www.editada.org

International Journal of Combinatorial Optimization Problems and Informatics, 11(1), Jan-April 2020, 9-25. ISSN: 2007-1558.

Dictionary-based super resolution for positron emission tomography images

Leandro José Rodríguez Hernández [1], Humberto de Jesús Ochoa Domínguez [2]
[1]Departamento de Ingeniería Industrial y Manufactura
[2]Departamento de Ingeniería Eléctrica y Computación
Universidad Autónoma de Ciudad Juárez, Ciudad Juárez, México
leandrojose87@gmail.com

Abstract. In this paper, a strategy to increase the resolution of positron emission tomography (PET) images, using a previously trained high resolution dictionary for the sinograms is proposed. The low resolution input sinogram is divided into patches of 5x5 samples. The sparse code of each patch is calculated and applied to the high resolution dictionary to obtain the best high resolution patch. The estimated high resolution sinogram is processed by the filtered backprojection (FBP) or by the ordered subsets expectation maximization (OSEM) reconstruction algorithm. Results show that, in both cases, the dictionary method outperforms the bicubic interpolation method by more 3% in PSNR. OSEM algorithm yields even better results than the FBP algorithm. However, the reconstruction time is exacerbated.

Keywords: Positron emission tomography, super resolution, sinogram, OSEM, FBP.

Article Info
Received Sep 11, 2018
Accepted Sep 11, 2019

1 Introduction

Medical imaging is the set of techniques used to inspect the human body, in order to diagnose, monitor or treat medical conditions. Medical images provide information about the areas of the body under study [1]. Most of the time, medical images are acquired with a poor resolution that needs to be increased or super resolved, in order to increase their clarity and to obtain defined details of the texture and edges of the structures inside the images.

Images can be divided into high resolution (HR) and low resolution (LR) [2]. HR is required in several areas such as medicine [3], [4] because they provide clarity and more defined details of the texture and edges of an object, which facilitates diagnosis. The increase in resolution can be obtained through improvements in the acquisition hardware of the equipment, which implies a high cost. In addition, the hardware has physical limits [2]. The devices, for image acquisition, use sensors and the resolution of the acquired images is directly related to the size of sensors. Resolution refers to the number of pixels per unit area (pixels/cm^2, pixels/mm^2, etc.) in an image. Therefore, it is also related to the number of sensors per unit area [2]. A promising and cheaper option is to use image processing techniques [4], through which the quality of LR images can be increased. This set of techniques is known as super resolution (SR). The SR is a mathematical term for a set of methods of upscaling images and video sequences. We will refer to a low resolution (LR) image as the image acquired with a positron emission tomography (PET) scanner and the high resolution image (HR) as the image whose amount of details per unit area is larger.

The PET acquires representative events of the metabolic activity of the human body. The patient is injected with a radio marker, which is a radioactive substance, which decays by radiation (β +) emitting a neutron, a positron and a neutrino [3], [4]. The positron is annihilated with an electron producing two gamma photons of high energy, traveling in opposite directions that are counted when reaching the scanner detectors within a time window. The point of annihilation is, ideally, located in a straight line

joining the two points, called the line of response (LOR) [3], [4]. If the data is acquired with a sufficiently large number of angles around the patient, these can be organized to form a set of projections called sinogram and the reconstruction can be done by filtered backprojection (FBP) method or iterative methods [3], [4] such as the ordered subset expectation maximization (OSEM).

The SR of images was proposed for the first time by Tsai et al. in 1984 [5]. Since then, a variety of SR algorithms have been proposed to solve this highly ill-posed problem. These algorithms can be classified mainly into three categories: interpolation-based, reconstruction-based or a-priory-based and dictionary-based, also known as learning-based methods.

The interpolation-based SR [6] is considered as a set of basic and simple methods. These techniques are used to estimate the unknown pixels and to obtain a high resolution image. The reconstruction-based SR [7], [8], [9], assumes that the LR image can be obtained from the HR image by a series of impairments, such as blurring, downsampling and noise (normally considered additive Gaussian noise of zero mean) with this assumption, the SR uses different types of prior knowledge to improve the LR image. Compared to the interpolation-based SR, this family of methods has the ability to filter edges and to suppress artifacts efficiently. However, images cannot be restored with true or optimal new details, especially when large scale factors are used.

The third category is learning-based SR [10], [11], [12], [13]. These methods capture prior knowledge of training data sets to learn the underlying mapping relationships between LR images and HR images. Such a mapping relationship is subsequently used to estimate new HR images from the LR entry. According to the mapping formulations, learning-based SR includes neighborhood-based, regression-based, and dispersion-based methods. The methods often require to search for a vast set of reference data for similar patterns.

In [14], a learning-based strategy through a Markov random field (MRF) is proposed and solved by the belief propagation method. In [15], the local and linear embedding (LLE) method was proposed based on the SR of embedded neighbors (NE). Model-based regression establishes the direct mapping relationships between the corresponding LR and HR images [16], [17]. However, the quality of the final estimate depends on a large extent on the number of neighbors and the qualities and number of prototypes. When the number of neighbors decreases, the blurring effects become unavoidable during reconstruction.

Yang et al. [18] proposed a new strategy based on sparsity. A couple of LR-HR dictionaries were trained. The results are optimal, however, the algorithms are time consuming. Some efforts were made in [19], [20] and [21] to save computational cost while preserving the quality of the SR estimate. Zeyde et al. [22] used dimensionality reduction based on principal component analysis (PCA) and singular value decomposition (K-SVD) dictionary training with orthogonal matching pursuit (OMP) to improve the computational cost of Yang et al. [18] model.

Timofte et al. [23] introduced an anchored regression model (ANR) and its extreme case called a global regression model (GR) supported by a dictionary. However, the dictionary is large and an overloaded dictionary, with high redundancy of information, is potentially unstable.

Dong et al. [24] proposed an adaptive sparse domain selection model (ASDS) for decomposing images in subspaces by K-means grouping and PCA for dictionary training. They retrieve the HR image using autoregressive parts (AG) models and a self-regulatory restriction. The non-local similarity was used as the regularization term. The same authors proposed an improved model called non-locally centralized sparse representation (NCSR) [25]. Recent researches extend conventional super resolution approaches with some new ideas, [26] include the sparse representation of each RGB channel of color images as a group of three atoms, instead of considering the luminance channel only.

In a PET study, the acquisition time of photons is reduced, which results in noisy and LR images. Several approaches have been proposed to increase the resolution [3], [2], [27]. Some include SR in the reconstruction process; others use SR to the images after reconstruction.

In [28], the use of the SR applied to PET images is described. The study is based on the approach of [29], where an HR image is estimated using the projection of several observed LR images. This work demonstrates the feasibility of applying SR techniques to PET images. The authors used images already reconstructed and displaced at the subpixel level to apply an SR algorithm of multiple images. The drawback is the need to have multiple images with a subvoxel displacement from each other, which makes its real application infeasible.

In [30] a scheme of improvement of the resolution in the domain of the sinogram is introduced. The data is acquired through the relative movement of the object. The algorithm estimates the HR sinogram by interpolation, followed by the reconstruction using

the expectation maximization (EM) algorithm. However, the object under study must be displaced in distances of subvoxels to be able to pick up the LR images.

Other works like [31], [32] and [33], are based on [28]. The researches have applied different techniques, such as a term of Tikhonov or a hybrid regularization term, to regularize the solution. Although this approach improves the resolution, the results are based on multiple images and the methods are applied to reconstructed images.

In [34], the authors use the dictionary-based super resolution to improve PET images with good results. They proposed to use random forest algorithms to extract the patches of the underlying images. However, this work is focused on the SR applied to reconstructed images, in a post-reconstruction stage. We believe that the results can be improved if this approach is applied directly to the sinograms, before reconstruction.

The idea of improving PET images before reconstruction, that is, in the sinogram domain, has been less explored in the literature. Therefore, if a HR sinogram is recovered, the results may improve, as more information is preserved during the reconstruction process.

