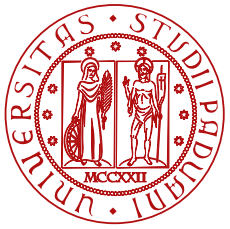


Università degli Studi di Padova

Department of Information Engineering
University of Padova

Master Thesis in Bioengineering

Spatio-temporal PET imaging reconstruction with learned diffeomorphism

Master Candidate

Marco Pinamonti

Student ID 2037014

Supervisor

Prof. Mattia Veronese

University of Padua

Co-supervisor

Prof. Massimiliano Colarieti-Tosti

KTH Royal Institute of Technology

Academic Year
2022/2023

"There is something irreversible about acquiring knowledge;
and the simulation of the search for it differs
in a most profound way from the reality."

— J. Robert Oppenheimer

Abstract

Positron Emission Tomography (PET) is a medical imaging modality to reconstruct the distribution of metabolic activity that is used to detect cancer lesions thanks to their peculiar metabolic fingerprints. However, since it requires long acquisition time, it is affected by motion artifacts and this leads to a difficult detection of small size tumours, that are the most important for early-stage diagnosis.

The Morphed Maximum Likelihood Activity and Attenuation (M-MLAA) algorithm has been developed to assess the motion artifact problem by gaining advantage of gated data and SynthMorph image registration network to reconstruct a motion corrected image.

This project's goal is to implement the M-MLAA algorithm on clinical data and to evaluate its performance; unfortunately, this was not achieved due to problems in the implementation of the Maximum Likelihood Activity and Attenuation (MLAA) algorithm on Synergistic Image Reconstruction Framework (SIRF) Python library. The results show that those problems might be caused by an incorrect definition of the Radon transform in the library.

Despite that, M-MLAA algorithm shows good performances when tested on synthetic data, suggesting that it could be a promising motion correction reconstruction method for PET images, capable of detecting early-stage cancer lesions.

Sommario

La Tomografia ad Emissione di Positroni (PET) è una modalità di imaging medico per ricostruire la distribuzione dell'attività metabolica, che viene utilizzata per rilevare lesioni tumorali grazie alle loro peculiari impronte metaboliche. Tuttavia, poiché richiede un lungo tempo di acquisizione, è soggetta ad artefatti da movimento e ciò porta ad una difficile individuazione dei tumori di piccole dimensioni, che sono i più importanti per una diagnosi precoce.

L'algoritmo Morphed Maximum Likelihood Activity and Attenuation (M-MLAA) è stato sviluppato per affrontare il problema degli artefatti da movimento sfruttando i dati suddivisi in gate e la rete neurale SynthMorph per la registrazione di immagini, al fine di ricostruire un'immagine corretta dagli artefatti di movimento.

L'obiettivo di questo progetto è l'implementazione su dati clinici dell'algoritmo M-MLAA e la valutazione delle sue prestazioni; purtroppo, ciò non è stato possibile a causa di problemi nell'implementazione dell'algoritmo Maximum Likelihood Activity and Attenuation (MLAA) sulla libreria Synergistic Image Reconstruction Framework (SIRF) Python. I risultati mostrano che tali problemi potrebbero essere causati da una definizione errata della trasformata di Radon nella libreria.

Nonostante ciò, l'algoritmo M-MLAA mostra buone prestazioni quando testato su dati sintetici, suggerendo che potrebbe rappresentare un promettente metodo di correzione dagli artefatti di movimento nella ricostruzione di immagini PET, consentendo di individuare lesioni tumorali in fase precoce.

Contents

List of Figures

List of Algorithms

List of Acronyms

CNN	Convolutional Neural Network
COM	Center of Mass
CRT	Coincidence Resolving Time
CT	Computed Tomography
FOM	Figure of Merit
FOV	Field of View
LOR	Line of Response
M-MLAA	Morphed Maximum Likelihood Activity and Attenuation
M-MLEM	Morphed Maximum Likelihood Expectation Maximisation
M-MLTR	Morphed Maximum Likelihood for Transmission Tomography
MC	Motion Correction
MLAA	Maximum Likelihood Activity and Attenuation
MLEM	Maximum Likelihood Expectation Maximisation
MLRR	Maximum Likelihood Reconstruction and Registration
MLTR	Maximum Likelihood for Transmission Tomography
MRI	Magnetic Resonance Imaging
MSE	Mean Squared Error
NN	Neural Network

ODL	Operator Discretization Library
OSEM	Ordered Subset Expectation Maximisation
PET	Positron Emission Tomography
PSNR	Peak Signal-to-Noise Ratio
RC	Recovery Coefficient
ROI	Region of Interest
SDNR	Signal Difference to Noise Ratio
SIRF	Synergistic Image Reconstruction Framework
SVF	Stationary Velocity Field
TOF	Time-of-Flight

1 Introduction

1.1 Background

Positron Emission Tomography (PET) is a medical imaging modality that reconstructs the 3D distribution of metabolic activity by detecting photons emitted during the in vivo annihilation of free electrons with positrons from an injected radioactive tracer. In principle, cancer lesions will be visible in the reconstructed image with high contrast compared to the surrounding healthy tissue thanks to their peculiar metabolic fingerprints (e.g. higher sugar metabolism) and PET is, in fact, one of the most powerful imaging modality for cancer diagnosis and staging.

1.2 Problem

PET modality requires long acquisition time to collect projection data with an acceptable noise level and this leads to motion artifacts that strongly affect contrast-to-noise ratio of lesions.

This is of particular interest when trying to detect tumours that are of size comparable to the system resolution (~ 5 mm) and that are continuously moving because of respiratory motion. In fact, those lesions are the most important ones to detect for early-stage diagnosis that can lead to a better prognosis for the patient.

1.3 Purpose

In order to improve the early-stage diagnosis of small lesions it is necessary to account for the motion during the reconstruction of the image and to correct based on that.

1.4 Methodology

One way of avoiding motion blurring is to divide the projection data in several subsets (gates) in which the patient (and therefore the activity distribution) can be considered stationary. Gating can be achieved by means of an external gate-triggering device or by analysing the data itself [18].

Once gated data is available one can choose to reconstruct each gate separately or just one of the gates, avoiding in this way motion artifacts. The drawback with the approach mentioned before is the higher noise level of the reconstructed images that can, especially for smaller lesions, result in scarce visibility and consequently uncertain diagnosis.

Recently, a new method has been proposed: by making use of the entire projection data-set, it is possible to simultaneously evaluate the deformations between gates and the activity distribution folded in one chosen template [10]. This increases PET diagnostic power for small lesions. The reconstruction of the activity distribution follows the line of the most used reconstruction algorithm for PET (Maximum Likelihood Expectation Maximisation (MLEM) [17]) with a modification accounting for the presence of different gates, while the morphing of all gates to the chosen template is performed via VoxelMorph [1, 2], a neural network for image registration. This algorithm is called M-MLEM.

The MLEM algorithm can be also modified to take into account for the tissue attenuation but it needs an attenuation image. If this image can be extracted from a Computed Tomography (CT) scan which is usually acquired before the PET acquisition, it is not possible to use the same image for all the different gates since there will be some misalignment in the activity and attenuation images due to the different time of acquisition. For this reason, the standard version of M-MLEM is not completely suitable for clinical application.

The attenuation correction can be performed for each individual gate by making use of the deformation fields estimated by an evolution of VoxelMorph called SynthMorph [5], a modality and contrast agnostic neural network for registration of the CT image to the different gates or by using a morphed version of algorithms that reconstruct activity distribution and attenuation from emission data like Maximum Likelihood Activity and Attenuation (MLAA) [13] and Maximum Likelihood Reconstruction and Registration (MLRR) [12].

The group of Prof. Massimiliano Colarieti-Tosti has proposed a morphed version of the MLAA algorithm, called M-MLAA that will be the object of this report.

The M-MLAA algorithm applies the idea behind the morphed version of the MLEM algorithm to the MLAA algorithm with the objective of being capable to perform motion correction PET reconstruction on clinical data by jointly estimating the activity and attenuation images for each gate and then using the information from multiple gates to reconstruct a motion corrected PET image.

1.5 Goal

The goal of this project is to apply the newly developed M-MLAA algorithm to clinical data acquired with the Biograph128 mCT-1104 PET scanner (Siemens Medical Solutions USA Inc., Hoffman Estate, Illinois, USA, VG60) at Karolinska University Hospital in Huddinge and compare the reconstructed images with the ones obtained with state-of-the-art algorithms, with the objective of evaluating the performance of M-MLAA.

1.6 Outline

The report is divided into the following six chapters:

1. The first chapter is this introduction to the work.
2. The second chapter gathers the theoretical information about PET imaging data acquisition and image reconstruction, then it presents the Morphed Maximum Likelihood Expectation Maximisation (M-MLEM) algorithm already developed for motion correction and it ends with a description of SynthMorph image registration Neural Network.

3. The Methodology chapter delineates the research methods followed in this project, starting from the procedure for decompressing the data acquired from the scanner, it follows with the presentation of three standard reconstruction techniques and subsequently reports the proposed reconstruction algorithm Morphed Maximum Likelihood Activity and Attenuation (M-MLAA).

4. In the fourth chapter the results obtained both on synthetic data - coming from the XCAT phantom - and on clinical data - coming from the Siemens Biograph mCT scanner - are displayed; here it's also shown an analysis of the SIRF library projectors about the creation of the attenuation and normalization correction matrices.

5. The Discussion chapter collects the author's reasoning on the results obtained in the previous chapter.

6. The final Conclusions chapter sums up the work and propose the future steps to continue developing this project.

2

Theoretical Background

This chapter presents the theoretical background necessary for a better understanding of the project's context. It encompasses the following subjects:

1. The first section provides a concise overview of how PET imaging techniques operate, how the data is acquired and presents some of the main reconstruction algorithms.
2. The subsequent section is titled "Motion correction image reconstruction" and offers a detailed description of the M-MLEM algorithm, developed to perform motion correction during PET image reconstruction.
3. The third and final section introduces the reader to SynthMorph, a convolutional neural network used to compute the diffeomorphism between gates. This step is crucial for deforming all the gates into a fixed one, a necessary process for motion correction.

2.1 PET IMAGING

Positron Emission Tomography (PET) is a medical imaging technique based on the detection of the coincidences of two annihilation photons with an energy of 511 keV, which are produced by a patient injected with a positron-emitting radiopharmaceutical.

The coincidences are detected using an electronic collimation technique, which selectively captures photons detected within a specific time window tailored for each particular scanner.

As shown in Figure 2.1 there are 4 types of coincident events that collectively go under the name of *prompts*:

True	The two photons detected by a detector pair come from the same annihilation event
Random	The two photons detected come from two different annihilation events
Scatter	One (or both) of the photons changes its direction due to Compton scattering
Multi-coincidence	More than two photons are detected in the same time window

Typically, multi-coincidence events are discarded due to the challenge of accurately positioning them.

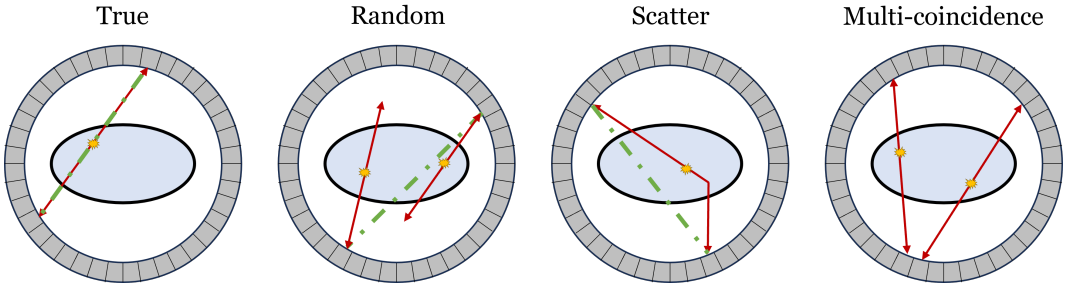


Figure 2.1: Types of coincident events. The annihilation events are represented with the yellow bursts, the red arrows describe the trajectories of the photons and the Line of Response (LOR) is drawn with a green dashed line.

2.1.1 Data acquisition in PET

Starting with the simple case of a single ring of detectors, when a coincidence is detected, it is possible to infer that the annihilation event occurred along the Line of Response (LOR) that connects the two detectors registering the coincidence. This line is characterized by two components:

- The angle ϕ of orientation of the LOR, ranging from $-90°$ to $90°$.
- The distance r of the LOR from the center of the field of view.

Each coincidence can be represented as a point in a diagram with distance r on the x-axis and angle ϕ on the y-axis. Photons emitted from the same point and detected by different LORs result in a graph with the shape of half a sine wave rotated by 90°. This type of graph is commonly known as a "sinogram" and is presented in Figure 2.2.

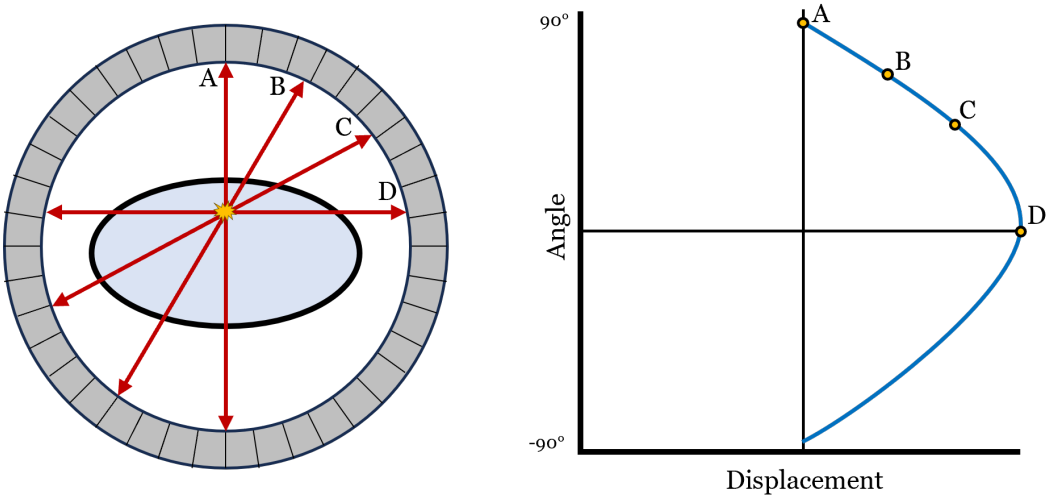


Figure 2.2: Creation of a sinogram curve. Adapted from Fahey [4]

Due to this specific structure, all parallel LORs have the same orientation angle, and they will be aligned on the same horizontal row in the sinogram. Hence, it is referred to as a "projection" [4].

During the acquisition process, each time a new coincidence is detected, the corresponding LOR is determined. After calculating its distance from the center of the Field of View (FOV) and its orientation angle, the associated pixel in the sinogram is incremented by one. At the end of the acquisition, each pixel will contain the number of events acquired along the corresponding LOR.

Time-of-Flight (TOF) Method

Time-of-Flight PET is an advancement in traditional PET imaging that takes advantage of the time difference between the arrival of the two 511 keV annihilation photons at the detectors. This time information is used to calculate the segment along the LOR where the annihilation event occurred.

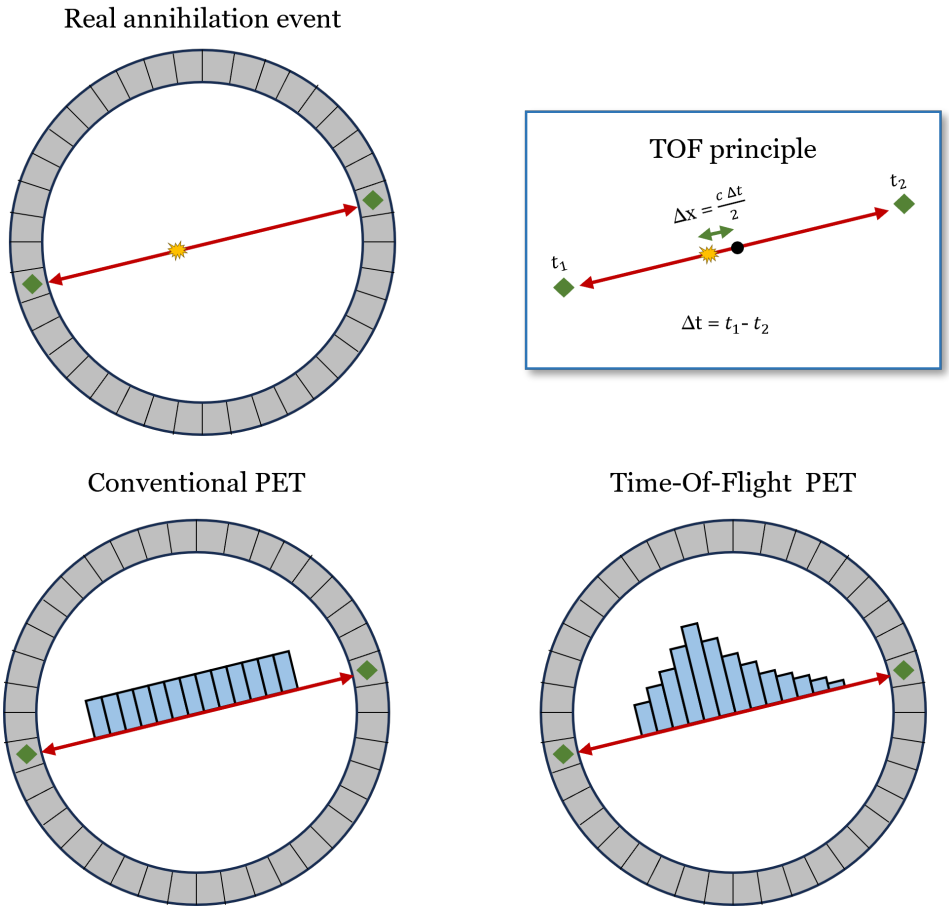


Figure 2.3: Time-of-Flight PET, adapted from Vandenberghe et al. [20]

As displayed in Figure 2.3, given two opposite detectors at a distance x from the center of the FOV and an event occurring at a distance Δx from the center, the two photons will travel different distances: $x - \Delta x$ and $x + \Delta x$. This leads to a time difference $\Delta t = \frac{2\Delta x}{c}$ [14]. By considering the Coincidence Resolution Time (CRT) of the scanner, which denotes the ability of a detector pair to resolve the time difference between the interactions of two detected photons in coincidence, the information Δt can be used to assign the annihilation event to a specific segment along the LOR instead

of assigning the event to the full LOR [15]. This is valuable for improving the quality of reconstructed images. However, in practice, the CRT of the current scanner is not small enough to assign the event to a single image voxel.

Sinogram Structure

In a 2D PET scanner, the detectors are allowed to be in coincidence only with the detectors on the opposite side of the same ring, and these events are defined as *direct coincidences*. In a 3D PET scanner, coincidences between detectors belonging to neighboring rings are also permitted, referred to as *cross coincidences*. The cross planes are characterized by the difference between the numbers of the rings of the detectors, and each scanner defines the maximum value of that difference.

To visualize the structure of the permitted cross-coincidences, a Michelogram is drawn (Figure 2.4), a specific diagram showing the detector ring numbers on the x and y axes, creating a table with cells that are either empty or populated with a dot or a circle to indicate whether coincidences across the corresponding rings are allowed. In this organization, a 2D PET scanner will only have dots along the main diagonal since only direct coincidences are allowed.

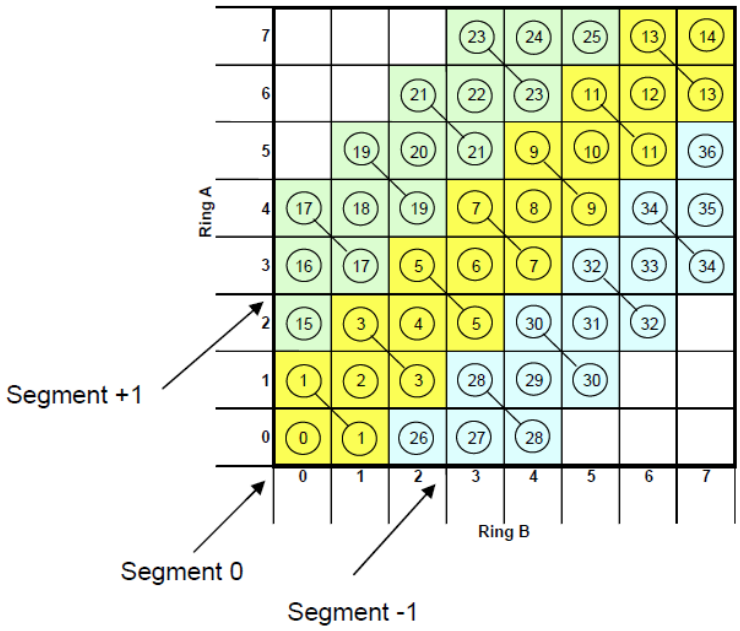


Figure 2.4: Example of a Michelogram organization.

Additionally, the data from two or more different cells of the Michelogram is usually combined to ensure a higher number of counts in the resulting plane and to reduce noise. The term *span* describes the extent of axial data combined, and it is defined as the sum of the number of cells in the Michelogram combined into an odd-numbered plane added to the number of cells combined into an even-numbered plane [4]. The obtained planes are then grouped into *segments*, which order the data starting from the inner diagonal of the Michelogram and going towards the outer diagonals.

2.1.2 PET image reconstruction

In absence of noise and any corrections, the process of acquisition of PET data can be expressed as the following forward problem:

$$g = A\, f$$

In the previous expression f represents the activity image that is unknown and that generates the data with A that is the forward operator that models the Radon transform that mathematically create the sinograms g starting from the image f.

The image reconstruction algorithms aim to solve the correspondent inverse problem: starting from the sinogram data g they estimate the original image f [19].

In this section two of the most used reconstruction algorithms are presented:

- The first is the Maximum Likelihood Expectation Maximisation (MLEM) algorithm, state-of-the-art in PET imaging to reconstruct the activity image.
- The second one is called Maximum Likelihood Activity and Attenuation (MLAA) and jointly reconstructs both the activity and attenuation images.

MLEM - Maximum Likelihood Expectation Maximisation

The MLEM algorithm estimates the emission density activity by maximising the likelihood function of the reconstructed image [17].

It starts from an initial estimates of the activity f^0 and updates the estimate based on this iterative formula that returns the new estimate $\hat{f}^{new}$ from and old estimate $\hat{f}^{old}$:

$$\hat{f}^{new}(b) = \hat{f}^{old}(b) \sum_{d=1}^{D} \frac{g^*(d)\, p(b,d)}{\sum_{b'=1}^{B} \hat{f}^{old}(b')\, p(b',d)} \tag{2.1}$$

where $g^*(d)$ is the number of emission registered in each detector d, D is the number of detector units, $p(b,d)$ is the probability that detector d detects an emission originated in box b and B is the number of boxes (pixels).

A more intuitive way to understand this algorithm is by considering the following statistical model:

$$g \sim Poisson(Af)$$

where f is the activity image, g is the acquired data and A is the forward operator that describes the physics of PET [10].

The MLEM update then becomes:

$$f^{(n+1)} = \frac{f^{(n)}}{A^*\mathbb{1}} A^* \left(\frac{g}{Af^{(n)}} \right) \tag{2.2}$$

where A^* is the backward operator.

The MLEM algorithm observes the following structure, also displayed in Figure 2.5 and in Algorithm 1:

1. Starts with an initial guess of the reconstructed image, usually initialized to a matrix filled with ones.
2. Forward project the estimated image, computing its projection.
3. Compares the computed projection with the measured ones, obtaining the discrepancies in Projection Space.
4. Back-project those discrepancies in Image Space.
5. Updates the estimated image and continues from point 2.

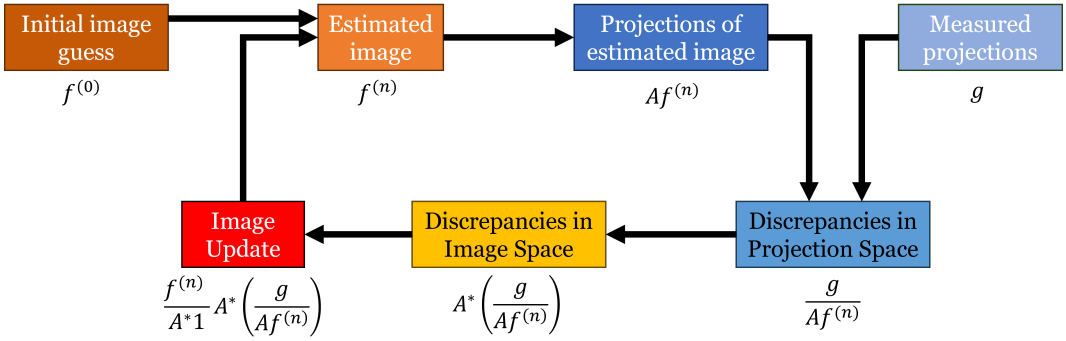


Figure 2.5: MLEM algorithm scheme, adapted from Ravi et al. [11].

Algorithm 1 MLEM

Require: $n \geq 0$

$A \leftarrow A$ ▷ Radon transform

$A^* \leftarrow A^*$ ▷ adjoint operator of Radon transform

$n \leftarrow MLEM_{iter}$ ▷ nb of MLEM iterations

$f^{(0)}$ ▷ init of activity image

g ▷ projection data

for $i \leftarrow 0, n$ **do**

$$f^{(n+1)} = \frac{f^{(n)}}{A^*\mathbb{1}} A^* \left(\frac{g}{Af^{(n)}}\right)$$

end for

MLAA - Maximum Likelihood Activity and Attenuation

The MLAA estimation algorithm is an image reconstruction algorithm that jointly estimates the activity and attenuation correction factors with the use of Time-of-Flight (TOF) data [13].

Each iteration of the algorithm can be divided in 2 parts:

- an initial MLEM step that explicitly uses the TOF PET data to update the activity estimate by keeping the attenuation coefficients constant:

$$\forall i: \quad a_i^n = e^{-\sum_j l_{ij}\mu_j^n} \tag{2.3}$$

$$\forall j: \quad \lambda_j^{n+1} = \frac{\lambda_j^n}{\sum_{it} a_i^n c_{ijt}} \sum_{it} a_i^n c_{ijt} \frac{y_{it}}{\sum_{\epsilon} a_i^n c_{i\epsilon t} \lambda_\epsilon^n + s_{it}} \tag{2.4}$$

- a subsequent Maximum Likelihood for Transmission Tomography (MLTR) step to calculate the new attenuation estimate:

$$\forall i: \quad \psi_i^n = a_i^n \sum_{jt} c_{ijt} \lambda_j^{n+1} \tag{2.5}$$

$$\forall j: \quad \mu_j^{n+1} = \mu_j^n + \frac{\sum_i l_{ij} \frac{\psi_i^n}{\psi_i^n + s_i} \left(\psi_i^n + s_i - y_i \right)}{\sum_i l_{ij} \frac{\left(\psi_i^n\right)^2}{\psi_i^n + s_i} \sum_\epsilon l_{i\epsilon}} \tag{2.6}$$

With reference to the formulas above: a^n represents the attenuation factors computed from the current attenuation map estimate μ^n, λ represents the activity estimate, c_{ijt} is the sensitivity of the detector, ψ_i^n are the expected TOF-integrated count for LOR i, y_i is the measured count and s_i is the additive contribution (randoms and scatter coincidences).

The MLAA algorithm observes the following structure (Figure 2.6):

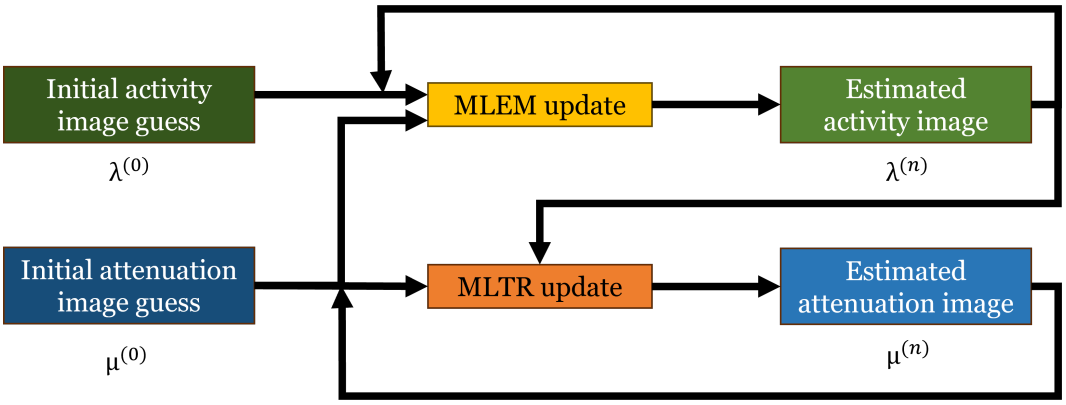


Figure 2.6: MLAA algorithm scheme.

The MLAA formulas can be rewritten with the use of mathematical operators A and A^* that model the forward and back-projection of the Radon transform; for the particular case of synthetic data that does not model the sensitivities of the detector and the randoms and scatter coincidences contributes, the resulting expression are:

- MLEM step:

$$f^{(n+1)} = f^{(n)} \frac{A^* \left[e^{-A\left(\mu^{(n)}\right)} \frac{g}{\left[e^{-A\left(\mu^{(n)}\right)} A\left(f^{(n)}\right) \right]} \right]}{A^* \left[\mathbb{1} \cdot e^{-A\left(\mu^{(n)}\right)} \right]} \tag{2.7}$$

- MLTR step:

$$\mu^{(n+1)} \leftarrow \mu^{(n)} + \frac{A^* \left[e^{-A\left(\mu^{(n)}\right)} A\left(f^{(n)}\right) \right]}{A^* \left[e^{-A\left(\mu^{(n)}\right)} \cdot A\left(\mathbb{1}\right) \right]} \tag{2.8}$$

The complete structure of the MLAA algorithm is presented in Algorithm 2.

Algorithm 2 MLAA

Require: $n \geq 0$
Require: $m \geq 0$
Require: $p \geq 0$
  $A \leftarrow A$ ▷ Radon transform
  $A^* \leftarrow A^*$ ▷ adjoint operator of Radon transform
  $n \leftarrow MLAA_{iter}$ ▷ nb of MLAA iterations
  $m \leftarrow MLEM_{iter}$ ▷ nb of MLEM sub-iterations
  $p \leftarrow MLTR_{iter}$ ▷ nb of MLTR sub-iterations
  $f^{(0)}$ ▷ init of activity image
  $\mu^{(0)}$ ▷ init of attenuation image
  g ▷ projection data
  for $i \leftarrow 1, n$ **do**
    $f^{(1)} \leftarrow f^{(i-1)}$
    $\mu^{(1)} \leftarrow \mu^{(i-1)}$
    for $j \leftarrow 1, m$ **do**

$$f^{(j+1)} \leftarrow f^{(j)} \frac{A^*\left[e^{-A\left(\mu^{(i)}\right)} \frac{g}{\left[e^{-A\left(\mu^{(i)}\right)} A\left(f^{(j)}\right)\right]}\right]}{A^*\left[\mathbb{1} \cdot e^{-A\left(\mu^{(i)}\right)}\right]}$$ ▷ MLEM step

    end for
    $f^{(i)} \leftarrow f^{(j)}$
    for $k \leftarrow 1, p$ **do**

$$\mu^{(k+1)} \leftarrow \mu^{(k)} + \frac{A^*\left[e^{-A\left(\mu^{(k)}\right)} A\left(f^{(k)}\right)\right]}{A^*\left[e^{-A\left(\mu^{(k)}\right)} \cdot A(\mathbb{1})\right]}$$ ▷ MLTR step

    end for
    $\mu^{(i)} \leftarrow \mu^{(k)}$
  end for

2.2 Motion Correction image reconstruction

The algorithms described in the previous section are the state-of-the-art for PET image reconstruction but they still present one major flaw: since the data acquisition is not instantaneous and it last several seconds, usually between 2 to 3 minutes per bed position, the patient cannot be considered stationary for the whole process. This causes the data to be affected by mo-

tion artifacts given the fact that the annihilation events occurring in the same relative position of the body can be localised in different spatial coordinates of the PET system.

In particular, during a thoracic PET acquisition, the principal motion artifact is related to the respiration process: the movement of the lungs caused by the alternation of inspiration and expiration leads to noticeable artifacts in the creation of the image that impact its final quality.

In order to account for this problem and correct for those type of artifacts some Motion Correction (MC) techniques have been developed for PET imaging. In the following section is presented the Morphed Maximum Likelihood Expectation Maximisation (M-MLEM) algorithm, that evolves the state-of-the-art Maximum Likelihood Expectation Maximisation (MLEM) in order to correct for the movement of the patient.

2.2.1 M-MLEM

M-MLEM is a motion correction algorithm that combines the MLEM algorithm with deep-learning based movement registration [10].

The basic idea behind this method is to divide the acquired data in different respiratory *gates* so that each gate contains events coming from a static position of the patient and then reconstruct the motion-corrected image by using the gated data and the diffeomorphisms between the gates calculated with the help of a Convolutional Neural Network (CNN) to merge together the different gates.

As a way to better understand how this works, let's start with the definition of the mathematical model behind. Given $N + 1$ different gates, suppose to have obtained g_i, $i = 0, \dots, N$ data from f_i images with the statistical model:

$$g_i \sim Poisson(Af_i) \quad i = 0, \dots, N$$

Two consecutive images f_{i-1} and f_i are related by the statistical model

$$f_i = W_{\psi_i} f_{i-1} + e_i$$

where $W_{\psi_i} : X \mapsto X$ represents the intensity-preserving action $W_{\psi_i} f(x) \coloneqq f\left(\psi^{-1}(x)\right)$ of the diffeomorphism $\psi \colon \Omega \to \Omega$ that binds the two images

together and models the movement and e_i is a X-valued random variable [10].

The algorithm then is divided in two steps:

- the *motion estimation*: for each new estimate of the images f_i^k the diffeomorphism is calculated by minimizing

$$\sum_{i=1}^{N} \left(d_2 \left(f_i^k, W_{\psi_i} f_{i-1}^k \right) + \lambda R \left(\psi_i \right) \right)$$

  where d_2 is the L^2-distance on X and R is a regularisation term with its regularisation parameter λ. The solution to this registration problem is found with the use of *VoxelMorph* [2] Neural Network that returns a mapping matching the two images it receives as input:

$$\gamma \left(f_i, f_{i-1} \right) \coloneqq \exp \left(\psi_i \left(f_i, f_{i-1} \right) \right)$$

- the *reconstruction*: after assuming $f_i^k \approx W_{\psi_i^{k+1}} f_{i-1}^k$ for $i = 1, \ldots, N$, the minimisation problem becomes

$$\arg\min_{f_i} \sum_{i=0}^{N} d_{KL} \left(g_i \parallel A \, f_i \right)$$

  By focusing on the first gate, number zero, the optimisation is reduced to the only variable f_0:

$$\arg\min_{f_0} \sum_{i=0}^{N} d_{KL} \left(g_i \parallel A_i \, f_0 \right)$$

  where $A_i \coloneqq A W_{\phi_i}$ with $\phi_i \coloneqq \psi_i \circ \cdots \circ \psi_1$ for $i = 1, \ldots, N$. Solving this it is possible to obtain the next estimate of the first gate with MLEM written for the compound operator $A = (A_0, \ldots, A_N)$:

$$f_0^{(n+1)} = \frac{f_0^{(n)}}{\sum_{i=0}^{N} A_i^T \mathbb{1}} \sum_{i=0}^{N} A_i^T \left(\frac{g_i}{A_i \, f_0^{(n)}} \right) \tag{2.9}$$

Image registration

The motion estimation problem is solved with use of the Neural Network (NN) *VoxelMorph* which is capable of performing 3D medical image registration [1, 2]. For this project, a variation of this NN has been selected, called SynthMorph [5] that is presented more thoroughly in Section 2.3.

The complete structure of M-MLEM is described in Algorithm 3.

Algorithm 3 M-MLEM

Require: $n \geq 0,\ m \geq 0,\ p \geq 0$
Require: $N \geq 1$
  $A \leftarrow A$ ▷ Radon transform
  $A^* \leftarrow A^*$ ▷ adjoint operator of Radon transform
  $n \leftarrow MLEM_{init}$ ▷ nb of initial MLEM iterations
  $m \leftarrow MMLEM_{iter}$ ▷ nb of M-MLEM iterations
  $p \leftarrow MLEM_{iter}$ ▷ nb of MLEM iterations
  $N \leftarrow nb_{gates} + 1$ ▷ nb total of gates
  $f_0^{(0)}, \ldots, f_N^{(0)} \leftarrow 1 \ldots 1$ ▷ init of the distribution estimation
  $g_0, \ldots, g_N$ ▷ data for every gates
  for $i \leftarrow 0, N$ **do**
    for $j \leftarrow 1, n$ **do**
      $f_i^{(j+1)} \leftarrow \frac{f_i^{(j)}}{A^*\mathbb{1}} A^* \left(\frac{g}{A f_i^{(n)}} \right)$ ▷ Initial MLEM reconstruction for each gate
    end for
  end for
  for $k \leftarrow 1, m$ **do**
    for $i \leftarrow 1, N$ **do**
      $\psi_i \leftarrow \gamma \left(f_{i-1}, f_i \right)$ ▷ Network registration
    end for
    $W_0 \leftarrow Id$
    for $i \leftarrow 1, N$ **do**
      $W_i \leftarrow W_{\psi_i} W_{i-1}$
      $A_i \leftarrow A W_i$
    end for
    $A_0 \leftarrow A$
    $f_{0-upd}^{(0)} \leftarrow f_0^{(k-1)}$
    for $j \leftarrow 1, p$ **do**
      $f_{0-upd}^{(j)} \leftarrow \frac{\sum_{i=0}^{N} A_i^* \left[\frac{g_i}{A_i (f_0^{(j)})} \right]}{\sum_{i=0}^{N} A_i^* \mathbb{1}}$ ▷ M-MLEM update to gate 0
    end for
    $f_0^{(k)} \leftarrow f_{0-upd}^{(j)}$ ▷ New estimate of gate 0
    for $i \leftarrow 1, N$ **do**
      $f_i^{(k)} \leftarrow W_i f_0^{(k)}$ ▷ New estimates of other gates
    end for
  end for

Limitations

This algorithm can be applied to synthetic data by creating different gates with the deformation of the same phantom but it is not so easily transposed

to clinical data because for the purpose of a good reconstruction of real data there's the need of correcting for the attenuation effect with the use of an attenuation map. Such map is usually fixed for the entire acquisition and so will be the same for all the different gates and inherently not aligned with any of them and this will result in a sub-optimal image reconstruction.

The proposed solution to this problem is to estimate the attenuation image at the same time of the activity with the use of the MLAA algorithm which will reconstruct a different attenuation map for each gate and that attenuation image will be already aligned with the activity image of that gate.

2.3 SynthMorph

SynthMorph is a Neural Network (NN) designed for image registration. It is trained on synthetic label maps and images to achieve robust generalization across various medical imaging contrasts [5].

Although initially developed for accommodating different MRI contrasts, SynthMorph can also be effectively applied to PET images.

2.3.1 Purpose

Classical registration techniques require optimizing the warping estimate for each new image pair. In contrast, learning-based registration utilizes datasets of images to learn a function that maps an image pair to a deformation field aligning the images. This leads to sub-second runtimes on a GPU [5]. However, a limitation of this approach is that it may not generalize well to different MRI contrasts, resulting in sub-optimal performance for new image types.

SynthMorph addresses this limitation as it presents itself as a contrast-agnostic registration technique. It is trained solely on synthetic data composed of label maps. A generative model randomly creates these maps, which are then synthesized with different contrasts, deformations, and artifacts [5].

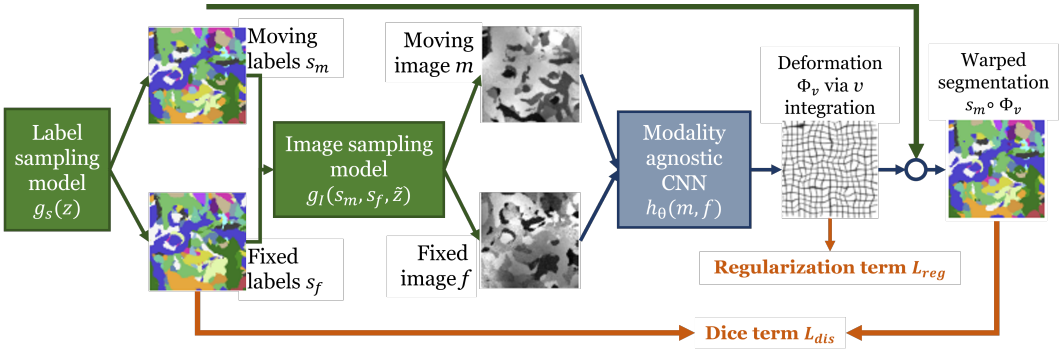


Figure 2.7: SynthMorph training scheme, adapted from Hoffmann et al. [5]

2.3.2 Operating principle

As shown in Figure 2.7, the training of the NN starts with a random label map s, which serves as the input for a generative model. This model creates two new label maps, s_f for the fixed labels and s_m for the moving labels, by deforming s with a random smooth diffeomorphic transformation.

Subsequently, the pair of s_f and s_m is used to synthesize gray-scale images m, f with generative models of MR images used for Bayesian segmentation [5]. The function involved in this process is defined as $g_I\left(s_m, s_f, \tilde{z}\right) = \left\{m, f\right\}$, which calculates two 3D intensity volumes based on the label maps and seed $\tilde{z}$.

These intensity volumes are further corrupted with a spatially varying intensity-bias field. The intermediate steps of the process of creation of the synthetic data are shown in figure 2.8.

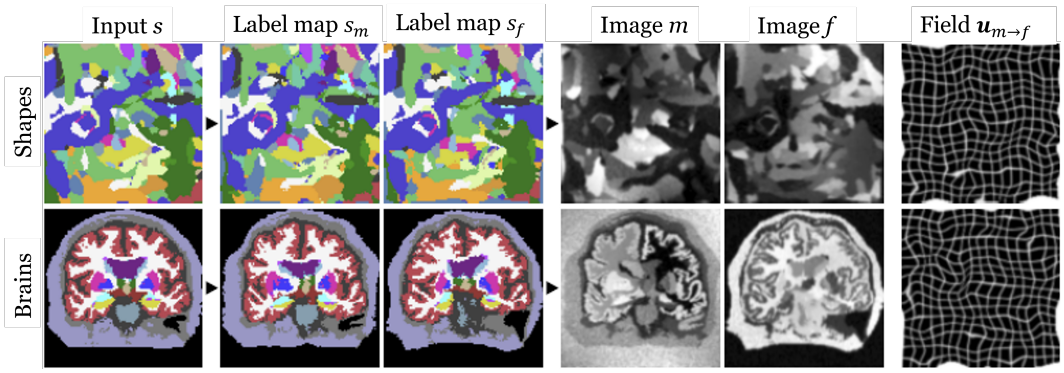


Figure 2.8: SynthMorph training data synthesis, adapted from Hoffmann et al. [5]

Once the pair $\left\{m, f\right\}$ of moving and fixed gray-scale images is obtained,

it is fed into a CNN that calculates the deformation Φ_v between the images. This deformation is then used to warp the moving labels s_m into the space of the fixed labels $s_m \circ \Phi_v$. The loss function used in the training combines a Dice term L_{dis} that accounts for the differences between the fixed labels s_f and the warped $s_m \circ \Phi_v$, and a regularization term L_{reg} that prevents the CNN from producing output deformations with high displacement.

When the SynthMorph network is trained with shape-agnostic images, it is referred to as the `sm-shapes` model. However, the same network can also be trained with anatomical labels of the brain, resulting in the `sm-brains` model.

2.3.3 Architecture of the Neural Network

The model utilizes a network structure from the VoxelMorph library: a convolutional U-Net predicts a spatially varying field (SVF) v_θ from the input m, f, where m and f represent the moving and fixed images, respectively.

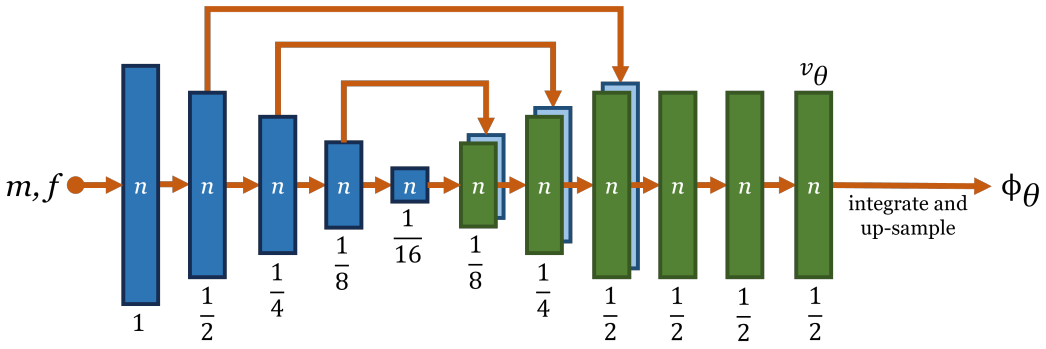


Figure 2.9: SynthMorph U-Net architecture, adapted from Hoffmann et al. [5]

Figure 2.9 illustrates that the encoder consists of 4 blocks, each composed of a stride-2 convolution and a LeakyReLU layer. These blocks halve the resolution of the inputs. On the other hand, the decoder has 3 blocks with a stride-1 convolution, an up-sampling layer, and a skip connection to the corresponding encoder block. The final SVF v_θ is obtained after 3 further convolutions at half resolution, whereas the warp ϕ_θ requires an integration and up-sample step [5].

The convolution kernels have dimensions of $3 \times 3 \times 3$ pixels, and the default network width is $n = 256$.

The network is implemented using TensorFlow and Keras, and the training is performed with an Adam optimizer with a batch size of one registration pair and an initial learning rate of 10^{-4}. The convergence is assessed via Dice metric and usually takes 4×10^5 iterations [5].

2.3.4 Performance

As presented in Figure 2.10, when compared to the most commonly used classical method for registration, the `sm-shapes` network matches the performance in all the MRI contrasts compared (T2-T2, T1-PD, T1-T2) except for T1-T1 registration, where it has slightly lower Dice score and higher surface distances. On the other hand, the `sm-brains` model outperforms all the classical methods in every MRI contrast, in particular it exceeds the state-of-the-art accuracy of VoxelMorph-Normalized Cross-Correlation network (`vm-ncc`) [5].

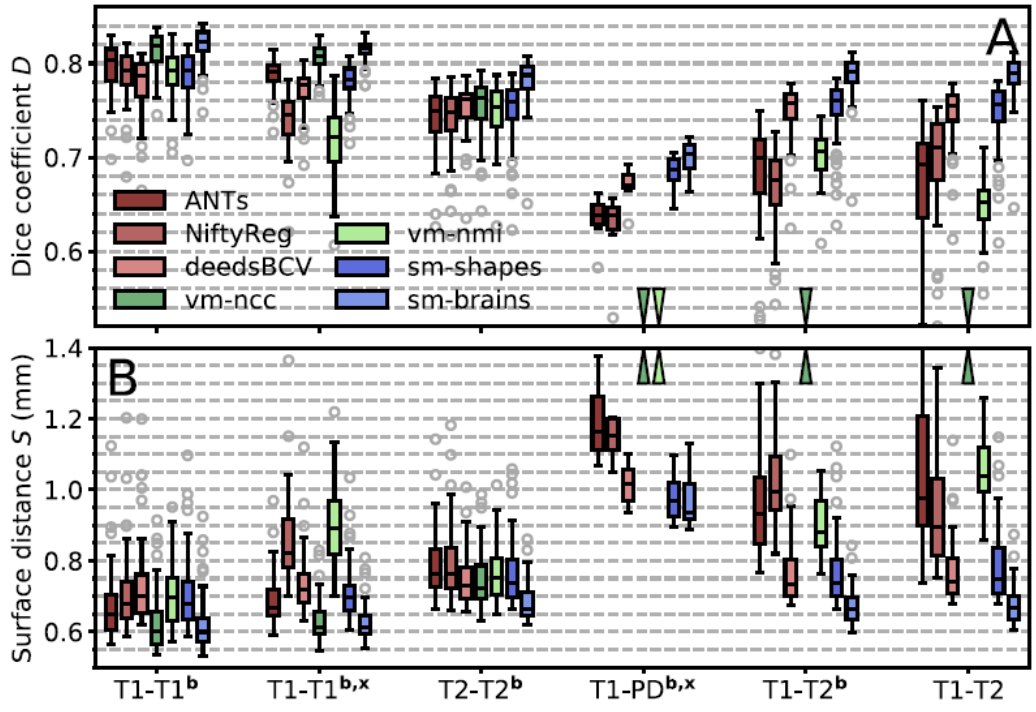


Figure 2.10: SynthMorph performances as (a) volume overlap D using Dice metric and (b) mean symmetric surface distance S between label contours, adapted from Hoffmann et al. [5]

2.3.5 Strengths of this NN

SynthMorph addresses the main weak point of VoxelMorph, which is its inadequate performance with MRI image contrasts different from the ones it has been trained with. Since SynthMorph is trained on non-MRI images that have a wide range of different contrasts, this NN can generalize well and becomes resilient to contrast changes.

3 Methodology

In this chapter are presented the methodologies employed in this research to process and reconstruct clinical PET data acquired with the Biograph128 mCT-1104 PET scanner (Siemens Medical Solutions USA Inc., Hoffman Estate, Illinois, USA,VG60) at Karolinska University Hospital in Huddinge.

The chapter begins with an overview of the data processing pipeline, which involves decompressing the raw data from the scanner and converting the physiological information related to respiratory motion. A custom Python script is implemented to automate these tasks and streamline the data organization. Subsequently, the description delves into the details of the Siemens mCT sinogram organization, discussing the scanner's crystal configuration and the corresponding sinogram structure.

Following this, two primary reconstruction methods are presented: the "standard" clinical data reconstruction using the proprietary e7-tools software or the state-of-the-art reconstruction algorithm MLEM and MLAA, and the proposed Morphed Maximum Likelihood Activity and Attenuation (M-MLAA) reconstruction algorithm, which is a novel approach developed within this research.

The image quality evaluation is conducted via visual inspection for clinical data and using also figures of merit for synthetic data, enabling a comprehensive assessment of the reconstructed images. Through the implementation of the M-MLAA algorithm, the aim is provide improved image quality and enhanced visualization of lesions and other critical anatomical

features. The results of the evaluation serve to validate the effectiveness of the proposed methodology in enhancing clinical PET image reconstruction.

The methodology pipeline followed to evaluate the proposed algorithm on clinical data is presented in Figure 3.1.

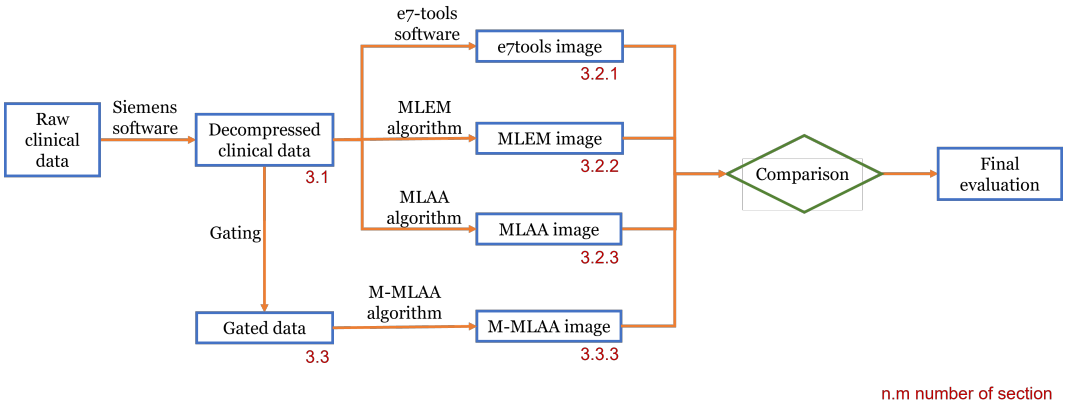


Figure 3.1: Methodology pipeline to evaluate the M-MLAA algorithm, the numbers in red represent the corresponding section in this chapter.

3.1 Clinical data decompression

The "Siemens Biograph 128 mCT" PET scanner stores the acquired patient data with the following folder structure:

- Patient folder
  - CT images folder - which contains 284 Dicom files, one for each CT axial slice
  - PET raw data folder - which contains a variable amount of Dicom files, one file for the sinogram and another one for the list-mode of each bed position and an eventual physiological file

3.1.1 Raw data decompression

The PET raw data Dicom files are stored in a compressed format, therefore, to obtain usable files, it is necessary to execute the proprietary software "JSRecon12" that will create a new folder containing the normalization file

and its header and multiple sub-folders, one for each sinogram and list-mode files. This sub-folders will contain a data file, its header and some batch scripts to autonomously perform the reconstruction based on Siemens proprietary software.

3.1.2 Physio-file conversion

In order to decompress the physiological file that registered the respiration motion via a respiratory belt, another script called "PhysioReader" is run and it returns a `.csv` file containing as many rows as many time intervals recorded and 5 columns labelled as:

Time	Representing the correspondent time of the data in the row
Amplitude	Containing the amplitude of the movement registered
Beam on	Flag to notice if the PET data acquisition is in progress or not
Inspiration & Expiration	Flags to distinguish between the two phases of the respiration process

3.1.3 File selection

Once the decompressed files are obtained, their structure has been reorganised to select only the useful ones and copy them in a new folder to simplify the next analysis.

To do all the steps defined in the previous paragraphs of this section the author has implemented a Python script that semi-autonomously runs the Siemens software and reorders the data.

3.1.4 Siemens mCT sinogram organization

The Siemens Biograph128 mCT-1104 scanner is composed by 32 447 crystals divided in 4 major rings each one containing 48 detector blocks. A detector block is defined by a square grid of 13×13 crystals of dimensions $4\,\text{mm} \times 4\,\text{mm} \times 20\,\text{mm}$.

The total number of single-detector rings of the scanner is 55 because there are 4 blocks × 13 crystals = 52 "real" detector rings, plus 3 "virtual" rings that model the physical empty volume between one block and the next one. Given this fact, the number of direct planes is 55 but the data is stored with span[1] equal to 11 and this means that there are other 54 cross-planes to add, for a total of 109 planes.

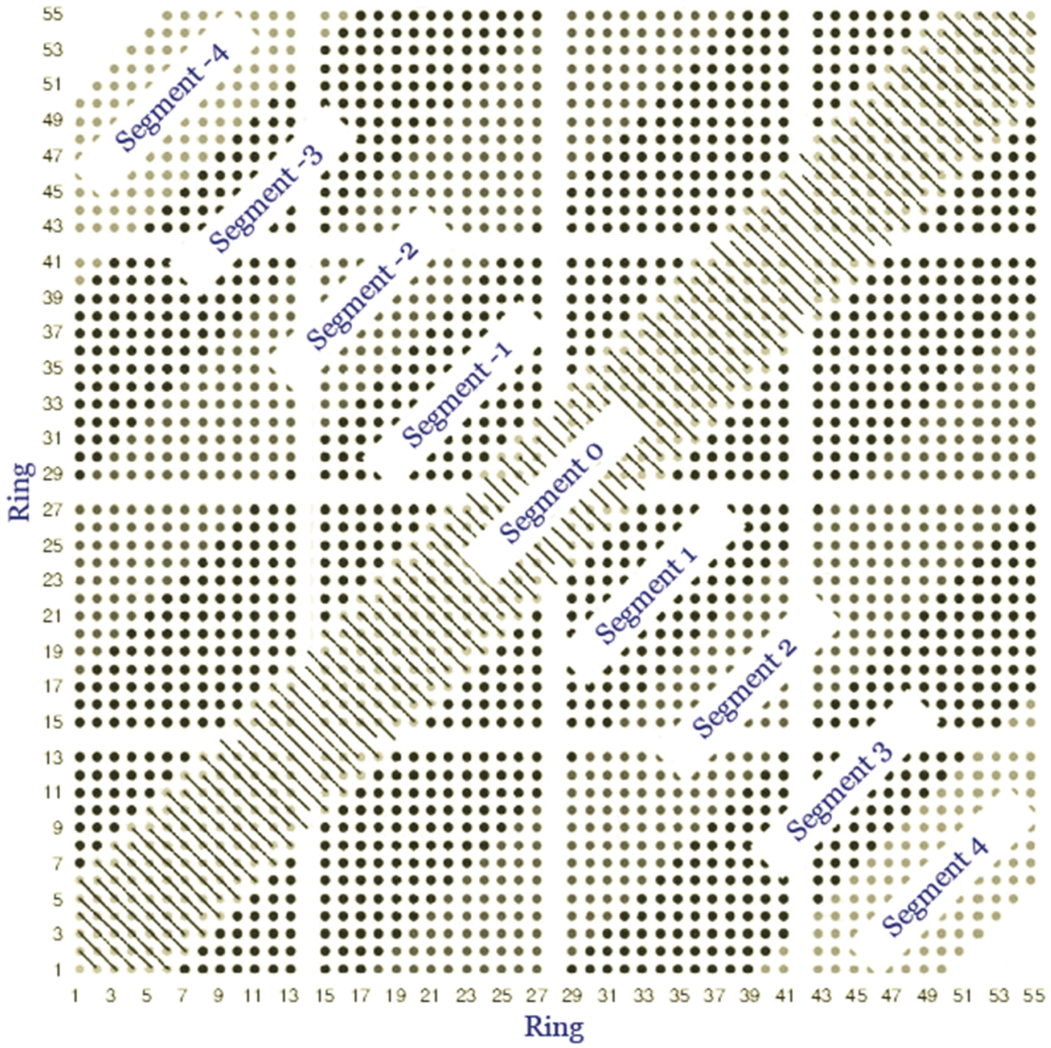


Figure 3.2: Michelogram of the Biograph mCT-1104 scanner, adapted from Martí-Climent et al. [9]

The acquired sinograms are 621 in total and they are grouped together

[1]The sum of the number of cells in the Michelogram combined into an odd-numbered plane added to the number of cells combined into an even-numbered plane [4].

as shown in Figure 3.2 in 9 segments referred as 0, ±1, ±2, ±3 and ±4 that respectively contain 109, 97, 75, 53 and 31 sinograms each. The sinograms themselves are 168×400 matrices where the first dimension corresponds to the number of angular samples and the second one to the number of radial samples. This results in an angular resolution of 1.071 °/pixel and a radial resolution equal to 2.005 mm/pixel, the axial resolution of the scanner is equal to 2.027 mm/pixel and given the 109 planes results in an axial FOV of 220.943 mm [9].

The described structure can be arranged in a three-dimensional matrix with size $621 \times 168 \times 400$ but since the data is TOF this matrix is repeated for each one of the TOF bins, so the final matrix will have size $13 \times 621 \times 168 \times 400$. An example of the different TOF bins of a single transversal plane is shown in Figure 3.3.

The scanner also registers the delayed coincidences and they are stored as the "fake" 14^{th} TOF bin, as can be seen in Figure 3.4 that presents an example of the random coincidences in the same plane as in Figure 3.3.

The data can be summed along the TOF dimension to obtain *NON-TOF* data that is later needed for the reconstruction algorithms; an example NON-TOF sinogram of a single plane is shown in Figure 3.5.

3.2 Standard clinical data reconstruction

The standard method used in clinics to reconstruct PET images from the data acquired through the Siemens mCT scanner is with its proprietary software *JSRecon* and *e7-tools* that are able to decompress the raw data obtained from the scan, then they create the sinograms and the attenuation maps and finally apply different types of reconstruction techniques to get a reconstructed image.

In this section are also presented other two state-of-the-art reconstruction algorithms used in clinics: the Maximum Likelihood Expectation Maximisation (MLEM) and the Maximum Likelihood Activity and Attenuation (MLAA). Those algorithms have been implemented by the author as Python scripts, following the procedures described in literature.

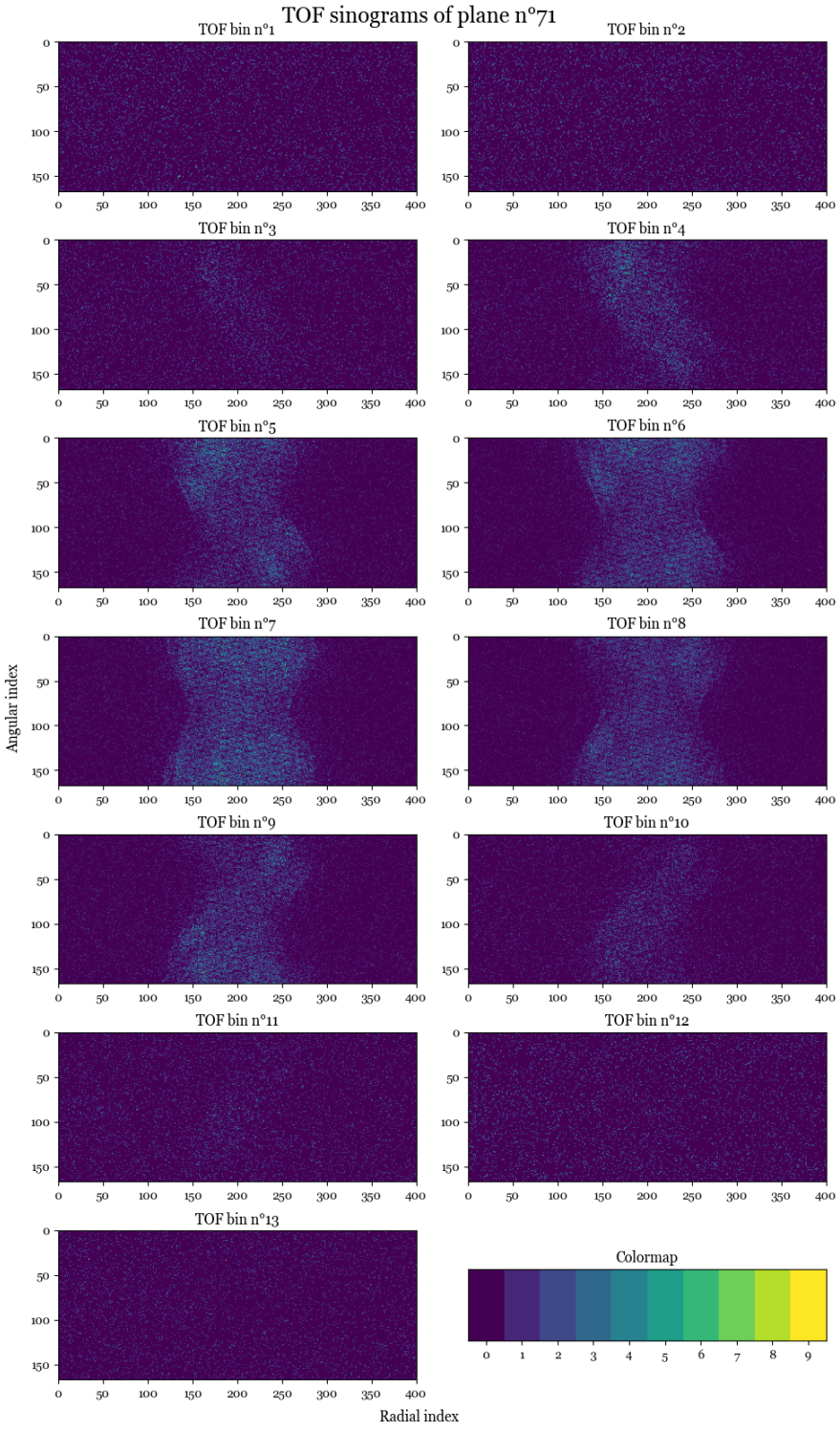


Figure 3.3: Example of the TOF sinograms of a single plane.

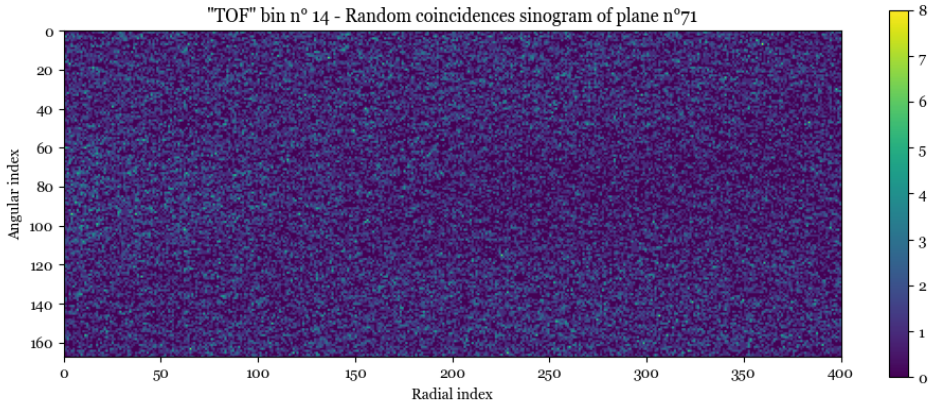


Figure 3.4: Example of the random coincidences sinogram of a single plane.

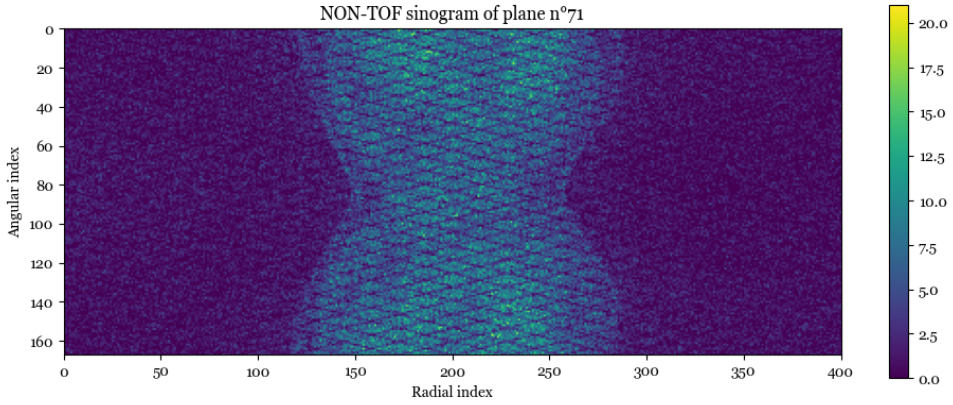


Figure 3.5: Example of a NON-TOF sinogram of a single plane.

3.2.1 e7-tools reconstruction

The proprietary reconstruction starts with the execution of the script `JSRecon12.bat` by dragging on it the folder containing the raw data. This process decompresses the clinical raw data and creates a number of batch files ready to perform the image reconstruction based on the proprietary algorithms. The chosen state-of-the-art reconstruction algorithm is the TOF Ordered Subset Expectation Maximisation (OSEM), a faster implementation of the MLEM algorithm. Instead of computing the forward and back-projections on the whole data space, at each sub-iteration those projections are computed only for a specific subset of views and the subset is changed for the next iteration; this reduces the computational power needed for each sub-iteration and provides an order of magnitude acceleration over MLEM [6].

By looking at the folder containing the list-mode data of a singular bed position, the following files are contained (`XX` stands for the number of bed position, usually between `00` and `05`):

- `PET-RAW-DATA-LM-XX.hdr`
- `PET-RAW-DATA-LM-XX.l`
- `Run-00-PET-RAW-DATA-LM-XX-Histogramming.bat`
- `Run-01-PET-RAW-DATA-LM-XX-Makeumap.bat`
- `Run-04-PET-RAW-DATA-LM-XX-OPTOF.bat`

The `.hdr` file contains the header of the list-mode data and points to the binary data contained inside the `.l` file.

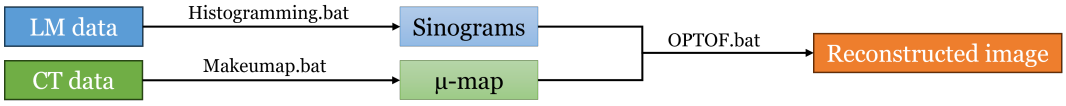


Figure 3.6: e7-tools reconstruction scheme.

The 3 `.bat` files are needed to perform the TOF OSEM reconstruction following the next steps, also shown in Figure 3.6:

1. Execute `Run-00-PET-RAW-DATA-LM-XX-Histogramming.bat` that calculates the sinograms starting from the list-mode data and stores them in the file `PET-RAW-DATA-LM-04-sino-0.s` and its corresponding header `PET-RAW-DATA-LM-04-sino-0.s.hdr`

2. Run the script `Run-01-PET-RAW-DATA-LM-XX-Makeumap.bat` that creates the attenuation map by converting the CT into attenuation factors and registering it in the PET image space.
   After this the files `PET-RAW-DATA-LM-XX-umap_000_000.v`, `PET-RAW-DATA-LM-XX-umap_000_000.v.hdr` and `PET-RAW-DATA-LM-XX-umap.mhdr` will be created, also with the correspondent ones where the bed of the scanner has been removed from the CT (`PET-RAW-DATA-LM-XX-umapBedRemoval_000_000.v`, `PET-RAW-DATA-LM-XX-umapBedRemoval_000_000.v.hdr` and `PET-RAW-DATA-LM-XX-umapBedRemoval.mhdr`)

3. Execute `Run-04-PET-RAW-DATA-LM-XX-OPTOF.bat` that performs the OSEM algorithm and saves the reconstructed images in the files `PET-RAW-DATA-LM-XX-OPTOF_000_000.v`, `PET-RAW-DATA-LM-XX-OPTOF_000_000.v.hdr` and `PET-RAW-DATA-LM-XX-OPTOF.mhdr`

The default settings of the reconstruction algorithm are the following:

- number of OSEM iteration = 3 and number of subsets = 21, for a total number of 63 iterations

- width of the reconstructed image = 200 pixels
- applied Gaussian filter of resolution 5 mm both in trans-axial and axial direction

As a way to obtain images comparable with the other ones, the width of the reconstructed image has been changed to 353 pixels and the Gaussian filter is removed; also, the scatter correction is not performed.

3.2.2 MLEM reconstruction

The implementation of the reconstruction algorithms has been based on Python notebooks with the aid of the SIRF[2] library developed by the CCP SyneRBI[3] and the ODL[4] library created by the Math department of KTH university in Stockholm, under the supervision of Prof. Ozan Öktem.

The SIRF library is the central point in the development of the reconstruction algorithm used in this work because it contains the models of the forward and back-projections of the Radon transform that are defined for the particular geometry of the Siemens mCT scanner.

On the other hand, the Operator Discretization Library (ODL) Python package has been used to wrap those forward and back-projector inside mathematical operators in order to make the implementation of the code much easier.

The just mentioned Python environments have been necessary to implement the chosen reconstruction algorithms and obtain images to compare with e7-tools.

When applying the MLEM reconstruction to clinical data, the formula 2.2 has to be adapted to account for the effect of some physical elements that affect the acquisition data process, those are normalization and attenuation factors and random and scatter coincidences.

[2]Synergistic Image Reconstruction Framework, an open source toolkit for image reconstruction of PET and MRI (`https://github.com/SyneRBI/SIRF`)

[3]Collaborative Computational Project in Synergistic Reconstruction for Biomedical Imaging (`https://www.ccpsynerbi.ac.uk/`)

[4]Operator Discretization Library, a Python library for fast prototyping focusing on inverse problems (`https://github.com/odlgroup/odl`)

The corrected data $\hat{g}$ can be represented with the following model:

$$\hat{g} = UNAf + r + s \tag{3.1}$$

where:

- f is the image and A the forward projector modelling the Radon transform
- U is the inverse of the attenuation correction factors obtained from *e7-tools* and attenuates the projected data based on the CT acquired
- N represents the inverse of the normalization factors derived from *e7-tools* and simulates the sensitivities of the different LOR based on the detectors efficiency and geometry
- r accounts for the random coincidences that are obtained during the acquisition
- s are the scatter coincidences

Since all the scatter estimation algorithms for TOF PET are protected by patents, the scatter coincidences are not accounted for in this analysis.

Keeping this in mind, the MLEM update step then becomes:

$$f^{(n+1)} = \frac{f^{(n)}}{UNA^*\mathbb{1}} A^* \left(\frac{g}{UNAf^{(n)} + r} \right) \tag{3.2}$$

where A^* is the adjoint operator of the Radon transform and $A^*\mathbb{1}$ represents the back-projection of a uniform scan.

The complete structure of the MLEM algorithm for clinical data is displayed in Algorithm 4.

3.2.3 MLAA reconstruction

Similarly to MLEM, when reconstructing PET clinical data with the MLAA algorithm, the model has to account for attenuation and sensitivity factors and also for the random and scatter coincidences.

The MLEM step of the algorithm remains the same as in formula 3.3 but the matrix U is not calculated from the CT, since it is not fixed for all the iterations, and it is computed as follows:

$$U^n = e^{-A(\mu^n)}$$

Algorithm 4 MLEM for clinical data

Require: $n \geq 0$
$A \leftarrow A$ ▷ Radon transform
$A^* \leftarrow A^*$ ▷ adjoint operator of Radon transform
$n \leftarrow MLEM_{iter}$ ▷ nb of MLEM iterations
$f^{(0)}$ ▷ init of activity image
g ▷ projection data
U ▷ attenuation matrix
N ▷ normalization matrix
r ▷ randoms data
for $i \leftarrow 0, n$ **do**
$f^{(i+1)} = \frac{f^{(i)}}{UNA^*\mathbb{1}} A^* \left(\frac{g}{UNAf^{(i)}+r} \right)$
end for

So, for each new estimate of the attenuation map, the MLEM update becomes:

$$f^{(n+1)} = \frac{f^{(n)}}{U^n N A^* \mathbb{1}} A^* \left(\frac{g}{U^n N A f^{(n)} + r} \right) \tag{3.3}$$

Moreover, the MLTR update accounting for the real world situation is the following:

- first, the sinogram of the current activity estimate is simulated accounting for sensitivity and attenuation
$$\psi^n = U^n N A f^{(n+1)}$$
- then, the attenuation map is updated with the formula
$$\mu^{n+1} = \mu^n + \frac{\frac{\psi^n}{\psi^n + r} \left(\psi^n + r - y \right)}{\frac{\left(\psi^n\right)^2}{\psi^n + r} A\mathbb{1}} \tag{3.4}$$

The complete structure of the MLAA algorithm for clinical data is displayed in Algorithm 5.

3.3 Proposed reconstruction

After having presented the state-of-the-art reconstruction techniques used in clinical practice for PET imaging, in this section is described the recently developed Morphed Maximum Likelihood Activity and Attenuation (M-MLAA) algorithm, that is proposed as a solution to correct for

Algorithm 5 MLAA for clinical data

Require: $n \geq 0$
Require: $m \geq 0$
Require: $p \geq 0$
  $A \leftarrow A$ ▷ Radon transform
  $A^* \leftarrow A^*$ ▷ adjoint operator of Radon transform
  $n \leftarrow MLAA_{iter}$ ▷ nb of MLAA iterations
  $m \leftarrow MLEM_{iter}$ ▷ nb of MLEM sub-iterations
  $p \leftarrow MLTR_{iter}$ ▷ nb of MLTR sub-iterations
  $f^{(0)}$ ▷ init of activity image
  $\mu^{(0)}$ ▷ init of attenuation image
  g ▷ projection data
  for $i \leftarrow 1, n$ **do**
    $f^{(1)} \leftarrow f^{(i-1)}$
    $\mu^{(1)} \leftarrow \mu^{(i-1)}$
    $U^{(1)} = e^{-A(\mu^{(1)})}$
    for $j \leftarrow 1, m$ **do**
      $f^{(j+1)} = \frac{f^{(j)}}{U^{(1)}NA^*\mathbb{1}} A^* \left(\frac{g}{U^{(1)}NAf^{(j)}+r} \right)$ ▷ MLEM step
    end for
    $f^{(i)} \leftarrow f^{(j)}$
    for $k \leftarrow 1, p$ **do**
      $U^k = e^{-A(\mu^k)}$
      $\psi^k = U^k NAf^{(k+1)}$
      $\mu^{k+1} = \mu^k + \frac{\frac{\psi^k}{\psi^k+r}\left(\psi^k+r-y\right)}{\frac{\left(\psi^k\right)^2}{\psi^k+r}A\mathbb{1}}$ ▷ MLTR step
    end for
    $\mu^{(i)} \leftarrow \mu^{(k)}$
  end for

motion artifacts present during the acquisition of the scan, in particular for the movement associated to breathing.

Since this process relies on the use of data divided into different respiratory gates, the method is divided into two major part:

- The creation of gated TOF sinograms from listmode data.
- The image reconstruction via M-MLAA algorithm.

3.3.1 Reconstruct sinogram from list-mode

The first step needed to be able to divide the acquired data in different motion gates is to reconstruct the sinograms from the list-mode files.

The list-mode data is stored with the Siemens PETLINK [8] format that stores together the events, the time elapsed and other information, each in a 32-bit packet.

After having divided the bytes between the different packets, the ones of interest are selected; they are the *events* - that are marked by having the first bit equal to `0` - and the *time markers* - which have `10` at the start of their binary sequence. The events packets contains both the *prompts* and the *delayed* coincidences that can be separated by looking at the second bit, that is 1 for the prompts and `0` for the delays. The remaining 30 bits stores the 4 sinogram coordinates:

- TOF_{bin} that contains the corresponding TOF bin
- mi that is the number of the michelogram
- tx_{ang} the trans-axial angle of orientation of the LOR
- ro_{bin} the radial distance of the LOR from the centre of the FOV

Those coordinates are then coded together with the formula:

$$event = ro_{bin} + c_1 \times tx_{ang} + c_1 c_2 \times mi + c_1 c_2 c_3 \times TOF_{bin} \tag{3.5}$$

where c_1, c_2 and c_3 are the dimension of the sinogram: c_1 is the number of radial distance bins, c_2 is the number of angular orientation bins and c_3 is the total number of sinograms.

The coordinate are then recovered by inverting Formula 3.5 (where the symbol "//" denotes the integer division) resulting in:

$$TOF_{bin} = event//(c_1\, c_2\, c_3)$$

$$mi = (event - c_1\, c_2\, c_3\, TOF_{bin})\,//(c_2\, c_3)$$

$$tx_{ang} = (event - c_1\, c_2\, c_3\, TOF_{bin} - c_1\, c_2\, mi)\,//c_1$$

$$ro_{bin} = event - c_1\, c_2\, c_3\, TOF_{bin} - c_1\, c_2\, mi - c_1\, tx_{ang}$$

For each prompt packet those 4 parameters are calculated and the number of counts of the corresponding cell in the 4D sinogram is incremented by one.

3.3.2 Gating

The proposed reconstruction algorithm needs the acquisition data to be divided into different gates, each one corresponding to an almost static position of the body. In order to obtain those gates there are two main strategies:

- use the information of the amplitude of the chest movement recorded with the respiratory belt;
- calculate the Center of Mass (COM) for each group of consecutive coincidence events.

Physio-file gating

The physiological file contains the data registered with a respiratory belt that can be summarize as a table that associates an amplitude to each time instant of the acquisition. That amplitude is related to the respiration motion and gives a signal with a cyclic pattern between the inspiration and expiration phases as displayed in Figure 3.7.

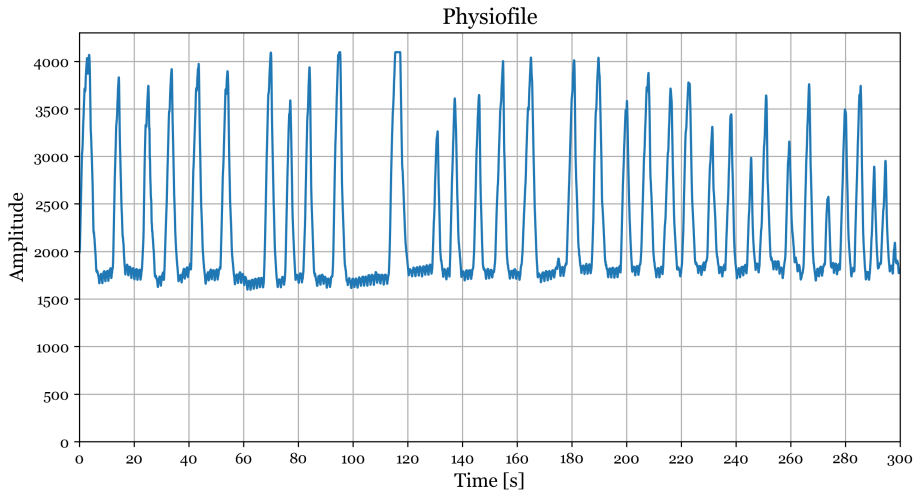


Figure 3.7: Example of a Physio-file correspondent to a single bed position.

The easiest way to use this signal to divide the PET data into different gates is to designate a set of thresholds to each gate and divide the acquisition times between those gates, assigning the corresponding coincidences to their gate. An example of this procedure is shown in Figure 3.8.

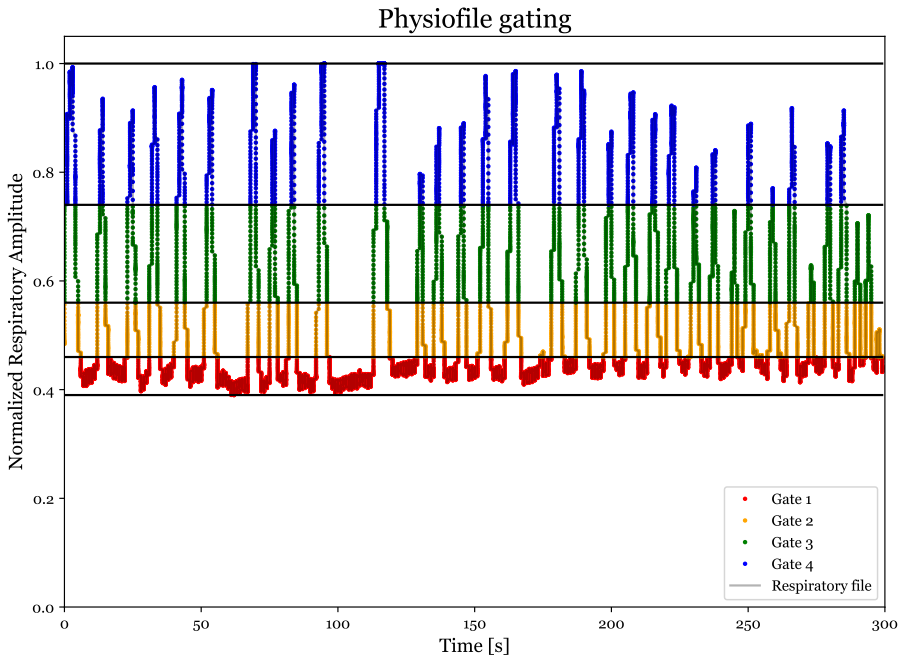


Figure 3.8: Example of gating based on amplitude of respiratory signal.

Center of Mass gating

A second way to divide the data between the gates is to calculate the Center of Mass (COM) of a restricted amount of consequent coincidence events acquired in a specific time window and group together the events which COM are close in the spatial volume.

The Center of Mass software had already been developed by Hugo Linder in his final project "Characterization of Center-of-Mass and Rebinning in Positron Emission Tomography with Motion" [7].

The idea behind this is to calculate the COM of a group of consequent events X as a weighted average position of that set X as shown in Equation

3.6.

$$COM\left(X,p,m\right)=\frac{\sum\limits_{x\in X}m\left(x\right)\cdot p\left(x\right)}{\sum\limits_{x\in X}m\left(x\right)} \tag{3.6}$$

where $m\left(x\right)$ is the mass of event x and $p\left(x\right)$ its position.

Then, those sets X are divided into different gates by grouping together the sets with the closest COM.

3.3.3 M-MLAA reconstruction

The proposed reconstruction algorithm is called M-MLAA, which stands for *Morphed Maximum Likelihood Activity and Attenuation estimate*, that has been developed inside the research team of Prof. Colarieti-Tosti, in particular by PhD exchange student Enza Cece from University of Naples Federico II [3].

This algorithm applies the idea of a morphed version of the MLEM reconstruction already presented in section 2.2.1 to the MLAA algorithm with the objective of obtaining a reconstructed PET image without motion artifacts and without the need of using an attenuation image coming from a different source.

The algorithm can be summed up in the following steps:

1. An initial number m of MLAA iteration is performed on each single gate, obtaining the initial estimates for the activity and attenuation images.

2. The initial deformation between the gates W_i are computed using SynthMorph Neural Network.

3. For each of the n iterations of M-MLAA:

   (a) The new f_0^{k+1} activity estimate for gate 0 is computed by using its previous estimate f_0^k and the contribute from the other gates deformed onto the first one with the M-MLEM formula:

$$f_0^{k+1}\leftarrow f_0^k\frac{\sum\limits_{j=1}^{N}A_{j-1}^{*}\left[e^{-A_j(\mu_{CT-0})}\frac{g_j}{e^{-A_j(\mu_{CT-0})}A_j(f_0^k)}\right]}{\sum\limits_{j=1}^{N}A_{j-1}^{*}\left[\mathbb{1}\cdot e^{-A_j(\mu_{CT-0})}\right]} \tag{3.7}$$

(b) The μ_0^{k+1} attenuation estimate for gate 0 is updated in a similar way as for the activity, by using the Morphed Maximum Likelihood for Transmission Tomography (M-MLTR) formula:

$$\mu_0^{k+1} \leftarrow \mu_0^k + \frac{\sum_{j=1}^{N} A_{j-1}^* \left[e^{-A_j(\mu_0^k)} A_j(f_0^k) \right]}{\sum_{j=1}^{N} A_{j-1}^* \left[e^{-A_j(\mu_0^k)} \cdot A_j(\mathbb{1}) \right]} \tag{3.8}$$

4. The algorithm ends when the maximum number of iterations is reached.

The complete structure of the M-MLAA algorithm is displayed in Algorithm 6.

3.3.4 Image quality evaluation

For the purpose of assessing the quality of the reconstructed image the easiest and only way used in this work for clinical data is via visual inspection, since the images are somehow complicated to be judged via objective measurements. Nevertheless, in order to have some data to support the personal inferences on the quality of the images, the following Figure of Merit (FOM) have been evaluated for the analysis on synthetic data: Peak Signal-to-Noise Ratio (PSNR), Recovery Coefficient (RC) and Signal Difference to Noise Ratio (SDNR).

In the case of the synthetic data, I_1 is defined as the original phantom and I_2 as the reconstructed image from noisy and attenuated sinogram, then Ω_l is the ROI around the lesion and Ω_b is the ROI within the background. Thus, FOMs are defined as follows:

- Peak Signal-to-Noise Ratio (PSNR):

$$PSNR = 10\, log_{10}\left(\frac{d^2}{MSE}\right) \tag{3.9}$$

Here, d is the maximum fluctuation in the input image data type and MSE is the Mean Squared Error, computed as follows:

$$MSE = \sum_{M,N} \frac{[I_1(m,n) - I_2(m,n)]^2}{M \cdot N} \tag{3.10}$$

where M and N are the dimensions of the image.

Algorithm 6 M-MLAA

Require: $n \geq 0$
Require: $m \geq 0$
Require: $N \geq 1$
$A \leftarrow A$ ▷ Radon transform
$A^* \leftarrow A^*$ ▷ adjoint operator of Radon transform
$\theta \leftarrow \theta$ ▷ learned parameters of SynthMorph network
$m \leftarrow MLAA_{iter}$ ▷ nb of MLAA iterations
$p \leftarrow M - MLAA_{iter}$ ▷ nb of M-MLAA iterations
$N \leftarrow nb_{gates} + 1$ ▷ nb total of gates
$f_0, \ldots, f_N \leftarrow 1 \ldots 1$ ▷ init of the distribution estimation
$g_0, \ldots, g_N$ ▷ data for every gates
$W_0, \ldots, W_N \leftarrow Id, \ldots, Id$ ▷ init registration fields
for $i \leftarrow 0, N$ **do**
for $k \leftarrow 1, n$ **do**

$$f_i^{k+1} \leftarrow f_i^k \frac{A^*\left[e^{-A(\mu_i^k)} \frac{g_i}{\left[e^{-A(\mu_i^k)} A(f_i^k)\right]}\right]}{A^*\left[\mathbb{1} \cdot e^{-A(\mu_i^k)}\right]}$$ ▷ MLAA iteration: MLEM step

$$\mu_i^{k+1} \leftarrow \mu_i^k + \frac{A^*\left[e^{-A(\mu_i^k)} A(f_i^k)\right]}{A^*\left[e^{-A(\mu_i^k)} \cdot A(\mathbb{1})\right]}$$ ▷ MLAA iteration: MLTR step

end for
if $i > 0$ **then**
$W_{i,i-1} \leftarrow H(f_{i-1}, f_i, \theta)$ ▷ vector field between two close gates
$W_{i,i-1} \leftarrow H(\mu_{i-1}, \mu_i, \theta)$ ▷ vector field between two close gates
$W_{i-1,i} \leftarrow -W_{i,i-1}$ ▷ approximate estimation of inverse
$W_i \leftarrow W_{i,i-1} \circ \ldots \circ W_{1,0}$
$W_i^{-1} \leftarrow W_{0,1} \circ \ldots \circ W_{i-1,i}$
end if
end for
for $k \leftarrow 1, m$ **do**

$$f_0^{k+1} \leftarrow f_0^k \frac{\sum_{j=1}^{N} A_{j-1}^*\left[e^{-A_j(\mu_{CT-0})} \frac{g_j}{e^{-A_j(\mu_{CT-0})} A_j(f_0^k)}\right]}{\sum_{j=1}^{N} A_{j-1}^*\left[\mathbb{1} \cdot e^{-A_j(\mu_{CT-0})}\right]}$$ ▷ M-MLAA iteration:

M-MLEM step

$$\mu_0^{k+1} \leftarrow \mu_0^k + \frac{\sum_{j=1}^{N} A_{j-1}^*\left[e^{-A_j(\mu_0^k)} A_j(f_0^k)\right]}{\sum_{j=1}^{N} A_{j-1}^*\left[e^{-A_j(\mu_0^k)} \cdot A_j(\mathbb{1})\right]}$$ ▷ M-MLAA iteration: M-MLTR step

end for

- Recovery Coefficient (RC):

$$RC = \frac{\sum\limits_{\Omega_l} I_2}{\sum\limits_{\Omega_l} I_1} \tag{3.11}$$

  and it represents the ratio of observed to true activity in a PET image [21].

- Signal Difference to Noise Ratio (SDNR):

$$SDNR = \frac{\frac{1}{N_{pixel_l}} \sum\limits_{\Omega_l} I_2 - \frac{1}{N_{pixel_b}} \sum\limits_{\Omega_b} I_1}{\sqrt{\sigma\left(I_1\left(\Omega_l\right)\right) + \sigma\left(I_1\left(\Omega_b\right)\right)}} \tag{3.12}$$

  where N_{pixel_l} and N_{pixel_b} are the number of pixels of the Region of Interest (ROI) selected l and the background b,

  Here, the numerator is the difference between the mean of the reconstructed lesion and the mean of the reconstructed background; the denominator is computed as the square root of the sum of the standard deviations σ of the original lesion and original background.

4 Results

In this chapter, the results obtained from evaluating various reconstruction algorithms using synthetic and clinical data are presented.

The chapter begins by discussing the outcomes of simulations conducted on synthetic data using the XCAT 4D phantom. The following algorithms for reconstructing the XCAT phantom have been evaluated: Maximum Likelihood Expectation Maximisation (MLEM), Maximum Likelihood Activity and Attenuation (MLAA), Morphed Maximum Likelihood Expectation Maximisation (M-MLEM) and Morphed Maximum Likelihood Activity and Attenuation (M-MLAA).

Next, the results obtained on clinical data are shown for the MLEM and MLAA algorithms in comparison to the reconstructed image obtained with the e7-tools software.

In the last section, the differences between attenuation and normalization correction factors obtained from e7-tools and the SIRF library are analyzed. Some visualizations of these differences in various planes are provided to identify potential discrepancies that may impact the performance of clinical data reconstruction.

4.1 Simulations on synthetic data

Before reporting the results obtained on clinical data, here are presented the simulations run on synthetic data using the XCAT 4D phantom [16].

The algorithms that were evaluated in the reconstruction of the XCAT phantom were the following:

- MLEM, state-of-the art in clinical data reconstruction, with the use of a simulated CT scan to correct for the attenuation.
- MLAA, to jointly reconstruct activity and attenuation images.
- M-MLEM, to perform motion correction with the use of gated data and a deformed attenuation image by applying SynthMorph to the CT image and the different gate reconstructions.
- M-MLAA, to perform motion correction without the use of the CT scan.

The reconstructions shown are the ones corresponding to the number of iteration that gives the maximum PSNR as later shown in Section 4.1.6.

The reconstructed images all are 3D matrices of dimensions $109 \times 345 \times 345$, where 109 is the number of transversal slices reconstructed and 345×345 is the transversal plane resolution, obtained directly with the SIRF library.

4.1.1 XCAT phantom

The XCAT 4D extended cardiac-torso phantom is a multi-modality imaging phantom developed to produce realistic imaging data when combined with accurate models of the imaging process [16].

In particular, it can simulate respiratory motion in the form of multiple gates, in this case it has been decided to divide the respiratory movement from inspiration to expiration in 4 different gates, numbered from 0 to 3.

The XCAT phantom defines a label map of the anatomy of a simulated patient and, to obtain the activity and attenuation phantoms, those labels have been mapped with a Python function onto the corresponding values of activity and attenuation, respectively shown in the Figures 4.1 and 4.2. The 3D matrices obtained have been resized to fit the dimension expected by the SIRF library of $109 \times 345 \times 345$ voxels.

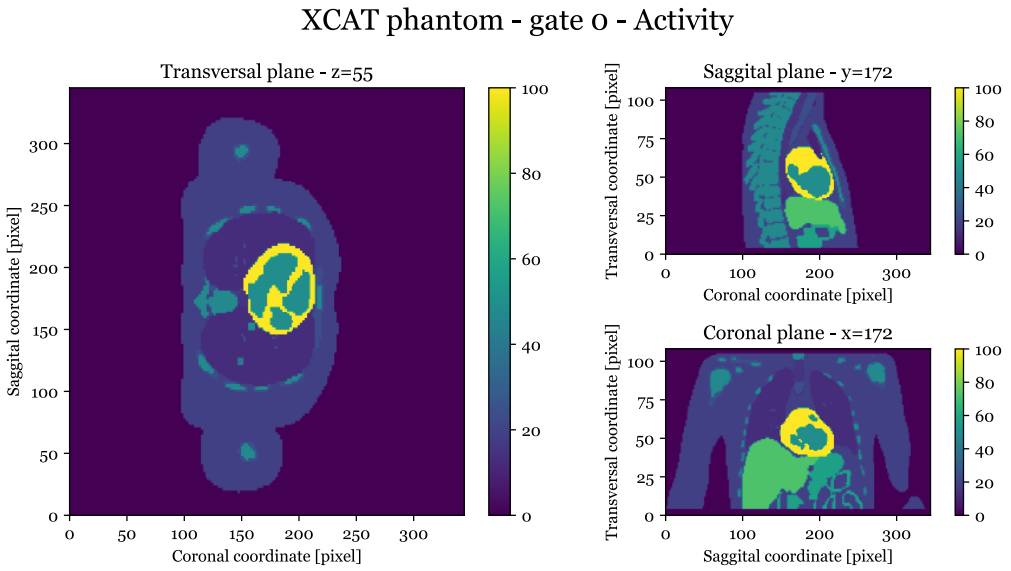


Figure 4.1: XCAT activity phantom.

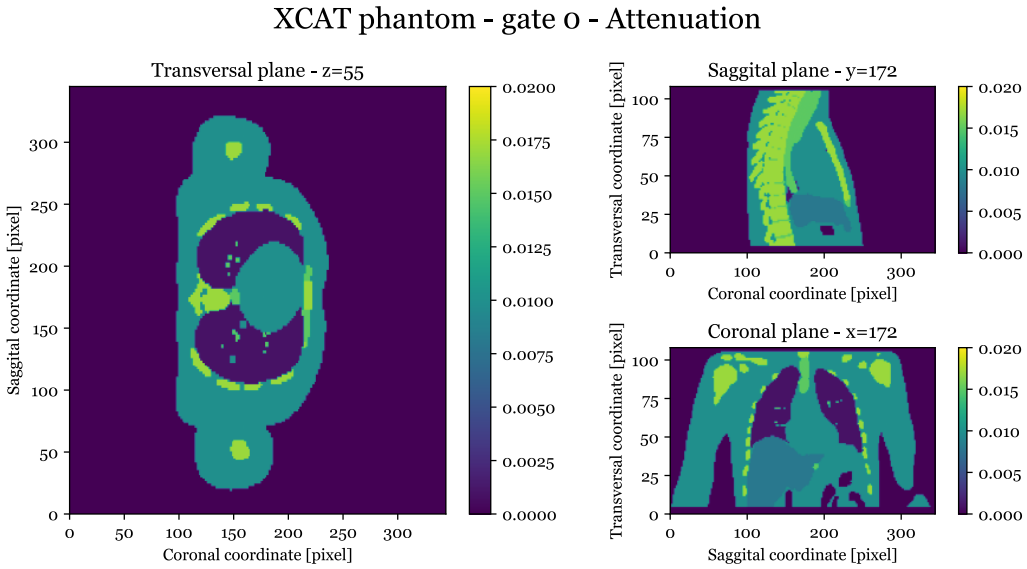


Figure 4.2: XCAT attenuation phantom.

4.1.2 MLEM simulation

The MLEM reconstruction of the phantom was performed by merging the data of the 4 gates obtained by forward projecting the different activity images of the gates, resulting in the correspondent sinogram and summing them together, with the objective of mimicking motion during the acquisition of the data. The attenuation factors come from the simulated CT image

derived from a single gate of the XCAT phantom, to simulate a difference in the respiratory gates between the acquisition of the CT and the PET scans.

The activity image shown in Figure 4.3 was reconstructed after 20 iterations of the MLEM algorithm.

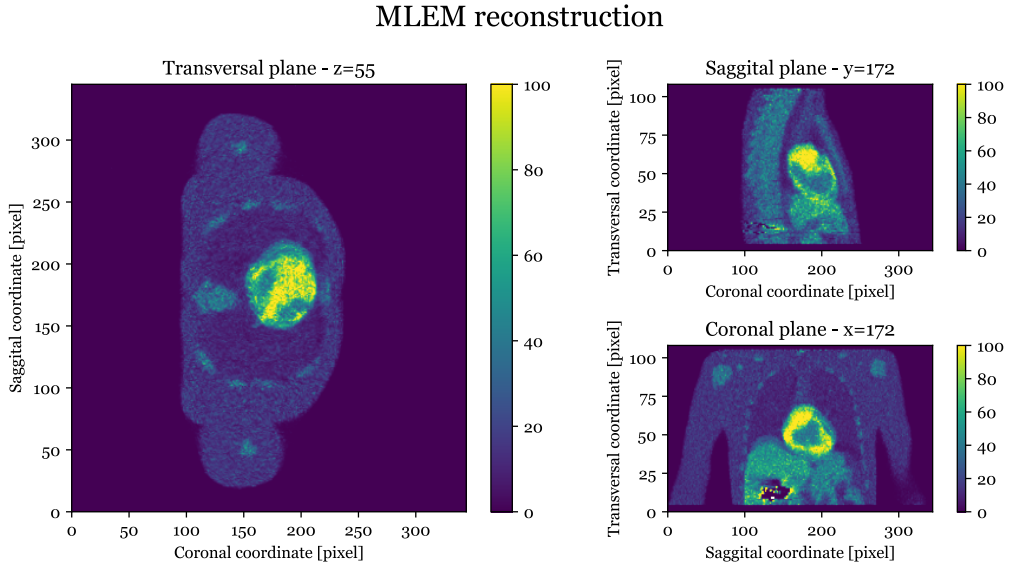


Figure 4.3: MLEM image reconstruction of synthetic data from XCAT phantom.

4.1.3 MLAA simulation

The MLAA reconstruction of the activity and attenuation images of the XCAT phantom uses the same data coming from multiple gates as in the MLEM simulation.

The ratio between the number of MLTR iteration each MLEM update is 5, as recommended from literature [13].

Figure 4.4 shows the activity reconstruction after 5 iterations of the MLAA algorithm and Figure 4.5 displays the reconstructed attenuation image obtained after the same number of total MLAA iterations, that corresponds to 25 MLTR updates for the attenuation.

In order to have a comparison with the best case scenario of absent respiratory movement, the same algorithm has also been applied to data coming from a single gate, gate number 0.

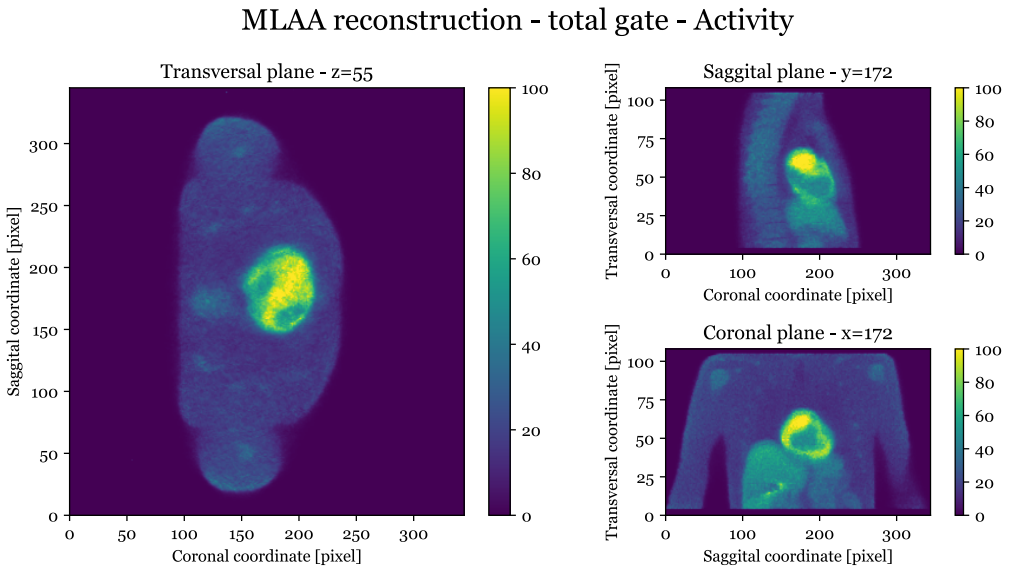


Figure 4.4: MLAA total gates activity reconstruction of synthetic data from XCAT phantom.

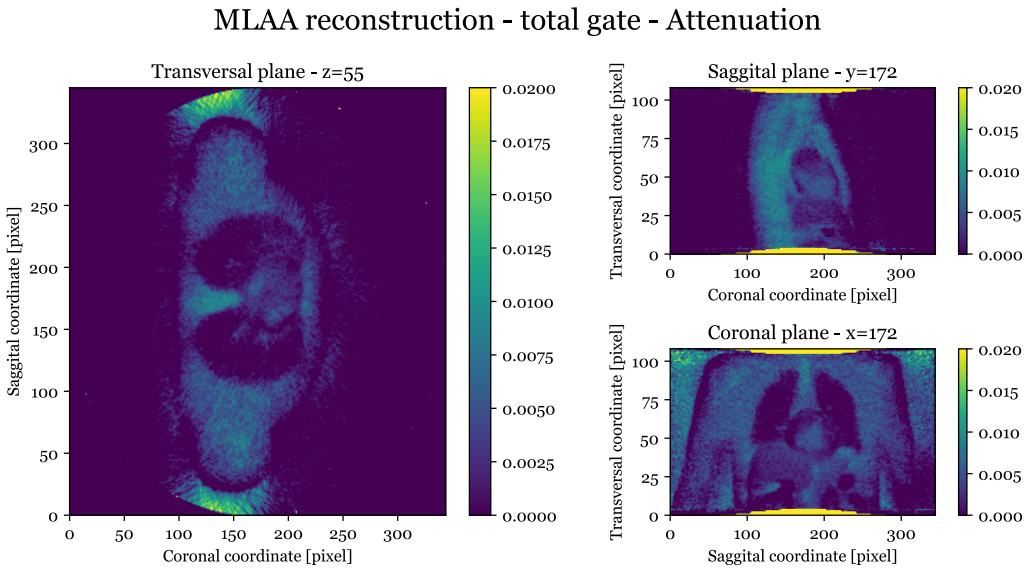


Figure 4.5: MLAA total gates attenuation reconstruction of synthetic data from XCAT phantom.

The number of MLAA iterations for this simulation is also 5 and the reconstructed images are shown in Figure 4.6 for the activity and Figure 4.7 for the attenuation.

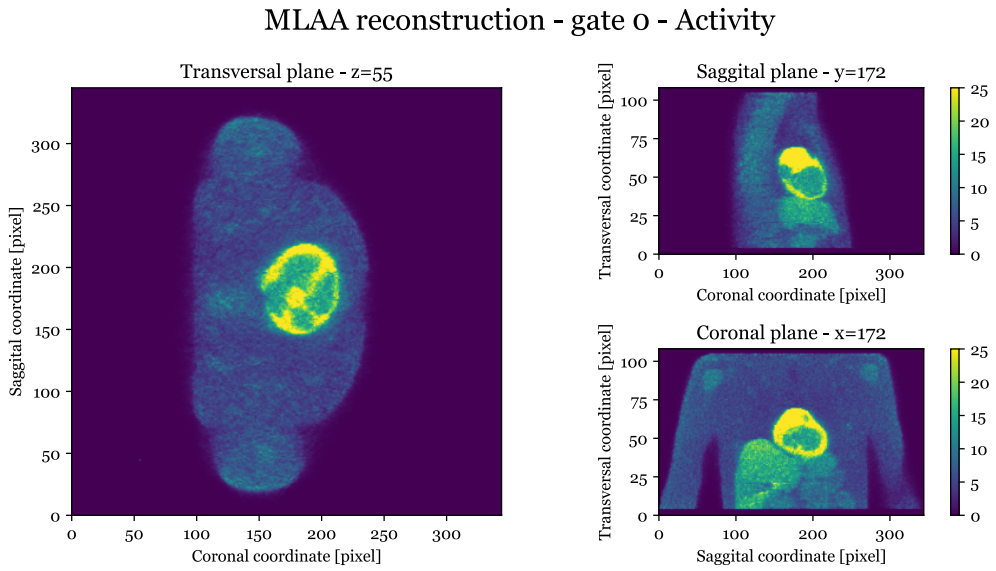


Figure 4.6: MLAA single gate activity reconstruction of synthetic data from XCAT phantom.

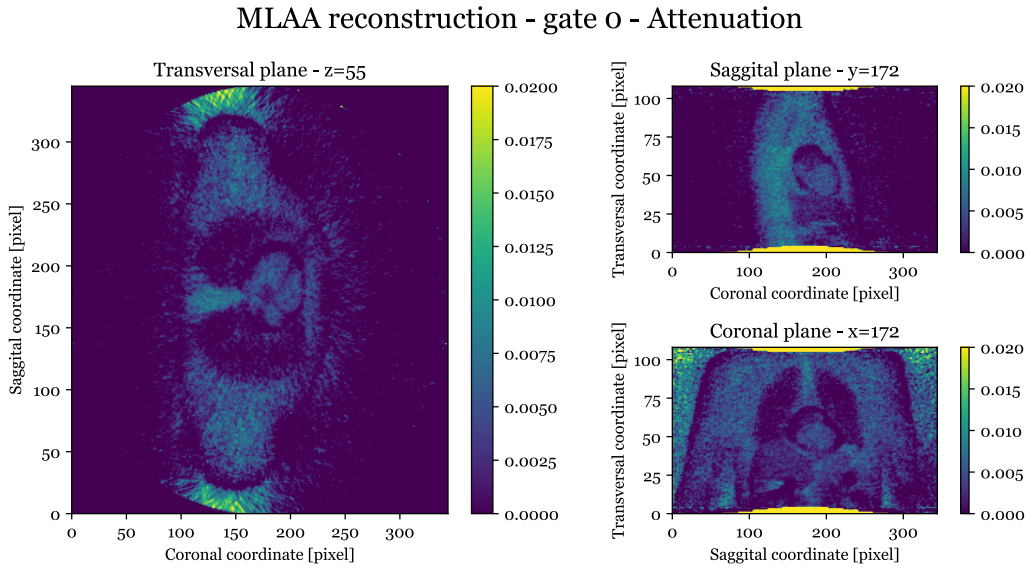


Figure 4.7: MLAA single gate attenuation reconstruction of synthetic data from XCAT phantom.

4.1.4 M-MLEM simulation

The M-MLEM algorithm has been applied to the gated data of the XCAT phantom with the goal of performing motion correction with the use of the information coming from different gates and with a different attenuation

map for each gate by deforming the simulated CT onto the gates.

The reconstructed activity displayed in Figure 4.8 is obtained after 9 iterations of the M-MLEM algorithm.

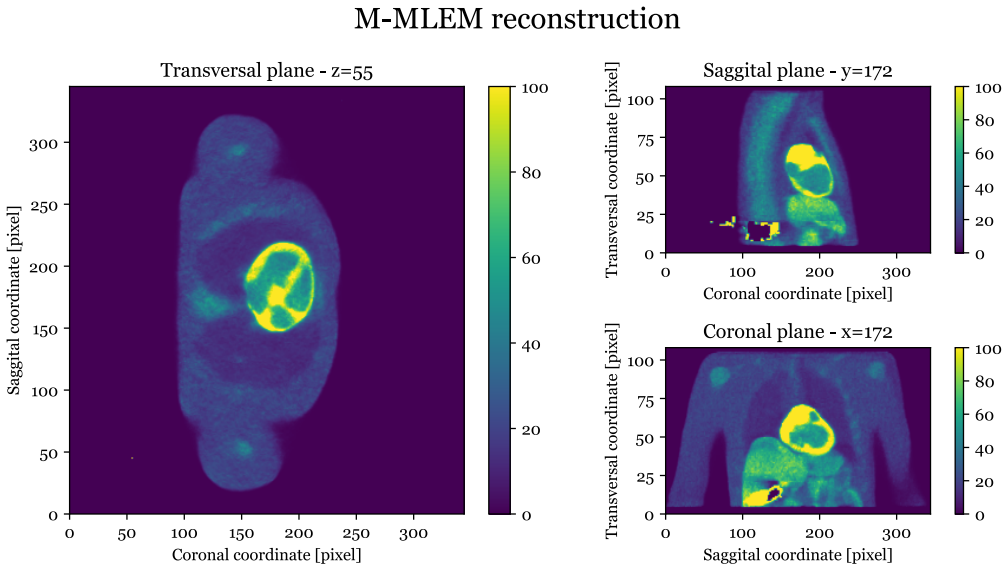


Figure 4.8: M-MLEM image reconstruction of synthetic data from XCAT phantom.

4.1.5 M-MLAA simulation

As for the M-MLEM case, the M-MLAA algorithm also has the objective of removing respiratory motion artifacts during the reconstruction of the images and it also takes advantage of the gated data.

This algorithm returns an estimate both for the activity and the attenuation image, in the particular case of Figure 4.9 and 4.10 the results shown come after 35 total iterations of the M-MLAA algorithm that corresponds to 35 MLEM and 175 MLTR updates, having used the usual 1:5 ratio between the two steps.

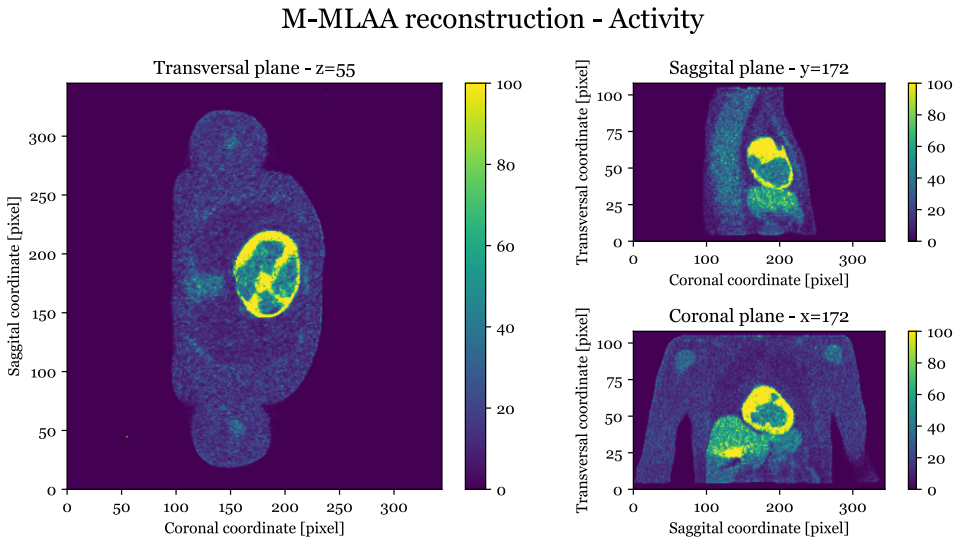


Figure 4.9: M-MLAA activity reconstruction of synthetic data from XCAT phantom.

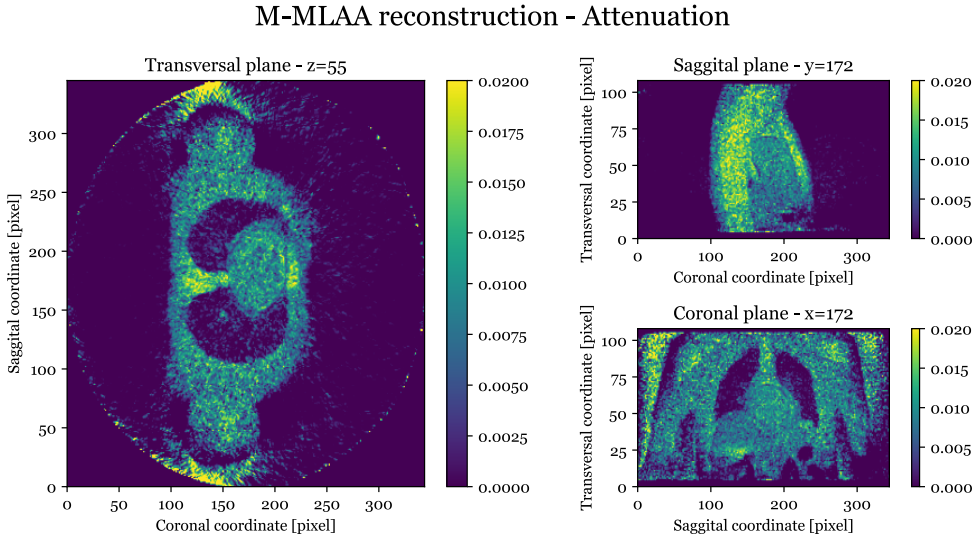


Figure 4.10: M-MLAA attenuation reconstruction of synthetic data from XCAT phantom.

4.1.6 Image quality evaluation

A first visual comparison between the XCAT original phantom and the reconstruction images obtain from different algorithms is displayed in Figure 4.11.

As explained in Section 3.3.4, after visual inspection, the following Figure

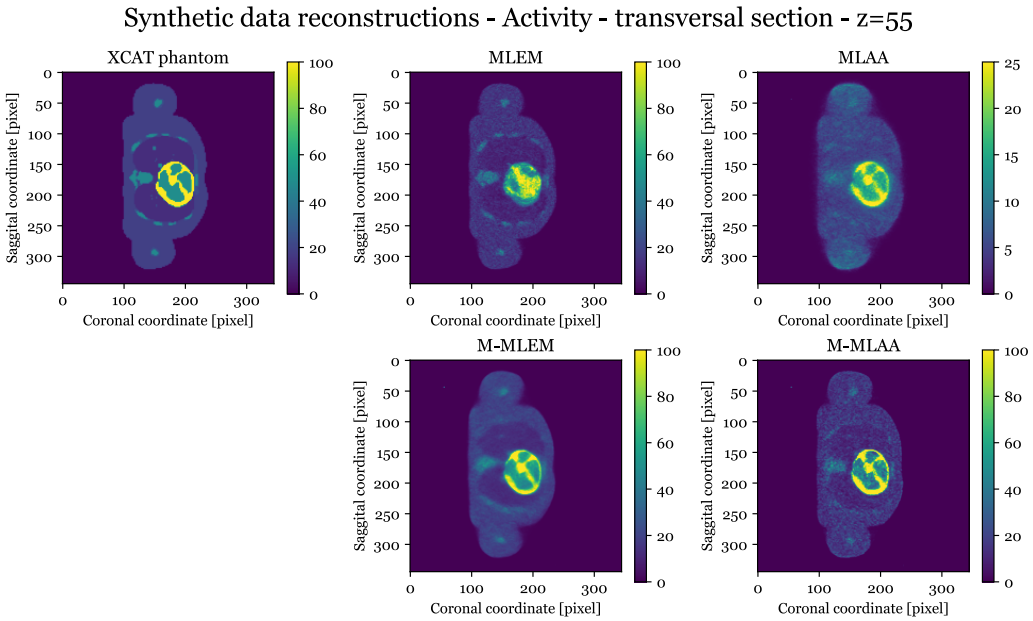


Figure 4.11: Visual comparison between different reconstruction algorithms.

of Merit have been evaluated on the reconstructed images of the XCAT phantom: Peak Signal-to-Noise Ratio (PSNR), Recovery Coefficient (RC) and Signal Difference to Noise Ratio (SDNR).

The PSNR is one of the standard method to evaluate the quality of images and shows in Figure 4.12 that:

- for a very limited amount of iterations run, the MLAA results in the highest value of this FOM;
- between 6 to 18 iterations, the M-MLEM algorithm obtains the highest values of PSNR;
- after 19 iterations, the M-MLAA constantly has the highest values of PSNR;
- the MLEM reconstruction always underperforms in this FOM with respect to the other reconstruction algorithms;
- the morphed versions of the algorithms (M-MLEM and M-MLAA) show a trend of higher PSNR values in comparison to their non-morphed versions (MLEM and MLAA).

Figure 4.13 shows the Recovery Coefficient and the Signal Difference to Noise Ratio evaluated on 5 different Region of Interest. Looking at the RC the main trends are the overall highest values for the MLEM algorithm and an improving trend, following the growth of the number of iteration, for the

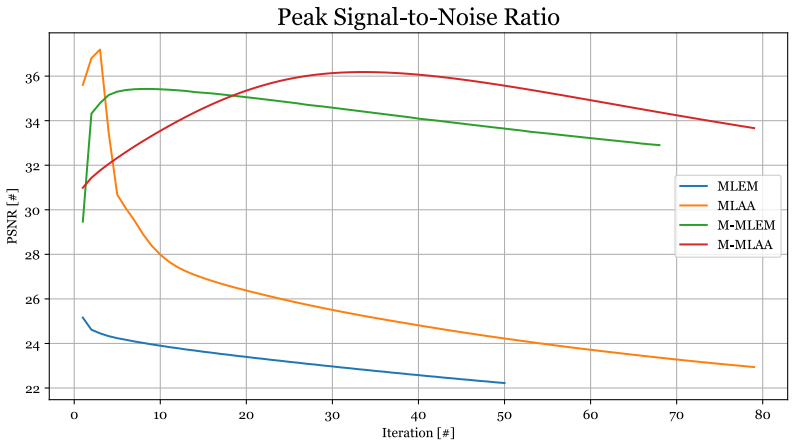
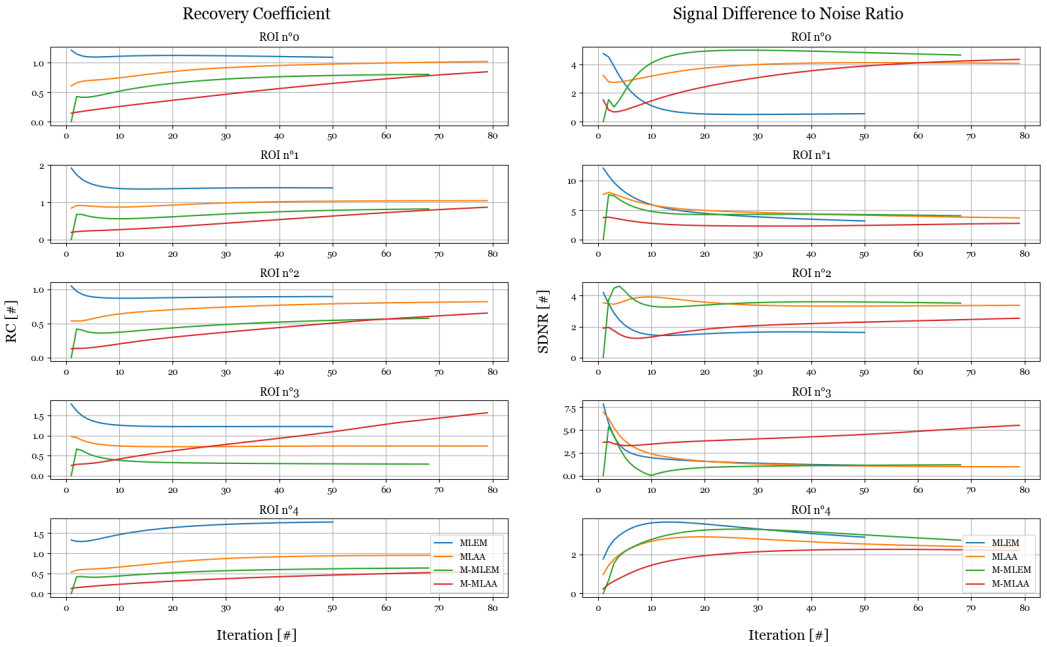


Figure 4.12: Peak Signal-to-Noise Ratio figure of merit between different type of reconstructions.

M-MLAA algorithm, which RC values are close to the M-MLEM and MLAA algorithms, between the different ROIs.

Moreover, the SDNR shows an unstable behaviour between the different ROIs.



(a) (b)

Figure 4.13: Recovery Coefficients (a) and Signal Difference to Noise Ratio (b) Figure of Merit comparison between different types of reconstruction algorithm.

4.2 Clinical data reconstruction

In this section are reported the results obtained from the reconstruction of clinical data acquired with the Biograph128 mCT-1104 PET scanner (Siemens Medical Solutions USA Inc., Hoffman Estate, Illinois, USA, VG60).

The reconstruction techniques tested are the "standard" reconstruction with e7-tools software, MLEM reconstruction and MLAA reconstruction. Due to implementation problems it was not possible to test the M-MLEM and M-MLAA algorithms on clinical data.

4.2.1 e7-tools reconstruction

The state-of-the art clinical reconstruction is achieved with the *e7-tools* reconstruction using the software provided by Siemens.

The output obtained from the software is presented in Figure 4.14 and is reconstructed with the OSEM algorithm on 21 subsets with 3 full iterations.

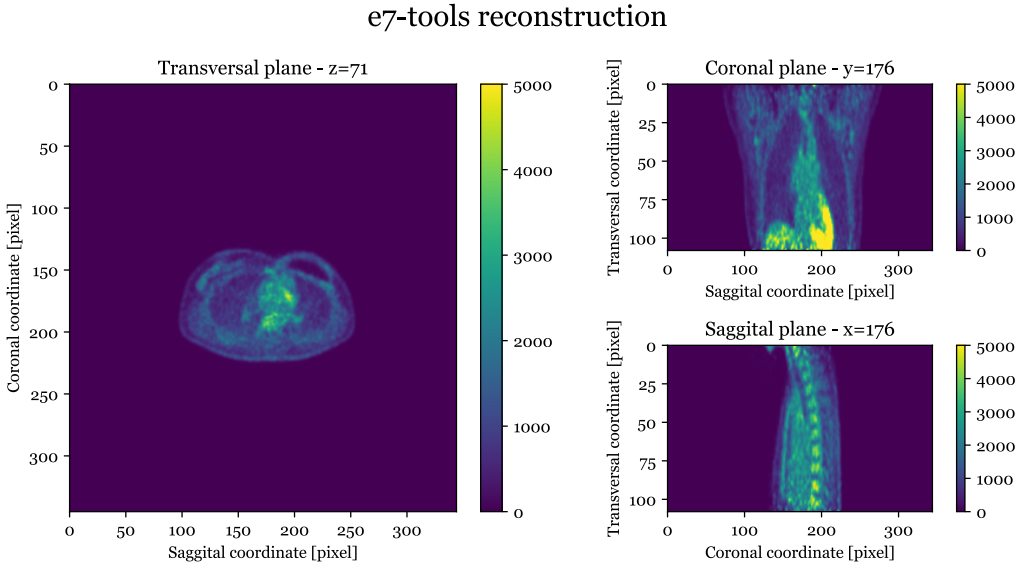


Figure 4.14: e7-tools standard reconstruction, including scatter correction and Gaussian filter smoothing.

To obtain images comparable to the ones reconstructed with the use of the SIRF library, the data has also been reconstructed using the *e7-tools* software without the application of a Gaussian filter and without performing the scatter correction, the results are shown in Figure 4.15.

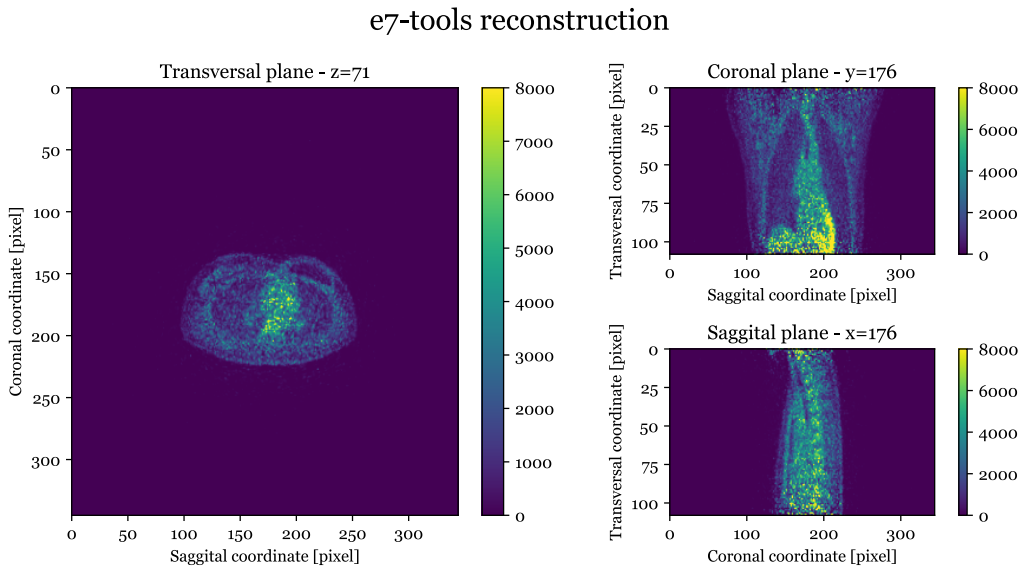


Figure 4.15: e7-tools reconstruction without scatter correction and without Gaussian filter smoothing.

4.2.2 MLEM RECONSTRUCTION

The clinical reconstruction with the MLEM algorithm was performed with the use of the inverse attenuation correction and normalization factors obtained from the *e7-tools* software to guarantee their correct alignment with the data.

The reconstruction shown in Figure 4.16 is obtained after 21 iterations.

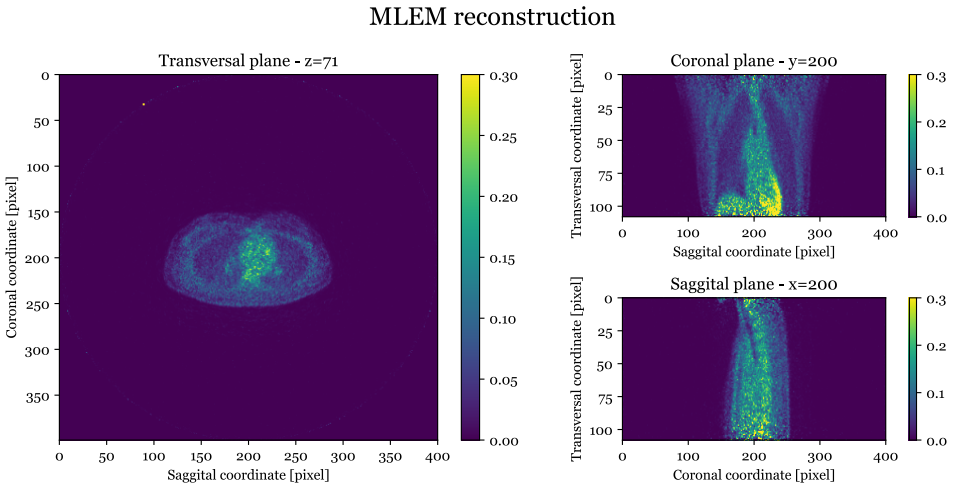


Figure 4.16: MLEM activity reconstruction without Gaussian filter and without scatter correction.

4.2.3 MLAA reconstruction

The MLAA reconstruction results are shown in Figures 4.17 and 4.18 after 10 iterations of the algorithm and in Figures 4.19 and 4.20 after 20 iterations.

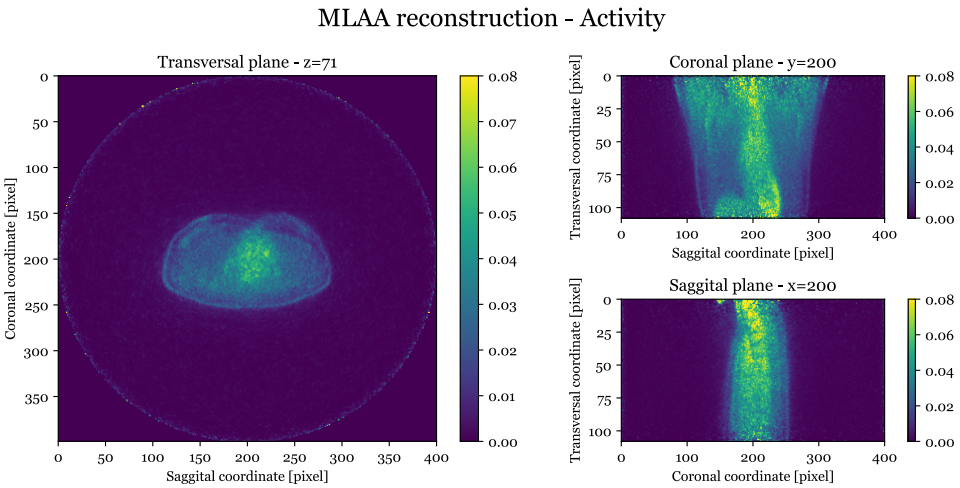


Figure 4.17: MLAA activity reconstruction after 10 iterations.

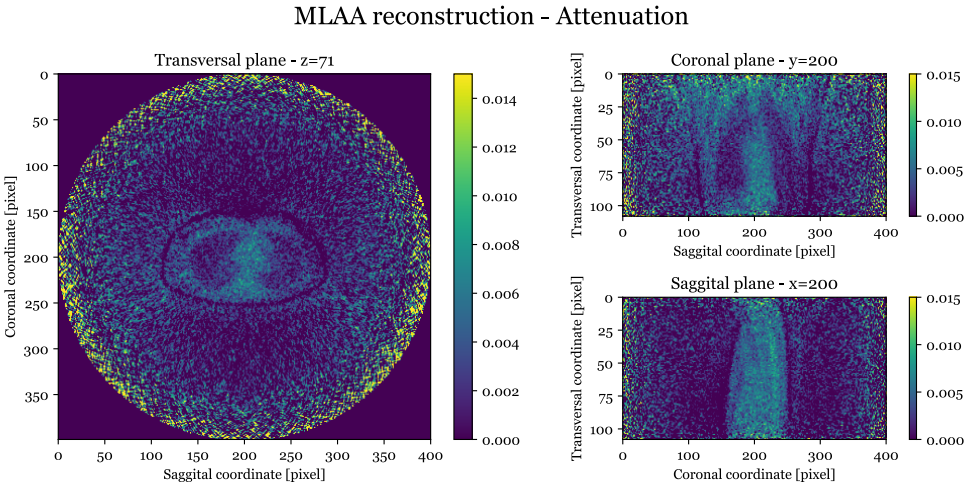


Figure 4.18: MLAA attenuation reconstruction after 10 iterations.

4.2.4 M-MLEM and M-MLAA reconstructions

Due to the problems encountered with the MLAA reconstruction that did not permit to reconstruct the clinical data with the MLAA algorithm, it was also not possible to obtain a reconstructed image with the use of

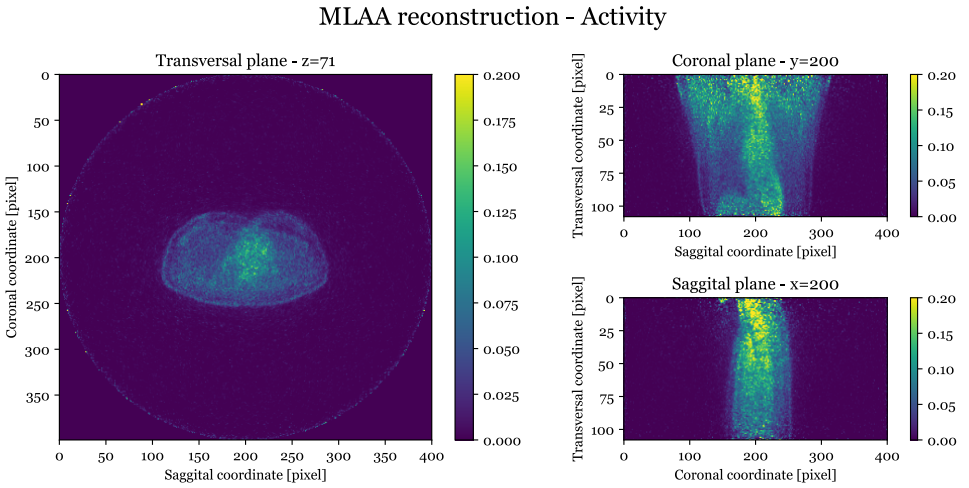


Figure 4.19: MLAA activity reconstruction after 20 iterations.

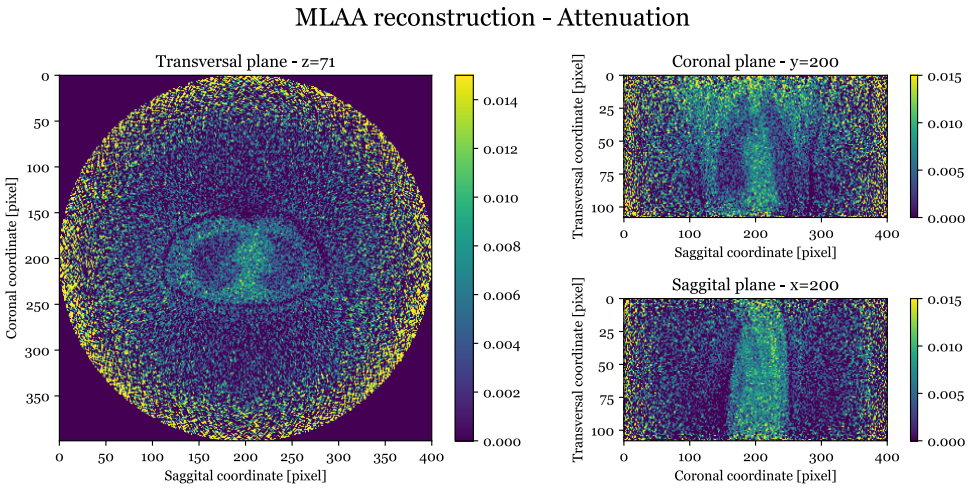


Figure 4.20: MLAA attenuation reconstruction after 20 iterations.

the M-MLAA algorithm and therefore there are no clinical results on that method.

On the other hand, regarding the M-MLEM algorithm, the problems regarding this technique was related to the registration between the CT attenuation image and the gated activity reconstruction after few iterations, that was the first step needed for obtain results with this algorithm. The issue was related to the estimated deformation between the images with SynthMorph that resulted in a too big distortion between the attenuation and activity images that was found not to be true and that would lead to intrinsic problems during the reconstruction process.

4.3 Analysis of SIRF library's projectors

With the aim of trying to understand why the clinical MLAA reconstruction fails, a brief analysis on the differences between the attenuation and normalization factors obtained from the SIRF library is presented in this section.

4.3.1 Attenuation correction factors

Figure 4.21 shows a comparison between the matrices containing the inverse of the attenuation correction factors given from e7-tools and the ones obtained through the use of the SIRF library.

4.3.2 Normalization factors

At the same way as above, Figure 4.22 shows a comparison between the matrices containing the inverse of the normalization factors given from e7-tools and the ones obtained through the use of the SIRF library.

Prior to the analysis, those factors have been normalized to have comparable values.

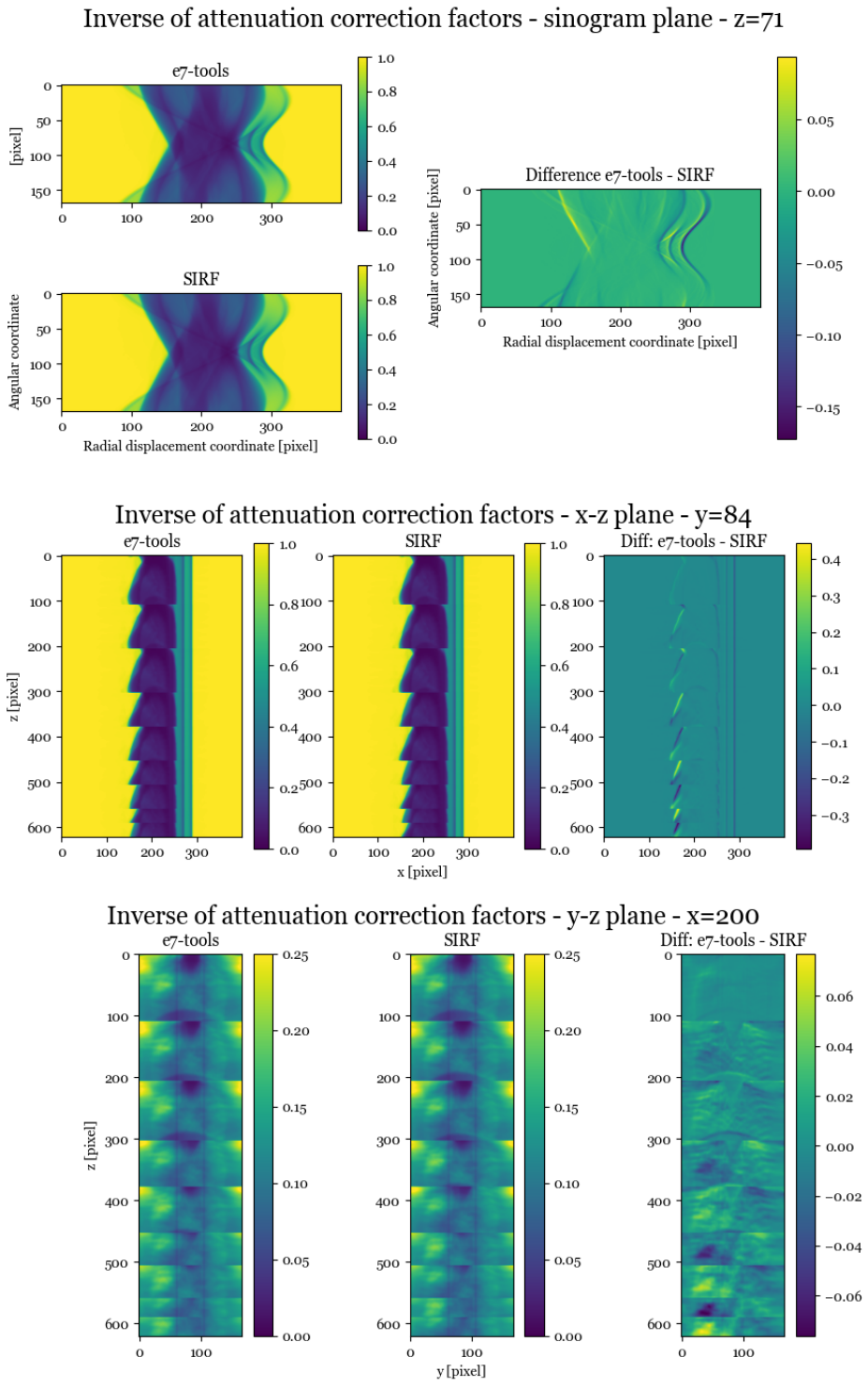


Figure 4.21: Differences between the inverse matrix of attenuation correction factors between e7-tools and SIRF.

