

Coded aperture optimization for compressive X-ray tomosynthesis

Angela P. Cuadros,¹ Christopher Peitsch,² Henry Arguello,³
and Gonzalo R. Arce^{1,*}

¹Department of Electrical and Computer Engineering, University of Delaware Newark, DE 19716, USA

²Chesapeake Testing Services, Inc. Belcamp, MD 21017, USA

³Department of System Engineering, Universidad Industrial de Santander, Bucaramanga, 680002, Colombia

*arce@udel.edu

Abstract: Radiation dose is a concern in X-ray tomographic imaging; coded aperture compressive X-ray tomosynthesis is an approach used to reduce radiation. It places a coded aperture in front of an X-ray source in order to obtain 2D patterned projections of a three-dimensional object onto a detector plane. By using different coded apertures in a multiple source system, multiplexed projections can be obtained instead of sequential projections as in conventional tomosynthesis systems. Compressed sensing (CS) reconstruction algorithms are then used to recover the three-dimensional data cube. An optimization approach to design the structure of the coded apertures in a multiple source compressive X-ray tomosynthesis imaging system is presented. A uniform energy criteria on the voxels and detector elements is used so that the object is uniformly sensed and the elements of the detector plane uniformly sense the information. Simulations and experimental results for optimized coded apertures are shown, and their performance is compared to the use of random coded apertures.

© 2015 Optical Society of America

OCIS codes: (340.7430) X-ray coded apertures; (340.7440) X-ray imaging; (110.6960) Tomography.

References and links

1. J. T. Dobbins III and D. J. Godfrey, "Digital X-ray tomosynthesis: current state of the art and clinical potential," *Phys. Med. Biol.* **48**(19), R65 (2003).
2. R. Smith-Bindman, J. Lipson, R. Marcus, K.P. Kim, M. Mahesh, R. Gould, A. Berrington de Gonzalez, and D. L. Miglioretti, "Radiation dose associated with common computed tomography examinations and the associated lifetime attributable risk of cancer," *Arch. Internal Med.* **169**(22), 2078–2086 (2009).
3. I. Reiser and S. Glick, *Tomosynthesis Imaging* (Taylor and Francis, 2014).
4. F. Natterer, *The Mathematics of Computerized Tomography* (Vieweg Teubner Verlag, 1986).
5. K. Hämäläinen, A. Kallonen, V. Kolehmainen, M. Lassas, K. Niinimäki, and S. Siltanen, "Sparse tomography," *Computational Methods in Science and Engineering*, SIAM, **35**, B644–B665 (2013).
6. K. Choi and D. J. Brady, "Coded aperture computed tomography," *Proc. SPIE* **7468**, 74680B (2009).
7. Y. Kaganovsky, D. Li, A. Holmgren, H. Jeon, K. MacCabe, D. Politte, J. O'Sullivan, L. Carin, and D. J. Brady, "Compressed Sampling Strategies for Tomography," *J. Opt. Soc. Am. A* **31**, 1369–1394 (2014).
8. M. Slaney and A. Kak, *Principles of Computerized Tomographic Imaging* (Society for Industrial and Applied Mathematics, 2001).
9. E. Candes and M. Wakin, "An introduction to compressive sampling," *IEEE Sig. Proc. Mag.* **25** (2), 21–30 (2008).
10. A. Cuadros, G. R. Arce, and H. Arguello, "Coded aperture design in compressive X-ray tomography," in *IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, 656–659 Dec (2014).

11. A. Cuadros, K. Wang, C. Peitch, H. Arguello, and G. R. Arce, "Coded aperture design for compressive X-ray tomosynthesis," in *Imaging and Applied Optics 2015*. Optical Society of America, 2015, p. CW2F.2.
12. J. P. Allebach, "DBS: retrospective and future directions," *Proc. SPIE* **4300**, 358–376 (2000).
13. D. L. Lau and G. R. Arce *Modern Digital Halftoning* (CRC Press; Taylor & Francis Group, 2008).
14. W. Xu, F. Xu, M. Jones, B. Keszthelyi, J. Sedat, D. Agard, and K. Mueller, "High-performance iterative electron tomography reconstruction with long-object compensation using graphics processing units," *J. Structural Biol.* **171** (2), 142–153 (2010).
15. D. J. Brady, *Optical Imaging and Spectroscopy* (Wiley; Optical Society of America, 2009).
16. N. Halko, P. G. Martinsson, and J. A. Tropp, "Finding Structure with Randomness: Probabilistic Algorithms for Constructing Approximate Matrix Decompositions," *SIAM Review*, **53**(2), 217–288, (2011).
17. M. Figueiredo, R. Nowak, and S. Wright, "Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems," *IEEE J. Sel. Top. Sig. Proc.* **1**(4), 586–597 (2007).
18. G. R. Arce, D. J. Brady, L. Carin, H. Arguello and D. S. Kittle, "Compressive Coded Aperture Spectral Imaging: An Introduction," *IEEE Signal Processing Magazine*, 105-115, January (2014).

1. Introduction

X-ray tomosynthesis imaging systems have become essential in medical imaging diagnostic tasks such as coronary angiography, dual energy imaging and mammography, among others [1]. Recent data suggest that medical radiation exposure may significantly increase the risk of adverse radiation effects, including damage of body cells and even DNA molecules [2]. In order to reduce damage that radiation can cause to patients, optimized hardware settings have been proposed by lowering the number of angles at which projections are taken [3]. In this sense, tomosynthesis can be considered a limited-angle computed tomography (CT) that results in less radiation exposure for the patient [3]. However, the reduction of measurements leads to a highly ill-posed inverse problem, sensitive to measurement and modeling errors. Filtered backprojection (FBP) image reconstructions with ill-posed systems of Eqs. produce artifacts and noise that make the reconstructions useless for medical diagnosis [4]. Sparsity-promoting and total variation regularization algorithms have been recently used to improve the ill-posed inverse problem, obtaining better image reconstructions [5]. Reducing the number of angle or projection rays, invariably leads to artifacts in reconstructions. Coded aperture X-ray tomography is one approach that can overcome these limitations.

In [6], Choi et al. introduced coded aperture X-ray tomosynthesis, which goes beyond sparse regularization since it allows the acquisition of compressive measurements. The physical coding in coded aperture X-ray tomosynthesis controls the correlation between the measurement vectors. The projections used in [6], however, used random coded apertures. No coded aperture optimization was considered. The optimization of coded aperture for the coded aperture compressive X-ray tomosynthesis system is introduced in the present work, further reducing the radiation exposure in compressive tomosynthesis. Furthermore, multi-frame measurements are obtained by taking sequential snapshots of the object, which leads to more degrees of freedom and improved results. The performance of the optimized codes is compared to that of random codes by means of the singular value decomposition (SVD) analysis of the forward operator.

Recently, in [7], Kaganovsky et al. introduced coded aperture projections for medical CT scanner geometries. Random coded apertures are used to modulate the measurements obtained by varying the angle and detector use for each projection [7]. The methods presented in this paper for compressive X-ray tomosynthesis can be extended to the third-generation CT scanners, in which a fan beam X-ray source rotates around the object.

2. Forward projection model

The X-ray transmission imaging model for a single source is given by the Beer-Lambert law [8]: $I = I_0 \cdot e^{-\int_0^{\infty} \mu(x) dx}$, where I_0 is the intensity of a particular X-ray originated from the X-ray source passing through the object, I is the measured intensity in the detector, and $\mu(x)$ is the

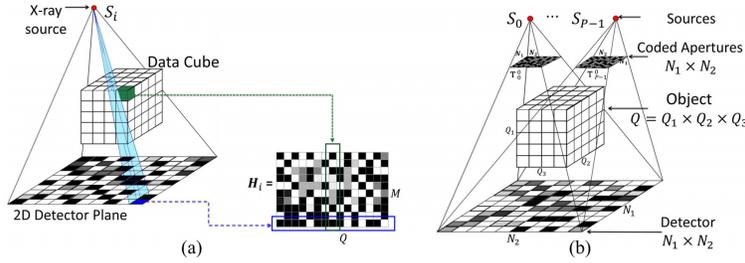


Fig. 1. (a) X-ray tomosynthesis. The system matrix \mathbf{H}_i determines the mapping of the X-ray cone beam sources to the detector. Each row describes the sensing for a particular detector element and each column corresponds to the sensing of a particular voxel. (b) Coded aperture compressive X-ray tomosynthesis. The energy of each source is modulated by means of a coded aperture.

linear attenuation coefficient varying in the location given by x . If such X-ray source is located at position \vec{s} and illuminates an object in direction $\hat{\theta}$, the data function for the imaging model is given by $y(\vec{s}, \hat{\theta}) = -\ln(I/I_0)$. Therefore, Beer-Lambert law can be rewritten as $y(\vec{s}, \hat{\theta}) = \int_0^\infty f(\vec{s} + x\hat{\theta})dx$, where f corresponds to the three-dimensional object function, i.e., the X-ray linear attenuation coefficient map. This continuous-to-continuous imaging model is known as the X-ray transform [3].

The imaging model needs to be discretized since only a discrete number of measurements can be taken. Thus, the three-dimensional data cube is represented by a vector formed by a discrete number of unknowns $[\mathbf{f}]_j$ with $j = 0, \dots, Q - 1$ that correspond to the attenuation coefficients of each of the voxels that constitute the object $\mathbf{f} \in \mathbb{R}^Q$, where $Q = Q_1 \times Q_2 \times Q_3$ corresponds to the number of voxels and Q_1 is the number of slices of dimensions $Q_2 \times Q_3$ each. The detector is designed to be a two-dimensional plane composed by $M = N_1 \times N_2$ detector elements placed under the object as shown in Fig. 1(a).

The projection measurements are recorded by each of the detector elements such that $[\mathbf{y}]_m \in \mathbb{R}^M$ for $m = 0, \dots, M - 1$, corresponds to the m^{th} detector measurement. Tomosynthesis sensing with a single source i can be written as a finite linear system of Eqs. of the form $\mathbf{y}_i = \mathbf{H}_i \mathbf{f}$, where the matrix \mathbf{H}_i of dimensions $M \times Q$ is the system matrix obtained by specifying the hardware settings. The weights correspond to the mapping of the cone-beam energy radiating from the X-ray source onto the detector. As shown in Fig. 1(a), each of the elements in the weighting matrix \mathbf{H}_i , i.e., $[\mathbf{H}_i]_{mj}$, correspond to the portion of the volume of voxel j that is irradiated by the X-ray associated with the detector element m . Moreover, each of the rows of \mathbf{H}_i corresponds to the information gathered by one detector and each of the columns corresponds to the information gathered from a single voxel.

Compressive X-ray tomosynthesis multiplexes measurements from multiple sources onto the detector. Coded apertures are placed in front of each of the cone-beam sources to modulate the energy of each X-ray source, producing a particular coded projection onto the detector plane [6]. The coded apertures have the same number of elements as the detector plane. The size of the elements of the coded apertures is fixed to obtain one-to-one correspondence with the detector elements. The coded aperture \mathbf{T}_i is paired to the corresponding i^{th} source, for $i = 0, \dots, P - 1$ with P being the number of sources with each (u, v) element in the code denoted by $(\mathbf{T}_i)_{uv} \in \{0, 1\}$, where 0 blocks the X-ray beam and 1 lets the X-ray beam pass. The configuration for the coded aperture compressive X-ray tomosynthesis is shown in Fig. 1(b). Each of the sources has a different projection \mathbf{y}_i and a different system matrix \mathbf{H}_i . To account for the coded apertures, the matrix \mathbf{C}_i is defined as a diagonal matrix whose diagonal elements are the elements of

the coded aperture \mathbf{T}_i , i.e $\mathbf{C}_i = \text{diag}((\mathbf{T}_i)_{00}, (\mathbf{T}_i)_{12}, \dots, (\mathbf{T}_i)_{(N-1)(N-1)})$. Therefore, the sensing process for a single source i is given by $\mathbf{y}_i = \mathbf{C}_i \mathbf{H}_i \mathbf{f}$.

To generalize the sensing process, \mathbf{C} is defined as the matrix concatenating the structures of the coded apertures of the P sources $\mathbf{C} = [\mathbf{C}_0 | \mathbf{C}_1 | \dots | \mathbf{C}_{P-1}]$, and \mathbf{H} is defined as $\mathbf{H} = [\mathbf{H}_0 | \mathbf{H}_1 | \dots | \mathbf{H}_{P-1}]^T$. Thus, the measurements for a multiple source system are described by

$$\mathbf{y} = \left(\sum_{i=0}^{P-1} \mathbf{C}_i \mathbf{H}_i \right) \mathbf{f} = \mathbf{C} \mathbf{H} \mathbf{f}. \quad (1)$$

The reconstruction of \mathbf{f} from \mathbf{y} describes an ill-posed problem; thus, it cannot be solved by the use of traditional least square approaches. In general, the solution is not unique [4]. However, compressive sensing (CS) asserts that the function \mathbf{f} can be recovered, provided two principles are met: 1) the function \mathbf{f} is sufficiently sparse in some basis Ψ , and 2) the basis used to represent the object and the system matrix used to sense the object are incoherent [9].

Let \mathbf{f} be represented by $\mathbf{f} = \Psi \theta \in \mathbb{R}^Q$, where θ is the sparse coefficient representation of the object, and Ψ is the basis representation. The cumulative sensing at the detector from all P sources is given by $\mathbf{y} = \mathbf{C} \mathbf{H} \Psi \theta = \mathbf{A} \Psi \theta$, where $\mathbf{A} \in \mathbb{R}^{M \times Q}$ is the sensing matrix, with $M \ll Q$. The mapping of the energy from all sources onto the detector \mathbf{y} captures the modulated energy of all X-ray sources by the coded apertures and the effect of the three-dimensional data-cube on the coded X-ray field.

