Noise- and Outlier-Resistant Tomographic Reconstruction under Unknown Viewing Parameters
Paper ID - 53

ABSTRACT

In this paper, we present an algorithm for effectively reconstructing an object from a set of its tomographic projections without any knowledge of the viewing directions or any prior structural information, in the presence of pathological amounts of noise, unknown shifts in the projections, and outliers among the projections. The outliers are mainly in the form of a number of projections of a completely different object, as compared to the object of interest. We introduce a novel approach of first processing the projections, then obtaining an initial estimate for the orientations and the shifts, and then define a refinement procedure to obtain the final reconstruction. Even in the presence of high noise variance (up to 50% of the average value of the (noiseless) projections) and presence of outliers, we are able to successfully reconstruct the object. We also provide interesting empirical comparisons of our method with the sparsity based optimization procedures that have been used earlier for image reconstruction tasks.

KEYWORDS

tomographic reconstruction, tomography under unknown viewing directions, outliers, image reconstruction

1 INTRODUCTION

Reconstructing the structure of an object from its tomographic projections (hereafter referred to as 'projections') is a fundamental research problem that arises in diverse fields, such as medical imaging [9] and reconstruction in cryo-electron-microscopy (referred to hereafter as 'Cryo-EM') [10]. If the viewing orientations are known *a priori*, standard algorithms such as Filtered Backprojection (FBP) or its variants adapted for different acquisition geometries [8], can be used to reconstruct the image. However, there are many scenarios where the viewing orientations are unknown. One such example is Cryo-EM where the objective is to determine the structure of a macro-molecule or a biological specimen such as a ribosome or virus, from its projections which essentially appear in various unknown orientations [10, 19]. Other examples include insect tomography [11], or tomography of objects performing unknown rigid motion [21]. Both these are equivalent to performing tomographic reconstruction on a fixed object, with the viewing directions being unknown. Uncertainty in viewing angles may also occur due to patient motion in medical imaging, even though to a lower degree. In recent times, significant research has emerged in the field of 'tomography under unknown viewing parameters'. Much of this research belongs to one of the following two categories:

(1) Machine learning based approaches, which are based on assumptions on the distribution of the unknown parameters, typically the uniform distribution. This essentially requires

that the number of projections is large. However, certain orientations might be more likely for a given structure [19], and thus would yield the aforementioned assumption unfounded. In fact, the orientation distribution may not even be known in advance. For example, in [1] a nearest neighbour search is performed over the projections to find an ordering of the angles, and then these ordered projections are assigned to angles chosen at uniform intervals on the unit half-circle (i.e. from 0 to 180 degrees). Similarly in [7], a variant of the popular Locally Linear Embedding (LLE) algorithm, termed 'Spherical LLE', is used to embed the projections on a 2D circle and finally the 2D projections are sorted and arranged in a uniform manner around the circle. A graph Laplacian based clustering approach with similar assumptions is used in [5].

(2) The other major class of algorithms have used the Moment based approach [18] which uses the Helgason Ludwig Consistency Conditions (HLCC). These conditions relate the geometric moments of the underlying image $z(x, y)$ and the geometric moments of its tomographic projections at any angle. An alternating optimization produces estimates of the unknown angles, as well as the image moments as a by-product [6].

It has even been proved that in the case where the projections are noiseless, a unique solution exists for the given HLCC [2] under some weak assumptions. However, if we apply this algorithm in practical applications it may yield very poor results. This is because practical applications often bring with them a whole new set of challenges, such as those enlisted below.

(1) In Cryo-EM, most biological specimens are extremely radio-sensitive, so they must be imaged with low-dose electron beams which leads to extremely high amounts of noise in the projections. The geometric moments are known to be highly sensitive to noise which compromises the entire procedure.

(2) Moreover, samples of the same biological specimen (eg: virus, ribosome, etc.) that are acquired on a single slide may often not be exactly identical due to contaminants such as ice particles as well as genuine conformational changes, and therefore there is an added complexity due to outliers [12].

(3) Some of the projections may even be shifted by a small random amount, which if not corrected, produces a very poor fidelity reconstruction. In medical tomography or insect tomography, this occurs due to subject motion.

(4) In Cryo-EM, the reason for the shifts is different. Here, a single slide contains potentially thousands of samples of the same biological specimen. The acquired image (often called a 'micrograph') thus consists of thousands of noisy projections against a very noisy background. This necessitates a pre-processing step where the projections of the

actual specimen (termed 'particles') are detected via a procedure called 'particle picking' [19]. Since such a detection procedure will usually not be perfect, there are bound to be small translational errors between the actual and predicted particle location. This leads to uncertainty in the positions of the particles.

