J Digit Imaging (2016) 29:394–402
DOI 10.1007/s10278-015-9853-x

Local and Non-local Regularization Techniques in Emission (PET/SPECT) Tomographic Image Reconstruction Methods

Munir Ahmad [1,3] · Tasawar Shahzad [3] · Khalid Masood [1] · Khalid Rashid [2] · Muhammad Tanveer [3] · Rabail Iqbal [3] · Nasir Hussain [1] · Abubakar Shahid [1] · Fazal-e-Aleem [3]

Published online: 29 December 2015

Abstract Emission tomographic image reconstruction is an ill-posed problem due to limited and noisy data and various image-degrading effects affecting the data and leads to noisy reconstructions. Explicit regularization, through iterative reconstruction methods, is considered better to compensate for reconstruction-based noise. Local smoothing and edge-preserving regularization methods can reduce reconstruction-based noise. However, these methods produce overly smoothed images or blocky artefacts in the final image because they can only exploit local image properties. Recently, non-local regularization techniques have been introduced, to overcome these problems, by incorporating geometrical global continuity and connectivity present in the objective image. These techniques can overcome drawbacks of local regularization methods; however, they also have certain limitations, such as choice of the regularization function, neighbourhood size or calibration of several empirical parameters involved. This work compares different local and non-local regularization techniques used in emission tomographic imaging in general and emission computed tomography in specific for improved quality of the resultant images.

✉ Munir Ahmad
munirahm@gmail.com

1 Institute of Nuclear Medicine and Oncology (INMOL), New Campus Road, Lahore PC 10068, Pakistan

2 Applied Physics, Pakistan Council of Science and Industrial Research (PCSIR), Ferozpur Road, Lahore, Pakistan

3 Department of Physics, The University of Lahore, Pakistan, 1-KM Raiwind Road, Lahore 54000, Punjab, Pakistan

Keywords Regularization · Tomographic image reconstruction · Ill-posedness · Maximum a posteriori (MAP) reconstruction · Non-local priors

Introduction

Emission tomographic data, in positron emission tomography (PET) and single-photon emission computed tomography (SPECT), are considered to be an ideal line integral of the underlying activity distribution inside an object being imaged and are used to reconstruct images to evaluate different anatomical or physical processes of the human body [1, 2]. However, real data from these scanners are incomplete and noisy and deviate from their true line integral model owing to various physical and detector-based effects [3, 4]. Iterative image reconstruction methods can easily model these effects, as opposed to analytical reconstruction techniques, along with system geometry response, detector response and emission object models. This is generally modelled in the form of an explicitly defined regularization function, which is added to the data modelling term [5–7]. These methods can reduce reconstruction-based noise through better conditioning of the reconstruction problem using an emission object model [1, 8]. A priori knowledge about object distribution properties can easily be included through regularization functions in order to obtain some user-defined characteristics of the resultant images, for example images bearing the highest required user-defined resolution, with an extra control over these properties through these functions [9–11].

Commonly used local regularization methods, such as local quadratic or non-quadratic regularization functions, are based on Gibbs' distribution functions, have a simple implementation form and can easily capture local image properties [8, 12, 13]. However, these functions produce overly smoothed image regions [8], induce staircase or piecewise blocky artefacts and result in contrast loss because they can provide indiscriminate local prior information available in the image and are less efficient due to their local behaviour [12, 14]. Recently, several non-local regularization schemes have been introduced to combat these issues of local regularization [2, 12, 14–20]. Non-local regularization makes use of the global image connectivity and continuity available in the objective image and can prevent, somehow, over-smoothness with robust edge preservation at the same time, without appreciable staircase effects or contrast loss [2, 8, 17–19, 21–23].

This work reviews various regularization techniques used to reduce ill-conditioning of the emission tomographic reconstruction problem with a focused comparison of the local and non-local regularization techniques used to reduce image and reconstruction-based noise and preserve salient image features.

Review

Emission tomographic measurement data, i.e. the number of coincident photon pairs detected by an opposite pair of detectors along a line in a PET system or the gamma ray detection data from a SPECT system, are approximately linearly proportional to the integral of the tracer density along that line as shown in Fig. 1. These data are used to reconstruct images to evaluate different anatomical or physical processes of the human body [24]. Analytical reconstruction methods, such as filtered backprojection (FBP) or Fourier reconstruction (FR), assume the data to have a true line integral relationship with the object distribution function. This line integral model, between the object emission density function $f(x, y)$,defined over a 2D real space $\mathbb{R}^2$ and its ideal projection data $p_\theta(r)$, can be modelled as a function of the radial distance r from the centre of the detector surface at an angle θ along the line $L(r, \theta)$ in polar coordinates as

$$p_\theta(r) = \int_{\mathrm{Line}(r,\ \theta)} f(x,y)\, ds$$
$$p_\theta(r) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\delta(x\cos\theta + y\sin\theta - r)dxdy \tag{1}$$

A complete collection of all projections around the object $\{p_\theta(r) : \theta \in [0, \pi], r \in (-\infty, \infty)\}$ is known as its Radon transform [1, 25]. This analytical model cannot incorporate various physical or detector-based image-degrading effects and the stochastic nature of the data easily and needs to compromise the reconstructed resolution for noise in the resultant images [24]. These effects include attenuation, scatter, scintillation process blur and background noise in PET and SPECT or accidental coincidences, crystal penetration effect and variation in individual detector performance in PET, which tend to make the data noisy and incomplete and force it to deviate from its true line integral representation. This drives the overall analytical reconstruction problem to be inconsistent and difficult to solve numerically [4, 26]. Additionally, all the above-mentioned effects are non-uniform across the spatial span of the system's field of view (FOV) and add further difficulty in defining the response of the system or its point spread function (PSF) [3, 27].