The number of compressive measurements obtained by one shot may not be sufficient for adequate reconstruction. Therefore, the sensing can be generalized to account for K 2D snapshot projections and P sources, located in a fixed position. The coded aperture for the i^{th} source and k^{th} shot is denoted by \mathbf{T}_i^k , for $k = 0, \dots, K-1$. The matrix \mathbf{C}_i^k is the diagonal matrix associated with \mathbf{T}_i^k . Define $\mathbf{C}^k = [\mathbf{C}_0^k | \mathbf{C}_1^k | \dots | \mathbf{C}_{P-1}^k]$, thus \mathbf{y}^k corresponds to the measurements for the k^{th} shot, which can be rewritten as $\mathbf{y}^k = \mathbf{C}^k \mathbf{H} \Psi \theta = \mathbf{A}^k \Psi \theta$. Defining $\tilde{\mathbf{y}} = [\mathbf{y}^0 | \mathbf{y}^1 | \dots | \mathbf{y}^{K-1}]^T$, the sensing process for K shots and P sources is described by:

$$\tilde{\mathbf{y}} = \tilde{\mathbf{C}} \mathbf{H} \mathbf{f} = \tilde{\mathbf{C}} \mathbf{H} \Psi \theta = \tilde{\mathbf{A}} \Psi \theta, \quad (2)$$

where $\tilde{\mathbf{C}} = [\mathbf{C}^0 | \mathbf{C}^1 | \dots | \mathbf{C}^{K-1}]^T$. In order to reconstruct the object \mathbf{f} , the under-determined system of Eqs. given in (2) is solved by minimizing the cost function $\|\tilde{\mathbf{y}} - \tilde{\mathbf{A}} \Psi \theta\|_2^2 + \lambda \|\theta\|_1$, where λ is a regularization constant and $\|\cdot\|_1$ and $\|\cdot\|_2$ correspond to the ℓ_1 and ℓ_2 norms, respectively. This method of data acquisition provides a means to attain multiplexed coded measurements.

3. Coded aperture optimization

Multiplexed tomosynthesis introduced by Choi et al. in [6] used random projections generated by coded apertures with entries randomly distributed. These codes are, in general, sub-optimal since they do not take into account the fixed geometry of the tomographic system. The coded aperture optimization framework is described next.

3.1. Optimization constraints

Given K tomosynthesis detector measurements, the goal is to design K distinct coded apertures for each of the X-ray sources. Let \mathbf{T}_i^k be the coded aperture assigned for the i^{th} source and the k^{th} shot. Note that the coded aperture does not depend on the object under inspection but on the structure of the system matrix \mathbf{H} . To achieve incoherent measurements and non-redundant sensing, the coded apertures can be designed such that uniform sensing is achieved under the following criteria.

Criterion 1 *Achieve uniform sensing in the detector*: Each detector element should measure approximately the same amount of information, indicating that the detector elements are sensing the data cube uniformly. Sensing matrix \mathbf{A}^k is binarized so that each entry \mathbf{A}_{mq}^k represents if the q^{th} voxel is sensed by the m^{th} detector element. Vector \mathbf{d}^k is defined as the matrix product between the matrix \mathbf{A}^k and a Q -long one-valued vector $\mu_Q = [1, \dots, 1]^T$ i.e. $\mathbf{d}^k = \mathbf{A}^k \mu_Q$, where \mathbf{d}^k represents the sum along the M rows of the sensing matrix \mathbf{A}^k , i.e. for the k^{th} shot. Since each of the rows of the sensing matrix corresponds to the information related to certain detector element, each of the elements of the vector \mathbf{d}^k represents the number of voxels measured by the aforementioned detector element [10, 11]. For multiple shots, the goal is to reduce the variance between entries of each vector \mathbf{d}^k , thus making the entries of the vector $\bar{\mathbf{d}} = \frac{1}{K} \sum_{k=0}^{K-1} \mathbf{d}^k$ uniformly distributed.

Criterion 2 *Uniformly sense the data-cube voxels*: The number of times a certain voxel is measured should be approximately the same for all voxels. To this end, \mathbf{r}^k is defined as the matrix product between the transpose of the sensing matrix \mathbf{A}^{kT} and an M -long one-valued vector $\mu_M = [1, \dots, 1]^T$ i.e. $\mathbf{r}^k = \mathbf{A}^{kT} \mu_M$, where \mathbf{r}^k represents the sum of the columns of the sensing matrix for the k^{th} shot. Each of the columns of the sensing matrix is related to a particular voxel of the three-dimensional object; hence, each of the elements of the vector \mathbf{r}^k represents the number of times a particular voxel is measured [10, 11]. For multiple shots, the goal is to reduce the variance between entries of each vector \mathbf{r}^k , thus making the entries of the vector $\bar{\mathbf{r}} = \frac{1}{K} \sum_{k=0}^{K-1} \mathbf{r}^k$ uniformly distributed.

Criterion 3 *Uncorrelated codes for multiple shots*: When $K \geq 2$, for a particular X-ray source a different set of coded apertures is used in each shot, and Constraint 3 is defined to assure complementary codes are obtained for each source. Specifically, the codes are designed such that for a fixed spatial location (u, v) in all the set of coded apertures of a particular source $(\mathbf{T}_i)_{uv}^k$, only one out of K coded apertures should contain a non-zero value. To this end, \mathcal{T}_i is defined as the sum of the K codes for the i^{th} source, i.e. $\mathcal{T}_i = \sum_{k=0}^{K-1} \mathbf{T}_i^k$. In order to make the codes uncorrelated, all the entries of \mathcal{T}_i should be 1. To that end, \mathbf{S}_T is defined as $\mathbf{S}_T = \mathcal{T} - \mathbb{U}_{N \times PN}$, where $\mathbb{U}_{N \times PN}$ corresponds to a one-valued matrix of dimensions $N \times PN$, $\mathcal{T} = [\mathcal{T}_1 | \dots | \mathcal{T}_P]$ and c_3 as the ℓ_0 norm of the vectorized matrix \mathbf{S}_T , i.e. $c_3 = \|\text{vect}(\mathbf{S}_T)\|_0$; by minimizing c_3 uncorrelated codes are obtained.

Based on the previous three constraints, a cost function that shapes the set of coded apertures such that the three-dimensional data cube and the detector plane are sensed as uniformly as possible while obtaining complementary codes for each source is defined. The cost function thus aims to minimize the variance of the average number of detector elements measuring each voxel, i.e., the entries of vector $\bar{\mathbf{d}}$, the variance of the average number of voxels that each detector element measures given K shots, i.e., the entries of vector $\bar{\mathbf{r}}$, and the error term defined as c_3 for the third constraint. Thus, the optimization of the coded apertures for multiple snapshots is determined by the minimization of the cost function:

$$\underset{[\mathbf{T}_0^k, \dots, \mathbf{T}_{P-1}^k]_{k=0}^{k=K-1}}{\text{argmin}} \quad \alpha \cdot \sum_{m=0}^{M-1} [(\bar{\mathbf{d}})_m - m_1]^2 + \beta \cdot \sum_{j=0}^{Q-1} [(\bar{\mathbf{r}})_j - m_2]^2 + \gamma \cdot c_3$$

Subject to $(\bar{\mathbf{d}})_m > 0$ and $(\bar{\mathbf{r}})_j > 0 \quad \forall m, j,$ (3)

where $\alpha \approx \frac{1}{M}$ since the first term corresponds to the sum of M elements, $\beta \approx \frac{1}{Q}$ since the second term corresponds to the sum of Q elements, and $\gamma \approx \frac{1}{MP}$ since the third term corresponds

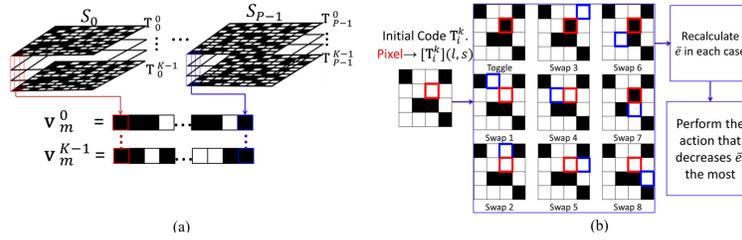


Fig. 2. (a) To generate the initial set of codes, vector \mathbf{v}_m^k is defined. It is formed by the values of the m^{th} elements of the P coded apertures used in the k^{th} shot. (b) Iteration Process for the DBS algorithm.

to the sum of MP elements. m_1 is the desired median of the number of voxels sensed in each detector element and m_2 is the desired median of the number of times each voxel is sensed. Both parameters depend on hardware settings. For each of the experiments, the median of the number of voxels sensed in each detector element and the median of the number of times each voxel is sensed are obtained from the vectors $\bar{\mathbf{d}}$ and $\bar{\mathbf{r}}$ obtained when using random codes. Therefore, for different number of shots there are different values for m_1 and m_2 . $(\bar{\mathbf{d}})_m$ corresponds to the m^{th} element of the average sum of the rows of the sensing matrix for K shots, and $(\bar{\mathbf{r}})_j$ corresponds to the j^{th} element of the average sum of the columns of the sensing matrix for K shots. In order to solve the optimization problem in (3), the following approach is proposed.