Summarily, there are four main challenges that need to be addressed, in order to accurately estimate the underlying structure from the given set of projections: (1) Severe levels of noise in the projections, (2) Unknown orientations of the projections, (3) Unknown or inaccurately known (albeit small) shifts in the projections, and (4) Presence of outliers.

In this paper, we present an algorithm which (1) systematically removes the outliers, (2) clusters similar projections together, (3) uses a moments based approach to obtain an initial estimate for the orientations and the shifts, and (4) finally optimizes for the structure of the unknown object along with a refinement of the viewing parameters (i.e., the angles and the shifts). As we will see, this algorithm can successfully determine the structure of the object from its projections taken in unknown orientations despite the presence of noise, outliers and random unknown shifts. In this paper, we work exclusively with 2D images (and hence simulated 1D parallel-beam projections) for reconstruction, even though actual objects are 3D. This follows previous work in the image processing community which has studied the 2D variant of this problem extensively [1, 20]. Nonetheless, the underlying principles remain the same, and the computational problem remains very challenging even for reconstructing 2D images.

2 ALGORITHM DESCRIPTION

2.1 Robust Clustering

Typically in the case of Cryo-EM, a large number of projections are available. But all of these projections have pathological amounts of noise and there may even be a fair number of completely erroneous projections which owe themselves, not to the actual particle, but to the presence of foreign objects and ice particles [12]. We henceforth refer to these completely erroneous projections as 'outliers of Class 1'. Reconstructing the object from these projections directly understandably yields very poor results. To combat this, we seek to cluster the projections into a small number of classes, based on orientation and structural similarity. The aim is to produce a representative set of projections that will be significantly less noisy, while simultaneously detecting and rejecting outliers of Class 1. We use the K-means algorithm [16] to cluster a large number of projections into a much smaller number of clusters K_c. The distance metric used is the ℓ_q quasi-norm ($0 < q \leq 1$) and therefore the cluster centroid is expected to be robust to outliers. The objective function that is minimized is as follows:

$$(L_{centroid})(\{\xi_j\}_{j=1}^{K_c}) = \sum_{j=1}^{K_c} \sum_{z_i \in \pi_j} \|z_i - \xi_j\|_q, \tag{1}$$

where there are K_c clusters, π_j represents the jth cluster and ξ_j represents the jth cluster centroid. In the case when $q = 1$, the cluster centroid would be the element-wise median of the points belonging to the cluster.

2.2 Removal of Class 1 Outliers

After clustering, we remove some $f\%$ of the projections based on their ℓ_2 distance from the closest cluster centroid. It is likely that since a completely erroneous projection is located far away from the other projections, a cluster will not be formed close to it. Therefore, removing $f\%$ of projections that are placed farthest from any cluster centroid will remove the Class 1 outliers. A reasonable estimate of f can be provided by a biologist upon eye-balling of the micrograph, and usually, a moderate over-estimate of f is not a problem.

2.3 Averaging to form a single cluster

Within each cluster, we define the processed projection image $\tilde{p}_j$ (for cluster index j) to be the average of all the projections assigned to that cluster, and which were not discarded by the previous step. This is mathematically represented as follows:

$$\tilde{p}_j = \frac{\sum_{z_i \in \pi_j} z_i(1 - I_j(z_i))}{\sum_{z_i \in \pi_j}(1 - I_j(z_i))} \tag{2}$$

where $I_j(z_i) = 1$, when the ith projection image belonging to the jth cluster is discarded, and 0 otherwise. Note that this procedure is very similar to M-estimators (such as the Huber estimator) for robust class means [13]. The only difference is that the Huber estimator computes a weighted linear combination of the sample points, whereas we use a 'harder' form of weights, i.e. we discard the outliers entirely.

2.4 Patch-Based Denoising

The processed cluster centers $\{\tilde{p}_j\}_{j=1}^{K_c}$ as obtained in the previous step are expected to be devoid of outliers. The averaging also induces a basic form of filtering to remove the noise. However, some residual noise still remains. Hence, we pass these cluster centers ($\{\bar{p}_j\}_{j=1}^{K_c}$) through a denoising algorithm described next. We use a patch-based PCA denoising method to reduce the noise in the projections. This algorithm is adapted from a similar algorithm for 2-D images, as described in [17]. In this algorithm, we extract small-sized patches from each cluster center. For each such patch, we find some L patches nearest to it in terms of the ℓ_2 distance. After performing PCA on this set of patches, we project each patch along the principal direction to produce eigen-coefficients. To denoise the patch, we manipulate these coefficients using Wiener-like updates of the form

$$\beta_{il} = \alpha_{il} \left(\frac{\sigma_l^2}{\sigma_l^2 + \sigma^2/\bar{K}} \right) \tag{3}$$

where β_{il} is an estimate of the lth denoised coefficient for patch i (part of cluster center $\bar{p}_j$), α_{il} is the corresponding noisy coefficient, σ^2 is the noise variance in the original projections which is assumed to be known, $\bar{K}$ is the average number of projection vectors assigned to a cluster, and σ_l^2 is the mean square value of the lth coefficient estimated as follows:

$$\sigma_l^2 = \max \left(0, \frac{1}{L} \sum_{i=1}^{L} \alpha_{il}^2 - \sigma^2/\bar{K} \right). \tag{4}$$

