Ultra-Fast List-Mode Reconstruction of Short PET Frames and Example Applications

Matthew G. Spangler-Bickell1, Timothy W. Deller2, Valentino Bettinardi3, and Floris Jansen2

1Department of Radiology, University of Wisconsin, Madison, Wisconsin; 2PET/MR Engineering, GE Healthcare, Waukesha, Wisconsin; and 3Nuclear Medicine Unit, IRCCS Ospedale San Raffaele, Milan, Italy

Standard clinical reconstructions usually require several minutes to complete, and this time is mostly independent of the duration of the data being reconstructed. Applications such as data-driven motion estimation, which require many short frames over the duration of the scan, become unfeasible with such long reconstruction times. In this work, we present an infrastructure whereby ultra-fast list-mode reconstructions of very short frames (≤1 s) are performed. With this infrastructure, it is possible to have a dynamic series of frames that can be used for various applications, such as data-driven motion estimation, whole-body surveys, quick reconstructions of gated data to select the optimal gate for a given attenuation map, and, if the infrastructure runs simultaneously with the scan, real-time display of the reconstructed data during the scan and automated alerts for patient motion.

Methods: A fast ray-tracing time-of-flight projector was implemented and parallelized. The reconstruction parameters were optimized to allow for fast performance: only a few iterations are performed, without point-spread-function modeling, and scatter correction is not used. The resulting reconstructions are thus not quantitative but are acceptable for motion estimation, and gating purposes and for motion estimation using image registration. Data-driven motion can be estimated using image registration, with the resultant motion data being used in a fully motion-corrected list-mode reconstruction.

Results: The infrastructure provided images that can be used for visualization and gating purposes and for motion estimation using image registration. Several case studies are presented, including data-driven motion estimation and correction for brain studies, abdominal studies in which respiratory and cardiac motion is visible, and a whole-body survey.

Conclusion: The presented infrastructure provides the capability to quickly create a series of very short frames for PET data that can be used in a variety of applications.

Key Words: fast reconstruction; list-mode reconstruction; real-time reconstruction; data-driven motion correction

DOI: 10.2967/jnumed.120.245597

Clinical image reconstruction for a single frame of PET data typically requires computation time on the order of minutes. Reconstructions are usually conducted using sinogram data; therefore, the reconstruction time is roughly constant, no matter the acquisition duration or number of coincidence counts in the dataset. Even if reconstruction corrections are disabled and a large pixel size is used, reconstruction computation remains on the order of tens of seconds. This constraint limits potential PET imaging applications. For example, visualization of nearly real-time images during an ongoing acquisition is made difficult by lengthy reconstruction processing times. Additionally, tracking changes in the activity distribution due to biologic processes or patient motion are also impractical. Data-driven motion tracking with high temporal resolution, such as 0.2-s frames to detect cardiac motion, could require a reconstruction processing time approximately 100 times longer than the actual acquisition, even without imaging corrections. This duration would be computationally prohibitive for a standard clinical protocol.

As a result, data-driven motion detection techniques often avoid the reconstruction step, instead processing on short-duration, coarsely sampled sinogram data. One prominent approach is to apply principal-component analysis to these coarse sinograms to detect periodic respiratory motion (1). Unfortunately, this approach discards meaningful information in the original dataset, such as the time-of-flight (TOF) dimension. The principal-component analysis approach outputs a waveform of motion, but it is not in spatial dimensions. Therefore, it is ineffective for measuring the spatial amplitude of motion, which would allow amplitude-based gating approaches. The time points of nonperiodic motion can be detected in short-duration sinogram data, but rigid or nonrigid motion vectors cannot be computed in a reliable way directly from sinogram data.

The multiple-acquisition frame technique is an approach that has previously been described to correct for occasional head motion (2). As originally described, cameras are used to identify time points of motion, and a new frame of sinogram data is started with each motion instance. The multiple frames of data are reconstructed independently, registered to a reference frame, and added together. This approach successfully reduces the motion artifacts, but summing independently reconstructed frames is not a statistically optimal method when using modern iterative reconstruction techniques, and the approach cannot reliably correct slow and extended motion. Preliminary results (3) demonstrate a method to estimate the motion parameters directly from the list-mode data—parameters that can then be used in a motion-corrected reconstruction. Similarly, centroid-of-distribution calculations estimate when motion occurred, followed by framewise reconstruction and registration to determine the motion parameters (4).