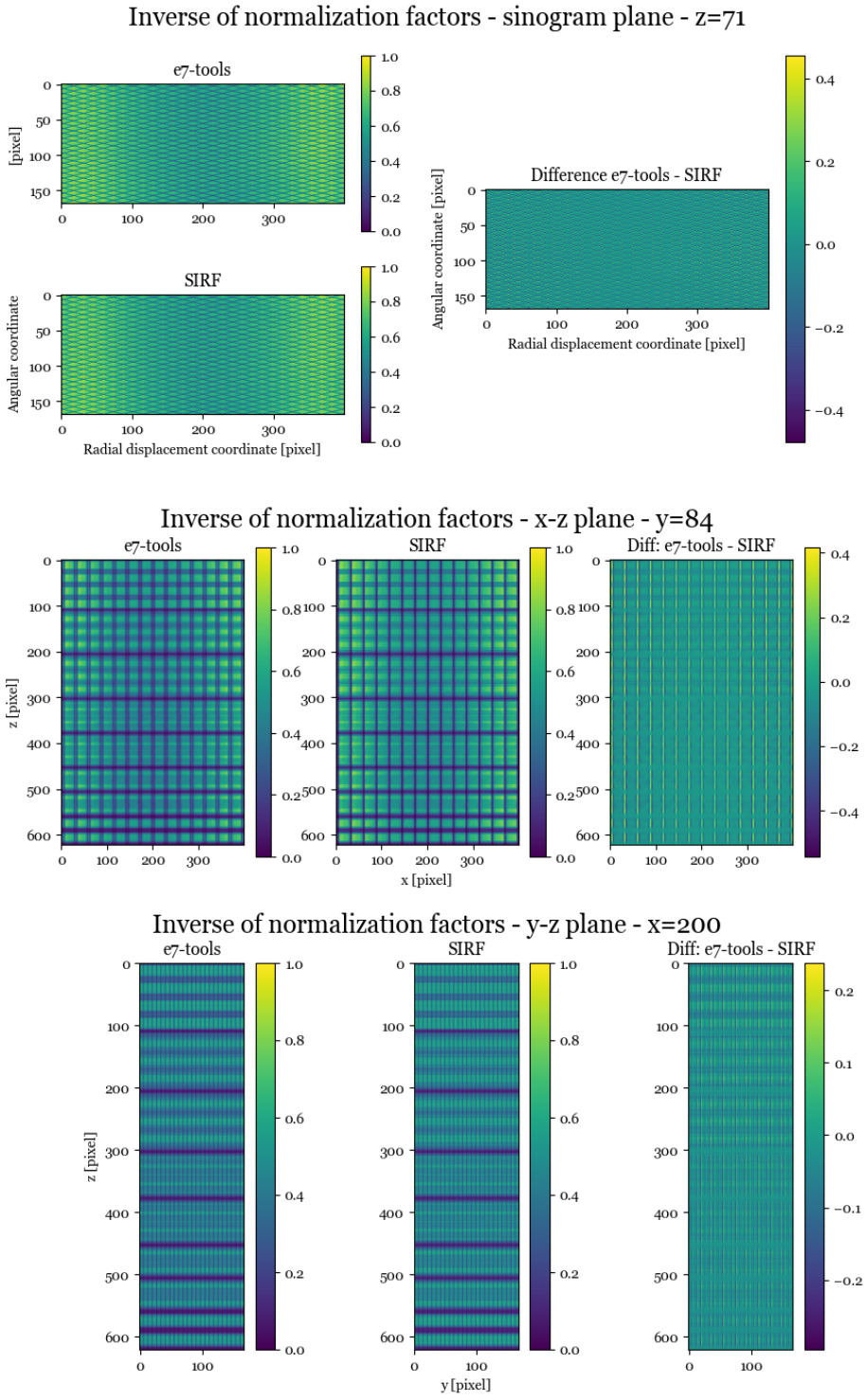


Figure 4.22: Differences between the inverse matrix of normalization factors between e7-tools and SIRF.

5 Discussion

5.1 Synthetic data simulations

Figure 4.11 displays the synthetic data reconstructions obtained with the use of all the algorithms presented; the visual inspection of this figure suggests that the reconstructed image with the M-MLAA algorithm (Figure 4.9) is the most similar to the original phantom (Figure 4.1). The M-MLEM reconstruction (Figure 4.8) also shows good reconstruction, with that reconstructed image being smoother compared to the M-MLAA one, which instead shows higher contrast. On the other hand, both MLEM and MLAA images (respectively Figures 4.3 and 4.4), reconstructed from the combined data of multiple gates, show higher level of noise and blurring that is expected since the do not correct the motion artifacts and this results in lower image quality and contrast.

Peak Signal-to-Noise Ratio (Figure 4.12) confirms visual inspection, so that for a number of iterations that is sufficient to obtain a good reconstruction (≥ 19) the M-MLAA algorithm has better performances than the other algorithms. For a limited number of iterations (< 19), even though other algorithms have higher PSNR values, the reconstructed image will present lower level of spatial resolution, due to the earliness of the reconstruction.

The Recovery Coefficient and Signal Difference to Noise Ratio indexes evaluated on different ROIs corresponding to various lesions in the phantom do not seem to reflect visual inspection and are not coherent for all ROIs as

shown in Figure 4.13. For this reason, the results from visual inspection and PSNR are found to be more significant.

The results obtained through simulations on synthetic data therefore show that M-MLAA improves the quality of the reconstructed image, reaching the goal of motion correction and obtaining better performances than the state-of-the-art reconstruction algorithms MLEM and MLAA and also beats the M-MLEM algorithm for motion correction.

5.2 Clinical data reconstruction

The reconstructions obtained through the Siemens proprietary e7-tools software are at visual inspection better than the ones obtained through MLEM and MLAA reconstruction with the use of the SIRF library.

In particular, comparing the e7-tools reconstructed image (Figure 4.15, without scatter correction and without Gaussian filtering) with the MLEM reconstruction (Figure 4.16), it is noticeable that the second image presents at visual inspection lower definition overall and especially on the body contours and on the spinal column. A possible explanation to this difference could be the higher volume of data used in the e7-tools reconstruction since the SIRF reconstruction could only use 559 out of the 621 sinogram planes due to memory constraints of the library objects. Another possible cause that is presented more thoroughly in Section 5.3 could be a non correct definition of the forward projector of the Radon transform modelling the data acquisition for the scanner used in this project.

Looking at the MLAA reconstructed image displayed in Figures 4.17, 4.18, 4.19 and 4.20, the results are not comparable with the other algorithms and present in fact lower quality in particular in the contrast between the lungs regions that are barely visible in the activity reconstruction. The principal possible cause can be found in the reconstructed attenuation image that starts to have a coherent shape during the first iterations but that stops improving after about 10 iterations, in favour of higher attenuation values put to the limits of the Field of View (FOV) and increasing noise levels. This does not change even with the application of a mask on the body contour and could also be explained by a non correct modelling of the forward and backward operators by SIRF library.

For the reasons presented above the results obtained on clinical data reveal that it was not possible to match the performances obtained through proprietary software *e7-tools* reconstruction probably because of an incorrect modelling of the Radon transform in SIRF library. This also impeded the achievement of reconstructed images with the MLAA, M-MLEM and M-MLAA algorithms, precluding an in-depth performance comparison.

5.3 Critical points for Siemens-mCT scanner in SIRF library

The observation of the different attenuation correction factors and normalization factors showed in Figures 4.21 and 4.22, between the true ones obtained from the e7-tools software and the ones created with the SIRF library, suggests that the library does not model the data acquisition process of the Siemens mCT Biograph scanner exactly as it is supposed to.

In particular, by looking especially at the sinogram plane z in both the figures it is noticeable a slight horizontal and vertical misalignment between the matrices from e7-tools and from SIRF, both in the attenuation correction factors and in the normalization factors. Regarding this last one, the image of the difference between the two shows a distinguishable pattern generated by opposing positive and negative values of the difference that suggests a vertical misalignment between the matrices. A plausible cause of that could be the incorrect modelling of the Radon transform in the SIRF library for the type of scanner used to acquire the data.

It is worth noting that the SIRF library does not fully support the Siemens Biograph mCT scanner and no complete testing on it has been conducted. Also, TOF data acquisition is not officially supported in the major release of the library and it is still work in progress. This two factors could also explain why the Radon transform is not modelled as expected.

6 Conclusions

This work aimed to test the performances of the M-MLAA algorithm for PET image reconstruction on clinical data. Unfortunately, this goal was not achieved because of an unsolved problem in the implementation of the Maximum Likelihood Activity and Attenuation (MLAA) algorithm with the Synergistic Image Reconstruction Framework (SIRF) Python library for the case of real data acquired with Siemens Biograph mCT scanner.

Before the testing on clinical data, four reconstruction algorithms (MLEM, MLAA, M-MLEM and M-MLAA) have been compared with each other on the task of reconstructing synthetic data generated with the use of the XCAT phantom. This analysis showed that visual inspection and Peak Signal-to-Noise Ratio suggest that both M-MLEM and M-MLAA reconstruction techniques are capable to perform motion correction on PET data, with M-MLAA giving higher contrast images and better PSNR values. On the other hand, the reconstructed images obtained from MLEM and MLAA algorithms had lower quality and contrast and presented evident blur due to the lack of motion correction.

Moving onto the trials on clinical data, the reconstruction with MLEM algorithm was possible but it resulted in lower overall image quality when compared with the image reconstructed via proprietary Siemens e7-tools software. In addition, the MLAA algorithm implemented with the SIRF library was not able to converge to a good image reconstruction both for the activity and attenuation images. This impeded the next analysis on

the performance of M-MLAA algorithm on clinical data. The implementation of M-MLEM on real data was also not possible due to problems with the correct registration between the CT attenuation image and the activity reconstructions.

Given the fact that both the MLAA and M-MLAA algorithm worked fine for the reconstruction of synthetic data generated with the forward operators of the SIRF library, the problems encountered during the clinical reconstruction could suggest that the cause behind the non-convergence of the MLAA activity and especially attenuation reconstruction is related to an incorrect definition of the Radon transform for the particular scanner used, also because it is not officially supported by the library. This is confirmed by an analysis on the creation of the matrices for the attenuation and normalization corrections with SIRF that leads to different results than the ones obtained from Siemens proprietary software, where the main dissimilarities lie in a translation between the two, suggesting once more of an incorrect modelling of the Radon transform for the Siemens Biograph mCT scanner in SIRF library.

6.1 Future work

In order to solve the problems affecting the clinical reconstruction with the SIRF library for data acquired with the Siemens Biograph mCT scanner, the first step, that is already in progress, is to get in touch with the developers of the Python library with the aim of defining the correct forward and backward operators matching the specifications of the above-mentioned scanner.

If this would not be possible then, in order to test the M-MLAA algorithm, there would be the need of moving to a different library that implements the Radon transform for PET imaging, one option could be the CASToR[1] platform but it has not yet been evaluated for this purpose.

Furthermore, when a working version of both MLAA and M-MLAA algorithm is found, in order to get better results, those algorithms could be

[1]Customizable and Advanced Software for Tomographic Reconstruction, an open-source multi-platform project for 4D emission (PET and SPECT) and transmission (CT) tomographic reconstruction, `https://castor-project.org/`

improved by making use of all the different bed position acquired during the scan process, merging them together and use the higher volume of information to obtain better image resolution and contrast.

6.2 Final words

Even though this work did not manage to successfully test the Morphed Maximum Likelihood Activity and Attenuation (M-MLAA) algorithm on clinical data, the M-MLAA algorithm shows promising performances on synthetic data and it is capable of performing motion correction during reconstruction, paving the way for improving the quality of PET images and it is expected to contribute to improve the early-stage diagnosis of cancer lesions.

References

[1] Guha Balakrishnan et al. "An Unsupervised Learning Model for Deformable Medical Image Registration". In: *CVPR: Computer Vision and Pattern Recognition* (2018), pp. 9252–9260.

[2] Guha Balakrishnan et al. "VoxelMorph: A Learning Framework for Deformable Medical Image Registration". In: *IEEE TMI: Transactions on Medical Imaging* 38 (8 2019), pp. 1788–1800.

[3] E. Cece and M. Colarieti-Tosti. "Spatiotemporal PET Reconstruction with Attenuation and Motion Correction". 2023.

[4] Frederic H. Fahey. "Data Acquisition in PET Imaging". In: *Journal of Nuclear Medicine Technology* 30.2 (2002), pp. 39–49. ISSN: 0091-4916. eprint: https://tech.snmjournals.org/content/30/2/39.full.pdf. URL: https://tech.snmjournals.org/content/30/2/39.

[5] Malte Hoffmann et al. "SynthMorph: Learning Contrast-Invariant Registration Without Acquired Images". In: *IEEE Transactions on Medical Imaging* 41.3 (2022), pp. 543–558. DOI: 10.1109/TMI.2021.3116879.

[6] H.M. Hudson and R.S. Larkin. "Accelerated image reconstruction using ordered subsets of projection data". In: *IEEE Transactions on Medical Imaging* 13.4 (1994), pp. 601–609. DOI: 10.1109/42.363108.

[7] Linder Hugo. *Characterization of Center-of-Mass and Rebinning in Positron Emission Tomography with Motion*. 2021. URL: https://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-297505.

[8] William F. Jones. *Guideline to the PETLINK Proposal*. URL: https://marketing.webassets.siemens-healthineers.com/1800000005420291/a2683b7d935a/petlink_guideline_j2-05420291_1800000005420291.pdf.

[9] J.M. Martí-Climent et al. "Contribution of time of flight and point spread function modeling to the performance characteristics of the PET/CT Biograph mCT scanner". In: *Revista Española de Medicina Nuclear e Imagen Molecular (English Edition)* 32.1 (2013), pp. 13–21. ISSN: 2253-8089. DOI: https://doi.org/10.1016/j.remnie.2012.11.012. URL: https://www.sciencedirect.com/science/article/pii/S2253808912001541.

[10] Ozan Öktem, Camille Pouchol, and Olivier Verdier. "Spatiotemporal PET Reconstruction Using ML-EM with Learned Diffeomorphic Deformation". In: *Machine Learning for Medical Image Reconstruction*. Ed. by Florian Knoll et al. Cham: Springer International Publishing, 2019, pp. 151–162. ISBN: 978-3-030-33843-5.

[11] Murali Ravi et al. "FPGA as a Hardware Accelerator for Computation Intensive Maximum Likelihood Expectation Maximization Medical Image Reconstruction Algorithm". In: *IEEE Access* PP (Aug. 2019), pp. 1–1. DOI: 10.1109/ACCESS.2019.2932647.

[12] Ahmadreza Rezaei and Johan Nuyts. "Simultaneous reconstruction of the activity image and registration of the CT image in TOF-PET". In: *2013 IEEE Nuclear Science Symposium and Medical Imaging Conference (2013 NSS/MIC)*. 2013, pp. 1–3. DOI: 10.1109/NSSMIC.2013.6829297.

[13] Ahmadreza Rezaei et al. "Simultaneous reconstruction of activity and attenuation in time-of-flight pet". In: *2011 IEEE Nuclear Science Symposium Conference Record* (2011). DOI: 10.1109/nssmic.2011.6153883.

[14] Gopal Saha. "Data Acquisition and Corrections". In: (Jan. 2016). DOI: 10.1007/978-1-4419-0805-6_3.

[15] Dennis R Schaart. "Physics and technology of time-of-flight PET detectors". In: *Physics in Medicine & Biology* 66.9 (Apr. 2021), 09TR01. DOI: 10.1088/1361-6560/abee56. URL: https://dx.doi.org/10.1088/1361-6560/abee56.

[16] W. P. Segars et al. "4D XCAT phantom for multimodality imaging research". In: *Medical Physics* 37.9 (2010), pp. 4902–4915. DOI: https://doi.org/10.1118/1.3480985. eprint: https://aapm.onlinelibrary.wiley.com/doi/pdf/10.1118/1.3480985. URL: https://aapm.onlinelibrary.wiley.com/doi/abs/10.1118/1.3480985.

[17] Y. Shepp L. A.; Vardi. "Maximum Likelihood Reconstruction for Emission Tomography". In: *IEEE Transactions on Medical Imaging 1982-oct vol. 1 iss. 2* 1 (2 Oct. 1982). DOI: 10.1109/tmi.1982.4307558.

[18] Kris Thielemans et al. "Comparison of different methods for data-driven respiratory gating of PET data". In: *2013 IEEE Nuclear Science Symposium and Medical Imaging Conference (2013 NSS/MIC)*. 2013, pp. 1–4. DOI: 10.1109/NSSMIC.2013.6829055.

[19] Shan Tong, Adam M Alessio, and Paul E Kinahan. "Image reconstruction for PET/CT scanners: past achievements and future challenges". In: *Imaging in medicine* 2.5 (2010), p. 529.

[20] Stefaan Vandenberghe et al. "Recent developments in time-of-flight PET". In: *EJNMMI physics* 3 (2016), pp. 1–30.

[21] Xuping Zhu, J.A. Parker, and M.R. Palmer. "Recovery coefficient in PET as a function of object size and respiratory motion trajectory". In: *IEEE Nuclear Science Symposium Conference Record, 2005*. Vol. 4. 2005, pp. 2170–2172. DOI: 10.1109/NSSMIC.2005.1596764.

Acknowledgments

I would first like to thank my thesis advisor Prof. Massimiliano Colarieti-Tosti at KTH Royal Institute of Technology. Prof. Colarieti-Tosti was my leading light in the five months that I spent in his research group in Stockholm, and I thank him for always believing in me.

I would also like to thank all the members of the above-mentioned research group: Nina, Enza, Aurélie, Hamidreza and Lena. They were an important support during the work on the project and were always available to assist me.

I would like to express the greatest gratitude to my supervisor Prof. Mattia Veronese who was always willing to help during the making of this work and encouraged me to arrive at the end of this part of my academic path and continue in the research environment.

Above all I must thank with all heart my family, my parents Cinzia and Nicola and my sister Anna, that always provided me with unfailing support and continuous encouragement throughout my years of study.

I would also like to thank my long-time friends from high school, that were always by my side, even though separated from distance.

Next, I cannot forget to thank my university colleagues with whom I shared the most beautiful years of my life, in particular I send a hug to Erik, Alessia and Costanza.

Finally, I don't know how to properly thank "Lo Magnifico Granducato" that was like a second family for me and with which I've shared many unforgettable memories of the years in Padua, a huge thanks goes to Enrico, Edoardo and Saverio.