This patch-based PCA denoising algorithm is better than the PCA denoising algorithm used in [20]. This is because, in [20], entire projections are compared instead of just patches. The advantage

of our patch-based approach is that it is easier to find a number of patches which are structurally similar to a given reference patch, but that is not true for entire projections. The second advantage is that we now have the freedom to find similar patches from within a projection vector, but from other projection vectors as well. Moreover, we performed the denoising in sliding window fashion with a pixel stride of 1, resulting in several potential denoised values per pixel. The final denoised value was selected using averaging. Hereafter, we use the symbol $\tilde{q}_i$ to refer to the denoised version of the cluster center $\tilde{p}_i$ after the outlier removal step.

2.5 Initialization of the orientations and shifts using Helgason Ludwig Consistency Conditions (HLCC)

Determining the object structure from projections with unknown viewing parameters is a highly non-convex optimization problem. Starting from a random initialization of orientations, the algorithm may mostly converge to a local optimum and fail to give us the right structure. This is why a good initialization of the orientations and the shifts is necessary. Initially, we attempted to correct the shifts by shifting all the projections such that their center of mass is at the origin as stated in [2]. In practice, however, in spite of correcting a few shifts, it still resulted in unsatisfactory reconstruction. So we modified our approach and harnessed the information available in the image moments and projection moments to estimate the shifts and the orientations simultaneously. The HLCC [18] give us a relationship between the geometric moments of the underlying image $z(x, y)$ and those of its unshifted projections at any angle. We use this to derive a good initial estimate of the unknown orientations and shifts. The basis behind the algorithm is as follows. The moments of order p, q of the image $z(x, y)$ are given by

$$v_{p,q} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} z(x,y) x^p y^q dx dy. \qquad (5)$$

The nth order moment of the projection $g(\rho, \theta) \triangleq \int_{-\infty}^{\infty} z(x,y)\delta(\rho - x\cos\theta - y\sin\theta)dxdy$ is given by

$$m_{\theta}^{(n)} = \int_{-\infty}^{\infty} g(\rho, \theta)\rho^n d\rho. \qquad (6)$$

If the projection is shifted by s_i to give a projection $g(\rho, \theta, s_i)$, its nth order moment after reverse shifting by an amount s_k can be written as

$$m_{\theta, s_k}^{(n)} = \int_{-\infty}^{\infty} \mathcal{S}\{g(\rho, \theta, s_i), s_k\}\rho^n d\rho \qquad (7)$$

where $\mathcal{S}\{., s_k\}$ denotes the reverse shift operation. The above evaluates to the same quantity as (6) if $s_k = s_i$. That is,

$$m_{\theta, s_i}^{(n)} = \int_{-\infty}^{\infty} \mathcal{S}\{g(\rho, \theta, s_i), s_k\}\rho^n d\rho = m_{\theta}^{(n)}. \qquad (8)$$

The HLCC give a relationship between $m_{\theta, s_i}^{(n)}$ and $v_{p,q}$, which is defined as

$$m_{\theta, s_i}^{(n)} = \sum_{j=0}^{n} \binom{n}{j} (cos\theta)^{n-j} (sin\theta)^j v_{n-j, j}. \qquad (9)$$

Thus for each order n, we can write the constraints in matrix form, $\mathbf{m}^{(n)} = \mathbf{A}^{(n)}\mathbf{v}^{(n)}$. Here, for a total of K_c projections and for the nth order equation, $\mathbf{A}^{(n)}$ is the $K_c \times (n+1)$ matrix defined by $A_{ij}^{(n)} =$

$\binom{n}{j}(cos\theta_i)^{n-j}(sin\theta_i)^j$, and $v^{(n)} \triangleq \{v_{p,q}|(p+q) = n, p, q \in \mathbb{Z}_{\geq 0}\}$. Since, in practice, the projections are noisy and the shifts unknown, Eqn. 9 will not be satisfied exactly. Instead we define an energy function as follows

$$E(\{\theta_i\}, \mathbf{v}, \{s_i\}) = \sum_{n=0}^{k} \sum_{i=1}^{K_c} \left(m_{\theta_i, s_i}^{(n)} - \sum_{j=0}^{n} A_{i,j}^{(n)} v_{n-j, j} \right)^2. \qquad (10)$$

Note that in this equation, the moments $m_{\theta_i, s_i}^{(n)}$ correspond to those of the ith cluster center $\tilde{q}_i$ (post-denoising) where $1 \leq i \leq K_c$. By minimizing this energy function, we derive an initial estimate of the angles and the shifts through an iterative coordinate descent strategy as in [6]. A small number of multi-starts (around 10), each with a different random initialization of the pose parameters, helped further combat the non-convexity of the objective function $E(\{\theta_i\}, \mathbf{v}, \{s_i\})$, and in fact yielded much superior results. In case of multiple starts, the solution which yielded the least value of the objective function, was selected.

2.6 Optimization strategy to obtain the structure of the object

After clustering and obtaining an initial estimate of the angles and shifts, we initially decided to use a sparsity based optimization technique due to the promising results delivered by compressed sensing [4]. The following optimization problem was selected:

$$\mathcal{L}(\{\theta_i\}, \beta, \{s_i\}) = \sum_{i=1}^{K_c} \|\tilde{q}_{i, s_i} - \mathcal{R}_{\theta_i}(U\beta)\|_2^2 + \lambda_1 \|\beta\|_1 \qquad (11)$$

where $\{\theta_i\}_{i=1}^{K_c}$, $\{s_i\}_{i=1}^{K_c}$ are the K_c unknown angles and shifts for the cluster centers $\{\tilde{q}_i\}_{i=1}^{K_c}$, $\tilde{q}_{i, s_i}$ denotes $\tilde{q}_i$ shifted by s_i, the matrix $\mathcal{R}_{\theta_i}$ represents the set of line integrals at different shifts along the direction θ_i (for computing the Radon transform), U denotes the inverse discrete cosine transform (DCT) operator or any other sparsifying operator, and β is the vector of DCT or other transform coefficients of the image to be reconstructed. That is, the image is represented as $z = U\beta$, where β is a sparse or compressible vector of transform coefficients. The function was minimized using an alternating method over the unknown angles $\{\theta_i\}_{i=1}^{K_c}$, the unknown shifts $\{s_i\}_{i=1}^{K_c}$, and image DCT coefficients β.

Solving this problem without the initial estimates is extremely ill-posed because of the large number of degrees of freedom. Therefore, we hypothesized that the initial estimates provided by the moment-based estimation would serve as a good initial estimate for this problem and we would converge onto the accurate structure of the object. However, upon solving this optimization problem, the obtained results exhibited severe artifacts, as we show in Section 3. We also noticed convergence problems. If the accurate angles and shifts were provided, the sparsity-based technique, however, provided excellent reconstruction and no convergence problems arose. Hence we concluded that this reconstruction method is indeed very sensitive to errors in the angles, and even the initial estimate from the moment based approach was not accurate enough for this optimization to converge.

Back to FBP: In light of the above observations, we decided to consider an alternative optimization problem which we observed was

significantly more robust to errors in the initial estimate. We hence considered the following optimization problem, and decided to minimize the following function using a stochastic gradient descent approach:

$$\mathcal{M}(\{\theta_i\}, z, \{s_i\}) = \sum_{i=1}^{K_c} \|\tilde{q}_{i,s_i} - \mathcal{R}_{\theta_i}(z)\|_2^2. \tag{12}$$

Given an initial estimate of the angles and the shifts, the gradient with respect to the image z, is given by

$$\nabla_z \mathcal{M}(\{\theta_i\}, z, \{s_i\}) = \sum_{i=1}^{K_c} -2\mathcal{R}_{\theta_i}^T(\tilde{q}_{i,s_i} - \mathcal{R}_{\theta_i}(z)), \tag{13}$$

where $\mathcal{R}_{\theta_i}^T$ is the adjoint operator for $\mathcal{R}_{\theta_i}$. Using the initial estimate of the orientations and the shifts provided by the moments based approach, the problem was solved in an alternating way. We first estimated the structure using the gradient calculated above, which effectively makes use of FBP-based reconstruction. Given an estimate of the structure, the orientation and the shift of each projection by coordinate descent with a single-dimensional brute-force search. The complete procedure is summarized below in Algorithm 1.

Algorithm 1 Complete algorithm

$C_i \leftarrow$ Cluster assignment of the ith projection
$\xi_j \leftarrow$ Centroid of the jth cluster
$I_i \leftarrow$ Indicator of ith projection being filtered or not

$\{C_i\}, \{\xi_j\} = Lq\_KMeans(\{y_i\}), 0 < q <= 1$
$\{I_i\} = Outlier\_Detection(\{y_i\}, \{C_i\}, \{\xi_j\})$
$\{\tilde{p}_i\} = Robust\_Average\_Projections(\{y_i\}, \{C_i\}, \{\pi_j\}, \{I_i\})$
$\{\tilde{q}_i\} = Denoise\_Projections(\{\tilde{p}_i\})$
$\{\theta_i\}, \{s_i\} = MomentsBasedSolver(\{\tilde{q}_i\}, k)$

$\alpha \leftarrow LearningRate$
$E \leftarrow \infty$
$\Delta E \leftarrow \infty$

while $\Delta E > \tau$ **do**
    Choose a random subset of projections of size $K' < K_c$
    $\{\tilde{q}_l\}_{l=1}^{K'}$

    Update the image using the K' projections
    $z \leftarrow z - \alpha(z - FBP(\{\tilde{q}_l\}, \{\theta_l\}, \{s_l\}))$

    Refine all the K_c angle and shift estimates
    for $i = 1 : K_c$ **do**
        $\{\theta_i\}, \{s_i\} = Best\_Theta\_And\_Shift(z, \{\tilde{q}_i\})$
    end for

    $\alpha \leftarrow \alpha - AnnealRate$
    $\Delta E = E - \sum_{i=1}^{K_c} \|\tilde{q}_{i,s_i} - \mathcal{R}_{\theta_i}(z)\|_2^2$
    $E = \sum_{i=1}^{K_c} \|\tilde{q}_{i,s_i} - \mathcal{R}_{\theta_i}(z)\|_2^2$
end while

---

3 RESULTS

In this section, we present a complete set of results on the algorithm described in the previous section. Using our algorithm, we demonstrate how to successfully tackle all the problems mentioned earlier and achieve good quality reconstructions of the original object. The images used for our experiments were taken from the Yale and ORL face databases and the image sizes used were 192×192 and 112×112 respectively. A total of $Q = 2 \times 10^4$ projections per image were simulated using angles from Uniform$(0, \pi)$[1]. A fraction f_1 of these projections were outliers of class 1, i.e. they were projections of non-face images taken from the CIFAR-10 dataset [15]. For another fraction f_2 of projections, we deliberately generated them from a copy of the same image, but with a small number (f_3%) of pixel values (at randomly selected locations) set to 0. We term the corresponding projections 'Outliers of Class 2'. These simulate projections of biological specimens corrupted by overlapping ice particles or minor structural changes. Some sample illustrative images are presented in Fig. 1. All projections were subjected to additive i.i.d. noise from $\mathcal{N}(0, \sigma^2)$, where we assume σ to be known in advance, even though there are techniques to estimate it directly from the noisy projections. The noisy projections were clustered into $K_c = 180$ angles. The outlier projections were then identified and removed using the procedure described earlier in Section 2.2. We observed that in many cases, the outlier projections lying farther away from the cluster centroid were filtered out, and only the true projections of the object were preserved for the consequent steps. Fig. 2 shows an example of outlier and noisy inlier projections as well as the average projection vector after removing outlier projections. This example is for a face image from the ORL database, with outliers from the CIFAR dataset. The relevant parameters are $f_1 = 10\%$, $f_2 = 10\%$, $f_3 = 10\%$ and $\sigma = 0.1 \times$ average value of noiseless projections.
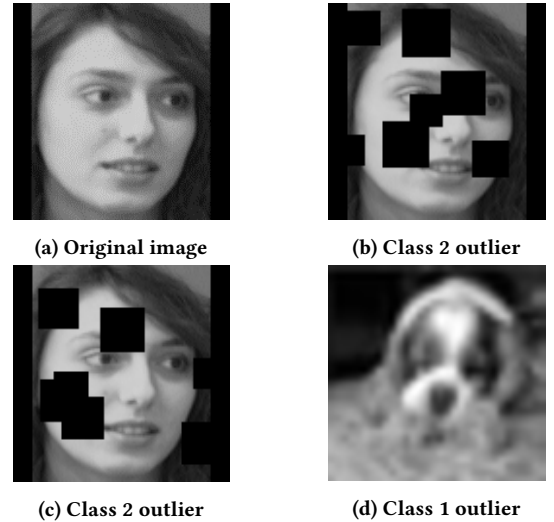

(a) Original image   (b) Class 2 outlier

(c) Class 2 outlier   (d) Class 1 outlier

Figure 1: Original image, and images which generated Outliers of class 1 and class 2

---

[1]Though we considered the Uniform distribution, our algorithm does not rely on this assumption, or knowledge of the distribution of the orientations.
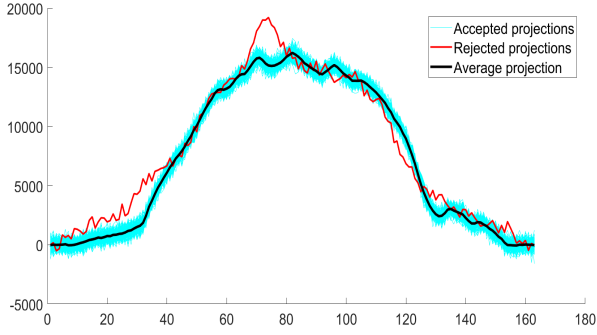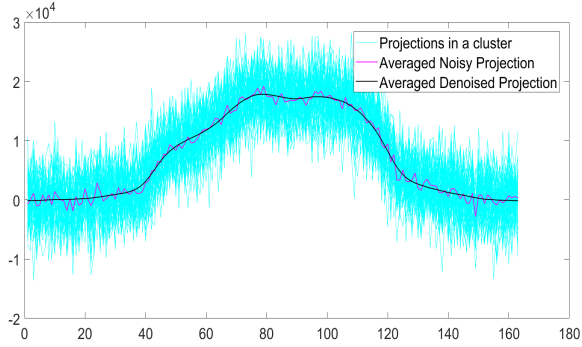
Figure 2: The red projection belonging to the class 1 outlier is filtered out in this step.

After outlier removal, the remaining projections in each cluster were averaged and passed on to the patch-based denoising step. An example of the projections after the denoising step are shown in Fig. 3. This is for a noise level of $\sigma = 0.5\times$ average value of noiseless projections.

(a) Denoising Example 1, 50% noise
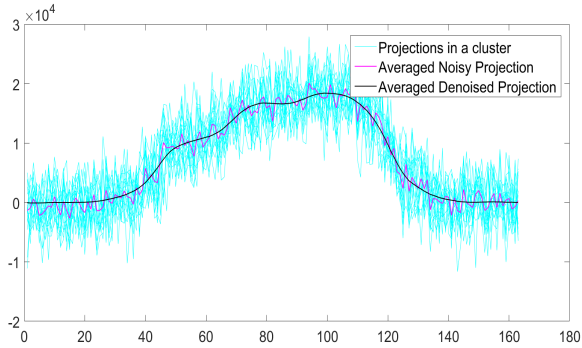
(b) Denoising Example 2, 50% noise

Figure 3: Patch-based denoising for two different clusters

The denoised projections were passed to the moments-based solver for an initial estimate of the orientations. However the problem

of tomography under unknown angles is inherently ambiguous and the solutions can be obtained only up to a global rotation. Hence the estimated orientations will in most cases be shifted from the original values by a *single global* offset δ. To quantify the accuracy of the estimates, the orientation estimates obtained using the HLCC-based method from Section 2.5 were corrected for this global rotational ambiguity and then displayed in Fig. 4. In Fig. 4, we show a scatter plot of the K_c projection angles corresponding to 'ground-truth cluster centers' (i.e. cluster centers obtained from noiseless projection vectors without outliers) and the corresponding values of the estimates of those angles after correcting for the offset δ. The ground truth angles turn out to have values from 0 to $180°$, because of the uniform distribution of the orientations of the projection vectors.
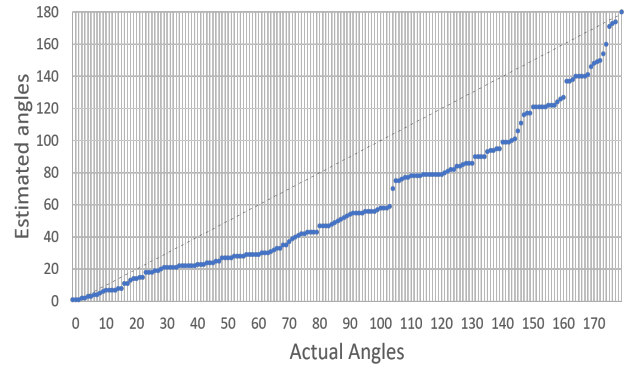

Figure 4: Initial estimate of orientations: $\sigma = 10\%$ **Noise,** $f_1 = 10\%$ **outliers of class 1,** $f_2 = 10\%$ **outliers of class 2**

A perfect set of estimates would have produced a plot aligned with the $45°$ line. However, as seen in Fig. 4, the initial estimate provided by the moments based approach deviated significantly from the actual orientations (even after correction for the global offset). Reconstruction of the object considering these orientations as the true values yielded poor results as seen in Fig. 5.
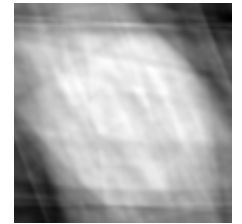
(a) Original Image   **(b) Reconstructed Image**

Figure 5: An example of the failure of the moments based approach for image reconstruction. Parameters: 10% Noise, 10% outliers of class 1, 10% outliers of class 2

This showed that there was a clear necessity to refine the results given by the moments based approach. We therefore explored two

ways to potentially improve the estimate obtained from the moment-based initialization method - the sparsity-based optimization framework, and the gradient descent based approach - both described in Section 2.6. The former was implemented using the $\ell_1 - \ell_s$ package [14]. As explained in Section 2.6, the sparsity-based reconstruction quite surprisingly failed to yield good results, an example of which is seen in Fig. 6. This was mainly due to errors in orientation estimates.
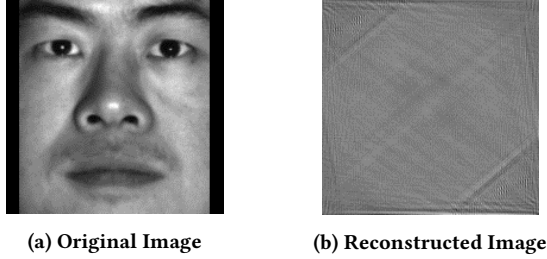

(a) Original Image　　　**(b) Reconstructed Image**

Figure 6: An example of the failure of sparsity-based optimization, with angle errors

3.1 Reconstruction using the gradient descent based approach

After the failure of the sparsity-based optimization framework, we altered our approach and adopted a gradient-based optimization framework using the FBP, and the results turned out to be very promising. Since the problem is ambiguous up to a global rotation, the reconstructed image might be a rotated version of the original image. Therefore the reconstruction was registered to the original test image, to obtain a quantitative error metric. The error metric used was the Relative Mean Squared Error between the registered reconstruction and the test image. This is defined as follows:
$\text{RMSE}(z, \hat{z}) = \dfrac{\|z - \hat{z}\|_2}{\|z\|_2}$ where $\hat{z}$ is the reconstructed estimate for z. Reconstruction without any noise and outliers produced nearly accurate results, an example of which can be seen in Fig. 7.

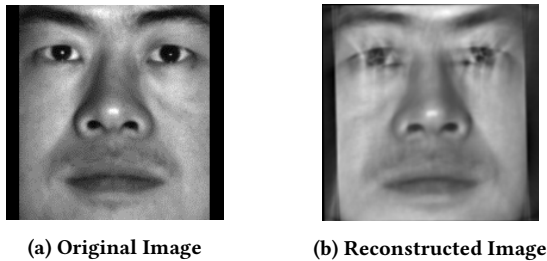**(a) Original Image**　　　**(b) Reconstructed Image**

Figure 7: Results of Algorithm 1 with 0% Noise and no outliers of class 1 or class 2, RMSE - 4.58%

However even in cases of noisy projections with 10% noise, and a significant percentage of outliers of class 1 and class 2 (10% each), we observed that our algorithm was able to successfully refine the initial estimate given by the moments based approach and obtain an

accurate reconstruction. On applying our algorithm to the moment-based orientation estimates that yielded the results shown in Fig. 4 and 5, we obtained significantly refined angle estimates, as shown in Fig. 9. A sample reconstruction is shown in Fig. 8.

(a) Original Image　　　**(b) Reconstructed Image**

Figure 8: 10% Noise, 10% outliers of class 1, 10% outliers of class 2, RMSE - 8.95%
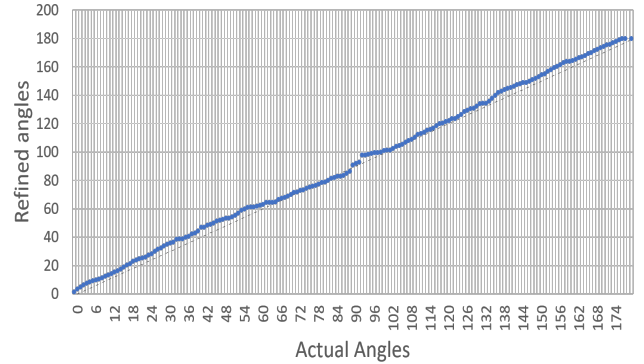

Figure 9: Refined estimate of orientations: 10% Noise, 10% outliers of class 1, 10% outliers of class 2

To check the robustness of the algorithm, we also experimented with reconstructions under higher noise levels and a higher percentage of outliers. The reconstructions in Fig. 10 and 11 were obtained for extremely high noise levels (50% of the average value of the projections).
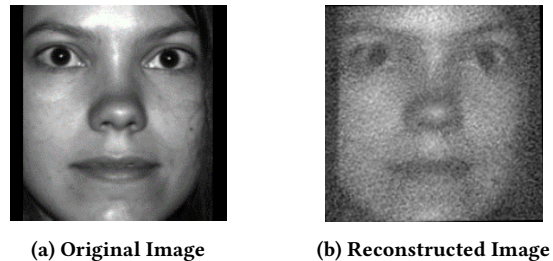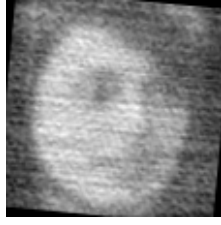