In this paper, we present an infrastructure to perform very-short-frame (i.e., ≤1 s) reconstructions faster than the duration of the frame, using list-mode reconstruction. Thus, a dynamic series of reconstructed images from very short frames of arbitrary duration is created, and the total time needed to produce the entire series is less
than the duration of the dataset and almost constant regardless of the chosen frame duration (although proportional to the total number of coincidence events in the dataset, as is typical for list-mode reconstruction).

In the future, this infrastructure could be implemented to run simultaneously with the PET acquisition, performing nearly real-time reconstruction, processing, and visualization of the data. The proposed infrastructure enables many interesting applications. It will be possible to quickly (or in real time) perform data-driven motion estimation using registration of the frames, for both rigid and nonrigid applications. After this estimation, a fully motion-corrected list-mode reconstruction can be performed. Data-driven gating information can be determined for respiratory and cardiac signals, using either phase-based or amplitude-based approaches. Reconstructed frames can be displayed in nearly real time while the patient is in the scanner for the technologist to see, allowing the technologist to perform such tasks as verifying patient positioning, ensuring injection success, and observing any patient motion and to intervene appropriately. Motion can be automatically monitored and an alert triggered for the technologist if motion exceeding a threshold is detected. Quick whole-body surveys can be performed and used for such purposes as to obtain an overview of areas of interest and adjust the bed durations accordingly to optimize use of the scan time. For gated cardiac imaging, quick reconstructions of gates can be used to select the gate that most closely matches the attenuation map, so that a gated reconstruction can be performed as soon as the acquisition is finished.

The current implementation of the infrastructure operates offline on top of GE Healthcare’s standard PET reconstruction research toolbox.

MATeRIALS AND METHODS

Frame Reconstruction

A list-mode reconstruction package was developed as a module that runs on top of the standard research toolbox distributed by GE Healthcare. The list-mode data file is read in short frames, which are then reconstructed. The reconstruction has been optimized to produce an image quickly—more quickly than the duration of the frame. A future aim is to run this infrastructure online during the data acquisition such that the reconstruction of each short frame would begin as soon as it has been acquired, and the reconstruction would be completed and processed before acquisition of the subsequent short frame has finished. This would constitute a nearly real-time (1 frame late) reconstruction and processing of the data.

A fast TOF-based ray-tracing projector similar to the Siddon projector (5) was implemented. For list-mode reconstruction, a TOF projection is faster than non-TOF since only those pixels inside the TOF kernel are visited. In fact, the speed of the projection is proportional to the TOF resolution. The TOF kernels were clipped at ±3 SDs. During the projection, any event that falls outside the image field of view (FOV) is ignored, and the projection weights for an event are calculated once per iteration and used for both the forward projection and the backprojection. List-mode reconstruction is naturally parallelizable since the projection of each event can be handled independently, with the final backprojected images from each thread being summed.

Thus, the projector was multithreaded. To reconstruct each frame, pure maximum-likelihood expectation maximization (MLEM) is used, without subsets, because of the low number of events in each frame. Only 2–5 iterations are performed, without point-spread-function modeling. The frame duration can be set arbitrarily and is usually governed by the activity level and the required task. Durations of 1 s or less are commonly used for most tasks.

For the applications presented here, accurate quantitation is not imperative, and therefore no scatter correction is performed. This avoids the computationally expensive step of estimating the scatter contribution. Randoms are estimated from the singles rate on a per-event basis, using the available per-second crystal singles histograms. If desired, attenuation correction is applied using the standard attenuation map generated from either the CT or the MR images. During the initialization of the reconstructions, a sensitivity image is generated and used for all frames.

The following list-mode reconstruction algorithm (6,7) is used:

\[
\lambda_{n+1}^i = \lambda_n^i \sum_k c_{ij} a_{ik} \frac{1}{m} \sum_m c_{im} \lambda_n^m + R_{ni} \sigma_n
\]

where \(\lambda_n^i\) is the image value at pixel \(j\) and iteration \(n\), \(c_{ij}\) is the system matrix, \(a_{ik}\) is the line-of-response factor associated with list-mode event \(m\), \(l\) is the total number of possible lines of response in the scanner, \(M\) is the total number of list-mode events, \(a_{ik}\) is the attenuation correction factor, \(R_{ni}\) is the randoms contribution, and \(\sigma_n\) is the scanner sensitivity factor for line-of-response \(i\).

Image initialization for each frame can be performed in 1 of 2 ways: either from a uniform image or using the result of the previous frame after smoothing. Using the previous frame can result in visually improved images that are nearer to convergence and might therefore be appropriate for visualization purposes. However, for motion estimation in the presence of significant motion, this method may introduce bias from one frame to the next. Therefore, the appropriate method should be selected on the basis of the task at hand.

Reconstruction Time

A standard Intel i9 central-processing-unit laptop computer with 8 cores was used for the benchmarking reported in this section. Using a FOV of 300 mm with pixel dimensions of 2.34 × 2.34 × 2.78 mm, image size of 128 × 128 × 89, and 2 iterations, the typical reconstruction time per frame is given in Table 1 for several standard clinical scenarios. The reconstruction time of each frame is consistently less than the frame duration, as is necessary to achieve real-time processing. The reconstruction time of each frame scales approximately linearly with each dimension of the image size (regardless of FOV or pixel size) and linearly with the number of iterations. Before each run, certain initialization steps were performed, taking approximately 20 s. Most of this initialization time is spent on calculating the sensitivity image, which can be precalculated as soon as the attenuation map is ready.

Additionally, a dynamic \(^{15}O-H_2O\) cardiac study was processed to observe the frame reconstruction time as the number of events in each frame varied throughout the scan. A constant frame duration of 0.2 s was used, and the events per frame and the reconstruction times are shown in Figure 1. The frame reconstruction time scales approximately linearly with the number of events in the frame. The deviation from linearity during the early frames just after the injection is due to a large number of random events in those frames, many of which do not pass through the image FOV, leading to a reconstruction time that is faster than normal. As can be seen in Figure 1, for frames with more than 900 × 10^4 events, the reconstruction time could surpass the frame duration. This issue is easily solved by limiting the number of events used for the reconstruction of those frames. This limitation was not necessary in any of the cases presented in this paper.

This linear relationship implies that for a given dataset, the total time taken to reconstruct all frames is approximately constant regardless of the chosen frame duration, except for a small but nonzero overhead contribution.
### Data-Driven Motion Correction

The dynamic series of short frames can be used to estimate the motion by performing image registration between each frame and a chosen reference frame. For brain scan data, rigid registration is performed and can be started as soon as a frame reconstruction is complete. A mean-square difference metric is used with a gradient-descent optimizer. Attenuation correction is not applied to avoid possible bias introduced by a mismatch between the PET data and attenuation map in the presence of motion. Initial tests indicate that both the frame reconstruction and registration can be performed more quickly than real time. Therefore, it is possible to fully estimate the motion concurrently with the scan, such that as soon as the acquisition is completed, a fully motion-corrected list-mode reconstruction can be performed (8).

It was previously reported (9) that a rigid registration with an accuracy (defined as the average error after registration in the brain position, which is calculated using multiple points scattered throughout the brain) of approximately 1 mm can be achieved for $^{18}$F-FDG brain scan data on frames containing approximately $40 \times 10^3$ true coincidence counts. For standard scans, this frame count level corresponds to a frame duration of about 0.26 s. Although such a short frame duration will translate to a high temporal resolution of the motion estimates, it is perhaps excessive in most cases and a frame duration of 1 s should provide adequate motion sampling while allowing for more robust image registration.