Iterative reconstruction methods can easily model different image-degrading effects, detector geometrical models, the stochastic nature of the emission process and modelling of the emission object [7]. These methods are the optimization algorithms and iteratively optimize an objective function to estimate the final image.

$$\varnothing\ (f) = \ln g\ (y)$$
$$\hat{f} = \max_{f\in\Omega} \varnothing(f) = \max_{f\in\Omega} \{L(g(y))\}. \tag{2}$$

Here, $g(y)$ is some function of the data and, generally, its log is maximized over the solution image set Ω. Two very popular choices for optimization are the least squares function in penalty framework and the likelihood function in Bayesian formulation [8]. Statistical reconstruction methods consider data and object as realizations of random vectors, such that if $y = \{y_i, i = 1, \ldots, M\}$ denotes the data vector and $f = \{f_j, j = 1, \ldots, N\}$ denotes the object, then conditional probability (log-likelihood) for $p(y|f)$, conditioned over the object f, may be considered by the independent Poisson distribution [28]. The maximum likelihood (ML) estimate is a maximizer of the probability of producing the measured data y, given the object f or maximizer of the log-likelihood function given in Eq. (4).

$$p\ (y|f) = \prod_i \frac{\bar{y}_i}{y_i!} e^{-\bar{y}}$$
$$L(y|f) = \sum_{i=1}^{M} y_i \log\left(\bar{y}_i\right) - \bar{y}_i - \log(y!) \tag{3}$$
$$\hat{f}_{\mathrm{MLEM}} = \max_{f\in\Omega} \{L(y|f)\}\ .$$

This requires the mean of the data modelled as $\bar{y}_i = \sum_j H_{ij} f_j$ with H_{ij} as the system matrix element defining probabilities of event detection and can also include other factors and models [4]. Though the ML solution may exist and achieve the lowest variance among all unbiased estimators, in general, the equation represents an inconsistent system and for finite data, the ML estimate, being a biased estimator, still exhibits high variance due to these problems and cannot acheive the least possible variance. Also with an increase in the number of voxels or the number of iterations, this variance increases

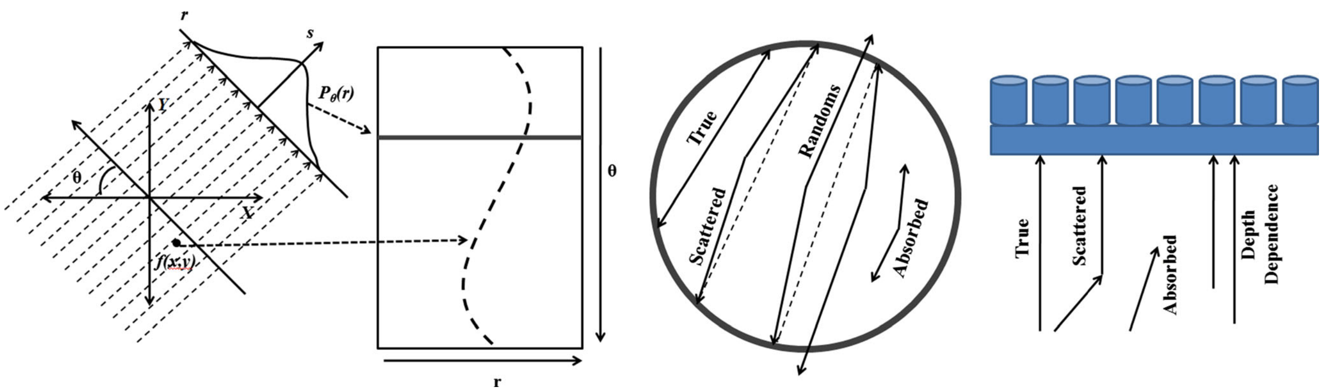

Fig. 1 A projection set of an object on a detector continuum is a function of radial distance [3]

continuously and indefinitely [7]. These equations represent an un-regularized form of the optimization method, and in ECT, an un-regularized image estimate turns out to be noisy due to its ill-conditioning, owing to limited and noisy data [10, 11, 29]. Hence, although an accurate system model may reduce spatial variations, further conditioning of the problem is needed to reduce reconstruction-based noise with a reduced ill-conditioned nature [10, 18, 20, 30].