In [35] a SR method in sinograms is proposed. In these works, it is required to have displaced sinograms at the subvoxel level in order to apply the algorithm, which, as we have mentioned, is difficult to carry out in real life. However, both works show the feasibility of increasing the resolution of the sinograms and the impact on the reconstructed final image.

In this paper, we propose a strategy to super resolve the sinograms of data from a scanner, using an SR algorithm based on dictionaries for PET images. The remaining of the paper is organized as follows: In section 2, methods are shown. In section 3, the SR strategy for PET sinograms is presented. In Section 4 the results are presented and discussed. Finally, Section 5 presents conclusions.

2 Methods

In PET imaging, the vast majority of SR techniques have been developed for multiple images. However, the acquisition of the LR images could be very difficult. In this section, we present the SR methods, based on dictionary learning, used to super resolve the sinograms and to recover the simulated images.

2.1 Positron emission tomography

Positron emission tomography is a non-invasive medical imaging technique. A PET tomograph detects photons that result of the annihilation of a positron and an electron [3] in the subject under study. Photons are counted and organized in a sinogram. The sinogram is further processed in order to obtain an image of the internal part of the body. The images obtained are useful for the diagnosis of diseases such as cancer, epilepsy and Alzheimer [36], [37] due to the ability to detect metabolic activity [4]. A PET tomograph is made up of several rings of detectors with different geometries as shown in Fig. 1 (a) and (b). Each ring is built-up of crystals that detect the interaction with the photons, that occur during annihilation inside the subject under study. The detectors consist of a scintillating crystalline material that serves as an interaction medium for high energy photons and that emits visible light when the energy is deposited inside them. Then, the light is detected by a photon detector of visible light and converted into electric current [3].

(a) (b)

Fig. 1. PET scanner (a) model and (b) three different shapes of detector rings.

The subject under study is injected with a permissible dose of a radioactive substance known as radio marker. Positrons are generated and interact with the electrons inside the body, producing a phenomenon known as annihilation (positron annihilate

with their antiparticle). The annihilation process generates two photons that travel in opposite directions, approximately 180 degrees. The detection and counting of these photons, traveling in opposite directions, is the foundation of the scanner operation [3]. Photons are detected within a small time interval (≈10ns) and located in the same line that connects the center of both detectors called the line of response (LOR) [4], as shown in Fig. 2.

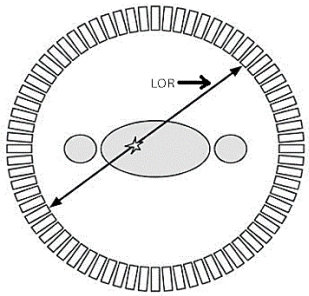


Fig. 2. Example of a LOR and an annihilation point *[3]*.

The last step is to store the information on the computer. The matching events are recorded as a sinogram or as a list mode file. The exact place where annihilation was generated is unknown. However, it is assumed that it occurred somewhere along the LOR. The events occurring along the same LOR are counted and accumulated in the same sinogram pixel.

2.2 Reconstruction methods for PET images

Reconstruction methods are used for the reconstruction of the images from the sinogram. There are two fundamental types of reconstruction algorithms: analytical algorithms such as the FBP and the iterative algorithms such as the OSEM. In this section, a brief review of these two algorithms is given.

Filtered backprojection

In the 2D acquisition, the activity through a LOR in a sinogram is the sum of all the activities detected by a detector pair [3]. The reconstruction matrix is selected of a specific size. The image matrix is in Cartesian coordinates and the sinogram data are in polar coordinates [3], [38]. An image in a rectangular coordinate (x, y) is related to its polar coordinates(r, ϕ) by Equation (***1***.

$$\mathrm{r} = \mathrm{y}\sin\phi + \mathrm{x}\cos\phi. \tag{1}$$

Where r is the distance from the center of the tomograph to the LOR and ϕ is the angle of the LOR and the x coordinate. Fig. 3 shows the relationship described in Equation (***1***.

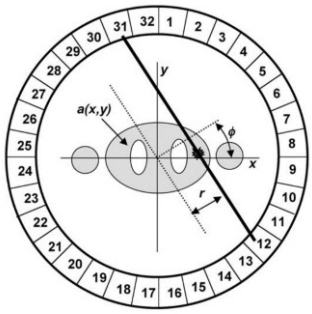


Fig. 3. Cross section of a PET tomograph and formation of a LOR *[3]*.

The magnitude of each pixel in a position (x, y) in the image, depends on the number of counts in the sinogram (r, ϕ) and is calculated using Equation (***2***. The counts measured in the LOR corresponding to the distance, are added to the pixel (x, y) in the reconstruction matrix. This is repeated for all projection angles. Therefore, the image of the form (x, y) in the reconstruction matrix is given by:

$$a(x,y) = \frac{1}{N}\sum_{n=1}^{N} p_n(r,\phi). \quad (2)$$

Where $p_n(r,\phi)$ is the number of counts in the sinogram and N is the number of projection angles. The image is reconstructed when all the pixels have been calculated.

Iterative methods

One of the most recent methods used for reconstruction techniques of tomography images is the so-called iterative methods. The recent increase in computational capacity allows the adoption of these types of algorithms, which are of high complexity but achieve better images. There are two main types of algorithms: deterministic interpretation algorithms and stochastic interpretation algorithms.

The algorithms of deterministic interpretation receive the generic name of algebraic reconstruction techniques (ART) in which an iterative method is used to solve the system of linear equations $\mathbf{Ax} = \mathbf{b}$. The values of the pixels are considered as variables collected in vector **x** and the process of image formation is described as matrix **A**. Projections acquisitions are stored in vector **b** [39]. The stochastic interpretation solves the system of equations $\mathbf{Ax} = \mathbf{b}$, where b is the data set, **A** is a known matrix and **x** the unknowns. The usual method of solving this problem is using the maximum likelihood estimator (MLE) that estimates the most likely image to generate the projections [40].

The maximum likelihood expectation maximization (MLEM) algorithm was introduced in the field of image reconstruction in 1982 by Shepp and Vardi [40] and remains a basis for the most popular statistical reconstruction method. The MLEM provides the foundation for many other methods. When the MLEM algorithm is applied to PET image reconstruction, it leads to the simple iterative equation:

$$\hat{x}_j^{(i+1)} = \frac{\hat{x}_j^{(i)}}{\sum_{i'} a_{i'j}} \sum_i a_{ij} \frac{y_i}{\sum_k a_{ik}\hat{x}_k^{(n)}} \quad (3)$$

Where y_i is the value of the ith projection, x_j is the value of the jth voxel in the image. Also, $\hat{x}_j^{(i+1)}$ is the next estimate of the voxel j based on the current estimate $\hat{x}_j^{(i)}$ and a_{ij} is a probability matrix that represents the probability that an event produced in pixel j be detected in projection i. This iteration is repeated until the final image is estimated. Photon detections are Poisson distributed, therefore, in the majority of methods, a Poisson model for the counts is used.

A variant of the MLEM method is the ordered subsets expectation maximization (OSEM) method that groups the detectors into subsets (or blocks) and makes the processing of each subset in iterations, making one block at a time [41]. OSEM reduces the reconstruction time of conventional MLEM. The grouping into subsets modifies Equation (***3***) as follows:

$$\hat{x}_j^{(i+1)} = \frac{\hat{x}_j^{(i)}}{\sum_{t\epsilon S_i} a_{tj}} \sum_{t\epsilon S_i} a_{tj} \frac{y_t}{\sum_k a_{ik}\hat{x}_k^{(i)}}. \quad (4)$$

Where S_i is a number of projections of a total of subsets. If there is only one subset, OSEM is the same as MLEM. The pseudo code of OSEM is shown in algorithm 1 of

Fig. *1*

Algorithm 1: Ordered Subsets Expectation Maximization (OSEM) [41]

Initialize

m=0 the initial iteration, x^m initial image (positive)

while no convergence of x^m **do**

$x^1 = x^m,\ m = m + 1$

for subsets i=1,...,n **do**

1. **Project**
   For detectors $t \in S_i$ calculate expected values for accumulated counts y as
   $$\mu_t^i = \sum_{j=1}^{J} a_{tj} x_j^i$$
2. **Backproject**
   For each pixel j=1,…,J, calculate
   $$x_j^{i+1} = x_j^i \sum_{t \in S_i} \frac{y_t a_{tj}}{\mu_t^i} / \sum_{t \in S_i} a_{tj}$$

end for
$x^m = x^{n+1}$ Output the optimized image
end while

Fig. 4. Pseudocode of OSEM algorithm.

