm 3.6	-1.5 \pm 2.2
Putamen	-20.9 \pm 8.7	-3.4 \pm 4.8	-0.8 \pm 4.2	-1.6 \pm 4.0	0.2 \pm 1.6
Temporal	-18.4 \pm 7.7	-4.2 \pm 4.3	-0.4 \pm 3.2	0.6 \pm 3.7	1.0 \pm 2.9
Thalamus	-15.4 \pm 8.7	-4.2 \pm 5.0	-1.0 \pm 2.3	-2.7 \pm 4.0	-0.5 \pm 1.8
Ave difference (%)	-14.9	-3.6	-0.6	-1.5	-0.3
Ave SD (%)	9.0	4.9	3.4	4.2	2.8

Mean \pm Standard deviation (SD) uptake error (%) compared to Vicia for real ¹¹C-RAC (30–60 min) study on HRRT scanner across 10 subjects.

Table 5

ROI	NMC	FIR1	FIR2	IDCOD	3DCOD
Caudate	-28.1 \pm 14.2	-15.4 \pm 13.7	-8.5 \pm 5.5	-3.9 \pm 4.8	-1.8 \pm 3.4
Putamen	-20.8 \pm 14.9	-11.4 \pm 14.8	-4.8 \pm 5.0	-0.4 \pm 3.4	1.0 \pm 2.9
Ave difference (%)	-24.5	-13.4	-6.7	-2.2	-0.4
Ave SD (%)	14.6	14.3	5.3	4.1	3.2

Table 6

Mean uptake error (%) compared to Vicra for the real 18 F-FDG (60–90 min) study for 3DCOD and COD*.

ROI	NMC	COD*	3DCOD
Amygdala	-13.4	8.1	3.9
Caudate	-21.2	-12.2	-2.2
Cerebellum	-14.2	3.3	-0.2
Frontal	-16.3	-0.7	-1.1
Hippocampus	-5.3	6.7	-2.6
Insula	-5.8	-5.9	0.0
Occipital	-19.7	-3.2	1.6
Parietal	-12.4	-0.3	-3.4
Putamen	-19.0	-0.4	-2.8
Temporal	-18.4	-2.4	1.0
Thalamus	-11.9	-5.9	-2.6
Ave difference (%)	-14.3	-1.7	-0.8
Ave SD (%)	5.3	6.0	2.3

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript