Whole Body Parametric Imaging on Clinical Scanner: Direct 4D Reconstruction with Simultaneous Attenuation Estimation and Time-Dependent Normalization

V. Y. Panin, Member, IEEE, H. Bal, M. Defrise, IEEE Fellow, M. E. Casey, Senior Member, IEEE, N. A. Karakatsanis, Member, IEEE, A. Rahmim, Senior Member, IEEE

Abstract—Whole body dynamic PET imaging has the potential to enhance detectability and quantification when assessing disease stage or progress. The same body region is repeatedly scanned within relatively short acquisition frames and with delays between time samples. Repeated scanning is sensitive to patient motion, which may cause mismatches between attenuation and activity maps and thus erroneous correction factors. Moreover, since count rates change with time, standard software correction factors can be time dependent.

The generation of parametric images requires proper physiological modeling and has been shown to benefit from so-called direct 4D reconstruction methods. In this work we extend the MLACF/MLAA algorithms for application to dynamic direct reconstruction. Handling time-dependent normalization requires a redesign of the existing algorithm as well.

The reconstruction methodology was verified on Siemens mCT scanner patient data using the standard Patlak model. Different number of frames and scan initiation time points were investigated. Initial results showed that direct 4D reconstruction outperformed the indirect approach. Available CT attenuation information can be corrected based on emission data.

I. INTRODUCTION

Whole body dynamic PET imaging [1, 2] has the potential to enhance diagnostic and prognostic assessment of disease due to the addition of quantitative information provided by the parametric images. The same body region or bed is repeatedly scanned within relatively short acquisition frames and with temporal delays between consecutive passes. The first encouraging results were obtained by indirect reconstruction [1], where time-dependent image reconstruction is performed first, followed by standard Patlak [3] kinetic model fit in time domain. Independent frame reconstruction can be performed by the best available methodology provided by the manufacturer.

Nevertheless, the generation of parametric images requires proper modeling of the noise, which may not be very accurate in the reconstructed image domain for the case of indirect kinetic modeling, and has been shown to benefit from so-called direct 4D reconstruction methods [2, 4-13]. The recently introduced ML-EM type reconstruction with a nested loop of kinetic parameters updates results in fast convergence for linear [14] and non-linear models [15, 16] by employing optimization transfer principles [17]. The nested EM algorithm independently updates each time frame of the activity, resembling an indirect approach. This separate data handling facilitates parallelization and leads to a lower computational burden. Despite the independent reconstruction step, activity images are always constrained by the underlying kinetic model through additional iterative step updates, which are performed in image space through a nested loop within each tomographic update.

Implementation of a direct approach requires dedicated clinical software. The standard data correction methodology, designed for independent frame reconstruction, must be adapted. For example, normalization factors are time-dependent (rather than just globally scaled) due to the variable count rate across the dynamic PET frames. In addition, repeated scanning can be sensitive to patient motion. For example, CT-based attenuation factors may not be fully compatible with the PET data due to patient motion, respiratory motion, and the fact that the short CT field-of-view yields only truncated data.

In this work we extend the MLACF/MLAA [18-21] algorithms for dynamic whole-body data direct reconstruction. In particular, the MLACF algorithm [18, 20] can be naturally adapted to the direct nested EM algorithm without explicit attenuation factor estimations. We will assume that attenuation is the same across all time data samples. Handling time-dependent normalization, due to dynamic dead-time effects, requires a redesign of the existing algorithm as well.

II. METHODS

A. Parametric Imaging

The dynamic data are acquired at various time points denoted by the index $m$. The activity in voxel $j$ at time $m$ is constrained by the parametric linear model:

$$f_{jm}(\theta) = \sum_k T_{mk} \theta_{jk}$$  \hspace{1cm} (1)
where $\theta_{jk}$ are kinetic parameters or coefficients over pre-selected temporal basis functions represented by the columns of matrix $T$. The vector $\{f_{j_1}(\theta), f_{j_2}(\theta), f_{j_3}(\theta), \ldots\}$ is the time activity curve (TAC) for voxel $j$.

### B. Parametric Image Reconstruction

Dynamic time-of-flight (TOF) prompt data $y = \{y_{im}\}$ with spatial projection (line of response, LOR) index $i$, TOF bin index $t$, and time index $m$ can be modeled by combining the modeled projection $p$ from the emission object $f$, corrected for scanner efficiency by a known time-dependent efficiency array (inverse of normalization) $e$ and for attenuation by an array of attenuation factors $a$. The background events have a known mean $b$, equal to the sum of the estimated scatter and randoms. The Poisson Likelihood objective function is then:

$$L(\theta, a) = \sum_{i,m} (y_{im} \ln(a_i e_{im} p_{itm} + b_{im}) - a_i e_{im} p_{itm} - b_{im}),$$

$$p_{itm} = \sum_j A_{it,j} f_{jm}(\theta)$$

where $A_{it,j}$ is the system matrix corresponding to the TOF geometric projector. In the following, quantities without index $t$ or $m$ denote quantities summed over the corresponding index:

$$A_{i,j} = \sum_t A_{it,j}, y_i = \sum_{t,m} y_{itm},$$

$$p_i = \sum_{t,m} e_{im} p_{itm}, b_i = \sum_{t,m} b_{im}$$

The MLACF algorithm is derived by applying a functional constraint on $a$. The attenuation factors that maximize the non-TOF, non-dynamic, i.e. static, likelihood of the TOF summed data $y_i$ for the given parametric images are estimated by:

$$a_i = \begin{cases} a_{i''}, \text{ known attenuation} \\ \frac{y_i - b_i}{p_i} > 0 \end{cases}$$

The primed spatial projection indices denote LORs that have known attenuation. Substituting (4) into (2) leads to a reduced log-likelihood objective function, which no longer contains the attenuation. The reduced log-likelihood is optimized using a surrogate function [22], which is separable in $j$. This surrogate can be optimized w.r.t. $\theta_{jk}$, leading to the following monotonic nested iterative procedure:

$$\theta_{jk}^{(n+1)} = \sum_m s_{jm} \sum_{mk} s_{jm} f_{jm}(\theta^{(n+1)}) \cdot \theta_{jk}^{(n+1)}$$

$$f_{jm}(\theta^{(n+1)}) = \sum_{i,t,m} A_{it,j} y_{itm} / \sum_{i,t,m} e_{im} p_{itm}(\theta^{(n+1)})^{b_{im} / p_{itm}(\theta^{(n+1)})} +$$

$$\sum_{t,m} b_{im}$$

$$\sum_{i,t} A_{it,j} y_{itm} / \sum_{i,t} e_{im} p_{itm}(\theta^{(n+1)})^{b_{im} / p_{itm}(\theta^{(n+1)})} +$$

$$\sum_{i,t} A_{it,j} y_{itm} / \sum_{i,t} e_{im} p_{itm}(\theta^{(n+1)})^{b_{im} / p_{itm}(\theta^{(n+1)})}$$

where $\tilde{y}_i$ are non-TOF and time integrated data corrected for background events, the index $l$ is the sub-iteration index for the update of the kinetic parameters, $n_i$ is the number of sub-iterations ($20$ were used in the presented results), $n$ is the iteration index of the nested algorithm as a whole, and $s$ is the sensitivity image, which is time-dependent:

$$s_{jm} = \sum_i A_{it,j} e_{im} a_{i''} + \sum_i A_{it,j} y_{im} e_{im} [p_i(\theta^{(n)})]$$

(7).

In cases of completely known attenuation (where only prime indexes are considered) and non-time-dependent normalization, (5) and (6) will reduce to an algorithm suggested in [14]. The MLACF algorithm was initialized with a single iteration of a direct nested algorithm, which uses available but potentially mismatched attenuation. This initialization also provides a total activity image, summed over dynamic frames and used for the MLACF rescaling. Since it was observed that the MLACF decreases in image value with an increase in iteration number and that the scaling parameter cannot be determined from data alone [23], the activity images were rescaled on a plane-by-plane basis after each subset update. Note that this scaling will not influence dynamic information, since it is a global TAC scaling for a given voxel. By this design, the first iteration of direct and parametric MLACF results in the same image.

The MLACF does not estimate an attenuation map ($\mu$-map). For illustrative proposes, we reconstruct $\mu$-maps from factors (4). Ten iterations of MLTR (the algorithm used in the MLAA $\mu$-map estimation step) with 21 subsets were used to produce the MLACF derived $\mu$-map image.

The TOF MLAA algorithm implementation is more straightforward, since (2) will be optimized with respect to both parametric and attenuation map images. The attenuation map reconstruction alternates with parametric image reconstruction and both update steps using established algorithms. The parametric images update is the same as (5) and (6) where attenuation factors are known (then the second and the third term of (6) are not considered) and iteration dependent. The attenuation factors are obtained from the attenuation map update step, which is a MLTR type algorithm:

$$a_{i''}(n) = e^{-\sum_i \mu_i^f(n)}, B_{itm} = e_{im} \sum_{j} \sum_{jm} n_{jm} f_{jm}(\theta^{(n)}) = e_{im} p_{itm}(\theta^{(n)})$$

$$\mu_i^{f(n)} = \mu_i^{f(n)} + \gamma \frac{1}{D} \sum_i \sum_{j} \sum_{jm} \frac{B_{itm} a_{i''}(n)}{B_{itm} a_{i''}(n) + B_{itm}}$$

(8).