2.3 Super resolution

In most electronic imaging, applications HR images are desired and frequently required because they show more information about the details. Therefore, these images may yield critical details useful for various applications. For example, HR is important in medical images for the correct diagnosis of diseases; in satellite images is important to distinguish an object from similar ones and to obtain better performance in pattern recognition programs. In the computer vision area, HR is useful to obtain images for more accurate recognition of specific characteristics. There are two main forms of SR: Multiple images SR and single image SR. The main goal of multiple image SR is to yield a high resolution image based on multiple LR representations of the same scene with a subpixel shift between them. Hence, the single image SR methodology recovers a high quality image from an LR resolution image by predicting the HR details.

Observational model

The observation model is the mathematical expression used to represent the image formation process and reflects the relationship between the desired HR image and the observed LR images. Consider the desired HR image $\mathbf{X}$ written in lexicographical order as $[\mathbf{X_1}, \mathbf{X_2}, \mathbf{X_3}, \ldots, \mathbf{X_N}]^T$, whose size is $D_1N_1 \times D_2N_2$, where $\mathbf{X}_i = x_1, x_2, \ldots, x_N$ represents each pixel in $\mathbf{X}$ and $N = D_1N_1 \times D_2N_2$, where $\mathbf{X}$ is an ideal high-resolution image sampled at the Shannon-Nyquist velocity of a continuous scene, and D_1, D_2 are considered as the decimation factors in the horizontal and vertical directions respectively. Therefore, the observed LR image $\mathbf{Y}$ has the size of $N_1 \times N_2$ and can be written in lexicographical notation as $[\mathbf{Y_1}, \mathbf{Y_2}, \mathbf{Y_3}, \ldots, \mathbf{Y_N}]^{\mathbf{T}}$, where $\mathbf{Y}_j = y_1, y_2, \ldots, y_M$ represents each pixel in $\mathbf{Y}$, and $M = N_1 \times N_2$.

During the acquisition of the observed LR image, defocusing is another parameter to be taken into account. It can be caused by the optical system or by the movement of the position of an object that occurs during the exposure time of the camera. Noise is another undesirable but inevitable distortion, produced by the acquisition device. If the probability density function (PDF) is known, noise can be categorized and treated differently; for example, Gaussian noise or salt and pepper noise. Gaussian noise is the most frequent noise found image processing. It is often caused by low lighting, high temperature and noise from an electronic circuit. To simplify, Gaussian noise in image processing is commonly considered additive Gaussian white noise (AGWN) with zero mean.

The single-image super resolution observation model can be established by combining the degradations mentioned previously. The observation model can be expressed mathematically in Equation (***5***) where $\mathbf{Y}$ and $\mathbf{X}$ represent the LR and HR images respectively; **B, D** and $\boldsymbol{\nu}$ represent the blurring matrix, the downsampling matrix and the additive noise respectively.

$$\mathbf{Y} = \mathbf{BDX} + \boldsymbol{\nu}. \quad (\mathbf{5})$$

Sparse model

The sparse model establishes a relationship between the LR and the HR image. It expresses that any signal is as a linear combination of several bases in **D**, as shown in Fig. 5., where $\mathbf{X} \in \mathbb{R}^N$ is a signal in a space of N dimensions. In the sparse model,

it can be represented as $\mathbf{X} = \mathbf{D}\boldsymbol{\alpha}$, with $\mathbf{D} \in \mathbb{R}^{N\times M}$ a set of redundant bases with parameters ($N < M$), and $\boldsymbol{\alpha} \in \mathbb{R}^{M}$ represents the sparse vector or the sparse code with the majority of the coefficients equal or close to zero. The sparse representation of $\mathbf{X}$ can be found by solving Equation (***6***).

$$\min \ \|\mathbf{X} - \mathbf{D}\boldsymbol{\alpha}\|_2^2 + \lambda\|\boldsymbol{\alpha}\|_1. \tag{6}$$

Where $\|.\|_1$ denotes the norm l_1 and λ is used to balance the relationship between the sparse vector and the fidelity of the approximate signal.

Similarly, an image $\mathbf{X} \in \mathbb{R}^{N}$ in a space of N dimensions can also be expressed as the product of $\mathbf{D}$ and $\boldsymbol{\alpha}$. Here, $\mathbf{D} \in \mathbb{R}^{N\times M}$ is called a dictionary, and it is usually over completed to represent generic images with different textures, ($M \gg N$).

$$\hat{\mathbf{X}} = \mathbf{D}\boldsymbol{\alpha}. \tag{7}$$

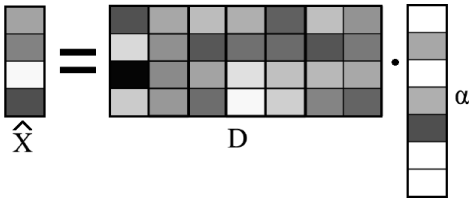


Fig. 5. Sparse model representation.

Each column of the dictionary is called an atom, $\boldsymbol{\alpha}$ is the sparse vector. Therefore, with the knowledge of the $\mathbf{D}$ dictionary and its sparse $\boldsymbol{\alpha}$ code, the image can be easily reconstructed using Equation (*7* and depicted in Fig. 5.

Dictionary training

Fig. *6*. shows the dictionary for (a) low resolution and (b) high resolution patches.

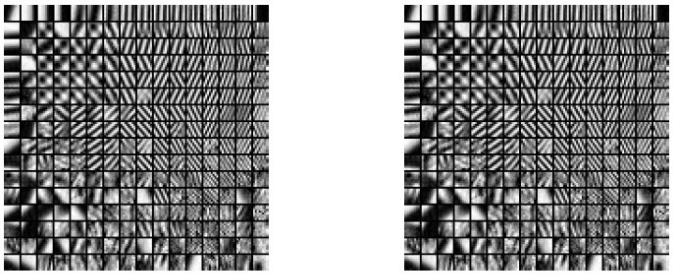


Fig. 6. Dictionaries of patches of (a) low resolution and (b) high resolution.

The following constrained minimization problem can be solved [19] to train the dictionary,

$$\mathbf{D} = \min\|\mathbf{X} - \mathbf{D}\boldsymbol{\alpha}\|_2^2 + \lambda\|\boldsymbol{\alpha}\|_1 \quad \text{s.t}\ \|\mathbf{D}_i\|_2^2 \le 1, \qquad i = 1, 2, 3, \ldots, K. \tag{8}$$

Where $\mathbf{D}$ is a dictionary with K atoms, $\|.\|_1$ is the norm l_1 used to reinforce the sparsity of $\boldsymbol{\alpha}$. $\|.\|_2$ is the norm l_2 and restriction of the columns of $\mathbf{D}$ to remove the ambiguity of the scale. To solve the optimization problem of Equation (***8***, the algorithm of **Fig. *7***. was implemented in this research.

Algorithm 2: Strategy to solve dictionary training

1. Initialize $\mathbf{D}$ and normalize each atom
2. Set $\mathbf{D}$ and update $\boldsymbol{\alpha}$ using
   $\boldsymbol{\alpha} = \min\|\mathbf{X} - \mathbf{D}\boldsymbol{\alpha}\|_2^2 + \lambda\|\boldsymbol{\alpha}\|_1$
3. Fix $\boldsymbol{\alpha}$ and update $\mathbf{D}$ using the following formulation

$$\mathbf{D} = \min \|\mathbf{X} - \mathbf{D}\boldsymbol{\alpha}\|_2^2 + \lambda\|\boldsymbol{\alpha}\|_1 \quad \text{s.t } \|\mathbf{D_i}\|_2^2 \leq 1,\ \mathrm{i} = 1, 2, 3, \ldots, \mathrm{K}$$

4.Iterate steps 2 and 3, until the convergence.