3.2. Optimization algorithm

The Direct Binary Search (DBS) algorithm is an iterative approach to evaluating the effect of trial changes for each pixel of a binary image for a particular search [12]. Using (3) as a cost function, optimal coded apertures are obtained using the DBS algorithm to perform a local search on each of the coded apertures by either swapping the current pixel with one of its eight nearest neighbors or toggling the coded aperture pixel from 1 to 0 or 0 to 1, keeping the changes that have positive effects in the cost function and ignoring the changes that have a negative effect. The algorithm stops when, after processing all the $K \times P$ coded apertures, no swaps or toggles occur. Being a steepest descent type of optimization, the DBS algorithm is susceptible to local minimum extrema. Thus the final codes depend on the initial set of coded apertures that are selected [13]. Therefore, an alternative algorithm that takes into account the three constraints is used to obtain a suitable initial first set of codes.

3.2.1. Initial set of codes

In order to produce an initial set of codes, a binary P long vector $\mathbf{v}_m^k = [(\mathbf{T}_0^k)_m, (\mathbf{T}_1^k)_m, \dots, (\mathbf{T}_{P-1}^k)_m]$ is defined as the concatenation of the values of the m^{th} elements of the P coded apertures used in the k^{th} shot. The binary vector could take one of $2^P - 1$ possible values. The matrix \mathbf{V} of dimensions $(2^P - 1) \times P$ is defined as the concatenation of all possible binary combinations for vector \mathbf{v}_m^k , such that each of the rows of the matrix corresponds to one possible value for the vector \mathbf{v}_m^k as shown in Fig. 2(a). For each location m , K rows of \mathbf{V} must be selected; to this end, W is defined as a matrix containing all the possible combinations that can be selected from the rows of the matrix \mathbf{V} .

To achieve uniform sensing in the detector, while having information only from the m^{th} pixel, the vector $\bar{\mathbf{d}}$ can be expanded as $\bar{\mathbf{d}} = \frac{1}{K} \sum_{k=0}^{K-1} \left[\sum_{i=0}^{P-1} \mathbf{C}_i^k \mathbf{H}_i \right] \mu_Q = \frac{1}{K} \sum_{k=0}^{K-1} \sum_{i=0}^{P-1} \mathbf{C}_i^k \mathbf{G}_i$, where $\mathbf{G}_i = \mathbf{H}_i \mu_Q$, i.e., the sum of the rows of the system matrices \mathbf{H}_i associated with each source.

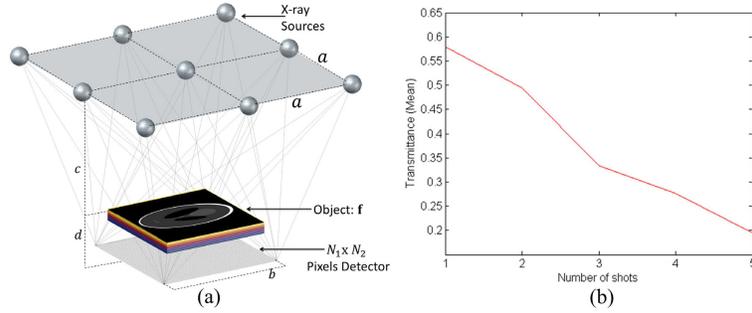


Fig. 3. (a) Configuration for X-ray tomosynthesis simulation. 9 sources placed uniformly over a 128×128 phantom with 16 slices. The dimensions for a general scenario are shown in (a), for the particular simulation scenario that was studied here $a = 128, b = 128, c = 675, d = 60, e = 150$. (b) Mean of the transmittance of the optimal coded apertures for each shot.

The m^{th} element of the vectors \mathbf{G}_i represents the information on how many voxels are measured in the m^{th} detector when illuminated by the i^{th} source, defining the aforementioned element as $(\mathbf{G}_i)_m$ and a P -long vector $\mathbf{g}_m = [(\mathbf{G}_1)_m, (\mathbf{G}_2)_m, \dots, (\mathbf{G}_{P-1})_m]$ the uniformity condition for the detector plane can be rewritten as:

$$\underset{[\mathbf{v}_m^0, \dots, \mathbf{v}_m^{K-1}]_{m=0}^{K-1}}{\text{argmin}} \left[\sum_{k=0}^{K-1} \mathbf{v}_m^k \mathbf{T} \mathbf{g}_m - m_1 \right]^2.$$

In order to have complementary codes in the K shots, the following relation has to be met:

$$\sum_{k=0}^{K-1} \mathbf{v}_m^k = 1, \text{ i.e., for the } m^{\text{th}} \text{ location in the } K \text{ coded apertures of a particular source, only one of them can have a value of 1.}$$

The algorithm starts selecting all the combinations in \mathbf{W} that obey $\sum_{k=0}^{K-1} \mathbf{v}_m^k = 1$ and discards all the entries of the matrix that do not meet the constraint. From the updated matrix \mathbf{W} , the combinations that achieve uniformity in the detector plane are kept, the other combinations are discarded. If the matrix \mathbf{W} has more than one combination after the two previous iterations, the combination that minimizes the variance of the m^{th} row of the sensing matrix \mathbf{A} is selected, then the summation of all the columns of the sensing matrix possess a very low variance, achieving uniform sensing of the object simultaneously.

3.2.2. Efficient DBS algorithm

The DBS algorithm takes an initial set of codes generated with the algorithm described in Section 3.2.1 and the cost function (3) is evaluated and defined as the current error \bar{e} , i.e. $\bar{e} = \alpha \cdot \sum_{m=0}^{M-1} [[\mathbf{d}]_m - m_1]^2 + \beta \cdot \sum_{j=0}^{Q-1} [[\mathbf{r}]_j - m_2]^2 + \gamma \cdot c_3$. Then each pixel from each of the $K \times P$ codes is visited in a random raster path. For each pixel, the effects of swapping or toggling that pixel's value is evaluated in terms of the error \bar{e} . If one of the nine operations results in a reduction of \bar{e} , such operation is performed and the error term \bar{e} is updated; otherwise no change in the codes is made. This process is illustrated in Fig. 2(b), where the pixel highlighted in red can be swapped with its 8 nearest neighbors or toggled to black. The results of each of the nine operations are shown. In Fig. 2(b), swap operations 1, 2, 4, 5 and 7 would not be considered since they do not alter the value of \bar{e} . Once the operation is completed, the process is repeated for the next pixel, which is chosen randomly. The process continues until no change in \bar{e} is

produced after evaluating all the pixels in all the codes.

Updating the error \bar{e} , implies calculating $\bar{\mathbf{d}}$, $\bar{\mathbf{r}}$ and c_3 for every toggle or swap of a pixel. However, the multiplication of matrices \mathbf{C} and \mathbf{H} for the computation of vectors $\bar{\mathbf{d}}$ and $\bar{\mathbf{r}}$ demands significant computational resources. To reduce the computational burden of the error calculation, instead of recalculating $\bar{\mathbf{d}}$ and $\bar{\mathbf{r}}$ by the matrix multiplications defined in Section 3.1, an alternative definition for the calculation of the constraints is proposed.

Criterion 1 $\bar{\mathbf{d}}$: using the previous definition developed for the initial codes, $\bar{\mathbf{d}} = \frac{1}{K} \sum_{k=0}^{K-1} \sum_{i=0}^{P-1} \mathbf{C}_i^k \mathbf{G}_i$ and given that \mathbf{C}_i^k is a diagonal matrix, each of the elements of the vector $\bar{\mathbf{d}}$

can be defined as $\bar{d}_m = \frac{1}{K} \sum_{k=0}^{K-1} \sum_{i=0}^{P-1} [\mathbf{C}_i^k]_m [\mathbf{G}_i]_m$, where $[\mathbf{C}_i^k]_m$ is the m^{th} element in the main diagonal of $[\mathbf{C}_i^k]$. When the m^{th} pixel of a particular \mathbf{T}_i^k code is changed or toggled, vector $\bar{\mathbf{d}}$ remains unchanged except for its m^{th} entry in case of a toggle, or m and the entry corresponding to the neighbor of the pixel implied in the swap. Note these changes do not imply the multiplication of the matrices \mathbf{C} and \mathbf{H} . Instead, they rely only on the multiplication of the entries involved in the change, i.e., $[\mathbf{C}_i^k]_m$ and $[\mathbf{G}_i]_m$.

Criterion 2 $\bar{\mathbf{r}}$: this constraint is related to the sum of the columns of matrix \mathbf{H}_i . Therefore, a swap or a toggle of one of the elements of the codes results in changing all the elements of the vector $\bar{\mathbf{r}}$ as opposed to the previous constraint. To obtain a simplification of the original expression for constraint 2, it is expanded as: $\bar{\mathbf{r}} = \frac{1}{K} \sum_{k=0}^{K-1} \sum_{i=0}^{P-1} [\mathbf{C}_i^k \mathbf{H}_i]^T \mu_M = \frac{1}{K} \sum_{k=0}^{K-1} \sum_{i=0}^{P-1} \mathbf{H}_i^T \mathbf{C}_i^k \mathbf{T}_i \mu_M =$

$\frac{1}{K} \sum_{k=0}^{K-1} \sum_{i=0}^{P-1} \mathbf{H}_i^T \mathbf{J}_i^k$, where $\mathbf{J}_i^k = \mathbf{C}_i^k \mu_M$ is a column vector composed by the components of code \mathbf{T}_i^k . From the previous expression it can be seen that the m^{th} element of the code multiplies all the elements of the m^{th} column of the matrix \mathbf{H}_i^T , i.e., all the elements of the m^{th} row of matrix \mathbf{H}_i . Since the elements of the coded apertures are binary 0,1, a toggle of the m^{th} pixel will result in the subtraction (change of the pixel from 1 to 0) or addition (change of the pixel from 0 to 1) of the elements of the m^{th} row of the matrix \mathbf{H}_i to the current vector $\bar{\mathbf{r}}$.

The efficient DBS optimal process is summarized as follows:

1. Generate the initial set of codes, and calculate the initial error \bar{e} .
2. For each pixel in the coded apertures, evaluate the effect of all possible trial changes using the modified constraints and the definition for c_3 . Perform the change that results in a lower \bar{e} .
3. Stop when, after processing all the $K \times P$ coded apertures, no swaps or toggles occur.

4. Simulations

To simulate the compressive X-ray tomosynthesis configuration, a scenario with a flat 2D detector plane composed by $N_1 \times N_2 = 150 \times 150$ elements, $P = 9$ cone-beam X-ray sources placed uniformly in a 3×3 geometry and an object of interest \mathbf{f} represented by a $Q_2 \times Q_3 \times Q_1 = 128 \times 128 \times 16$ are used, each of the pixels in the coded aperture corresponds to a particular detector element as detailed in Fig. 3(a). Therefore, the coded apertures placed in front of each of the sources are also composed by 150×150 elements. The ASTRA Tomography Toolbox (“All Scale Tomographic Reconstruction Antwerp”) [14] was used to obtain the system matrices \mathbf{H}_i as well as the projection measurements \mathbf{y}_i of each of the X-ray cone beam sources. Using the algorithm described in Section III, optimal coded apertures for $K = 1$ and $K = 2$ shots are obtained. The performance of random coded apertures and the optimal codes is compared using the singular value analysis.

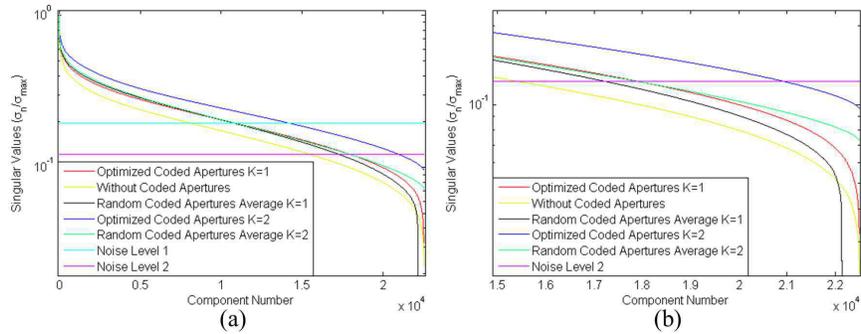


Fig. 4. (a) Singular Value Decomposition of the tomosynthesis matrix without coding, optimized codes and random codes for $K=1$ and $K=2$ shots (b) Singular Value Decomposition for the last 6900 components.

4.1. Singular value analysis

When two different measurement strategies are used to sense an object, the singular value decomposition (SVD) analysis can provide a simple mechanism for comparison [15]. The SVD of the matrix $\hat{\mathbf{A}}$ for the compressive X-ray tomosynthesis system showed in Fig. 4(a) is calculated. Scenarios of $K = 1$ and $K = 2$ shots are considered. For the latter case, a randomized method for computing an approximate singular value decomposition [16] is used due to the size of the matrix $\hat{\mathbf{A}}$. Therefore, only the first 22500 nonzero singular values of the matrix $\hat{\mathbf{A}}$ are obtained. Three different cases are analyzed for $K=1$ and $K=2$: (A) No coding ($\hat{\mathbf{A}} = \mathbf{H}$), which is equivalent to setting all the pixel elements of all the coded apertures to 1, for $K = 2$ the singular value decomposition is equivalent to $K = 1$ for this particular case; (B) Optimized codes using the algorithm previously described and the parameters describing the hardware settings $\hat{\mathbf{A}} = \mathbf{CH}$ are obtained and; (C) The coded aperture elements are generated randomly. For the latter, the mean for 20 different selections is obtained. Figure 4(a) presents the singular value decomposition for the cases previously discussed. Considering there is no prior information about the object under inspection, the measurement strategy that has more singular value components lying above certain noise level is considered to outperform the others, since it would capture more orthogonal components of the object. Thus, note that both random coding and optimized codes outperform, for any noise level, the case when no coding is used for both $K = 1$ and $K = 2$ shots. Additionally, the singular value spread for the curves corresponding to $K = 1$ is larger than for $K = 2$ thus showing that an increase in the number of shots results in a better measurement strategy to sense the object.

Two different noise levels are used in Fig. 4(a). It can be seen that for the higher noise level in the case of a single shot both optimized codes and random codes have similar behavior. However, for lower noise levels the optimized codes show better performance than the random codes, as it is shown in Fig. 4(b). For $K = 2$, Fig. 4(a) shows that optimized codes outperform random codes for both noise levels. From the SVD analysis, it can be concluded that optimized codes can provide advantage over random codes even under noisy conditions. This will be demonstrated in Section V for real data results.

For $K = 1$ and $K = 2$, the problem is very ill-conditioned. It can be noted that the number of measurements is much lower than the number of unknowns (voxels). The condition number (ratio of greatest singular value to the least nonzero singular value κ) measures how ill-conditioned the problem is [15]. The condition number (κ) for the three cases studied in this section for $K = 1$ show that when using optimized codes the sensing matrix becomes less ill-conditioned

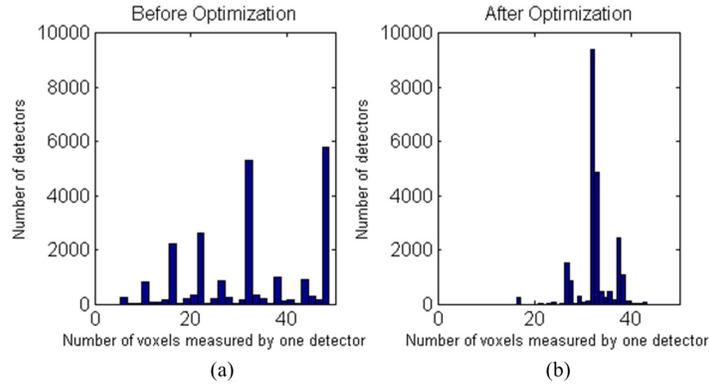


Fig. 5. Histogram of the number of voxels measured by a detector element, \bar{d} . (a) Before the optimization, (b) After the optimization.

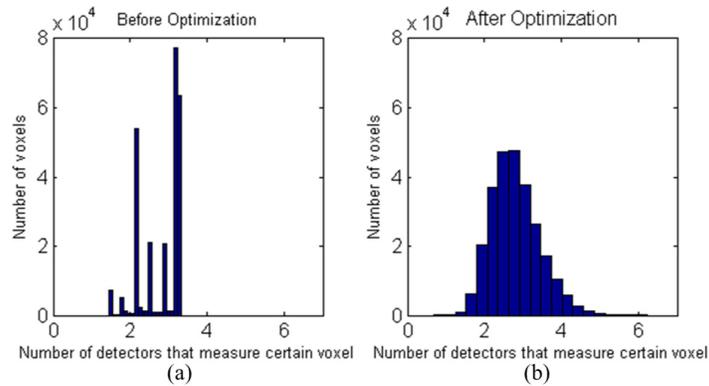


Fig. 6. (a) Histogram of the number of detectors that measure a certain voxel, \bar{r} . (a) Before the optimization, (b) After the optimization.

($\kappa = 15.60$) compared to using random codes ($\kappa = 562.12$) or no coding ($\kappa = 20.49$), showing that uniformly sensing of the detector plane and the data cube leads to better conditioning of the forward operator.

4.2. Results

Experimental tomography data was obtained at Chesapeake Testing Inc., with a Nikon metrology 225/450kV Vault CT scanning system with a 450kV micro-focus X-ray source capable of producing a spot-size down to $80\mu\text{m}$. The detector is a $16\text{in} \times 16\text{in}$ square plane and the detector pitch is $200\mu\text{m}$. Multiple X-ray projections over 360 degrees around an object, in this particular case a vivofit watch, are acquired. These projection images are then reconstructed into a full 3D volumetric data set.

This 3D data cube is re-sampled to obtain the data cube of size $128 \times 128 \times 16$ described at the beginning of Section IV. Assuming that the line integrals are measured directly and the hardware settings previously described, the measurements \mathbf{y} and the matrix \mathbf{H} are obtained as described in (1). The set of coded apertures \mathbf{T}_i^k with $i = 0, \dots, P - 1$ and $k = 0, \dots, K - 1$ was acquired using the algorithm for the coded aperture design introduced in Section III. It can be seen in Fig. 3(b) that the transmittance (τ_λ) decreases as the number of shots increase given

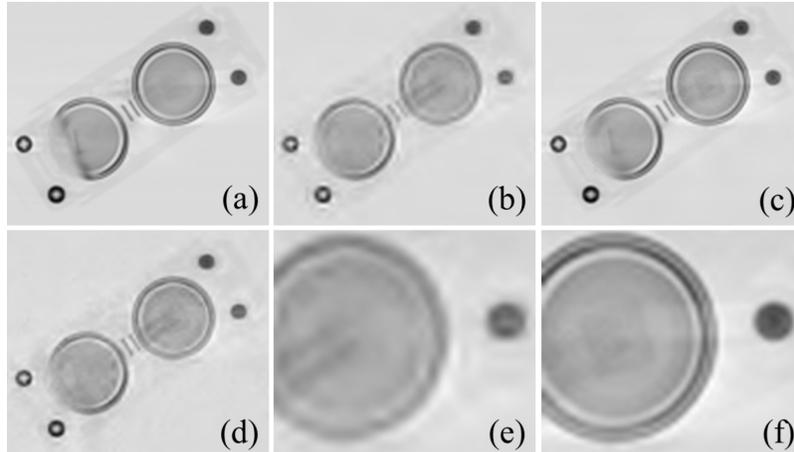


Fig. 7. (a) Thirteenth slice of the data cube. Sparse regularized reconstructions from: (b) Random coded X-ray projections using 3 snapshots (PSNR=25.96 dB); (c) Optimized coded apertures using 3 snapshots (PSNR=29.68 dB); (d) Uncoded X-ray projections using 1 snapshot (PSNR= 23.60 dB). Zoomed versions of: (e) Random coded X-ray projections; (f) Optimized coded apertures.

the constraint that the codes have to be complementary. The reconstruction algorithm used to recover the data cube is the GPSR (Gradient Projection for Sparse Reconstruction) [17]. The signal representation basis used to represent the three-dimensional data cube is a Kronecker product of a 2D wavelet transform and a 1D discrete Fourier transform (DCT) [18]. For $K = 3$, Figs. 5 and 6 show the histograms corresponding to the number of voxels measured by one detector and the number of detectors that measure certain voxel respectively. Figure 5(a) shows the distribution of the entries of the vector $\bar{\mathbf{d}}$ before the optimization. Figure 5(b) shows that the distribution of the entries of vector $\bar{\mathbf{d}}$ become concentrated around $m_1 = 32.47$; thus an average of 32 voxels are measured per detector, after the optimization. The initial distribution of the entries of the vector $\bar{\mathbf{r}}$ is shown in Fig. 6(a). After the optimization, a more uniform distribution concentrated around $m_2 = 3.21$ is obtained, as shown in Fig. 6(b); thus, every voxel is sensed an average of 3 times. The peak signal-to-noise ratio (PSNR) is used to compare the reconstructions obtained since it is suitable for comparing restoration results as it does not depend strongly on the image intensity scaling. For a scenario with an image I and a reconstruction R of size $N \times N$ it is defined as $PSNR = 10 \log_{10} \left(\frac{Max_I^2}{MSE} \right)$, where Max_I is the maximum possible pixel value of the image I and $MSE = \frac{1}{N^2} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} [I(i, j) - R(i, j)]^2$. Table 1 shows the PSNR of the reconstructions of the thirteenth slice for $K = 1, 2, 3, 4$ and 5 and for optimized codes and random codes. The elements of the random coded apertures used are random realizations of Bernoulli random variables, with different levels of transmittance. For the multi-frame scheme, the transmittance of the codes is fixed depending on K for one case and for comparison, another scenario is analyzed when the transmittance is fixed to $\tau_\lambda = 0.5$. The compression ratio, will be given in each case by $\rho = 1 - (M \times K)/Q$. Therefore, maximum compression is obtained when a single shot is used. It can be seen that as the number of shots increases the reconstruction quality improves. Nonetheless, the improvement is not significant after 3 shots, for the scenario used for the simulations, since the number of unknowns is limited. For a data cube composed by more slices, increasing the number of shots would lead to further improvement. Clearly, the best results are obtained using the optimized coded apertures (first column in Table 1).

Table 1. PSNR of the reconstructed image of the 13th slice for different number of shots (K)

Shots	PSNR (dB)			ρ
	Optimal Codes	$\tau_\lambda^* = 0.5$	$\tau_\lambda = 1/K$	
1	26.36	24.83	24.67	91.41%
2	28.29	25.82	25.92	82.83%
3	29.68	25.96	26.14	74.25%
4	29.76	26.45	26.48	65.67%
5	29.89	27.17	27.57	57.08%

Figure 7(a) shows the thirteenth slice of the three-dimensional data cube used for the simulations. By acquiring uncoded measurements from 1 snapshot ($\mathbf{A} = \mathbf{H}$), the reconstruction shown in Fig. 7(d) is obtained. Coded X-ray projections are next used in the measurements where random binary patterns with transmittance $\tau_\lambda = 0.5$ are used as coded apertures. For this scenario, 3 snapshots are used. The measurement set is now less correlated, such that improved reconstructions are obtained as depicted in Fig. 7(b). As stated in previous sections, random codes do not exploit the known geometry of the tomographic system. By applying the optimization algorithm described in Section 3, with $m_1 = 32.47$ and $m_2 = 3.21$, the improvement in the reconstruction PSNR can be observed in Fig. 7(c). Moreover, Figs 7(e) and 7(f) show zoomed versions of the reconstructions obtained when using random codes and optimized coded apertures respectively. The PSNR gain is evident in the zoomed versions of the reconstructions. The PSNR for slice 13 for random codes and $K = 3$ shots is 25.96 dB, for optimized codes is 29.68 dB and for the least squares approach is 29.10 dB. For the latter, least squares estimation is used to reconstruct the X-ray tomosynthesis problem, i.e. when each source produces a set of measurements on the detector. Note that the traditional least squares reconstruction uses 3 times the amount of measurements than the compressive X-ray tomosynthesis approach with $K = 3$ shots. Furthermore, the latter reduces the radiation exposure of the patient/sample.

For the simulation results slice 1 is the closest slice to the sources and slice 16 is the slice located farthest away from the sources. Table 2 shows the PSNR of the reconstruction of the 16 slices of the data cube for $K = 3$ shots, for both optimized codes and random codes. Furthermore, Figs. 8(a) and 8(d) depict the slices 1 and 16 of the original data cube respectively, and the reconstructions obtained when using optimized coded apertures and $K = 3$ snapshots are shown in Figs. 8(b) and 8(e) for each of the slices. Figures 8(c) and 8(f) depict the least squares reconstructions for slices 1 and 16 respectively, where least squares estimation is used to reconstruct the X-ray tomosynthesis problem when the full set of X-ray projections are used.

Table 2. PSNR of the reconstructed image of the 16 slices of the data-cube for $K = 3$ shots

Slice	1	2	3	4	5	6	7	8
Optimized (PSNR dB)	28.5	27.6	29.2	27.8	27.7	27.3	27.7	26.8
Random (PSNR dB)	24.7	25.5	25.7	25.7	26.3	26.4	26.1	26.1
Slice	9	10	11	12	13	14	15	16
Optimized (PSNR dB)	26.4	27	28	29.3	29.7	29.7	29.6	25.4
Random (PSNR dB)	25.9	26.3	26.5	24.9	26	24.5	24.1	22.6

The optimized coded apertures used for the central source and the source located in the lower right corner for $K = 2$ are depicted in Fig. 9. Figures 9(a) and 9(b) depict the coded

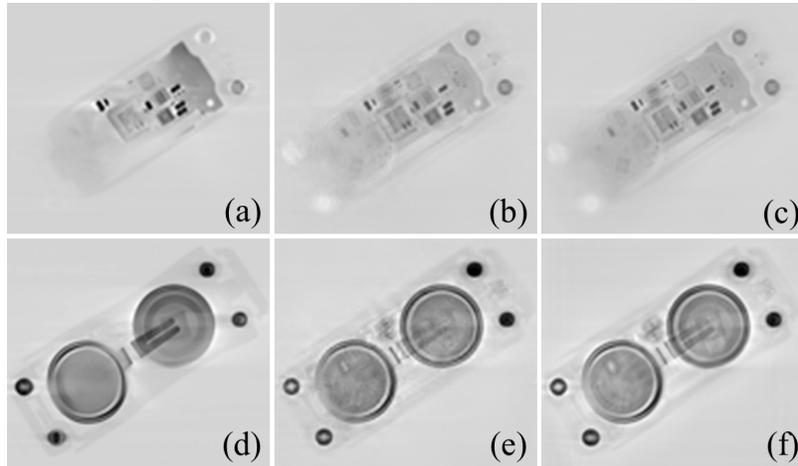


Fig. 8. (a) First slice of the data cube. (b) Sparse regularized reconstructions from optimized coded apertures using 3 snapshots (PSNR=28.47 dB). (c) Least squares reconstruction using the full matrix (PSNR=28.27 dB); (d) 16th slice of the data cube. (e) Sparse regularized reconstructions from optimized coded apertures using 3 snapshots (PSNR=25.45 dB). (f) Least squares reconstruction using the full matrix (PSNR=25.46 dB). Note that LS reconstructions uses 3 times the amount of measurements than the compressive X-ray tomosynthesis.

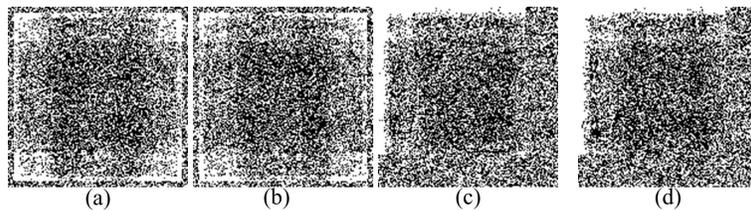


Fig. 9. Optimal coded apertures for: Two snapshots ($K=2$), (a) the central source and first snapshot, (b) the central source and second snapshot, (c) the source located in the lower right corner and first snapshot, (d) the source located in the lower right corner and second snapshot.

apertures corresponding to the source located in the center and for $K=2$, Figs. 9(c) and 9(d) correspond to the coded apertures used for the source located in the lower right corner for $K=2$. Notice the non-uniform density of the designed codes as well as the structured patterns, which also vary from location to location. Optimized coded apertures for $K = 3$ and the central source are depicted in Fig. 10; the decrease in the transmittance (τ_λ) is evident between the coded apertures used for $K = 2$ (Fig. 9) and $K = 3$ (Fig. 10). A one-dimensional cross-section of coded aperture elements in column 50, rows 130 to 140 in each of the codes used for the central source is shown in Fig. 10. Note for 9 of the 10 cases analyzed, only one out of 3 coded apertures contains a non-zero value for a specific spatial location in the codes, which was the condition to assure complimentary codes.