Instead of simple deterministic iterative methods, regularized iterative methods are used to solve ill-posed reconstruction problems because the behaviour of the deterministic methods is often unsatisfactory for the noisy data. The use of regularization improves the performance of the reconstruction algorithm using local irregularity measurement to restrict the solution image set. An explicit form of the regularization is preferably implied to overcome ill-conditioning of the problem and to gain flexibility over the reconstructed image properties [31]. The objective function to be maximized will become in this case

$$\begin{aligned} \varnothing(f) &= \ln g(y) + \ln g(f) \\ \hat{f} &= \max_{f \in \Omega} \{L(g(y)) + \ln g(f)\}. \end{aligned} \tag{4}$$

Here, $g(f)$ is the regularization function. In Bayesian formulation, the data term $g(y)=p(y|f)$ and the object prior function $g(f)=p(f)$ are combined through Bayes' rule, which turns out to be computing the maximum a posteriori (MAP) estimate in its logarithmic form [8]:

$$\begin{aligned} p\left(f\middle|y\right) &= \frac{p\left(y\middle|f\right)p(f)}{p(y)} \varnothing(f) = \ln p\left(f\middle|y\right) = \ln p\left(y\middle|f\right) \\ &+ \ln p(f) \hat{f}_{\mathrm{PLEM}} = \max_{f \in \Omega} \{L(y(f)) + \beta\, U(f)\} \\ &\text{where } U(f) = \ln g(f) \hat{f}_{\mathrm{MAP}} = \max_{f \in \Omega} \left\{L\left(y\middle|f\right) + \beta\, U(f)\right\} \\ &\text{where } U(f) = \ln p(f) \end{aligned} \tag{5}$$

In this equation, $\hat{f}$ is known as the MAP estimate or penalized likelihood (PL) estimate. In the penalty framework, the first term is called the data term, whereas the second one is termed as the penalty function and the estimator is called the PL estimator [7]. In both these forms, the purpose of this term is to impose additional constraints on the solution image set Ω to improve ill-conditioning of the reconstruction problem and to select those solutions only, which satisfy the data term $L(y(f))$ and the prior term $U(f)$ at the same time, with its associated weight β.

Prior Design

The purpose of the prior or the penalty term is to select those solutions which satisfy the data and which are the most probable in our prior expectations [5, 29, 32–34]. The simplest form of these functions assumes statistical independence between voxels so that the images are products of uni-variate grey levels, such as independent gamma prior [7]:

$$p(f) = \prod_i \frac{1}{\Gamma(\alpha_i)} \left({}^{\alpha_i}/_{\beta_i}\right)^{\alpha_i} f_i^{\alpha_i - 1} \exp\left(-{}^{\alpha_i f_i}/_{\beta_i}\right), \tag{6}$$

where β_j is the mean and α_i is the coefficient of variation for the ith voxel f_i. These are smoothing functions and due to no coupling between voxels, we get a closed-form solution. However, they operate through the mean image values, which are unfortunately not available in image reconstruction, beforehand, and computation of the mean image also induces coupling [35–37]. These priors work on the basic description of the medical images that they are locally smooth and random noise can be tackled, if images are locally smoothed. A quadratic Gaussian model works well to overcome noise and mathematically has very nice forms with continuous derivatives to make the objective function easier to optimize; however, it also blurs salient image features such as edges [8, 21].

Local Prior Design

Instead of independent prior models, images may be drawn from a bigger class of Markov random fields (MRFs), in the form of Gibbs distribution functions, to empirically capture local image properties [12, 28]. These are exponential functions in a simple mathematical formulation, and the exponents are the sum of functions of neighbouring voxel differences. Potential function is a very critical choice in defining the final image properties, and the form of the prior term will ultimately decide the behaviour of the overall objective function to be optimized [8, 12]. These priors are commonly defined as

$$p(f) = \frac{1}{Z} e^{-\beta U(f)} \text{ where } U(f) = \sum_{j=1}^{N} \sum_{k \in \text{ neigh}} w_{jk} V\left(f_j - f_k\right) \quad (7)$$

Here, Z is the normalization constant and $U(f)$ is the Gibbs energy function. Various forms of these potential functions attempt to capture characteristics of the tracer distribution in a local vicinity of an image voxel, while allowing abrupt changes in the intensity at regional edges at the same time, for example the Huber prior or the log-cosh prior [8]. A convex form of the functions is used commonly for easiness of optimization [28]. Weights w_{jk} define the level of interaction of the jth voxel with the kth voxel, and in locally defined neighbourhoods, these are taken constants [21]. Parameter β defines the overall weight of the prior function, and for low β values, the prior does have a very weak influence on the solution image, while large β values mean that the solution satisfies the prior constraints with more force [34]. Higher order neighbourhoods may be included in a local prior form; however, modelling and use of the higher order neighbourhoods and the computational burden are computationally prohibitive [7].

In local priors, the value assigned to each voxel depends on the voxel intensities within a small fixed neighbourhood. Another choice for the prior function could also be a generalized image description-based function or an anatomically defined image of the object [38]. The difficulty with this approach is that generally the prior information about the object is not known beforehand. However, some sort of relational information about the object could be used; for example, in emission tomography, the simplest known information about the emission object is its non-negativity. In local priors, potential functions are selected to reflect the local smoothness and to preserve regional boundaries, which are two contradicting properties, and different forms of the prior functions have been introduced in the literature to capture both of them together [7, 8, 12, 14, 32, 37, 39, 40]. Functions with higher order neighbourhoods have also been proposed to capture more complex correlation structures, with having difficulty of choosing and defining such model, which is mathematically not so convenient, for example the thin-plate spline model. Instead of an implicit boundary model, the compound MRFs or the weak membrane priors are attempts to model the boundaries explicitly by a dual lattice of line sites. Higher resolution imaging modalities can also be used to incorporate anatomical priors into the reconstruction framework [14]. Most of these prior functions are convex for mathematical elegance and have their well-defined maxima, whereas non-convex prior functions have also been introduced. When the prior function is incorporated into the objective function in a convex form, the objective function becomes concave and a Kuhn-Tucker necessary condition is also sufficient to find the maxima [3]. Another major nuisance with these methods is the absence of any general criterion to select a weight for the regularization function as compared to the data likelihood function for optimization. This weight is commonly defined in the form of a regularizing parameter whose value defines the influence of the prior function. This parameter, unlike the cut-off frequency in FBP, does not have any physical units and is difficult to evaluate. Several methods have been proposed to tune this parameter much like a tabulation of cut-off frequencies against resolution [34]. In smoothing priors, a convex formulation of the cost functions is implemented to better condition the reconstruction problem, for example quadratic priors (QPs) below, where a linear solver may be opted to solve the problem with faster convergence rates.

$$p(f) = \frac{1}{Z} e^{-\beta U(f)} \text{ with } U(f) = \sum_{j=1}^{N} \sum_{k \in \mathcal{N}_j} w_{jk} \frac{\left(f_j - f_k\right)^2}{2} \quad (8)$$

These priors assume that images are locally smooth, and plenty of work is based on quadratic regularization due to their simple implementation and good smoothing properties to reduce noise. However, their smoothing is anisotropic and depends on various factors such as activity concentration distribution, system geometry and non-uniform attenuation distribution. Few methods have been proposed to recover for their non-uniform response; however, smoothness of edges and high-count areas still remains a problem due to local behaviour of the regularizing priors [3]. Various non-quadratic or edge-preserving priors try to empirically change the behaviour of the prior function near the edges, using some additional parameter to define edge thresholds, for example the Huber prior or the Geman and McClure penalty functions [8, 28]; however, they make the objective function non-convex and cannot improve the ill-conditioning much as the quadratic priors and their convergence properties are generally not well defined and their results are much sensitive to the hyper-parameter used. Several such edge-preserving priors have been proposed; however, they are computationally intensive and use complex edge-defining techniques such as deterministic annealing or the method of level sets. These methods use computationally complex techniques to somehow preserve the edges empirically and can only incorporate local image properties [7].

Quadratic and non-quadratic regularizing functions (except MRPs) work on a basic description of the images that they are locally smooth, whereas MRP- and total variation (TV)-based priors work on local mono-tonicity (voxel values are only increasing or decreasing in a local neighbourhood) and image irregularity (image derivatives in a local neighbourhood) measure, respectively [3, 14]. These priors automatically preserve edges without any additional parameter used and produce smoothing with the least amount of blurring induced, and their non-linear response is less sensitive to the hyper-parameter values. However, their convergence properties are not known due to their undefined analytical derivatives and resultant images suffer from staircase artefacts and contrast loss due to their local behaviour. TV regularization functions have also been suggested for electrical impedance tomography (EIT) to preserve sharp discontinuities at edges, which are generally smoothed out by standard quadratic regularization [3]. However, for non-differentiability of the leading objective function, the lagged diffusivity and primal dual-interior point methods have been proposed to solve it, which showed improved results, though with additional mathematical complexity [40, 41].

All the above-mentioned local regularization prior functions perform very well to reduce reconstruction-based noise and to improve conditioning of the problem; however, they have limitations. For example, local quadratic priors over smooth edges and high-count regions have hyper-parameter sensitivity problems and produce anisotropic and asymmetrical responses. Non-quadratic priors can retain edges; however, they cannot improve the ill-conditioned problem very well, need additional parameters to define edges empirically, induce locally unpredictable smoothing behaviour and have convergence issues. MRP- and TV-based priors have robust edge preservation properties. However, their analytical derivatives are not available; hence, their convergence properties are not defined and they produce staircase artefacts, piecewise blocky artefacts and contrast loss as shown in Fig. 2. Most of these problems are due to the local dependence of the prior, constant prior weighting in local neighbourhood, non-convexity of the mathematical formulation or hyper-parameter sensitivity of the response of the system. To overcome these problems, non-local regularization priors have been proposed recently.

Non-local Regularization

Local priors provide only local information, where image roughness or variational information of the structures or texture available in the image is calculated considering the voxel differences in a local neighbourhood only, which is not robust [8, 42]. Non-local regularization techniques have been developed to include image regularization based on voxels exterior to the local neighbourhoods and are theoretically meant to include global connectivity and continuity present in the

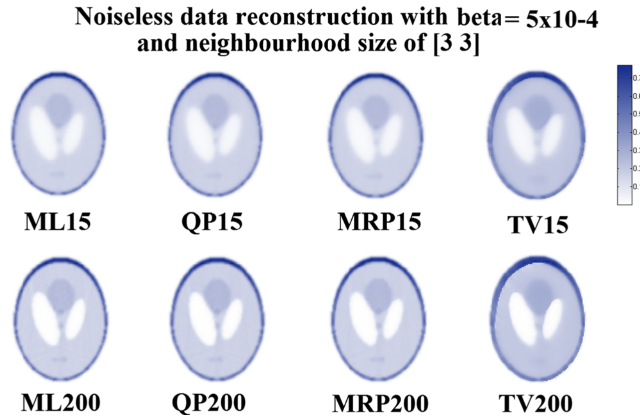

Fig. 2 Images of a thorax phantom reconstructed with ML, PL with QPs, MP and TV priors for 15 and 200 iterations for the given *0* value. QPs produce blurry images whereas TV priors produce blocky artefacts [3]

objective image [12, 13, 17, 18, 28, 32, 38, 39, 42]. The start of the non-local regularization techniques was mainly inspired by Buades and colleagues with an introduction of the non-local means filtering for image de-noising [21]. Many non-local regularization functions have been introduced for image de-noising [23] and image reconstruction [12] afterwards with efforts devoted to find a generalized solution for inverse problems using non-local regularization [18, 43]. Theoretically, three basic choices need to be made for local or non-local regularization to be implemented, i.e. mathematical formulation of the prior energy function, the size of the neighbourhood used and the form of the weight function used to evaluate weights for different pixels or neighbourhoods. Prior energy function, given in Eq. (6), can be re-written for non-local regularization as the sum of the potential functions with non-local weighting introduced as follows:

$$U_{NL}(f) = \sum_{j=1}^{N} \sum_{k \in \mathcal{N}_j} \varnothing_{jk} V\left(f_j - f_k\right) \tag{9}$$

Here, $\varnothing_{jk}$ are the weight functions and define the contribution of the kth patch or neighbourhood with respect to the jth patch and $V(f_j - f_k)$ are the potential functions of the voxel intensity differences in various patches or non-local regions in the image as shown in Figs. 3 and 4.

Non-local prior design may use energy functions in the form of quadratic functions of the voxel intensity differences [12, 33] or quadratic function with the Huber prior function [8], the TV roughness function [44], non-local median priors [14] or some proposed optimization transfer functions [31]. For example, a dynamic PET image cluster-based non-local prior, with a quadratic potential function, has been proposed with equally normalized weights using the total number of voxels in a cluster, and it was shown that the prior resulted in a quantitatively better parametric image in dynamic PET [42]. Normalized standard deviation (NSD) versus bias trade-

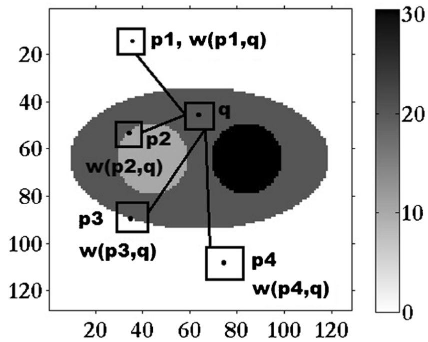

Fig. 3 Different pixel neighbourhoods get different weights based on the similarity of the pixel neighbourhoods. For example, point p2 has higher weight associated with it than pixels p1 and p4

off method was used to optimize the number of clusters used. Assuming that intensity difference-based local priors cannot calculate image roughness accurately, non-local patch-based regularization has been suggested with a non-quadratic function, achieving better contrast levels for smaller objects with lower parameter dependency as compared to MRPs [31]. Authors claim that their method is more robust in distinguishing edge pixels from noisy pixels as compared to other local non-quadratic priors; however, it is based on a new optimization transfer algorithm, similar to that of PLEM reconstruction, and is computationally complex. A non-local TV prior has been proposed, which minimizes a linear combination of non-local TV and least squares fitting term with the split Bregman algorithm and has been shown to produce a better signal to noise ratio and recovers staircase artefacts and contrast loss artefacts in MRI image reconstruction [32].

An extension of the variational framework to a non-local TV prior with a convex form of the prior to allow flexibility has been proposed [22] and for MR [32] to evade their staircase artefacts for local TV regularization. A coherence regularization prior has been introduced, with an energy function based on the diffusion equation of adaptive image filtering, to directly extract image features from noisy data, and structure detection was carried out using a non-local means filter [38]. This way, the intra-structure intensity changes, which contribute towards local noise, are penalized and salient features such as edges are preserved. A non-local prior for Bayesian image reconstruction in PET and SPECT has been proposed to improve the reconstructed images by exploiting information in larger neighbourhoods in the image [12, 31]. Though various functional forms have been used as the prior energy function, the simple forms make the overall objective function convex and easy to optimize, whereas it is difficult to mathematically model the dependence of the complex functions on the image.

Conventionally, the value of the prior function is calculated, as a weighted sum of the potential functions of voxel differences in a local neighbourhood [14]. These weights denote the degree of interaction between neighbouring voxels, and for local priors, they are assumed to be constants equal to the inverse of the distance between voxel centers. It was observed that constant prior weighting in a local neighbourhood produces anisotropic and asymmetric responses of the system with local priors. A non-local prior, mentioned above, has further been extended and applied to low-dose X-ray computed tomography with an improved adaptive weighting approach [8]. It has been shown that this prior can exploit selectively global image information with effective resolution recovery and noise removal as compared to the local priors and can remove negative regularization by excluding irrelevant neighbourhood pixels. Improved adaptive weighting presented a similarity measure between neighbouring patches and was calculated using the latest estimated image as a decreasing Euclidean distance between two neighbouring patches. Strategies to calculate weights are broadly based on similarity measures between neighbourhoods [14, 31], maximum entropy measure [28], non-local means affinity function [21] or some simple measure of the inverse of the distance between voxels [12]. For example, the following equation presents a method to evaluate weight function in a non-local neighbourhood based on similarity measured by a decreasing function of the