Return **D**

Fig. 7. Algorithm to train dictionary **D**.

3 Proposed method: Super resolution applied to PET sinograms

First, a series of ideal sinograms were generated, necessary for the training of high and low resolution dictionaries. These sinograms were generated by the Radon transform in Matlab. At the same time, they were degraded and downsampled in order to obtain pairs of HR sinograms and LR sinograms. Fig. 8 shows an example of some pairs of sinograms used to train the dictionary.

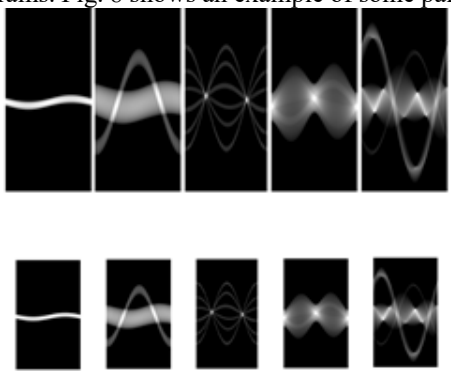


Fig. 8. Examples of pairs of sinograms used to train the dictionary. The upper row is the HR training set and the lower row is the LR training set.

The dictionaries are trained following the procedure shown in **Fig.** *7*, after training, one can proceed to super resolve the sinograms. With a given LR sinogram and the LR dictionary $\mathbf{D_l}$, the best sparse code ($\boldsymbol{\alpha}$) can be calculated by the minimization of Equation (***9***) using the feature-sign search algorithm of Fig. 9.

$$\boldsymbol{\alpha} = \min_{\alpha} \|\mathbf{X} - \mathbf{D_l}\boldsymbol{\alpha}\|_2^2 + \lambda\|\boldsymbol{\alpha}\|_1 \qquad (\mathbf{9})$$

Algorithm 3: Feature-sign search [42]

Initialize

α = 0, θ = 0, A={}

while $1 \leq j \leq n$ **do**

1. For zero coefficients in α, compute $\mathrm{i} = \operatorname{argmax}\left|\frac{\mathrm{d}\|\mathrm{X}-\mathrm{D}\alpha\|_2^2}{\mathrm{d}\alpha_j^i}\right|$

   if $\frac{\mathrm{d}\|\mathrm{X}-\mathrm{D}\alpha\|_2^2}{\mathrm{d}\alpha_j^i} > \lambda$, set $\theta_j^i = -1$, and $\mathrm{A} = \{\mathrm{i}\} \cup \mathrm{A}$

   if $\frac{\mathrm{d}\|\mathrm{X}-\mathrm{D}\alpha\|_2^2}{\mathrm{d}\alpha_j^i} > -\lambda$, set $\theta_j^i = 1$, and $\mathrm{A} = \{\mathrm{i}\} \cup \mathrm{A}$

3. Compute analytical sol. $\hat{\alpha}_{\mathrm{new}}$ by solving an unconstrained quadratic problem form formulated as:

$$\hat{\alpha}_{\mathrm{new}} = \left(\hat{\mathrm{D}}^{\mathrm{T}}\hat{\mathrm{D}}\right)^{-1}\left(\hat{\mathrm{D}}^{\mathrm{T}}\mathrm{X} - \hat{\lambda}\theta/2\right)$$

4. Update $\hat{\alpha}$, A, and θ by performing a discrete line search according to $\hat{\alpha}_{\mathrm{new}}$

5. Check the optimality conditions:

   if $\forall\alpha_j^i \neq 0$, $\frac{\mathrm{d}\|\mathrm{X}-\mathrm{D}\alpha\|_2^2}{\mathrm{d}\alpha_j^i} + \lambda\,\mathrm{sign}\left(\alpha_j^i\right) \neq 0$, back to step 4

   if $\forall\alpha_j^i = 0$, $\frac{\mathrm{d}\|\mathrm{X}-\mathrm{D}\alpha\|_2^2}{\mathrm{d}\alpha_j^i} > \lambda$, back to step 3

end while

```
Return α (the sparse code)
```

Fig. 9. Feature-sign search algorithm.

It is considered that the sparse representations of the HR and LR images are the same. Therefore, with the knowledge of $\boldsymbol{\alpha}$ and $\mathbf{D_h}$, HR image can be estimated by

$$\hat{\mathbf{X}} = \mathbf{D_h}\boldsymbol{\alpha} \tag{10}$$

Fig. 10 represents the process to super resolve the PET sinogram using the proposed dictionary technique. The LORs of the detected events are converted into a magnitude-angle (sinogram) file of LR. Each sinogram represents a 2D slice of the 3D scene. The sparse code of each 2D sinogram is calculated using the algorithm of Fig. 9 and the HR sinogram is calculated by using Equation (***10***. Finally, an iterative (or an algebraic) method is applied on the sinograms to estimate the volumetric image.

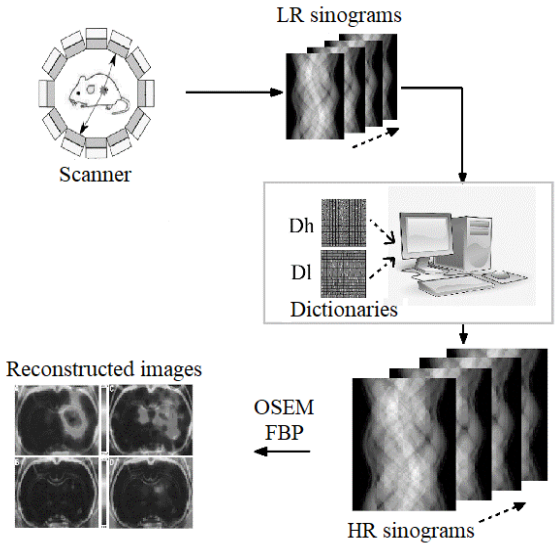


Fig. 10. Super resolution process for PET sinograms.

Fig. 11 shows the proposed algorithm implemented in this research.

```
Algorithm 4: Sinogram3D super resolution
in: sinogram3D
    β= number of sinograms in sinogram3D
for sinograms i=1,...,β do
      γ= extract_sinogram(sinogram3D,i)
      ψ= super_resolution(γ)
      SRsinogram3D= update_sinogram(ψ)
end for
Return SRsinogram3D
```

Fig. 11. Sinogram3D super resolution algorithm.

4 Results

Quantitative results were obtained using the maximum intensity projection (MIP) of coronal and transversal views. The peak-signal to noise-ratio (PSNR) and the structural similarity index measure (SSIM) metrics were used. In addition, the contrast C was also calculated as in [28]. The averaged in the area of activity S was divided by the average in the background B close to the same area. The contrast ratio is expressed as $C = S/B$. To verify the efficiency of the proposed method, a series of case studies were tested. A phantom is a mannequin design to simulate body structures. The phantoms in this work were designed by using the SimSET tool [43], one million events were generated for every experiment. The first phantom of rods with different diameters was designed and placed in different positions inside the tomograph. The reconstruction of the rods was carried out using the FBP and the OSEM algorithms. The second phantom of cold chambers filled with air and water was also simulated. Results were compared against the commonly used bicubic interpolation method. Fig. 12 illustrates the process carried out for comparison purposes. Sinograms are reconstructed and interpolated by a factor of 2 by using bicubic interpolation. Also, the sinograms are super resolved by a factor of 2 using the proposed dictionary method and then reconstructed. Both images were compared.