Convergence of the DBS algorithm as stated in Section 3.2 depends on the initial set of coded apertures that are selected, Fig. 11(a) presents the convergence of the DBS algorithm when using random codes and a set of checkerboard codes as opposed to the optimized initialization

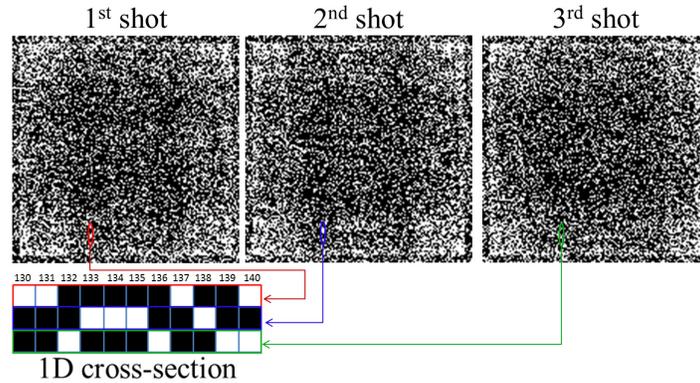


Fig. 10. Optimal coded apertures for three snapshots ($K=3$) and the central source and a 1D cross section of the coded aperture elements in column 50, rows from 130 to 140.

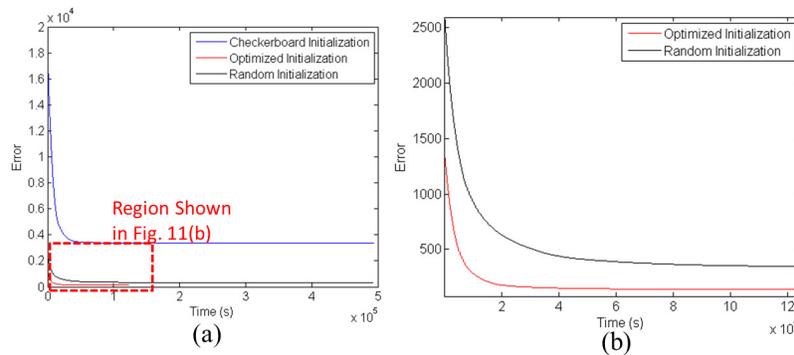


Fig. 11. (a) Convergence of DBS algorithm for different initial set of codes: (blue) checkerboard, (red) optimized set of codes, (black) random set of codes. (b) Convergence of DBS algorithm for the first 1.42 days (122,500 seconds)

proposed, for the simulation scenario previously discussed. Note that using a predefined pattern (checkerboard) results in a higher error compared to both initial random and optimized coded apertures. Furthermore, the initial error, the time of convergence and the final error are higher for random coded apertures compared to the initial optimized set of coded apertures as depicted in Fig. 11(b).

5. Testbed implementation

Experimental tomosynthesis data was obtained at Chesapeake Testing Inc. with the system described in Section IV. The energy used for the source is 245keV and five projections corresponding to five different locations of the same source were obtained. The source is moved along one line due to constraints of the testbed system. To obtain the measurements, the source is moved 4 times, 5cm at a time from the center position, which is aligned with the center of the detector. The detector remains static for all the measurements. The sources were located 892mm away from the object and the detector was placed 1100mm away from the source. The object imaged is a RJ45 cable, discretized as 12 slices of 128×128 pixels with voxels of dimensions $0.4\text{mm} \times 0.4\text{mm} \times 0.8\text{mm}$. The detector was composed of 140×240 elements and the size of the detector elements was 0.4mm. Figures 12(a)(b) depict the projections obtained from the central

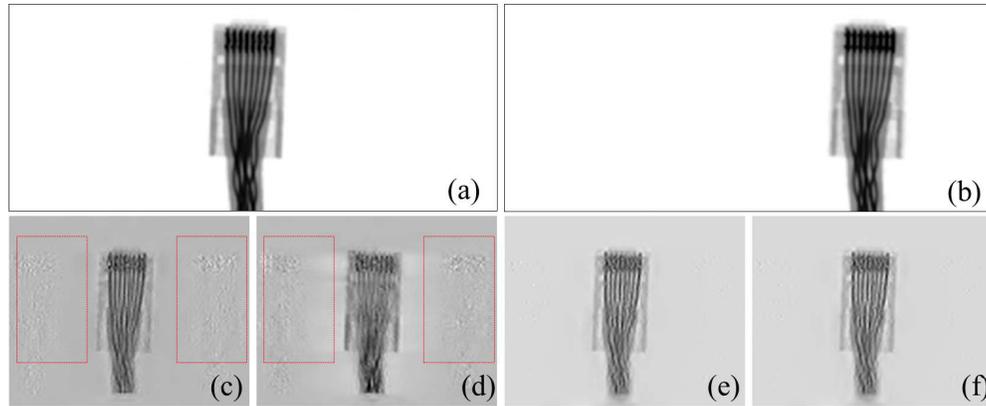


Fig. 12. Projections obtained from: (a) the central source, and (b) a source located 10 cm to the left of the center source. Reconstructions obtained using 2 shots and random coded apertures for: (c) the 6th and (d) 12th slice. Reconstructions obtained using 2 shots and optimized coded apertures for: (e) the 6th and (f) 12th slice.

source and the adjacent to it. Using the ASTRA Tomography Toolbox, the matrix \mathbf{H} is obtained for the hardware settings specified before. The coded apertures used in the simulations are assumed to match the pixels in the detector, using the data from the projections and the simulated optimized codes for the configuration the projections are superimposed to obtain the multiple source system proposed. Notice the artifacts highlighted in the reconstructions obtained when using random coded apertures in Figs. 12(c) and (d). Furthermore, these reconstructions show less image quality than those obtained using optimized coded apertures, as can be seen in Figs. 12(e) and (f).

6. Conclusions

A new algorithm for the coded aperture design for compressive X-ray tomosynthesis has been introduced. Simulations show an improvement of up to 3dB in PSNR for reconstructions obtained from optimized codes compared to random codes. The optimization does not depend on the object under inspection. Instead it is based on the criteria to achieve uniform sensing of the object and the detector plane while obtaining complementary codes. Experimental results demonstrate the spatial and spectral accuracy of the system. It has also been shown that increasing the number of shots while reducing the transmittance of the coded aperture, due to the complementary nature of the coded apertures, leads to improved image quality in the reconstructions. The optimization yields improved results since the three-dimensional object and the detector plane are uniformly sensed. This conclusion is based on the singular value decomposition (SVD) analysis and the condition number of the forward operator for each case. Additionally, a test bed implementation was presented with reconstructions for real data acquired from a high-resolution XCT system. Source location optimization and the fabrication of the coded apertures are under development. Calibration procedures will be used in order to mitigate the mismatching errors that might occur. Additionally, the angular collimation produced by the implemented coded apertures can be accounted for in the sensing matrix \mathbf{A} . The optimized codes would take into account this phenomenon since the optimization is based on \mathbf{A} .