### RESULTS

#### Real-Time Reconstruction

**Brain Study.** Informed consent was received for all studies in accordance with the institutions’ review boards.

A dementia patient underwent an $^{18}$F-FDG brain study in a Discovery-MI PET/CT scanner (GE Healthcare) with a 20-cm axial FOV at UZ Leuven, Belgium. During the scan, the patient’s hand moved into the FOV and scratched the patient’s forehead. This movement was not evident until the data were reconstructed into short frames using the methodology presented above. The data were reconstructed into 1-s frames, and each frame contained approximately $400 \times 10^3$ list-mode events. Two MLEM iterations were performed using pixel dimensions of $3.26 \times 3.26 \times 2.78$ mm and a FOV of 300 mm; a uniform image was used for the initialization, and postsmoothing of 8 mm in full width at half maximum (FWHM) was applied. Each frame was reconstructed in 0.2 s. A time range in which the patient’s hand can be seen entering the FOV is shown in Figure 2A. Figure 2B is an $^{18}$F-FDG study showing cardiac motion at the bed position over the heart. Figure 2C is another $^{18}$F-FDG study; here, transport of the urine from the kidneys into the bladder can be seen. Videos of these frames can be seen as Supplemental Videos 2–4. Such images could be displayed in nearly real time for the technologist to see but could also be used for non-rigid motion correction, for data-driven gating, or for selection of the appropriate gate to match the attenuation map from the CT.

#### Abdominal Studies

In Figure 3 are shown multiple examples of reconstructed frames over the abdomen from standard clinical scans on the SIGNA PET/MR scanner (GE Healthcare) at Ospedale San Raffaele, Milan, Italy. All studies used pixel dimensions of $2.34 \times 2.34 \times 2.78$ mm, a FOV of 300 mm, 2 iterations, and isotropic postsmoothing of 14 mm in FWHM. The frames were $0.1–0.3$ s, contained $50–100 \times 10^3$ events, and were reconstructed more quickly than real time. Figure 3A is a $^{68}$Ga-DOTATOC study showing respiratory motion of the abdominal organs. Figure 3B is an $^{18}$F-FDG study showing cardiac motion at the bed position over the heart. Figure 3C is another $^{18}$F-FDG study; here, transport of the urine from the kidneys into the bladder can be seen. Videos of these frames can be seen as Supplemental Videos 2–4. Such images could be displayed in nearly real time for the technologist to see but could also be used for non-rigid motion correction, for data-driven gating, or for selection of the appropriate gate to match the attenuation map from the CT.

### Data-Driven Motion Correction

**Case 1.** A clinical $^{18}$F-FDG brain study was conducted on a SIGNA PET/MR scanner on frames containing approximately $40 \times 10^3$ true coincidence counts. For standard scans, this frame count level corresponds to a frame duration of about 0.26 s. Although such a short frame duration will translate to a high temporal resolution of the motion estimates, it is perhaps excessive in most cases and a frame duration of 1 s should provide adequate motion sampling while allowing for more robust image registration.

### Table 1

Frame-Based Execution Times for Various Clinical Scenarios Using a Standard 8-Core Intel i9 Central-Processing-Unit System

<table>
<thead>
<tr>
<th>Scan type</th>
<th>Scanner</th>
<th>Tracer</th>
<th>Injected activity (MBq)</th>
<th>Time after injection (min)</th>
<th>Events per frame</th>
<th>Frame duration (s)</th>
<th>Frame reconstruction time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brain</td>
<td>D-MI</td>
<td>$^{18}$F-FDG</td>
<td>231</td>
<td>80</td>
<td>$400 \times 10^3$</td>
<td>1.0</td>
<td>0.18</td>
</tr>
<tr>
<td>Cardiac</td>
<td>SIGNA</td>
<td>$^{18}$F-FDG</td>
<td>259</td>
<td>143</td>
<td>$55 \times 10^3$</td>
<td>0.1</td>
<td>0.07</td>
</tr>
<tr>
<td>Abdomen</td>
<td>SIGNA</td>
<td>$^{68}$Ga-DOTATOC</td>
<td>137</td>
<td>121</td>
<td>$60 \times 10^3$</td>
<td>0.3</td>
<td>0.07</td>
</tr>
</tbody>
</table>

Initialization of each run took ~20 s.

---

**FIGURE 1.** (Left) Plot of events per 0.2-s frame recorded during dynamic $^{18}$O-H$_2$O cardiac study. (Right) Plot illustrating linear relationship between the reconstruction time per frame and number of events in the frame. Dashed line indicates real duration of frames. Red points are for those frames before the peak in events per frame. These frames occur immediately after injection and contain a high fraction of random events that do not pass through the image FOV, leading to reconstruction times that are faster than normal. For frames with high number of events ($>9 \times 10^5$), reconstruction times were longer than frame durations. Recon = reconstruction. (Data courtesy of Uppsala University Hospital.)