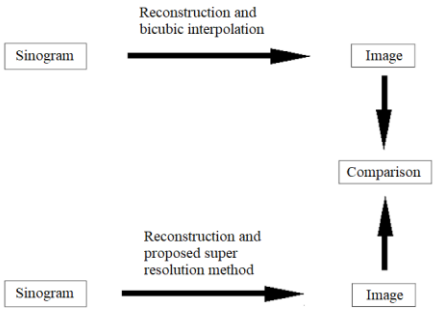
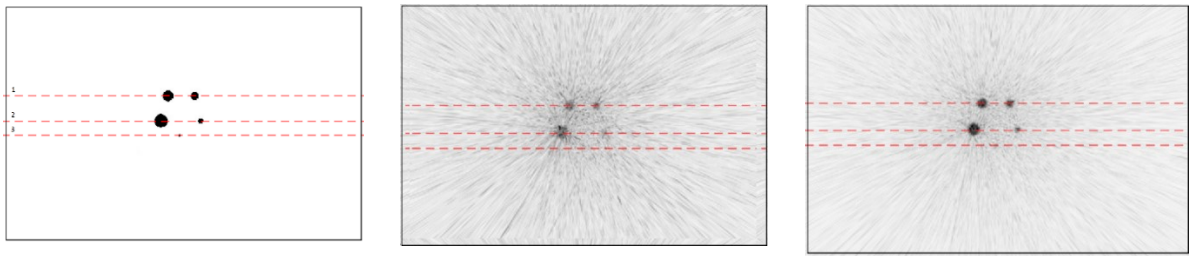


Fig. 12. Comparison strategy.

The first phantom was designed according to the following: Five rods filled with ^{18}F radio marker, spatially distributed inside the tomograph with angles of seventy-two degrees from each other. The diameter of each rod was 1mm, 2mm, 3mm, 4mm and 5mm respectively. The length of each rod was 20cm.

Results of transversal view of the rods

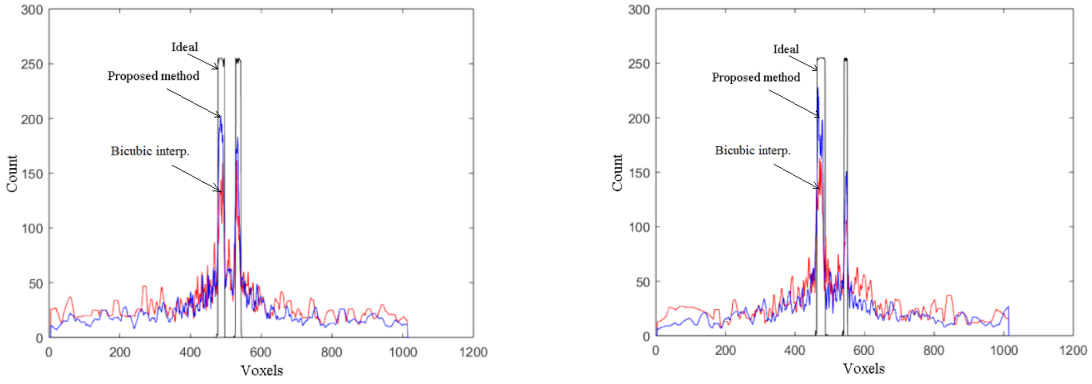
Fig. 13 shows the MIPs of the ideal phantom and the reconstructed phantoms using FBP. The dashed lines horizontal profiles of the MIPs used to plot Fig. 14. Here, a gain of the proposed method of about 2.7 dBs in PSNR with respect to the bicubic method was obtained.



(a) (b) (c)

Fig. 13. MIPs of the transversal view of the ^{18}F simulated phantom. (a) Ideal phantom. (b) Reconstructed phantom using FBP and bicubic interpolation (PSNR = 20,761), (SSIM = 0.224), (C = 1,124). (c) Reconstructed phantom using FBP and the proposed method (PSNR = 23.489), (SSIM = 0.235), (C = 1.398).

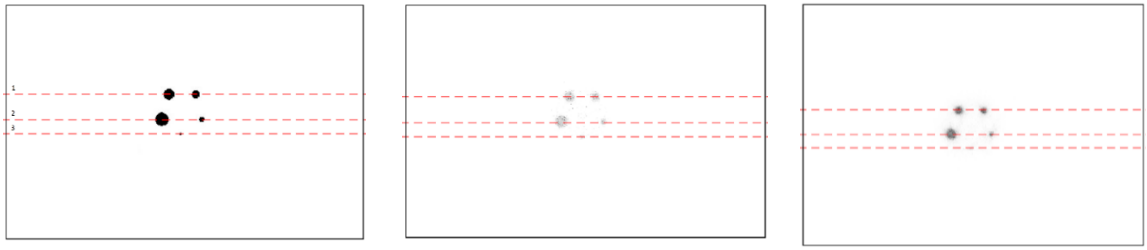
Fig. 14 shows the profiles of the MIPs of Fig. 13. It can be seen how the counts increase with the proposed method.



(a) (b)

Fig. 14. Horizontal profiles of the previous figure for the maximum intensity projections of the ^{18}F phantom. (a) shows comparison for profile 1 and (b) shows the comparison for profile 2 shown in Fig. 13.

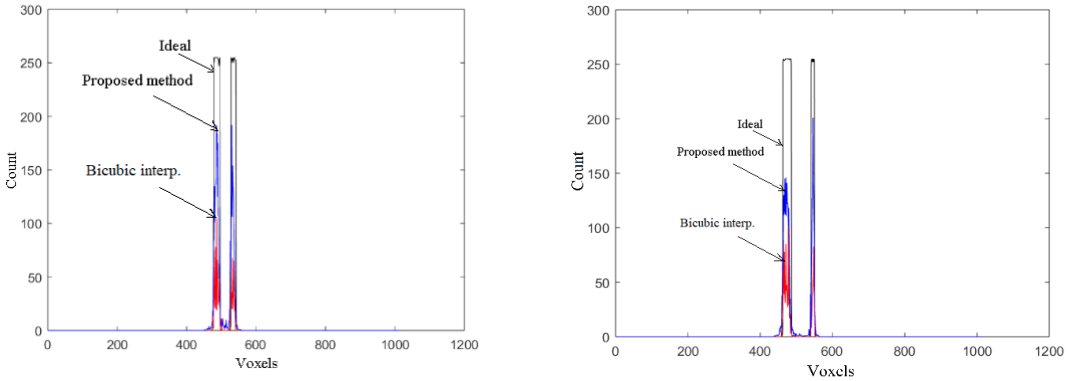
Fig. 15 shows the MIPs of the ideal phantom and the reconstructed phantoms using OSEM. A gain of about 1.7 dBs in PSNR of the proposed method with respect to the bicubic method was obtained. Also, the SSIM and contrast measure both are higher.



(a) (b) (c)

Fig. 15. MIPs of the transverse view of the ^{18}F simulated phantom. (a) Ideal phantom. (b) Reconstruction using OSEM and bicubic interpolation (PSNR = 31.091), (SSIM = 0.990), (C = 2.961). (c) Reconstruction using OSEM and proposed method (PSNR = 32.798), (SSIM = 0.992), (C = 3.338).

Fig. 16 shows the profiles of the MIPs of Fig. 15**Fig. 13**. It can be seen how the counts increase with the proposed method. In both cases, the counts of the proposed method are larger than the bicubic method.