Here, $B$ is “blank” data derived from current TAC (constrained by parametric images), $\gamma$ is the relaxation parameter used to accelerate convergence (we used a 1.5 value), and $D$ is the transaxial diameter of the reconstructed image support. The summation over the voxel $j$ index is a forward projector operation used in the computation of attenuation factor $a_i$. It can be modeled by a different geometrical system matrix $A^\mu$. The summation over projection index $i$ is a backprojection operation. Since the geometrical
functions
C. same subset, where the same modeled projection data (8) first, followed by a parametric image update (5-6) for the backprojection operation. We used an attenuation update combinations are summed over a TOF and time index before the system matrix of attenuation update is non-TOF, the data acquisition step lasted 45 seconds.

approximately from 10 to 90 minutes post injection, each bed step was acquired fifteen times, scanner at Johns Hopkins University. The data set consisted of six bed steps. Each bed step was acquired fifteen times, approximately from 10 to 90 minutes post injection, each bed acquisition step lasted 45 seconds.

For comparison purposes, frame-by-frame activity TOF PSF OS-EM reconstructions were performed to obtain TAC. The parametric images were estimated as a solution of the Least Squares fit (non-iterative solution) of model (1) to this image time sequence. This method commonly is referred to as an indirect approach. TAC reconstructions were summed over the time index to obtain SUV (standardized uptake value) images, which are commonly used in clinical practice for diagnostic tasks.

C. Patient Data
A set of FDG patient data was acquired on a Siemens mCT scanner at Johns Hopkins University. The data set consisted of six bed steps. Each bed step was acquired fifteen times, approximately from 10 to 90 minutes post injection, each bed acquisition step lasted 45 seconds.

A two parameter standard Patlak [3] model with slope $K_i$ and intercept $V$ was considered:

$$f_{jm} = \int_0^{t_m} C_p(t) dt' K_{ij} + C_p(t_m) V_j \equiv T_{m2} \theta_{j2} + T_{m1} \theta_{j1} \quad (9),$$

where $C_p(t)$ is the blood input function (BIF). The basis functions $T_{m1}$ and $T_{m2}$ (eq. 5) are constructed from the time samples of $C_p(t)$ and from its integral, respectively.

The BIF was derived from fully corrected (decay, frame length, calibration factor, etc. corrections) reconstructed images. The six minutes duration list mode file of the cardiac bed was acquired immediately after injection and had a duration of six minutes. Then it was histogrammed into 12 frames of ten seconds and 12 frames of twenty seconds duration (similar to [1]). A region of Interest (ROI) was drawn inside the left ventricles in the reconstructed images to define the BIF peak region. The fifteen dynamic acquisitions of the cardiac bed completed sampling of the BIF decaying tail using the same ROI.

We used up to six frames in the reconstructions. This defined the upper limit for a clinically feasible acquisition time of about 30 minutes. The liver ROI variability, defined as the standard deviation of this region divided by its mean value, served as a measure of noise. Various suspected liver tumors were used to evaluate contrast. As contrast value we considered the ratio of the mean value in the tumor ROI to the mean value of the liver ROI.

In the following results, two patient data sets were used. One patient had small lesions in the liver. This data set was used to assess image quality when various frames were used in the reconstruction. The second patient had advanced disease with multiple tumors in the liver and other organs. This patient reconstruction displayed CT and PET data mismatch and was used in the evaluation of the MLACF and MLAA algorithms performance.

III. RESULTS

A. Blood Input Function

Fig. 1 presents an example of image derived BIF of the first patient. The peak of BIF was well defined. Linear interpolation was used to produce BIF as a function of time.

B. Acquisition Time Window

The first six (~10-40 mins PI) and the last six (~60-90 PI) frames were used to produce parametric and SUV images, see Fig. 2. It was found that the first frames use resulted in the best contrast in the indirect approach reconstructions [26]. While the presented direct approach results showed a similar trend, a more uniform liver appearance can be achieved in the later frames, at least when the same iteration number is used. The indirect approach visually led to lesser tumor detectability at later frames, probably due to lower contrast. The tumor on the top of the liver was detectable in the SUV image in the later frames reconstruction, but it was not visible in earlier frames (corresponding image is omitted). Note that the slightly different cross-sections are shown for each time window, since the tumor was relocated over an extended period of time, presumably due to patient motion.
Fig. 3 provides a deeper understanding of presented image quality. According to Fig. 3a, direct and indirect approaches were in agreement, quantitatively, for the liver value. Nevertheless, earlier and later frames reconstruction resulted in a different liver value. In Fig. 3b, the contrast versus noise curve shows the superiority of direct reconstruction. The better contrast can be eventually achieved in earlier frames reconstruction; however, the contrast-noise tradeoff was similar when a relatively small (but commonly applied in clinical setting) iteration number was used. The SUV image resulted in lower tumor contrast overall; however its contrast – noise tradeoff curve was similar at a lower iteration number.

Fig. 4 shows the patient reconstruction from four frames (11-14) of 20 minutes total acquisition time. While the direct reconstruction did not degrade in image quality (the contrast – noise trade-off curve was similar), the indirect reconstruction quality suffered from shorter time acquisition. The liver ROI decreased its value compared to that of a six frame reconstruction. This can explain the better eventual tumor contrast in Fig. 4d.

Fig. 5 presents MLACF and MLAA attenuation maps of the second patient, along with a CT derived \( \mu \)-map. Patient motion was evident in the head and arms. The most distinguished artifact was due to a difference in position of liver dome. MLACF/MLAA derived \( \mu \)-maps showed liver dome repositioning of about 2 cm. This is likely due to a free breathing pattern during the PET scan, while the CT image was obtained at the state of inspiration.

Fig. 6 shows the corresponding artifact in SUV and the direct reconstruction \( K_i \). The liver was uniform in MLACF \( K_i \) image. Fig 6d showed that the MLAA resulted in a different liver value compared to that of MLACF and direct reconstruction, especially at the later iterations. The contrast-noise curve was similar for the tumor in the artifact free, lower, region of the liver for Direct, MLACF and MLAA of earlier iterations, Fig. 6e. In the upper region of the liver tumors have higher value in MLACF and MLAA \( K_i \) images, this can be attributed to the CT motion mismatch. This was reflected in better contrast with respect to the artifact absent liver even at the earlier iterations, see Fig. 6f. These observations were similar to ones we derived from dynamic MLACF computer simulations [22], where similar liver artifacts were modeled. The contrast-noise trade-off curve shows that MLACF converges more slowly compared to a
IV. DISCUSSION

The presented results showed the superiority of a whole body study direct approach in terms of the contrast-noise trade-off. There was no evidence of bias with this approach (due to non-negativity constraint) compared to the indirect method. However, the dependency of liver values on the time window position and duration choice requires a better understanding of suitability of the FDG Patlak model for this specific organ. The modification can be required, such as direct reconstruction Logan model [27], generalized Patlak model [28] or “self-chosen” organ specific model [29], to account for tracer reversibility. Another potential source of error can be the imprecision of the BIF definition from each frame reconstructed images. The investigations with population-based BIF can help clarify this issue.

While the earlier frame reconstruction showed better tumor contrast, a detectability study can clarify if later frames reconstruction can be more beneficial, resulting in less false positives. The later frame dynamic acquisitions will result in production of an SUV image as well, making this approach practically appealing [30].

We used time-dependent normalization factors and modified the algorithm accordingly. The motivation was in the existing requirement for clinical scan. For the presented studies, normalization can be different by about 10% between the first and last frames due to decaying activity and correspondingly lesser dead time effects. We did not investigate the significance of such a correction, but it can be assumed to be insignificant (ignoring quantification importance) in the presented oncology studies. It is left to future work to understand its significance for the high count rate dynamic studies.

The presented results showed that motion artifacts from CT and PET mismatch can be corrected by MLACF or MLAA techniques. The scatter correction was performed based on available CT information; therefore this approach will work, assuming relatively modest patient motion. The attenuation was assumed to be the same for all dynamic frames. While the approach can be generalized for the case of time-dependent attenuation, it will manifest the presence of non-negligible motion that should be corrected in the activity images as well by estimation the motion field.

Fig. 4. Reconstruction from frames 11-14, three iterations and 6 mm Gaussian filter post-smoothing. (a) Direct $K_i$, (b) Indirect $K_i$ and (c) SUV image. (d) Mean value of liver ROI versus liver ROI variability and tumor contrast versus liver ROI variability as functions of iteration number. Green curve represents direct $K_i$, red curve represents indirect $K_i$, and black curve corresponds to SUV. Circles denote frames 11-14 reconstruction, while crosses denote frames 10-15 reconstruction from Fig. 2-3.

Fig. 5. Attenuation maps: (a) CT derived, (b) 10 iterations MLACF derived, 6 mm Gaussian filter post-smoothing, and (c) 10 iterations MLAA estimation, 6 mm Gaussian filter post-smoothing.
V. CONCLUSIONS

The direct reconstruction outperformed an indirect one in whole body clinical data reconstruction. This approach is promising for maintaining acceptable image quality in case of clinically used acquisition duration at the one hour post injection. Time-dependent (due to dead time effects) normalization was incorporated into the existing direct reconstruction algorithm. Algorithms of attenuation estimation based on emission data were extended for dynamic data direct reconstruction. It was shown that CT information mismatch can be corrected and associated artifacts can be removed from parametric images.