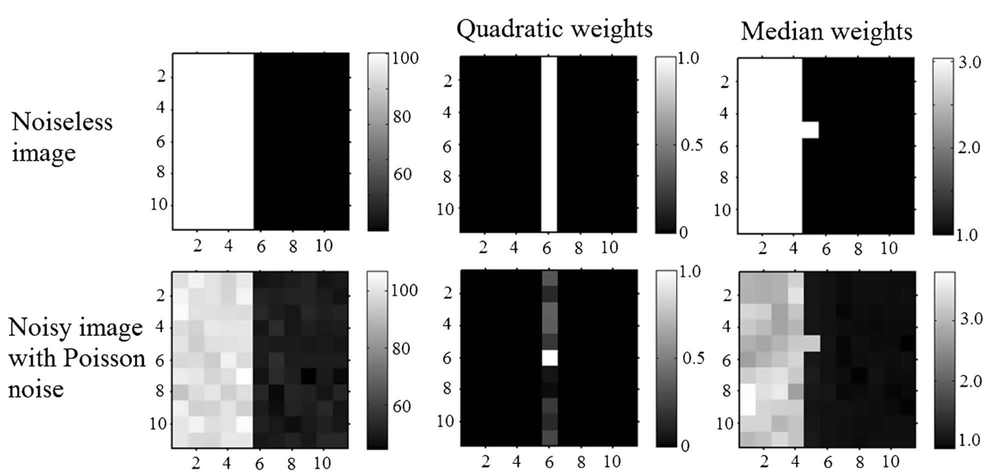

Fig. 4 A 11 × 11-pixel edge image (*top left*) with its weight distribution calculated using a smaller window of size 3 × 3 (*bottom left*). Weights are reasonably distributed along the edge. However, in a noisy version (*top right*) where Poisson noise has been added, the weights (*bottom right*) are distributed randomly. This indicates that noisy patches (*smaller windows*) may not bear well-distributed weights and may cause further noise

Gaussian-weighted Euclidian distance (L_2 norm convolved with a Gaussian kernel with standard deviation a) between two smaller neighbourhood vectors x_j and x_k of size N in a larger window of size M.

$$w_{jk} = \frac{1}{Z(j)} e^{-\|x_j - x_k\|_{2,a}^2 / h^2} \tag{10}$$

Here, $Z(j)$ is the normalization factor and h^2 controls the decay of the function and is a user-defined empirical parameter [14]. A cluster-based non-local prior for dynamic PET defines weights derived from similar kinetics of clusters where equally normalized weights have been used with respect to the total number of voxels in a cluster [42]. However, Euclidean distance-based weighting has also been introduced to overcome bias due to noise. A non-local TV (NLTV) function uses a non-local weight function based on graph function with an assumption that the weight function should represent a similarity of two points with their significance represented by the weight [32]. A patch similarity mixture (PSM) prior model, as a high-level MRF prior to model higher order neighbourhood patterns, has been proposed for CT data extending the same non-local prior [28], where prior weights are calculated using a constrained entropy maximization and an automatic balance between data and prior was also introduced [28].

A basic problem with the local priors is to smooth out the uniform regions while recovering the edges at the same time. Practically, adaptive prior weights are found to be aligned along the edges inside the image and hence better define priors' behaviour for edge preservation and uniform region smoothing. This signifies the edge-preserving nature of the non-local priors using adaptive non-local image information. Similarly, there is not a consensus to use the size of the larger neighbourhood or the search window [12] used around the voxel of interest. Technically, smaller neighbourhoods cannot capture global image information, while larger neighbourhoods increase computational burden a lot, though an empirically defined compromise has been introduced [28, 31]. In patch-based non-local priors, suggestively, few of the patches corrupted by noise receive smaller weighting as compared to the less noisy patches. A non-linear anisotropic structure tensor (NAST)-based prior has been introduced to correct for this problem [13]. Figure 2 shows a comparison of normalized means weights and median weight distribution for an image corrupted by Poisson noise.

Non-local priors may include geometric and continuity information in the image, while they can better define edges or texture and can overcome limitations of local priors; still, they have certain problems associated with their design. Non-local priors have complex mathematical formulations, and generally, no straightforward model of the prior and its dependence on the image is defined [8]. Mathematical complexity comes in two different ways, i.e. dependency of the prior on the image intensities is not well modelled and results in an inconsistent posterior function. Similarly, their analytical derivatives are not well defined and the convexity of the prior function may not be known [45, 46]. It is, theoretically, not easy to address convergence properties of some non-local priors due to these reasons [18, 37]. Non-local prior formulations are also computationally expensive, and almost in every proposed non-local prior, some method to speed up the reconstructions is suggested as a correction. Computational requirements increase with larger patch sizes, and one has to resort to a compromise between patch size and computational times [8, 28]. Non-local priors, similar to many local priors, require several parameter adjustments and face parameter sensitivity [8, 28, 34]. For example, non-local priors introduced for image denoising or for Bayesian tomographic reconstruction require several parameters which require manual adjustments [8].

Most of these recently introduced regularization techniques are non-quadratic and non-linear and are computationally expensive due to their complex and heavy computational requirements. Implementation of these algorithms has been made possible by increasing the speed and storage capabilities of the recent advances in computer systems. All these regularization methods are meant to achieve low noise and lesser artefacts, whereas a task-based comprehensive and multi-application-based analysis of these techniques is required to fully exploit their abilities, because an empirical and limited analysis for a limited number of applications has been carried out till now.