---

**FIGURE 2.**

**A.** Plot of events per 0.2-s frame recorded during dynamic $^{18}$O-H$_2$O cardiac study. **B.** Plot illustrating linear relationship between the reconstruction time per frame and number of events in the frame. Dashed line indicates real duration of frames. Red points are for those frames before the peak in events per frame. These frames occur immediately after injection and contain a high fraction of random events that do not pass through the image FOV, leading to reconstruction times that are faster than normal. For frames with high number of events ($>9 \times 10^5$), reconstruction times were longer than frame durations. Recon = reconstruction. (Data courtesy of Uppsala University Hospital.)

---

**FIGURE 3.** (A) Plot of events per 0.2-s frame recorded during dynamic $^{18}$O-H$_2$O cardiac study. (B) Plot illustrating linear relationship between the reconstruction time per frame and number of events in the frame. Dashed line indicates real duration of frames. Red points are for those frames before the peak in events per frame. These frames occur immediately after injection and contain a high fraction of random events that do not pass through the image FOV, leading to reconstruction times that are faster than normal. For frames with high number of events ($>9 \times 10^5$), reconstruction times were longer than frame durations. Recon = reconstruction. (Data courtesy of Uppsala University Hospital.)
scanner at Ospedale San Raffaele. During the scan, the patient spoke to the technologist, causing significant motion of the head, and the head did not return to its initial position. The movement of the patient’s jaw was visible in the short reconstructed frames, as can be seen in Figure 4 and in Supplemental Video 5. Reconstructions were performed of 1-s frames containing approximately $400 \times 10^3$ events using pixel dimensions of $2.34 \times 2.34 \times 2.78$ mm, a FOV of $300$ mm, 2 iterations, and isotropic postsmoothing of $8$ mm in FWHM, and the motion of the head was estimated using image registration. The plots of the 6 degrees of freedom are shown in Figure 5. Had the technologist been alerted that the patient’s head did not return to its initial position, the technologist could possibly have taken appropriate measures (such as restarting the scan).

Nonetheless, with these motion estimates, a fully motion-corrected list-mode reconstruction was performed, the result of which is shown in Figure 5, using pixel dimensions of $1.17 \times 1.17 \times 2.78$ mm, an FOV of 300 mm, 3 iterations with 47 subsets, point-spread-function modeling, and postsmoothing of 4 mm in FWHM.

Case 2. A 20-min clinical $^{11}$C-methionine study was conducted on the SIGNA PET/MR scanner at Ospedale San Raffaele. In the last 5 min of the scan, the patient’s head moved to one side and then to the other side. This motion was estimated using the described methodology, with 2-s frames containing approximately $600 \times 10^3$ events, and the estimated motion was used to perform a fully motion-corrected reconstruction with pixel dimensions of $1.17 \times 1.17 \times 2.78$ mm, an FOV of $300$ mm, 5 iterations with 28 subsets, point-spread-function modeling, and postsmoothing of $4$ mm in FWHM. The result of this reconstruction is shown in Figure 6.

Whole-Body Survey

A rapid whole-body survey could be performed by moving the patient through the scanner in less than 5 s (the limiting factor being the maximum bed speed and patient comfort) and quickly reconstructing the data. Such data could be used to identify areas of interest and thereby guide how the subsequent scan time is divided among the bed positions.

A standard whole-body dataset from a patient undergoing an $^{18}$F-FDG scan on the SIGNA PET/MR scanner at Ospedale San Raffaele was used to visualize what such a survey might look like. A single frame of 0.3 s from each of 5 bed positions was reconstructed, containing $50-100 \times 10^3$ events, and the frames were stitched together. Each frame was reconstructed using 2 MLEM iterations, with pixel dimensions of $2.34 \times 2.34 \times 2.78$ mm and with a postsmoothing filter of $14$ mm in FWHM. The maximum-intensity projection of the resulting reconstruction is shown in Figure 7. This image comprises only $1.5$ s of data, demonstrating how quickly the whole-body survey could be acquired. A video of multiple consecutive frames from this study can be seen as Supplemental Video 6.

