

RESEARCH ARTICLE

A Dictionary Learning Approach with Overlap for the Low Dose Computed Tomography Reconstruction and Its Vectorial Application to Differential Phase Tomography

Alessandro Mirone¹, Emmanuel Brun^{1,2*}, Paola Coan^{2,3}

1. European Synchrotron Radiation Facility, BP 220, F-38043, Grenoble, France, 2. Ludwig Maximilians University, Physics Department, 85748, Garching, Germany, 3. Institute for Clinical Radiology, Ludwig-Maximilians-University Hospital Munich, Munich, Germany

*emmanuel.brun@esrf.fr



CrossMark
click for updates

 OPEN ACCESS

Citation: Mirone A, Brun E, Coan P (2014) A Dictionary Learning Approach with Overlap for the Low Dose Computed Tomography Reconstruction and Its Vectorial Application to Differential Phase Tomography. PLoS ONE 9(12): e114325. doi:10.1371/journal.pone.0114325

Editor: Joseph Najbauer, University of Pécs Medical School, Hungary

Received: March 3, 2014

Accepted: November 6, 2014

Published: December 22, 2014

Copyright: © 2014 Mirone et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was partially supported by the DFG-Cluster of Excellence Munich-Centre for Advanced Photonics EXE158. Other funding is internal ESRF funding. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

Abstract

X-ray based Phase-Contrast Imaging (PCI) techniques have been demonstrated to enhance the visualization of soft tissues in comparison to conventional imaging methods. Nevertheless the delivered dose as reported in the literature of biomedical PCI applications often equals or exceeds the limits prescribed in clinical diagnostics. The optimization of new computed tomography strategies which include the development and implementation of advanced image reconstruction procedures is thus a key aspect. In this scenario, we implemented a dictionary learning method with a new form of convex functional. This functional contains in addition to the usual sparsity inducing and fidelity terms, a new term which forces similarity between overlapping patches in the superimposed regions. The functional depends on two free regularization parameters: a coefficient multiplying the sparsity-inducing L_1 norm of the patch basis functions coefficients, and a coefficient multiplying the L_2 norm of the differences between patches in the overlapping regions. The solution is found by applying the iterative proximal gradient descent method with FISTA acceleration. The gradient is computed by calculating projection of the solution and its error backprojection at each iterative step. We study the quality of the solution, as a function of the regularization parameters and noise, on synthetic data for which the solution is a-priori known. We apply the method on experimental data in the case of Differential Phase Tomography. For this case we use an original approach which consists in using vectorial patches, each patch having two components: one per each gradient component. The resulting algorithm, implemented in the European Synchrotron Radiation Facility tomography

reconstruction code PyHST, has proven to be efficient and well-adapted to strongly reduce the required dose and the number of projections in medical tomography.

Introduction

During the past two decades X-ray Phase Contrast Imaging (PCI) has shown a remarkable enhancement of image contrast and sensitivity for soft tissue. Reducing the deposited dose during PCI-CT is a crucial step towards an eventual clinical implementation of the technique.

A solution to this problem consists in applying iterative CT reconstruction schemes with a-priori knowledge of the solution. The signals occurring in Nature, when cleaned from noise, present most of the time an intrinsic sparsity when expressed in the proper basis. An image is intrinsically sparse when it can be approximated as a linear combination of a small number n of basis functions, with $n \ll N$, where N is the image dimensionality. Piece-wise constant images when they are expressed by their gradient are examples of sparse signal: they have non-zero signal only at the borders of flat regions. For piece-wise constant images one can apply very efficient methods based on minimization of a convex functional, called also convex objective function, which contains a total variation penalty term. For other classes of images, such as medical images, one has to choose different solutions which are adapted to the intrinsic sparsity of the case under study (depending on the specific organ and imaging modality). There are mainly two ways: either the sparsity structure is a-priori known and an appropriate basis of functions can be built from the beginning, or it must be automatically learned from a set with the dictionary learning technique [1]. This method consists in building an over-complete basis of functions, over an $m \times m$ domain, such that, taken an $m \times m$ patch from the studied image, the patch can be approximated with good precision as a linear combination of a small number $N \ll m^2$ of basis functions. The rationale for using an over-complete basis is that by increasing the basis dimension one increases the number of different patterns that can be fitted using just one or few basis functions. We can think as an example of images containing isolated and weakly curved lines: in this case we could use a basis where each function represents a line with a given intercept and slope, but other functions could be further introduced to fit other shapes. When we fit, patch by patch, a noisy image using the appropriate basis, the features of the original images will be accurately fitted with a small number of components. The noise instead has in general no intrinsic sparsity, and if it happens to have one, it is with high probability very different from the sparsity structure of the original images. Therefore the noise will be reproduced only if we allow a large number of components (the patch basis is over-complete so it can represent the noise) while it will be effectively filtered out if we approximate the noisy image with a small number of components.

The dictionary learning technique has recently been applied to tomography reconstruction using the Orthogonal Matching Pursuit (OMP) denoising procedure [2]. This procedure consists in obtaining first an over-complete basis of functions and then in least-square fitting every patch of the image using at most N_{omp} components selected from this basis. The components are heuristically selected choosing, each time, the one having the maximum overlap with the remaining error. This optimization method cannot be implemented as a convex objective function optimization problem, because the linear combination of two candidate solutions can have more than N_{omp} components. In other words, the optimization domain is not convex.

In this paper we present an advanced formalism which implements overlapping patches into a new convex functional that we describe in the section Materials and Methods. For the solution of our functional minimization problem we have applied the recently developed tools taken from the field of convex optimization [3]. Results on both synthetic and experimental data are compared to the state of the art reconstruction methods. We compared the obtained images to Equally Sloped Tomography (EST) [4], TV minimization [5]. Moreover we applied convex optimisation also to another formulation, for the dictionary learning technique, of the objective function, which has been applied, using the non-convex OMP procedure, by Xu et al. [1].

Methods

Dictionary Learning

In this section we introduce first the decomposition of an image into non-overlapping patches and the related objective function for denoising. Then we introduce our original formalism which ensures, using overlapping patches, a smooth transition at the patches borders, and finally we apply this formalism to CT reconstruction. In order to make clearer the text we added the following sentences in the introduction of the method: In this approach an iterative loop between the sinogram space and the real space is used. A fidelity term is imposed in the sinogram space while a sparsity inducing term is introduced in the real space.

We denote by $\mathbf{1}_p$ the indicator function of patch p , which is equal to 1 over the patch support (typically an $m \times m$ square) and is zero elsewhere. For non-overlapping patches, covering the whole domain, we have:

$$\sum_p \mathbf{1}_p(i) = 1 \quad \forall i. \tag{1}$$

where i denotes the pixel position and can be thought as a two-dimensional vector. We are looking for the ideal solution \mathbf{x} that we express by the vector, \mathbf{w} , of its coefficients in the basis of patch functions:

$$x_i = \sum_p \mathbf{1}_p(i) \sum_k w_{kp} \phi_k(i - r_p). \tag{2}$$

where the set $\{\varphi_k\}$ is an over-complete basis of functions over the patch support; r_p is the closest to the origin corner of the patch p , and w_{kp} is the component k,p of vector \mathbf{w} which multiplies the basis function φ_k in the patch p . The denoising problem, given an image \mathbf{y} , consists in finding the minimum of a functional $F(\mathbf{w})=f(\mathbf{w})+g(\mathbf{w})$ which is sum of two terms. The term $f(\mathbf{w})=\|\mathbf{y}-\mathbf{x}\|_2^2$ links the solution to the the data \mathbf{y} . The other term, $g(\mathbf{w})$, contains the a-priori knowledge about the solution. This way of breaking the functional in two terms has his roots in the Bayesian theorem. From a probabilistic point of view the denoising problem consists in finding, given a noisy image \mathbf{y} of an object, the most probable object \mathbf{x} that can generate that image. We represent the object \mathbf{x} through the patches coefficients \mathbf{w} . The Bayes theorem, applied to denoising, states that the conditional probability of \mathbf{w} being the exact object given a measurement \mathbf{y} , is the product of the probability of \mathbf{y} being the measure given the exact solution \mathbf{w} times the a-priori probability of \mathbf{w} .

Assuming gaussian noise, the conditional probability of \mathbf{y} being the measure given the exact solution \mathbf{w} is $exp(-\|\mathbf{y}-\mathbf{x}\|_2^2/(2\sigma^2))$, where \mathbf{x} is expressed through the patches coefficients \mathbf{w} by [equation 2](#). The exact a-priori probability of \mathbf{w} is unknown but we approximate it as $exp(-g(\mathbf{w})/(2\sigma^2))$. This function expresses our a-priori knowledge that a non sparse solution having a high value of the L_1 penalization term $g(\mathbf{w})=\beta\|\mathbf{w}\|_1$ (which is a sparsity-inducing term [\[6\]](#)) has low probability. The most probable solution \mathbf{w}^* is obtained by finding the $F(\mathbf{w})$ minimum:

$$\begin{aligned} \mathbf{w}^* &= \operatorname{argmin}_{\mathbf{w}}(f(\mathbf{w})+g(\mathbf{w})); \\ f(\mathbf{w}) &= \|\mathbf{y}-\mathbf{x}\|_2^2; \quad g(\mathbf{w})=\beta\|\mathbf{w}\|_1. \end{aligned} \tag{3}$$

The solution can be obtained by using the iterative shrinkage thresholding algorithm [\[3\]](#) (ISTA):

$$\mathbf{w}_{n+1}=T_{\beta\gamma}(\mathbf{w}_n-\gamma\nabla f(\mathbf{w}_n)); \quad \mathbf{w}^*=\mathbf{w}_\infty \tag{4}$$

where T_α is the shrinkage operator defined as

$$T_\alpha(\mathbf{w})=\frac{\mathbf{w}}{\|\mathbf{w}\|_2} \max(\|\mathbf{w}\|_2-\alpha,0) \tag{5}$$

and γ is a positive number lesser than the inverse of the Lipschitz condition number L :

$$\gamma\in]0,1/L]. \tag{6}$$

The Lipschitz number L is such that:

$$\|\nabla f(\mathbf{w}_2)-\nabla f(\mathbf{w}_1)\|_2\leq L\|\mathbf{w}_2-\mathbf{w}_1\|_2; \quad \forall\mathbf{w}_1,\mathbf{w}_2. \tag{7}$$

The ISTA algorithm can be accelerated by the Fast Iterative Shrinkage Thresholding [7](FISTA) method.

In its non-overlapping version, the image denoising with patches is able to detect features that are within the field of the patch: if a line crosses the central region of a patch, it will be detected if the basis of functions has been trained to detect such lines. But in the situation where a line intersects only one point in a corner of the patch square, the signal of this point is indistinguishable from that of a noisy point, no matter the dictionary training.

For this reason the patches denoising technique is often used with overlapping patches using post-process averaging [8]. In this case the minimization problem is solved for each patch separately first, and then the averaging is performed in the overlapped regions.

In this study we do not follow this procedure but we add an overlap term into the objective function. We choose a system of patches which covers the whole domain, and we allow for overlapping. In our implementation of the algorithm the set of all patches is formed by a set of non-overlapping $m \times m$ square patches covering the image plus the translated copies that we obtain by (i_x, i_y, s) translational vectors, where i_x, i_y, s are positive integers, s is a constant step size selected by the user and $i_x, i_y < m/s$.

In the case of overlapping patches the sum of all indicator functions is greater or equal to one:

$$\sum_p \mathbf{1}_p(i) \geq 1; \forall i. \tag{8}$$

We define the core indicator functions $\mathbf{1}_p^c$, which indicate the core of the patches, and make a non-overlapping covering:

$$\mathbf{1}_p^c(i) \leq \mathbf{1}_p(i); \quad \sum_p \mathbf{1}_p^c(i) = 1; \forall i. \tag{9}$$

For our implementation, when the translational step size s is equal to one, the core region is a pixel at the center of the patch. The core region gets larger when the step size increases.

For a given point i , $\mathbf{1}_p^c(i)$ indicates which patch p has its center C_p closest to point i :

$$\sum_p \mathbf{1}_p^c(i) \|i - C_p\|_1 \leq \|i - C_{p'}\|_1; \quad \forall p', i. \tag{10}$$

The solution \mathbf{x} is composed as a function of \mathbf{w} using the central part of the patches as indicated by the functions $\mathbf{1}_p^c$:

$$x_i(\mathbf{w}) = \sum_p \mathbf{1}_p^c(i) \sum_k w_{kp} \varphi_k(i - r_p). \tag{11}$$

Now we introduce the P operator which is the projection operator, for tomography reconstruction, and is the identity for image denoising. The functional $F(\mathbf{w})$ whose minimum gives the optimal solution is written, for both applications, as:

$$\begin{aligned}
 F(\mathbf{w}) &= f(\mathbf{w}) + g(\mathbf{w}); \quad g(\mathbf{w}) = \beta \|\mathbf{w}\|_1; \\
 f(\mathbf{w}) &= \|\mathbf{y} - \mathbf{P}(\mathbf{x}(\mathbf{w}))\|_2^2 + \\
 &\quad \rho \sum_{p_i} \mathbf{1}_p(i) \left(x_i(\mathbf{w}) - \sum_k w_{kp} \varphi_k(i - r_p) \right)^2.
 \end{aligned} \tag{12}$$

where the ρ factor weights a similarity-inducing term which pushes all the overlapping patches touching a point i , towards the value $x_i(\mathbf{w})$ of the global solution $\mathbf{x}(\mathbf{w})$ in that point. For future reference we call $\|\mathbf{y} - \mathbf{P}(\mathbf{x}(\mathbf{w}))\|_2^2$ the fidelity term. The factor ρ has also the role of a regularisation parameter: as an example we consider the case where the set of overlapping patches is generated with a translational step $s = 1$. In this case the core indicator function has a 1×1 pixel domain and, without the ρ term, we could get a perfect fit, for an arbitrary image, by using for each patch an arbitrary component chosen randomly amongst those which are not zero at the core pixel. The solution is found with the FISTA method, using the gradient of $f(\mathbf{w})$ which is easily written as:

$$\begin{aligned}
 \frac{\partial f(\mathbf{w})}{\partial w_{kp}} &= \sum_i 2\varphi_k(i - r_p) \mathbf{1}_p(i) \left\{ (\mathbf{P}^T(\mathbf{P}(\mathbf{x}(\mathbf{w})) - \mathbf{y}))_i + \right. \\
 &\quad \left. \rho \sum_{p'} \mathbf{1}_{p'}(i) \left(x_i(\mathbf{w}) - \sum_{k'} w_{k'p'} \varphi_{k'}(i - r_{p'}) \right) \right\} + \\
 &\quad \sum_i 2\varphi_k(i - r_p) \rho \mathbf{1}_p(i) \left(\sum_{k'} w_{k'p} \varphi_{k'}(i - r_p) - x_i(\mathbf{w}) \right)
 \end{aligned} \tag{13}$$

where P^T is the adjoint operator of P , called back-projection operator in the case of tomography, which is again the identity for image denoising.

Other iterative methods

Floating Solution Functional

Xu et al. [1] have recently used an objective function which differs from ours for the fact that their global solution x is a free variable, while ours is a function of $\mathbf{w}:\mathbf{x}(\mathbf{w})$. Their objective function is

$$\begin{aligned}
 F(\mathbf{w}, \mathbf{x}) &= f(\mathbf{w}, \mathbf{x}) + g(\mathbf{w}); \\
 f(\mathbf{w}, \mathbf{x}) &= \|\mathbf{y} - \mathbf{P}(\mathbf{x})\|_2^2 + \\
 &\quad \rho \sum_{pi} \mathbf{1}_p(i) \left(x_i - \sum_k w_{kp} \phi_k(i - r_p) \right)^2.
 \end{aligned} \tag{14}$$

Xu et al. used the non-convex OMP procedure for minimisation of the functional. In this paper we will also compare their functional to our, using for both functionals the FISTA optimization.

Total Variation penalisation

In the total variation (TV) method [5] one minimises a convex functional given by the sum of the fidelity term $\|\mathbf{y} - \mathbf{P}(\mathbf{x})\|_2^2$ and of a gradient-sparsity inducing term $\beta_{TV} TV(\mathbf{x})$, where β_{TV} is a regularisation parameter and TV gives the isotropic total variation of the image \mathbf{x} :

$$TV(\mathbf{x}) = \sum_{\text{pixels}} \sqrt{(\partial_1 \mathbf{x})^2 + (\partial_2 \mathbf{x})^2}. \tag{15}$$

Equally Sloped Tomography

Equally Sloped Tomography (EST) is a Fourier-based iterative reconstruction method that iterates back and forth between real and Fourier space, utilizing an algebraically exact Pseudo-Polar fast Fourier transform [4, 9]. Measured projections (in this case obtained by the radon transform) are inserted into the Fourier space thanks to the fractional Fourier transform (FrFFT). Then the Pseudo-Polar Fast Fourier transform (PPFFT) and its adjoint are utilized to transform the images back and forth between Fourier and object space. During each iteration, physical constraints including sample boundary and the positivity of the coefficient are enforced in real space, while the measured data are applied in the Fourier space. The algorithm, monitored by an error metric, is guided towards the minimum that is consistent with the experimental data. To prevent any human intervention, the algorithm is automatically terminated when no further improvement can be made. In this case the number of iterations was 50 for a good convergence of the algorithm using 80 projections.

The EST algorithm has proved to allow a drastic reduction of the number of projections for conventional CT [9] and phase contrast imaging [4]. Fahimian et al. [9] demonstrated that the image quality and contrast obtained with EST is comparable with other iterative reconstruction schemes such as TV minimization or expected minimization statistical reconstruction with a faster convergence.

Results

Numerical Experiment

According to the Shannon-Nyquist criterion, to achieve a proper reconstruction in conventional CT, the number of angular projections N required is determined by

$$N \geq \frac{\pi D}{2P}$$

where D is the thickness of the sample and P the detector pixel size. One scenario for reducing the deposited dose during a CT scan is to reduce the number of projections.

To investigate the potential of the Dictionary Learning method on synthetic data, we use the standard 512×512 pixel Lena image as phantom. The dictionary is learnt from a different image that the one to reconstruct. In this case we used the boy image showed in [fig. 1 a](#)). The dictionary is shown in [fig. 1b](#). Note that we intentionally did not use the standard Shepp-Logan phantom in this study as it is a piece-wise constant and therefore it does not reflect the complexity of a phase contrast medical image.

The sinogram is obtained by projecting the image at 80 angles between 0 and 180 degrees using the radon transform.

We have optimized the regularisation parameters maximizing the following improvement factor, which quantifies the improvement obtained with the TV or $patches$ methods, with respect to a simple FBP reconstruction:

$$Q_{TV,patches} = (1 - S(\mathbf{x}_{fbp}, \hat{\mathbf{x}})) / (1 - S(\mathbf{x}_{TV,patches}, \hat{\mathbf{x}})) \quad (16)$$

where $\hat{\mathbf{x}}$ is the exact solution (unnoised Lena) and $\mathbf{x}_{fbp,TV,patches}$ represents the reconstruction result from one of the three methods, and where S is the Structural SIMilarity index [10] which is equal to 1 when the images are identical. In [fig. 2](#) we show the quality factor dependency versus the regularisation parameter β for our overlapping patches method (squares), and versus β_{TV} (dots) for the total variation method. For our method we have used a similarity-inducing term (ρ) value fixed to 1000, the patches basis shown in [fig. 1](#) and a step size of 3.

We have performed this reconstruction also with the floating solution function. We have used the same optimisation method, FISTA, used for our functional. We optimized β and ρ values by scanning over a 2D grid and comparing to the ground-truth. If the ground truth is not available, statistical methods such as the discrepancy principle [11] or generalized cross-validation [12] can be used to select the optimal regularization parameter in future application of the method. We obtained no significant difference between the results obtained with our functional and the floating solution: the SSIM is the same up to the third significant digit, and no significant difference can be detected in the final images. The convergence rate of FISTA, shown for both functional (our method (blue) and [1] (red)) in [fig. 3](#) is instead much faster using our functional form.

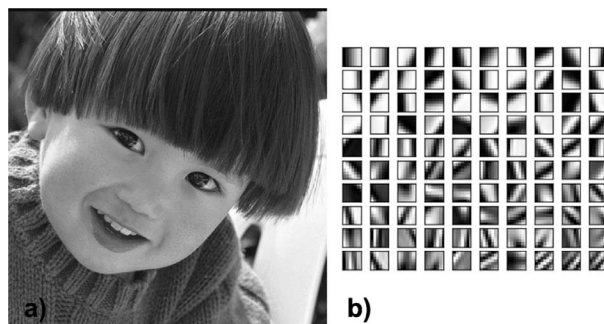


Fig. 1. The training image used in our work (a) and the dictionary elements obtained by using the K-SVD algorithm (b). This basis is then used for the numerical experiment.

doi:10.1371/journal.pone.0114325.g001

The comparison of the results obtained by the different reconstruction algorithms is shown in [fig. 4](#). We present the reconstruction results with four methods: filtered backprojection a), EST [9] b), the total variation penalisation [5] c) and our overlapping patches method d). These results have to be compared to the original cropped image shown in subfigure e). In this figure it is clear that the FBP images suffer a lot of streaking artefacts due to the number of projections well below the one required by the Shannon-Nyquist sampling theorem. The EST reconstruction removes some of these artefacts with the price of a blurred image with weak spatial resolution. Despite the fact that our technique gives a Q factor which is only slightly better than the one from the TV method, the obtained result looks much better by human eye inspection. The TV result shows a strong and irregular skin tessellation of those regions which have an illumination gradient.

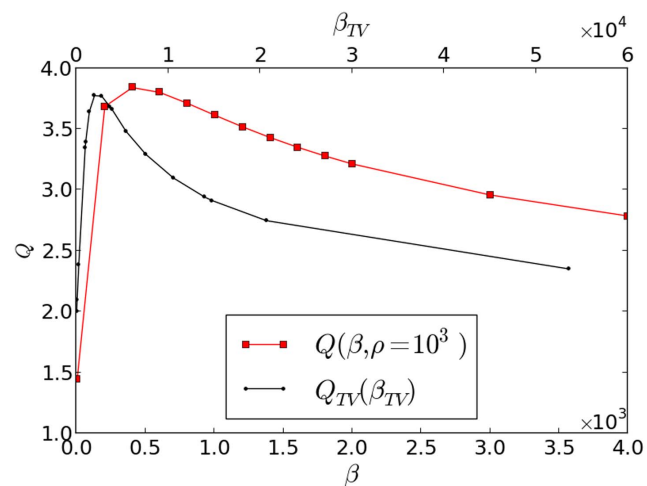


Fig. 2. The quality improvement factor Q versus the regularisation parameter β (squares) for our overlapping patches method, and versus β_{TV} (dots) for the total variation method. The used phantom is the 512×512 image of Lena, the reconstruction is performed using only 80 noised projections. For our method we have used a ρ value fixed to 1000, the patches basis shown in [figure 1](#) and a step size of 3.

doi:10.1371/journal.pone.0114325.g002

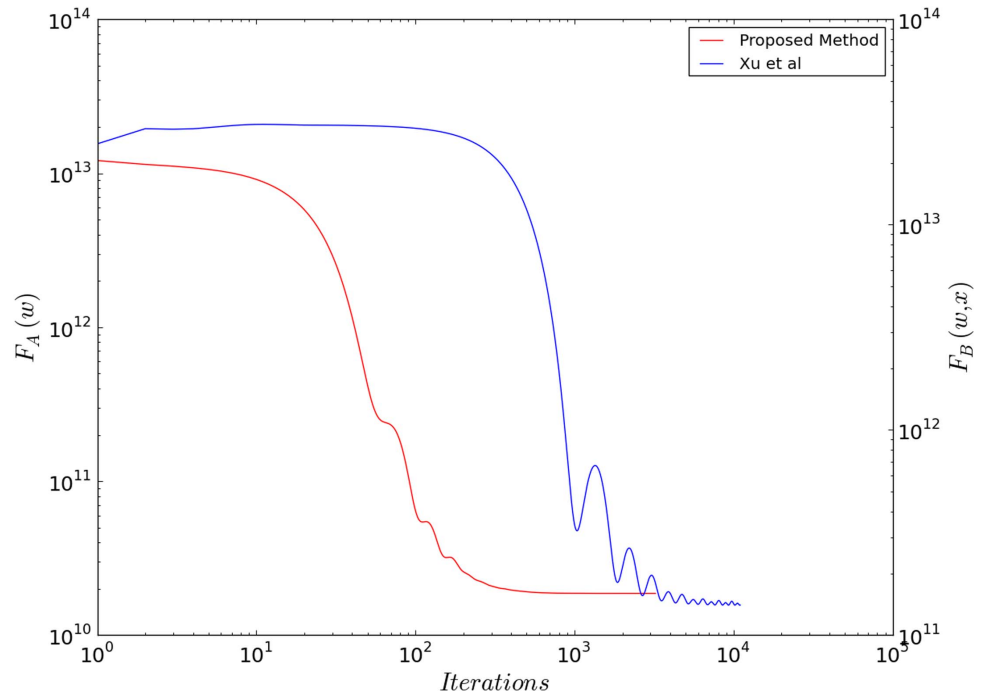


Fig. 3. The objective functions $F(w)$ and $F(w,x)$ (floating solution form) versus iteration number for FISTA optimisation.

doi:10.1371/journal.pone.0114325.g003

The hat feathers region (second row) is better resolved with the DL method. They look natural in the *patches* result, while the *TV* result produces strong grey levels distortions which vary irregularly along the feathers. *EST* results show a really good preservation of the tiny structure of the feathers but is noisier. The hat itself looks well preserved in the *patches* result while, in the *TV* image, the hat borders have irregular shapes. The *SSIM* values with the original images are reported for each subfigure. These values confirm the observation. Note that the visual difference between *TV* and *DL* seems greater than their *SSIM* values.

1000 iterations were used for the *DL* method. The computation time is 27 s on a Tesla k20m gnu card using the open source *PyHST* code [13]. On the same GPU card the computation time for the *FBP* is less than 1 s and 27 s for the *TV* method using the same number of iterations. The *EST* method has not yet been implemented on GPU and is therefore much slower (5 min).

Another strategy to further reduce the dose in tomography one can acquire with fewer number of photons onto the detector. To simulate this lack of photons we added a Poisson noise onto the sinogram data with a standard deviation $\sqrt{\lambda}$ equal to 0.3% of the sinogram value. We show in [fig. 5](#) the reconstruction results for the different algorithms.

The same conclusions for the not noisy case test can be drawn for the additive Poisson Noise: the dictionary learning method gives the best results. The *FBP*

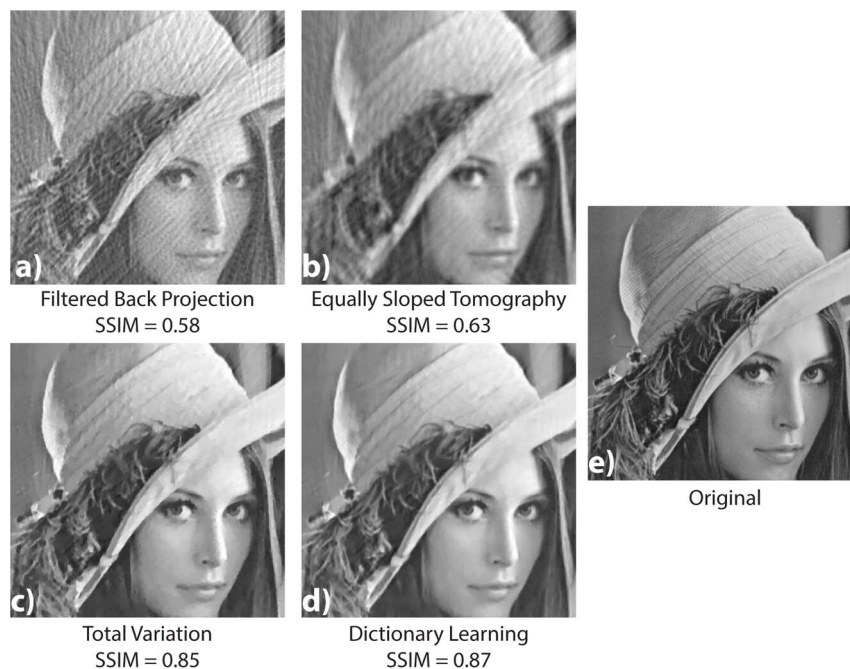


Fig. 4. Results from the numerical experiment on the Lena image 512×512 . The sinogram is obtained by projecting the image at 80 angles between 0 and 180 degree. A comparison of the images reconstructed with the FBP (a), the EST (b), the TV algorithm (c), and our DL method (d) is shown. The final image (e) is the same cropped zone in the original image for the sake of comparison (e).

doi:10.1371/journal.pone.0114325.g004

method is more sensitive to additive noise. With additive noise, the TV method shows a stronger and irregular skin tessellation of the skin surface of Lena.

Phase Contrast Tomography

In this section we apply our method to medical tomography of a human sample imaged using X-ray Phase Contrast Imaging (PCI).

PCI has shown an enhancement of soft tissue visualization in comparison to conventional imaging modalities [14]. It employs the dual property of X-rays of being simultaneously absorbed and refracted while passing through tissue. Among all the phase contrast techniques, we chose to test our method on analyzer based PCI [15, 16] because of the high sensitivity of the modality. Moreover, to the best of our knowledge, it is the only modality that showed results for investigating large and highly absorbing biological tissues (i.e. full human breasts) at a clinically compatible dose [4].

In the analyzer based PCI technique, the projection data contain a signal which is proportional to the gradient of the X-ray phase in one direction (i.e. the direction perpendicular to the plane formed by the incoming and diffracted X-rays on a perfect Bragg crystal which is used for analyzing the radiation passing through the sample). More details on the principles and technical aspects of PCI are available in [14]. Briefly the analyzer based imaging approach produces a

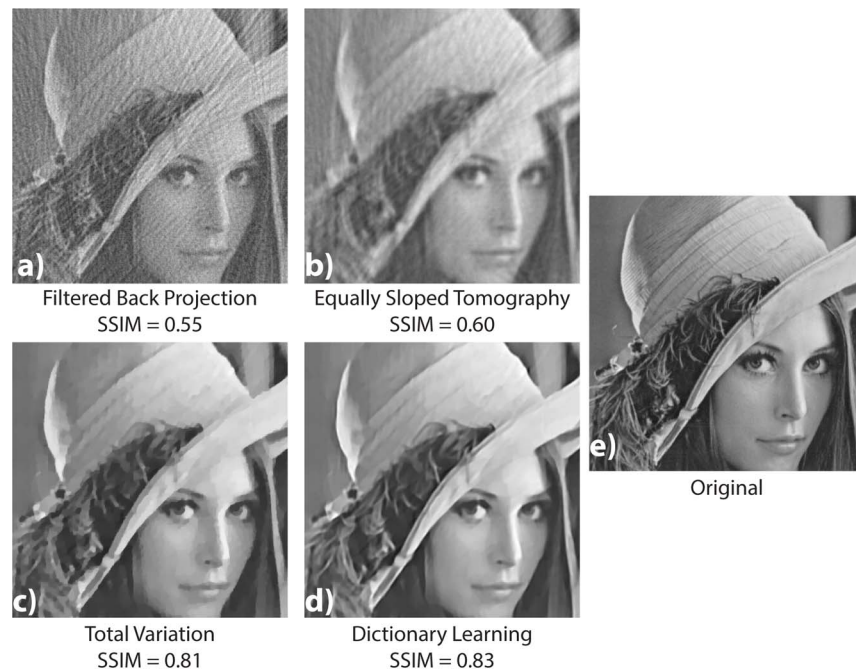


Fig. 5. Results from the numerical experiment on the Lena image 512×512 with additive Poisson noise. The sinogram is obtained by projecting the image at 80 angles between 0 and 180 degree, and adding a Poisson noise with a standard deviation equal to 0.3% of the maximal sinogram value. A comparison of the images reconstructed with the FBP (a), the EST (b), the *TV* algorithm (c), and our DL method (d) is shown. The final image (e) is the same cropped zone in the original image for the sake of comparison (e).

doi:10.1371/journal.pone.0114325.g005

mixed signal which originates from both X-ray absorption and refraction (i.e. phase derivative) [17]. The signal recorded by the detector is therefore very close to those recorded with other PCI techniques such as Grating Interferometry (GI) [18] or Edge Illumination (EI) [19]. All these methods are differential PCI methods and produce similar signals. Therefore the proposed approach can in principle be generalized.

When the object is rotated around an axis (*Z*-axis, for instance), this signal contains contributions from the *X* and *Y* gradient components, where *X* and *Y* axes co-rotate with the sample. The two components are de-phased by a rotation angle of 90 degrees and can be reconstructed separately by multiplying beforehand the sinogram with the cosine and sine of the rotation angle. We apply our formalism considering that the reconstructed and learning images are vectorial objects: the value associated with a pixel is not a scalar but a two-component vector.

The studied sample is a 7 cm human breast imaged with a pixel size of $100 \mu\text{m}$. The experiment was conducted at the biomedical beam line of the European Synchrotron Radiation Facility (ESRF). The sample was a human breast mastectomy specimen. The study was performed in accordance with the Declaration of Helsinki. A monochromatic X-ray beam with energy of 60 keV was used.

The training set is obtained from another breast sample imaged with the same technique but with high quality reconstruction. We consider a slice image for which the phase retrieval has been performed [20]. Then we apply a Sobel filter to extract the two derivative components and use the KSVD algorithm [21]. In this experiment 5 iterations were used to obtain the 100 atoms

Fig. 6 shows the patches basis functions that we use to fit both components at the same time. The patches size is 7×7 pixels and each basis function is displayed as a 14×7 rectangle whose upper 7×7 part is the X component and the lower one in the Y component.

In a previous work [20], it was demonstrated that the CT reconstruction of the refractive index obtained by first reconstructing the CT gradient field images and then applying a phase retrieval procedure, yields a better image quality than performing phase retrieval first and then reconstruction. The method is more robust with respect to noise, which may be a critical aspect in low dose tomography. In this case, the noise level may be such as it covers the information in the region where one gradient component of the refractive index has values close to zero. On the contrary, in those same regions the other component of the gradient of the refractive index has high values and it is thus less sensitive to noise. As a result, the information which is lost in one direction may be somehow retrieved by using the signal contained in the gradient image corresponding to the perpendicular direction. Additionally, when we use the vectorial approach, the information for the two reconstructed gradient components are intrinsically correlated by the dictionary and thus it increases the robustness of the method.

The result of reconstruction obtained by using filtered back projection algorithm with 1000 projections is shown in fig. 7a. In this image, radiologists could easily identify the skin, fat and glandular tissue. Fig. 7 is the reconstruction of a 765×765 pixel slice, using only 200 projections over the 1000 available. The upper left square is a zoom in the region marked in subfigure 7. The used projections cover, with constant spacing, a 180 degree range. The right column is the reconstruction with our method for X and Y components, while the left column (subfigure 7c) and d) is reconstructed with the standard FBP using all 1000 available projections. Using our method, we can still generate a high quality image with only one fifth of projections which would otherwise be necessary to generate a high quality reconstruction with the standard FBP method. Visually, the difference between the FBP results obtained with the full data set and our method with a five-fold reduction of the data is barely noticeable. The different borders of structures like skin layers, fatty tissues, and collagen strands are easily identified. The obtained results are very promising and a systematic evaluation for clinical application is under-way. The radiation dose absorbed by the sample during 200 projections is comparable to that of a standard clinical dual view (2D) mammography (3.5 mGy).

For the sake of comparison we report in fig. 8 the reconstruction obtained using the same number of projections using FBP (subfigure b,g), EST(subfigure d,i) and our method(subfigure f and k) using 200 projections in comparison with the full dose image (subfigure f and k). We report also the results obtaining

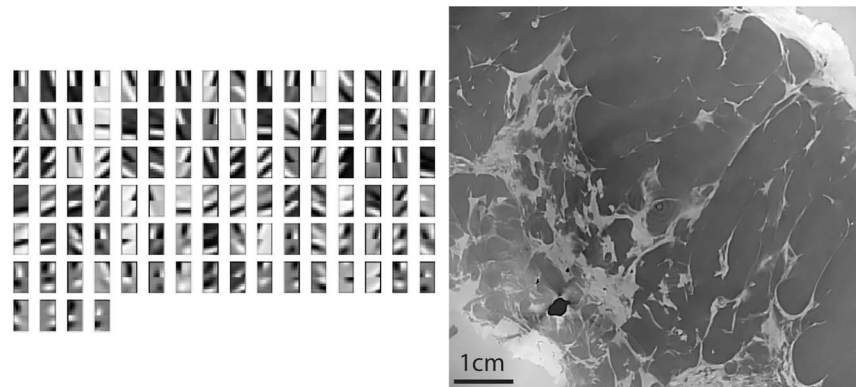


Fig. 6. The vectorial basis of patches (left) learnt from a high quality tomographic reconstruction of the phase image of a human breast(right). In each atom of the dictionary the upper 7×7 part is the X component while the lower part is the Y component.

doi:10.1371/journal.pone.0114325.g006

penalizing the L_1 of the reconstructed result(subfigure c and h). Our signal is a derivative, therefore penalizing the derivative modulus is similar to applying the TV method to the non derived object. The top inset is in a zone close to the skin with a blood vessel. The bottom insets are zoom in a zone with micro-calcifications. Note that micro-calcifications are of high interest for medical diagnoses because it may help identifying malignant masses. We report the $ssim$ values obtained by comparing the images with the FBP reconstruction with the full set of projections

In this figure it is clear that the overall image quality of the TV minimization is poor as well as the FBP with 200 images. The image quality of the EST 200 is lower than DL especially in terms of spatial resolution and sharpness. The EST image is indeed more blurred and the DL looks more similar the original full dose image. The DL image is indeed sharper with a clear delineation of the small micro-calcifications or blood vessel. The EST reconstructed image does not show the little micro-calcification in the middle of the image. Moreover on the top inset small round structures disappeared with the EST reconstruction whilst they are preserved in the DL image. The $SSIM$ values confirm the visual inspection.

Conclusion

For a decade Iterative CT reconstruction algorithms have demonstrated a possible dose reduction in conventional CT data. To the best of our knowledge, few works dealt with applying those algorithms to phase contrast tomography [4, 22, 23].

We have presented a new convex functional which implements in a mathematically pure form the concept of overlapping-patches-averaging, which was used so-far with a non-convex formalism. The resulting algorithm is efficient and well adapted to strongly reduce the noise in a natural image. A comparison with other iterative algorithms has been carried out on the Lena image showing

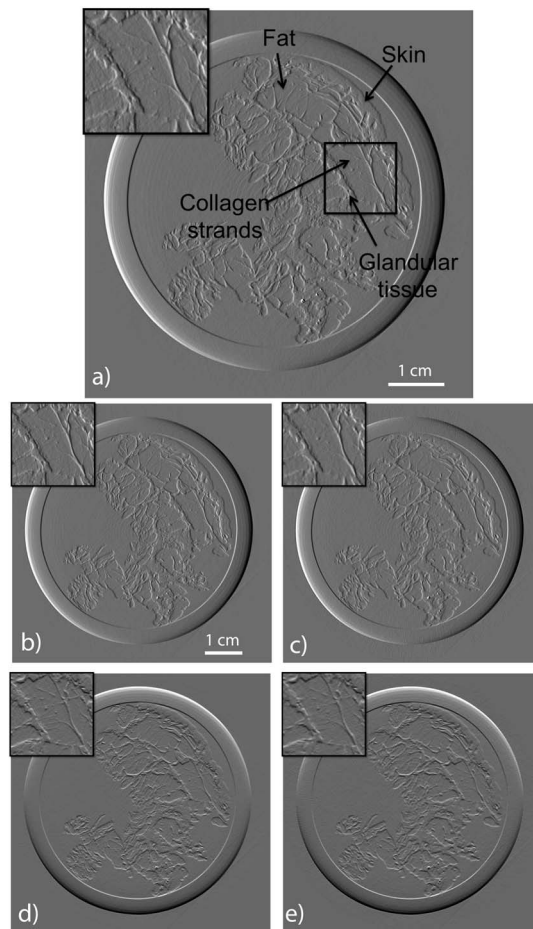


Fig. 7. Reconstruction of a computed tomographic slice of the breast. The images on the first and second row are the X and Y phase gradients, respectively. In the left column the results of the reconstruction obtained with the FBP method using the full set of data are reported. In the right column the results of our method using one projection over five are shown. For these reconstructions we set $\beta = 3 \times 10^{-6}$ and $\rho = 10$.

doi:10.1371/journal.pone.0114325.g007

that our method outperforms TV minimization and Equally Sloped Tomography. The method gives the best results with few projections and is less sensitive to additional noise. Compared to the state of the art dictionary learning method [1] our proposed approach converges faster to an equivalent image quality.

The method was applied to a medical diagnostic case by considering phase contrast tomographic data of a whole cancer-bearing human breast acquired with phase contrast imaging. A vectorial approach consisting of reconstructing gradients of the index of refraction was adopted. We demonstrated that thanks to this approach it is possible to reduce the deposited dose in breast CT by a factor of 5 compared to the standard filtered backprojection while keeping a comparable image quality.

Although we used this specific example as a proof of principle in this study, the method we developed and described can be easily applied to other tomography fields where a limited dose or a rapid acquisition time is a requirement. The

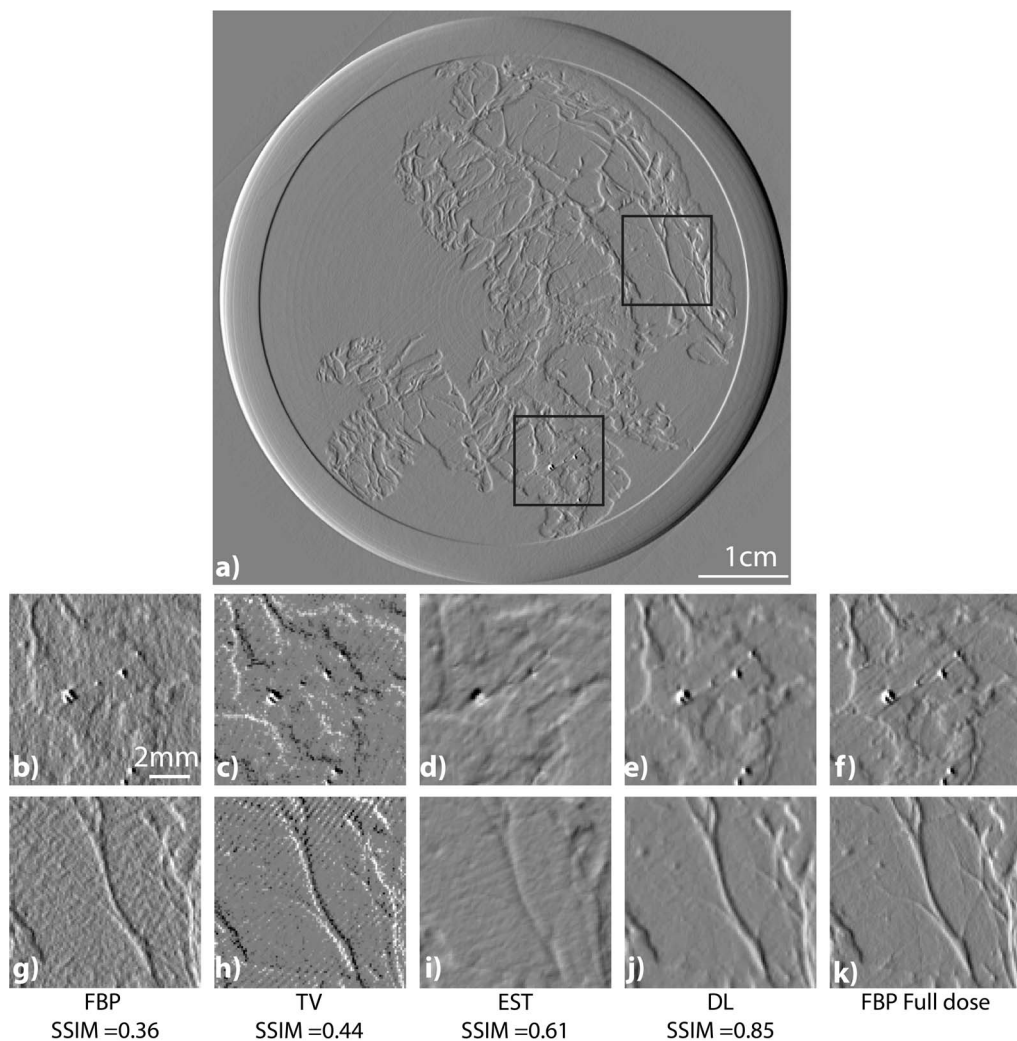


Fig. 8. Comparison of the tomographic image reconstructions of the breast obtained with the FBP (b and g), the TV minimization (c and h), the EST (d and i), and our method (e and j) using 200 projections over the 1000 available. The right images (f and k) are the images computed with the FBP using the entire set of projections. The top image (a) is the result obtained by the FBP using the entire set of projections. It is reported for showing the location of the insets. The SSIM values are reported based on the FBP full dose image.

doi:10.1371/journal.pone.0114325.g008

numerical results have been generated with PyHST [13,24], the ESRF tomography reconstruction code which uses the GPU implementation of the presented methods.