Conclusions

Tomographic image reconstruction problems are an ill-posed mathematical problem and are affected by noise and various image quality-degrading effects. Explicit regularization is implied to alleviate the problem, in the form of locally defined prior functions. However, local priors can only accommodate local image properties and are mainly categorized as smoothing or edge-preserving priors. Smoothing priors produce overly smoothed images with smoothed edges, whereas edge-preserving local priors result in staircase artefacts with piecewise patchy blocks. To overcome these issues of local priors, recently, non-local priors have been proposed. Non-local priors address the global image connectivity and continuity and lead to images with better characteristics. Certain choices need to be made and optimized to implement non-local priors, such as prior energy function, form of the weight function, size of the neighbourhood and the hyper-parameter used. Unfortunately, non-local priors are based on computationally expensive and mathematically complex forms and have undefined convergence properties. Similarly, their dependence on image intensities is not well modelled, and in case of larger non-local neighbourhoods, the problem may become intractable in real-time studies. The design considerations of non-local priors are random, and comparison of various techniques, found in literature, is empirical and limited to a few particular applications or methods. A comparative analysis of these regularization methods to find their commonalities, other than being non-local, is required. Non-local regularization seems to promise better image quality and enhanced accuracy; however, the comparative analysis is limited to few noise types and applications. We may conclude that non-local, non-quadratic and non-linear regularization is more favoured, in recent works, due to fast computing facilities available now, where various regularization schemes have been introduced in order to combat artefacts produced by the previous local regularization methods. However, an option for a universal technique to provide the best image quality with the fastest speed and to work for different imaging systems is still missing.

References

1. Bardsley JM, Calvetti D, Somerslab E: Hierarchical regularization for edge-preserving reconstruction of PET images. Inverse Probl 26:035010, 2010
2. Joyita D, Richard ML, Quanzheng L: Non-local means denoising of dynamic PET images. PLoS One 8:1–15, 2013
3. Ahmad M: Non-uniform resolution and partial volume recovery in tomographic image reconstruction methods. PhD Thesis, University College London (UCL), London, UK, 2014
4. Chong C, Guoliang X: A new reconstruction algorithm in tomography with geometric feature-preserving regularization. 3rd International Conference on Biomedical Engineering and Informatics (BMEI), Yantai, China: 41–45, 2010
5. Funama Y, Taguchi K, Utsunomiya D, Oda S, Katahira K, Tokuyasu S, Yamashita Y: Image quality assessment of an iterative reconstruction algorithm applied to abdominal CT imaging. Phys Med 30:527–534, 2014
6. Ghetti C, Ortenzia O, Serreli G: CT iterative reconstruction in image space: A phantom study. Phys Med 28:161–165, 2012
7. Qi J, Leahy R: Iterative reconstruction techniques in emission computed tomography. Phys Med Biol 51:R541, 2006
8. Chen Y, Gao D, Nie C, Luo L, Chen W, Yin X, Lin Y: Bayesian statistical reconstruction for low-dose X-ray computed tomography using an adaptive-weighting nonlocal prior. Comput Med Imaging Graph 33:495–500, 2009
9. Karali E, Koutsouris D: Towards novel regularization approaches to PET image reconstruction. J Biosci Med 1:6–9, 2013
10. Titarenko V, Bradley R, Martin C, Withers PJ, Titarenko S: Regularization methods for inverse problems in x-ray tomography. Proceedings of SPIE San Diago, CA: 78040Z-780410, 2010
11. Weihui D: Approach to Solve Ill-posed Inverse Problem in Image Reconstruction from Projections. Proceedings of the International Symposium on Information Processing Huangshan, China: 338–341, 2009
12. Chen Y, Ma J, Feng Q, Luo L, Shi P, Chen W: Nonlocal Prior Bayesian Tomographic Reconstruction. J Math Imaging Vis 30: 133–146, 2008
13. Li Y: PET Bayesian Tomography Improved Nonlocal Prior. Bull Adv Technol Res 5:61–64, 2014
14. Chan C, Fultin R, Feng DD, Meikle S: Median non-local means filtering for low SNR image denoising: Application to PET with anatomical knowledge. IEEE Nuclear Science Symposium and Medical Imaging Conference Knoxville, TN, USA: 3613–3618, 2010
15. Chun SY, Fessler JA, Dewaraja YK: Post-reconstruction non-local means filtering methods using CT side information for quantitative SPECT. Phys Med Biol 58:6225–6240, 2013
16. Dutta J, El Fakhri G, Yanguang L, Chuan H, Petibon Y, Reese TG, Leahy RM, Quanzheng L: Spatially varying regularization for motion compensated PET reconstruction. IEEE Nuclear Science Symposium and Medical Imaging Conference Anaheim, CA: 2156–2160, 2012
17. Gilboa G, Osher S: Nonlocal Linear Image Regularization and Supervised Segmentation. Multiscale Model Simul 6:595–630, 2007
18. Peyre G, Bougleux S, Cohen L: Non-local regularization of inverse problems. Inverse Probl Imaging 5:511–530, 2011
19. Xin L: Nonlocal Regularized Algebraic Reconstruction Techniques for MRI: An Experimental Study. Math Probl Eng 2013:1–11, 2013
20. Zhang X, Burger M, Bresson X, Osher S: Bregmanized Nonlocal Regularization for Deconvolution and Sparse Reconstruction. SIAM J Imaging Sci 3:253–276, 2010
21. Buades A, Coll B, Morel JM: A non-local algorithm for image denoising. Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on 60-65
22. G. Gilboa, J. Darbon, S. Osher, and T. F. Chan, Nonlocal convex functionals for image regularization, UCLA,CAM Report 06-57, 2006
23. Gilboa G, Osher S: Nonlocal Operators with Applications to Image Processing. Multiscale Model Simul 7:1005–1028, 2008
24. Rahmim A, Zaidi H: PET versus SPECT: strengths, limitations and challenges. Nucl Med Commun 29:193–207, 2008
25. Tianfang L, Jiangsheng Y, Junhai W, Zhengrong L: Inversion of the attenuated Radon transform for non-parallel geometries. IEEE Nuclear Science Symposium and Medical Imaging Conference Rome: 2525–2529, 2004
26. Wang Y, Yang J, Yin W, Zhang Y: A New Alternating Minimization Algorithm for Total Variation Image Reconstruction. SIAM J Imaging Sci 1:248–272, 2008
27. Ahmad M, Todd-pokropek A: Non-uniform resolution compensation in tomographic image reconstruction methods using median based priors. Nuclear Science Symposium Conference Record, 2007. NSS ‘07. IEEE 4447–4451
28. Chen Y, Li Y, Yu W, Luo L, Chen W, Toumoulin C: Joint-MAP Tomographic Reconstruction with Patch Similarity Based Mixture Prior Model. Multiscale Model Simul 9:1399–1419, 2011
29. Beister M, Kolditz D, Kalender WA: Iterative reconstruction methods in X-ray CT. Phys Med 28:94–108, 2012
30. Thurman ST, Fienup JR: Noise histogram regularization for iterative image reconstruction algorithms. J Opt Soc Am A Opt Image Sci Vis 24:608–617, 2007
31. Guobao W, Jinyi Q: Penalized Likelihood PET Image Reconstruction Using Patch-Based Edge-Preserving Regularization. IEEE Trans Med Imaging 31:2194–2204, 2012
32. Gopi VP, Palanisamy P, Wahid KA, Babyn P: MR image reconstruction based on iterative Split Bregman algorithm and nonlocal total variation. Comput Math Methods Med 2013:985819, 2013
33. Guobao W, Jinyi Q: Patch-based regularization for iterative PET image reconstruction. International Symposium on Biomedical Imaging University of California, USA: 1508–1511
34. Ramani S, Zhihao L, Rosen J, Nielsen JF, Fessler JA: Regularization Parameter Selection for Nonlinear Iterative Image Restoration and MRI Reconstruction Using GCV and SURE-Based Methods. IEEE Trans Image Process 21:3659–3672, 2012