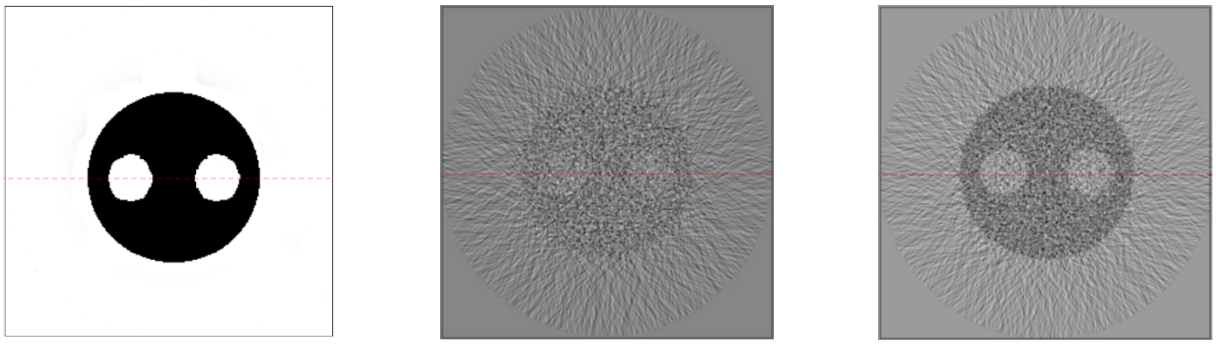
(a) (b)

Fig. 16. Comparisons of the horizontal profiles of the previous figure for the maximum intensity projections of the ^{18}F phantom. (a) shows comparison for profile 1 and (b) shows the comparison for profile 2 shown in Fig. 15.

The second phantom was designed according to the following: A large cavity of 3 cm in length and 3.3 cms in diameter filled with the isotope ^{18}F. Inside the cavity, there are two cold chambers of 0.1 cms in diameter filled with water and air respectively [44].

Results of the cold chambers transversal view

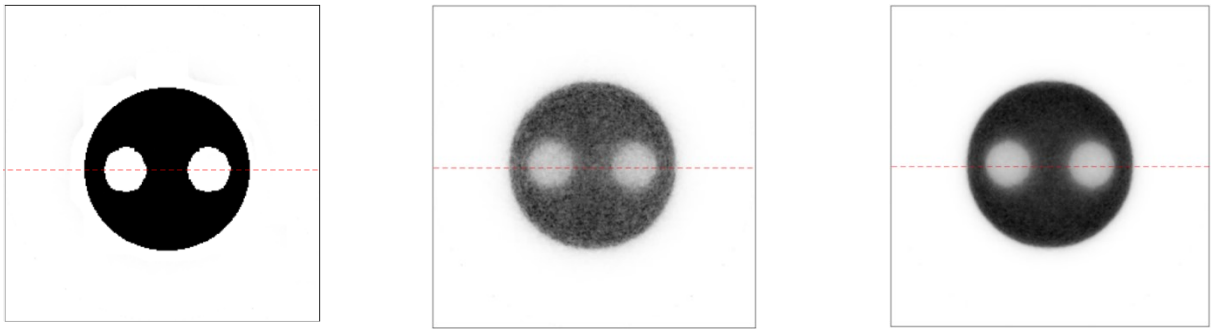
Fig. 17 shows the MIPs of the ideal and the reconstructed phantoms using FBP. The dashed line indicates the place where a horizontal profile of the MIP is plotted. Notice that a gain of about 1.1 dBs in PSNR is obtained with the proposed method as compared to the bicubic method. SSIM and contrast measures are also higher.



(a) (b) (c)

Fig. 17. MIPs of the transverse view. (a) Ideal phantom. (b) Reconstruction using FBP and bicubic interpolation (PSNR = 9.882), (SSIM = 0.590), (C = 1.095). (c) Reconstruction using FBP and the proposed method (PSNR = 10.954), (SSIM = 0.595), (C = 1.272).

Fig. 18. MIPS of the transverse view of the ^{18}F simulated phantom. (a) Ideal phantom. (b) Reconstruction using OSEM and bicubic interpolation (PSNR = 17.243), (SSIM = 0.926), (C = 3.160). (c) Reconstruction using OSEM and proposed method (PSNR = 18.236), (SSIM = 0.942), (C = 3.734).shows the MIPs of the ideal and the reconstructed phantoms using OSEM. In this case, a gain of about 1.0 dB in PSNR of the proposed method with respect to the bicubic method was obtained. Also, the proposed method shows a higher SSIM and contrast value.



(a) (b) (c)

Fig. 18. MIPS of the transverse view of the ^{18}F simulated phantom. (a) Ideal phantom. (b) Reconstruction using OSEM and bicubic interpolation (PSNR = 17.243), (SSIM = 0.926), (C = 3.160). (c) Reconstruction using OSEM and proposed method (PSNR = 18.236), (SSIM = 0.942), (C = 3.734).

Fig. 19 shows the horizontal profiles of the MIPs of Fig. 18**Fig. 13**. It can be seen how the count increases with the proposed method. In both cases, the count of the proposed method is larger than the bicubic method.

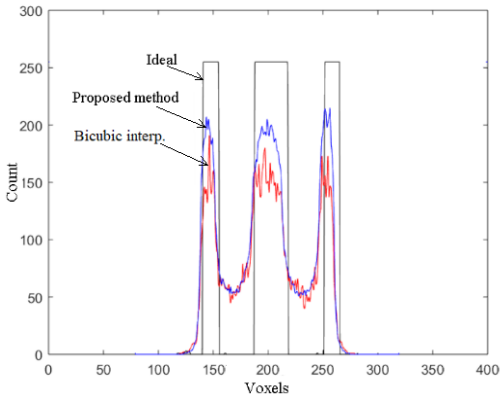
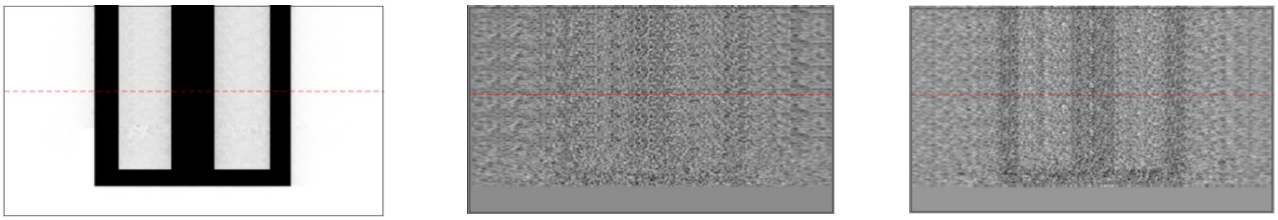


Fig. 19. Comparisons of the horizontal profile of Fig. 18. for the MIPs of the cold chambers.

Results of the cold chambers coronal view

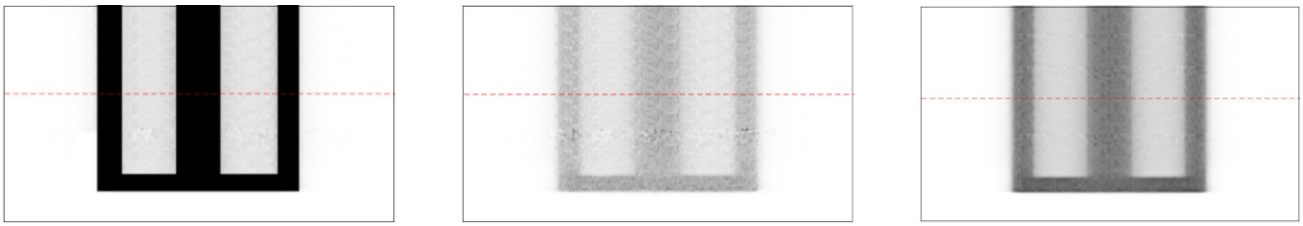
Fig. 20 shows the MIPs of the ideal phantom and the reconstructed phantoms using FBP. The dashed lines are the place where a horizontal profile of the MIPs is plotted. It was obtained a gain of about 0.7 dBs in PSNR in our method with respect to the bicubic method. SSIM and contrast measures are also higher.



(a) (b) (c)

Fig. 20. MIPs of the coronal view of the cold chambers phantom. (a) Ideal phantom. (b) Reconstructed phantom using FBP and bicubic interpolation (PSNR = 10,324), (SSIM = 0.620), (C = 1.066). (c) Reconstruction using FBP and the proposed method (PSNR = 10.965), (SSIM = 0.630), (C = 1.194).

Fig. 21**Fig. *20*** shows the MIPs of the ideal phantom and the reconstructed phantoms using OSEM. The dashed lines are used to plot a horizontal profile of the MIPs. It was obtained a gain of about 1.8 dBs in PSNR with our method with respect to the bicubic method. SSIM and contrast measures are also higher.