Ethics Statement

The study was performed in accordance with the Declaration of Helsinki. IRB-approval was granted by the ethics committee of the Ludwig-Maximilians University. Written informed consent was gathered before enrollment within the study

Acknowledgments

We thank Emmanuelle Gouillart and Gael Varoquaux for interesting discussions and for pointing out to us the possibilities of dictionary learning techniques and the examples contained in the scikit-learn package [25]. We thank ESRF for providing the experimental facilities, the ESRF ID17 team for assistance in operating the facilities. We thank Alberto Bravin, Sergei Gasilov, Alberto Mittone for their help during experiments. This work was partially supported by the DFG-Cluster of Excellence Munich-Centre for Advanced Photonics EXE158. Authors would like to thank the support of the Extrema COST action MP1207.

Author Contributions

Conceived and designed the experiments: AM EB PC. Performed the experiments: AM EB PC. Analyzed the data: AM EB PC. Contributed reagents/materials/analysis tools: AM EB PC. Wrote the paper: AM EB PC.

References

1. Xu Q, Yu H, Mou X, Zhang L, Hsieh J, et al. (2012) Low-dose x-ray ct reconstruction via dictionary learning. *Medical Imaging, IEEE Transactions on* 31: 1682–1697.
2. Zhao B, Ding H, Lu Y, Wang G, Zhao J, et al. (2012) Dual-dictionary learning-based iterative image reconstruction for spectral computed tomography application. *Physics in Medicine and Biology* 57: 8217.
3. Combettes PL, Pesquet JC (2011) Proximal Splitting Methods in Signal Processing. In: Bauschke H, Burachik R, Combettes P, Elser V, Luke DR, Wolkowicz H, editors. *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*. Springer. pp. 185–212.
4. Zhao Y, Brun E, Coan P, Huang Z, Sztrókay A, et al. (2012) High-resolution, low-dose phase contrast X-ray tomography for 3D diagnosis of human breast cancers. *Proceedings of the National Academy of Sciences of the United States of America* 109: 18290–18294.
5. Chambolle A (2004) An algorithm for total variation minimization and applications. *Journal of Mathematical Imaging and Vision* 20: 89–97.
6. Bach F, Jenatton R, Mairal J, Obozinski G (2011) Optimization for Machine Learning. Chapter Convex Optimization with Sparsity-Inducing Norms. Cambridge, MA: MIT Press.
7. Beck A, Teboulle M (2009) A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM Journal of Imaging Science* 2: 183–202.
8. Elad M, Aharon M (2006) Image denoising via sparse and redundant representations over learned dictionaries. *Image Processing, IEEE Transactions on* 15: 3736–3745.
9. Fahimian BP, Zhao Y, Huang Z, Fung R, Mao Y, et al. (2013) Radiation dose reduction in medical x-ray CT via Fourier-based iterative reconstruction. *Medical physics* 40: 031914.
10. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP (2004) Image quality assessment: from error visibility to structural similarity. *Image Processing, IEEE Transactions on* 13: 600–612.
11. Wen YW, Chan RH (2012) Parameter selection for total-variation-based image restoration using discrepancy principle. *IEEE Trans Image Process* 21: 1770–81.
12. Liao H, Li F, Ng MK (2009) Selection of regularization parameter in total variation image restoration. *Journal of the Optical Society of America A: Optics, Image Science and Vision* 26: 2311–20.
13. Mirone A (2013). Gpl release of pyhst2 in preparation. <http://forge.epn-campus.eu/projects/pyhst2>.
14. Bravin A, Coan P, Suortti P (2013) X-ray phase-contrast imaging: from pre-clinical applications towards clinics. *Physics in Medicine and Biology* 58: R1.

15. **Chapman D, Thomlinson W, Johnston RE, Washburn D, Pisano E, et al.** (1997) Diffraction enhanced x-ray imaging. *Physics in Medicine and Biology* 42: 2015–2025.
16. **Dilmanian FA, Zhong Z, Ren B, Wu XY, Chapman LD, et al.** (2000) Computed tomography of x-ray index of refraction using the diffraction enhanced imaging method. *Physics in Medicine and Biology* 45: 933–946.
17. **Diemoz PC, Bravin A, Coan P** (2012) Theoretical comparison of three X-ray phase-contrast imaging techniques: propagation-based imaging, analyzer-based imaging and grating interferometry. *Optics Express* 20: 2789.
18. **Weitkamp T, Diaz A, David C, Pfeiffer F, Stampanoni M, et al.** (2005) X-ray phase imaging with a grating interferometer. *Optics Express* 13: 6296–6304.
19. **Olivo A, Speller R** (2008) Image formation principles in coded-aperture based x-ray phase contrast imaging. *Physics in Medicine and Biology* 53: 6461–6474.
20. **Gasilov S, Mittone A, Brun E, Bravin A, Grandl S, et al.** (2014) An efficient tomographic reconstruction of the hard x-ray refractive index using the gradient vector-field approach. *Optics Express* 22: 5216.
21. **Aharon M, Elad M, Bruckstein A** (2006) KSVD: An Algorithm for Designing Overcomplete Dictionaries for Sparse Representation. *IEEE Transactions on Signal Processing* 54: 4311–4322.
22. **Sunaguchi N, Yuasa T, Ando M** (2013) Iterative reconstruction algorithm for analyzer-based phase-contrast computed tomography of hard and soft tissue. *Applied Physics Letters* 103: 143702.
23. **Xu Q, Sidky EY, Pan X, Stampanoni M, Modregger P, et al.** (2012) Investigation of discrete imaging models and iterative image reconstruction in differential X-ray phase-contrast tomography. *Optics Express* 20: 10724.
24. **Mirone A, Brun E, Gouillart E, Tafforeau P, Kieffer J** (2014) Pyhst2: an hybrid distributed code for high speed tomographic reconstruction with iterative reconstruction and a priori knowledge capabilities. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms* 324: 41–48.
25. **Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, et al.** (2011) Scikit-learn: Machine learning in python. *Journal of Machine Learning Research* 12: 2825–2830.