DISCUSSION

The presented methodology allows for ultra-fast reconstruction, processing, or visualization of acquired PET data, producing a dynamic series of images with very short frame-durations. The fast reconstructions are achieved using an optimized TOF ray-tracing projector within a multithreaded central-processing-unit setting and efficient reconstruction parameters such as only 2–5 MLEM iterations and no scatter correction. If this infrastructure
were to be used on the scanner during acquisition, then the applications presented in this report could be performed in nearly real time. As demonstrated, such an infrastructure enables, for example, fast and automatic data-driven motion estimation using image registration, real-time visualization of the tracer distribution for the technologist to see, quick whole-body surveys, and quick reconstructions of gated data to select the optimal gate given the attenuation map. We expect that once this infrastructure is available to users and researchers, other applications will be developed that take advantage of the dynamic series of short-frame reconstructed images.

Scanner capabilities play a role in the image quality and usefulness of short-duration PET frames. A PET detector with a large axial FOV provides additional anatomic coverage for input to registration algorithms. Timing resolution has been previously demonstrated to improve edge definition \cite{10,11} and provide robustness \cite{12} when attenuation correction is not applied. Detector sensitivity provides a signal-to-noise benefit that allows for shorter frames and therefore improved temporal resolution. The scanners used in this study were the SIGNA PET/MR \cite{13} and the 4-ring Discovery MI \cite{14}. Both these scanners use modern silicon photomultipliers \cite{15} and have favorable specifications among clinical PET systems in each of these categories.

For rigid motion estimation, the methodology presented here uses fully reconstructed images and image registration to estimate the motion more quickly than the real time of the data, with a temporal resolution \( \geq 1 \text{ Hz} \) that is sufficient for almost any possible patient head motion. A fully list-mode event-by-event motion correction can then be performed using these estimated motion parameters \cite{8}. This correction could be applied retrospectively to any list-mode dataset. Estimating the motion directly using reconstructed images results in the actual 6 degrees of freedom of the motion parameters rather than indicators of when motion occurred.

The cardiac and respiratory motion can be tracked and quantified for gating purposes or for nonrigid motion estimation. Deformation parameters could be estimated from the short frames and used in a motion-corrected reconstruction. This possibility will be further investigated in the near future. Whether this motion estimation could be performed more quickly than real time will need to be evaluated.

The current implementation results in nonquantitative reconstructions since scatter correction is not applied. Although quantitative reconstruction is not imperative for the applications that we have demonstrated, there are applications, such as kinetic modeling, for which quantitative reconstruction would be important. Additionally, the reconstructions are not iterated until convergence; thus, we intend to investigate other algorithms that converge more quickly than MLEM. Lastly, for data-driven motion estimation using image registration, if the tracer dynamics cause a significant change in the distribution during the acquisition, then the registration to the chosen reference frame may fail. Methods to handle such situations will be implemented.

CONCLUSION

An infrastructure to perform ultra-fast reconstructions of very short frames of list-mode data has been presented. Subsecond
frames can be reconstructed more quickly than real time. The resulting series of reconstructions can be used for various applications, such as data-driven motion estimation. Once this infrastructure is made available online during the PET data acquisition, the nearly real-time frames could be displayed for the technologist to see during the scan. Other possible applications have been demonstrated here for various clinical case studies for brain, abdomen, and whole-body scans.

DISCLOSURE

Timothy Deller and Floris Jansen are employees of GE Healthcare. Matthew Spangler-Bickell was funded by GE Healthcare at the time of this study. No other potential conflict of interest relevant to this article was reported.

ACKNOWLEDGMENTS

We gratefully acknowledge Prof. Dr. Koen Van Laere of UZ Leuven, Belgium, and Uppsala University Hospital, Sweden, for providing clinical data; Mohammad Mehdi Khalighi of Stanford University, Palo Alto, CA, for assistance with data; and Charlotte Hoo of GE Healthcare, Chicago, IL, for acceleration contributions to the software.

KEY POINTS

QUESTION: What applications become possible if ultra-fast reconstructions of very short PET data frames are available?

PERTINENT FINDINGS: An infrastructure for reconstructing very short frames (≤1 s) of list-mode PET data more quickly than the frame duration is presented. Example applications are demonstrated of real-time visualization of the PET data during acquisition, data-driven image-based motion estimation, and quick whole-body surveys.

IMPLICATIONS FOR PATIENT CARE: This work has many significant clinical implications: improved patient monitoring using the real-time tracer distribution, improved diagnostic value of images through motion estimation and correction, and optimized use of scan time by taking whole-body surveys, among others.

REFERENCES