(a) (b) (c)

Fig. 21. MIPs of the coronal view of the cold chambers phantom. (a) Ideal phantom. (b) Reconstructed using OSEM and bicubic interpolation (PSNR = 12,654), (SSIM = 0.934), (C = 2.622). (c) Reconstructed using OSEM and the proposed method (PSNR = 14.461), (SSIM = 0.937), (C = 3.104).

Fig. 22 shows the profiles of the MIPs of Fig. 21**Fig. 18Fig. 13**. It can be seen how the counts increase with the proposed method. In both cases, the counts of the proposed method are larger than the bicubic method.

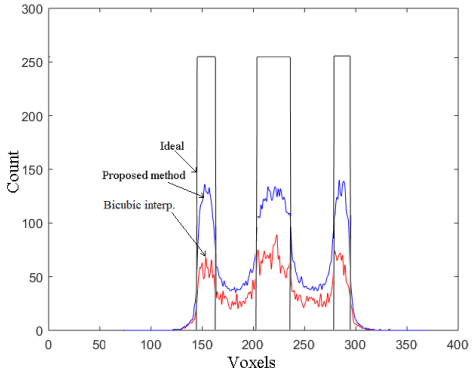


Fig. 22. Comparisons of the horizontal profile in **Fig. 20**Fig. 21 for the MIPs of the cold chambers phantom.

Table 1, Table 2 and Table 3 shows a summary of the quantitative results obtained in the experiments. As can be seen, in all cases the proposed method outperforms the bicubic interpolation for the three metrics used, PSNR, SSIM and contrast C.

Table 1. PSNR values obtained from experiments (Higher value is indicated in bold).

Experiment	Bicubic interpolation	Proposed method
Phantom 1-Transversal view with FBP	20,761	**23,489**
Phantom 1-Transversal view with OSEM	31,091	**32,798**
Phantom 2-Transversal view with FBP	9,882	**10,954**
Phantom 2-Transversal view with OSEM	17,243	**18,236**
Phantom 2-Coronal view with FBP	10,324	**10,965**
Phantom 3-Coronal view with OSEM	12,654	**14,461**

Table 2. SSIM values obtained from experiments (Higher value is indicated in bold).

Experiment	Bicubic interpolation	Proposed method
Phantom 1-Transversal view with FBP	0,224	**0,235**
Phantom 1-Transversal view with OSEM	0,990	**0.992**
Phantom 2-Transversal view with FBP	0,590	**0,595**
Phantom 2-Transversal view with OSEM	0,926	**0,942**
Phantom 2-Coronal view with FBP	0,620	**0,630**
Phantom 3-Coronal view with OSEM	0,934	**0,937**

Table 3. Contrast C values obtained from experiments (Higher value is indicated in bold).

Experiment	Bicubic interpolation	Proposed method
Phantom 1-Transversal view with FBP.	1,124	**1,398**
Phantom 1-Transversal view with OSEM	2,961	**3,338**
Phantom 2-Transversal view with FBP	1,095	**1,272**
Phantom 2-Transversal view with OSEM	3,160	**3,734**
Phantom 2-Coronal view with FBP	1,066	**1,194**
Phantom 3-Coronal view with OSEM	2,622	**3,104**

The SR strategy applied to the PET sinograms increased the resolution in the final images, which is evidenced of the improvements obtained by the method proposed in the comparison metrics used in this investigation. Table 1 shows how the PSNR values increased on average by more than 3% for both reconstruction algorithms (FBP and OSEM) compared to the bicubic interpolation method. For the two simulated phantoms, similar rates of improvement were obtained. Table 2 also shows an improvement of the SSIM, in both phantoms and for the two reconstruction methods. On average, the proposed method improves in about 1% in SSIM

the results obtained through bicubic interpolation. The levels of detail in the simulated rods also improve with the application of our method, Figs. 14 and 16 show a gain of counts. Fig. 14 shows the rise in the counts in the 2mm and 5mm rods and Fig. 16 shows an increase in counts in the 3mm and 4mm rods. These results imply a better definition of the edges of rods, a greater contrast and therefore, easier to visualize. The results of OSEM show a better quality of reconstruction compared to the results of the FBP. Figs. 19 and 22 show a significant gain in the number of counts for both FBP and OSEM algorithms. A better definition of the limits between the area containing the radiotracer and cold chambers is achieved. Table 3 summarizes the contrast results measured between the rods and the surrounding area. It is shown that the proposed method always improves, on average, more than 3% of the results obtained by bicubic interpolation. In addition, the results obtained in this research are superior to other methods of the literature that were also compared to the bicubic interpolation method.

5 Conclusions

In this paper, a strategy for super resolution using a dictionary learning method for PET images was proposed. The acquired sinograms were super resolved and reconstructed. Sets of HR and their correspondent LR ideal sinograms were generated by the Radon transform. Patches of 5x5 samples were taken from both sets of sinograms to train the dictionary. The images were reconstructed using FBP and OSEM algorithms. The quality metrics PSNR, SSIM and C were used to evaluate the results. The results show the superiority of the proposed method with respect to the bicubic interpolation method. The PSNR reaches, on average, improvement rates above 3% and the SSIM about 1%. In the same way, the contrast C shows improvements, allowing to preserve more faithful the details of the reconstructed images. The proposed strategy outperforms the bicubic method and it allows to recover more details than in a post reconstruction stage. In addition, the SR of an image allows our method to be more realistic to be applied in real tomography studies because it avoids using multiple images where a set of LR images with subvoxel displacement are needed. The proposed method and OSEM yields the best quality images, however, this combination is time consuming.

Future works are aimed to train dictionaries in an adaptive way, which could increase the performance of the method since it would be able to "learn" from the underlying image. Another direction of the method is to include machine learning techniques to improve the time and quality of reconstructed images.

References

1. K. Iniewski and T. Farcombe, Medical Imaging: Technology and Application, 1st ed., CRC Press, 2017.
2. J. Yang and T. Huang, "Image super-resolution : Historical overview and future challenges," in *Super-resolution imaging*, CRC Press, 2011, pp. 1-24.
3. G. B. Saha, "Chapter 1 - pet scanning systems," in *Basics of PET Imaging*, 2nd ed., New York, Springer-Verlag, 2016.
4. M. E. Phelps, PET Physics, Instrumentation and Scanners, Springer, 2006.
5. R. Y. Tsai and T. S. Huang, "Multiframe image restoration and registration," *Advances in computer vision and Image Processing,* vol. 1, no. 2, pp. 317-339, 1984.
6. Y. Zhang, Q. Fan, F. Bao, Y. Liu and C. Zhang, "Single-Image Super-Resolution Based on Rational Fractal Interpolation," *IEEE Transactions on Image Processing,* vol. 27, no. 8, pp. 3782-3797, 2018.
7. X. Liu, L. Chen, W. Wang and J. Zhao, "Robust Multi-Frame Super-Resolution Based on Spatially Weighted Half-Quadratic Estimation and Adaptive BTV Regularization," *IEEE Transactions on Image Processing,* vol. 27, no. 10, pp. 4971-4986, 2018.
8. S. Zhao, Z. Lyu, H. Liang and M. Sarem, "A Mixed Non-local Prior Model for Image Super-resolution Reconstruction," *Chinese Journal of Electronics,* vol. 26, no. 4, pp. 778-783, 2017.
9. M. Rossi and P. Frossard, "Geometry-Consistent Light Field Super-Resolution via Graph-Based Regularization," *IEEE Transactions on Image Processing,* vol. 27, no. 9, pp. 4207-4218, 2018.
10. K. Zhang, D. Tao, X. Gao, X. Li and J. Li, "Coarse-to-Fine Learning for Single-Image Super-Resolution," *IEEE Transactions on Neural Networks and Learning Systems,* vol. 28, no. 5, pp. 1109-1122, 2017.