35. Ahmad M, Todd-pokropek A: Comparison of noise performance for Penalized-Likelihood image reconstruction methods with uniform resolution estimators. IEEE Nuclear Science Symposium and Medical Imaging Conference Dresden, Germany: 4169–4171, 2008
36. Ahn S, Leahy RM: Analysis of resolution and noise properties of non-quadratically regularized image reconstruction methods for PET. IEEE Trans Med Imaging 27:413–424, 2008
37. Willett RM, Harmany ZT, Marcia RF: Poisson image reconstruction with total variation regularization. IEEE International Conference on Image Processing Orlando, FL: 4177–4180, 2009
38. Fang S, Ying K, Zhao L, Cheng J: Coherence regularization for SENSE reconstruction with a nonlocal operator (CORNOL). Magn Reson Med 64:1413–1425, 2010
39. Daniel FY, Fessler JA: Edge-Preserving Tomographic Reconstruction with Nonlocal Regularization. IEEE Trans Med Imaging 21:159–173, 2002
40. Tang J, Nett BE, Chen GH: Performance comparison between total variation (TV)-based compressed sensing and statistical iterative reconstruction algorithms. Phys Med Biol 54:5781–5804, 2009
41. Babacan SD, Molina R, Katsaggelos AK: Total variation super resolution using a variational approach. IEEE International Conference on Image Processing Dresden, Germany: 641–644, 2008
42. Lu L, Karakatsanis NA, Tang J, Chen W, Rahmim A: 3.5D dynamic PET image reconstruction incorporating kinetics-based clusters. Phys Med Biol 57:5035–5055, 2012
43. Bougleux S, Rahmim A, Cohen L: Non-local regularization of inverse problems. European Conference on Computer Vision (ECCV'08) Marseilles, France
44. Borsic A, Graham BM, Adler A, Lionheart WR: In vivo impedance imaging with total variation regularization. IEEE Trans Med Imaging 29:44–54, 2010
45. Ahmad M, Todd-pokropek A: Evaluation of resolution properties of median priors in list-mode reconstruction methods for PET. IEEE Nuclear Science Symposium and Medical Imaging Conference Orlando, FL: 2853–2856, 2009
46. Alenius S, Ruotsalainen U, Astola J: Using local median as the location of the prior distribution in iterative emission tomography image reconstruction. IEEE Trans Nucl Sci 45:3097–3104, 1998