11. M. Yang, K. Liu and S. Chien, "A Real-Time FHD Learning-Based Super-Resolution System Without a Frame Buffer," *IEEE Transactions on Circuits and Systems,* vol. 64, no. 12, pp. 1407-1411, 2017.
12. H. Song, B. Huang, Q. Liu and K. Zhang, "Improving the Spatial Resolution of Landsat TM/ETM+ Through Fusion With SPOT5 Images via Learning-Based Super-Resolution," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 53, no. 3, pp. 1195-1204, 2015.
13. Z. Wei, B. Xiaofeng, H. Fang, W. Jun and A. A. Mongi, "Fast image super-resolution algorithm based on multi-resolution dictionary learning and sparse representation," *Journal of Systems Engineering and Electronics,* vol. 29, no. 3, pp. 471-482, 2018.
14. W. Freeman, T. Jones and E. Pasztor, "Example-based super-resolution," *IEEE Computer Graphics and Applications Magazine,* vol. 22, no. 2, p. 56–65, 2002.
15. J. Jiang, R. Hu, Z. Wang and Z. Han, "Face Super-Resolution via Multilayer Locality-Constrained Iterative Neighbor Embedding and Intermediate Dictionary Learning," *IEEE Transactions on Image Processing,* vol. 23, no. 10, pp. 4220-4231, 2014.
16. H. Wang, X. Gao, K. Zhang and J. Li, "Single Image Super-Resolution Using Gaussian Process Regression With Dictionary-Based Sampling and Student- t Likelihood," *IEEE Transactions on Image Processing,* vol. 26, no. 7, pp. 3556-3568, 2017.
17. K. I. Kim and Y. Kwon, "Single-image super-resolution using sparse regression and natural image prior," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 32, no. 6, p. 1127–1133, 2010.
18. J. Yang, J. Wright, T. Huang and Y. Ma, "Image super-resolution as sparse representation of raw image patches," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR),* pp. 1-8, 2008.
19. J. Yang, Z. Wang, Z. Lin, S. Cohen and T. Huang, "Coupled dictionary training for image super-resolution," *IEEE Transactions on Image Processing,* vol. 21, no. 8, p. 3467–3478, 2012.
[20] Q. Song, R. Xiong, D. Liu, Z. Xiong, F. Wu and W. Gao, "Fast Image Super-Resolution via Local Adaptive Gradient Field Sharpening Transform," *IEEE Transactions on Image Processing,* vol. 27, no. 4, pp. 1966-1980, 2018.
21. C. Dang and H. Radha, "Fast Single-Image Super-Resolution Via Tangent Space Learning of High-Resolution-Patch Manifold," *IEEE Transactions on Computational Imaging,* vol. 3, no. 4, pp. 605-616, 2017.
22. R. Zeyde, M. Elad and M. Protter, "On single image scale-up using sparse representations," *Lecture Notes Comput. Sci,* vol. 6920, pp. 711-730, 2010.
23. R. Timofte, V. De and L. Van Gool, "Anchored neighborhood regression for fast example-based super-resolution," *IEEE International Conference on Computer Vision (ICCV),* p. 1920–1927, 2013.
24. W. Dong, L. Zhang, G. Shi and X. Wu, "Image deblurring and super-resolution by adaptive sparse domain selection and adaptive regularization," *IEEE Transactions on Image Processing,* vol. 20, no. 7, p. 1838–1857, 2011.
25. W. Dong, L. Zhang, G. Shi and X. Li, "Nonlocally centralized sparse representation for image restoration," *IEEE Transactions on Image Processing,* vol. 22, no. 4, p. 1620–1630, 2013.
26. M. Cheng, C. Wang and J. Li, "Single-image super-resolution in rgb space via group sparse representation," *IET Image Processing,* vol. 9, no. 6, pp. 461-467, 2014.
27. G. Akamatsu, K. Ishikawa, K. Mitsumoto, T. Taniguchi, N. Ohya, S. Baba, K. Abe and M. Sasak, "Improvement in pet/ct image quality with a combination of point-spread function and time-of-flight in relation to reconstruction parameters," *J NuclMed,* vol. 53, no. 11, p. 1716–1722, 2012.
28. J. Kennedy, O. Israel, A. Frenkel, R. Bar-Shalom and H. Azhari, "Super-resolution in pet imaging," *IEEE Transactions on Medical Imaging,* vol. 25, no. 2, pp. 137-147, 2006.
29. M. Irani and S. Peleg, "Motion analysis for image enhancement: Resolution, occlusion,and transparency," *Journal of Visual Communication and Image Representation,* vol. 4, no. 4, p. 324–335, 1993.
30. K. Y. Jeong, J. H. Kim, K. M. Kim and J. B. Ra, "Sinogram super-resolution using a space-variant blur matrix in pet," *IEEE Transactions on Nuclear Science,* vol. 60, no. 1, p. 158–165, 2013.
31. P. Yan, M. Yang, H. Huang and J. Li, "An Improved Approach to Super Resolution Based on PET Imaging," *Neural Information Processing,* 2012.

32. X. Ren and S. . Lee, "High-Resolution Image Reconstruction for PET Using Local and Non-local Regularizations," *Electronic Imaging Computational Imaging XV,* vol. 1, pp. 174-178, 2017.
33. Y. Li, S. Matej and S. D. Metzler, "Image reconstructions from super-sampled datasets with resolution modeling in pet imaging," *Medical Physics,* vol. 41, no. 12, 2014.
34. H. Zhanli, Y. Wang, X. Zhang, M. Zhang, Y. Yang, X. Liu, H. Zheng and D. Liang, "Super-resolution of PET image based on dictionary learning and random forests," *Nuclear Instruments and Methods in Physics Research,* vol. 927, 2019.
35. K. Y. Jeong, K. Choi, W. H. Nam and J. B. Ra, "Sinogram-based super-resolution in PET," *Physics in Medicine & Biology,* vol. 56, no. 15, pp. 4881-4894, 2011.
36. C. Dang and H. Radha, "Fast Single-Image Super-Resolution Via Tangent Space Learning of High-Resolution-Patch Manifold," *IEEE Transactions on Computational Imaging,* vol. 3, no. 4, pp. 605-616, 2017.
37. T. DeGrado, B. Kemp, M. Pandey, H. Jiang, T. Gunderson, L. Linscheid, A. Woodwick, D. McConnell, J. Fletcher, G. Johnson, R. Petersen, D. Knopman and V. Lowe, "First-in-human PET imaging of 63Zn-zinc citrate in healthy elderly subjects and patients with Alzheimer's disease," *The Journal of Nuclear Medicine,* vol. 57, no. 2, 2016.
38. G. J. Gang, J. W. Stayman, W. Zbijewski and J. H. Siewerdsen, "Task-based detectability in CT image reconstruction by filtered backprojection and penalized likelihood estimation," *Med. Phys,* vol. 41, no. 8, 2016.
39. A. C. Kak and M. Slaney, Principles of Computerized Tomographic Imaging, IEEE Press, 1988.
40. L. A. Shepp and Y. Vardi, "Maximum Likelihood Reconstruction for Emission Tomography," *IEEE Transactions on Medical Imaging,* vol. 1, no. 2, pp. 113 - 122, 1982.
41. H. M. Hudson and R. S. Larkin, "Accelerated Image Reconstruction using Ordered Subsets of Projection Data," *IEEE Transactions on Medical Imaging,* vol. 13, no. 4, pp. 601 - 609, 1994.
42. H. Lee, A. Battle, R. Raina and A. Ng, "Efficient sparse coding algorithms," *Proc. Advances in neural information processing systems,* p. 801–808, 2006.
43. U. o. Washington, "Simulation System for Emission Tomography SimSET," , 2019. [Online]. Available: http://depts.washington.edu/simset/html/contact_info/contact_info_index.html.
44. "QRM Quality Assurance in Radiology and Medicine," 2019. [Online]. Available: http://www.qrm.de/content/kontaktformular/kontaktformular.htm.