(a) Original Image　　　**(b) Reconstructed Image**

Figure 10: 50% Noise, 0% outliers of class 1, 0% outliers of class 2, RMSE - 11.99%

(a) Original Image   (b) Reconstructed Image

Figure 11: 50% Noise, 5% outliers of class 1, 5% outliers of class 2, RMSE - 18.39%

(a) Original Image   (b) Reconstructed Image

Figure 13: 20% Noise, 10% outliers of class 1, 0% outliers of class 2, Non-Uniform distribution of angles, RMSE - 17.69%

Experiments with non-uniform orientation distributions: Unlike previous algorithms, our algorithm doesn't assume that the orientations of the projections are uniformly distributed, nor does it require knowledge of the underlying distribution. To test this, we conducted the following experiment. Instead of angles being taken from the Uniform$(0, \pi)$ distribution, we considered the following distribution for the projection orientations: Uniform$(0, \pi/9) \cup$ Uniform$(2\pi/9, \pi/3) \cup$Uniform$(4\pi/9, 2\pi/3)\cup$Uniform$(7\pi/9, 8\pi/9)$. Further to re-emphasize the robustness of our algorithm, the projections were subjected to high amounts of noise (20% of the average value of the projections) and 10% outliers of class 1. The distribution of the original angles along with our estimates of the angles after the moment based solver and refinement are shown in Fig. 12. The final reconstruction is shown in Fig. 13. Note that in generating this result, the algorithm did not exploit any knowledge whatsoever of the distribution of the orientation.

Experiments with unknown shifts: In cases where projections have random unknown shifts, our algorithm was able to accurately estimate the shift in each projection and correct it to produce an accurate reconstruction. The following has been tested on a 86×86 image, with 50 projections, 10% Noise and random unknown shifts up to ± 2. Note that the limit of ± 2 on the shifts is not overly restrictive, as indeed the range of possible shifts is very small in these applications. The plot in Fig. 14 shows the absolute error between the actual shifts in the projection and the shifts we estimated (after correction for an inevitable global shift ambiguity). A sample reconstruction is shown in Fig. 15

(a) Shifts estimated by the moment based approach

(b) Refined estimates of the shifts

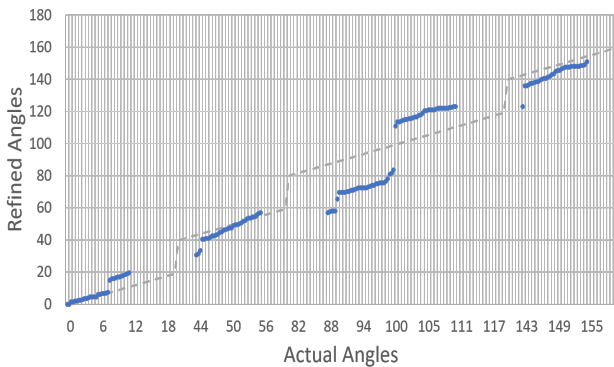**Figure 14: Estimation of shifts**

Figure 12: Refined estimate of orientations in case of non-uniform distribution of angles: 20% Noise, 10% outliers of class 1, 0% outliers of class 2
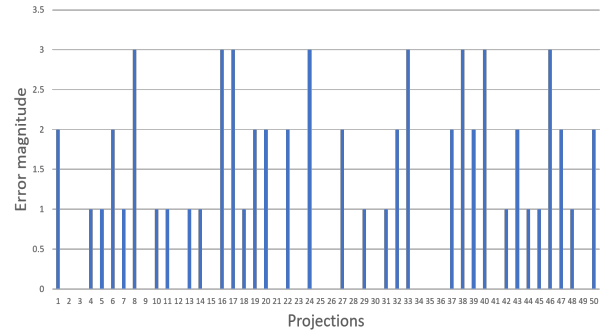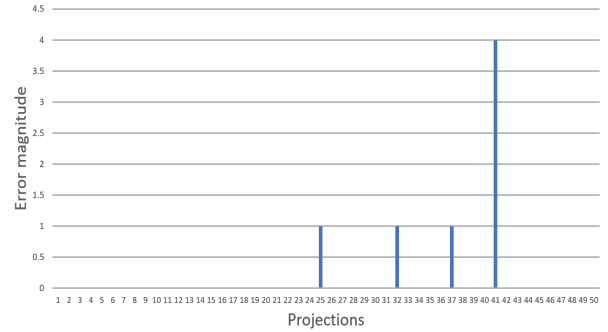
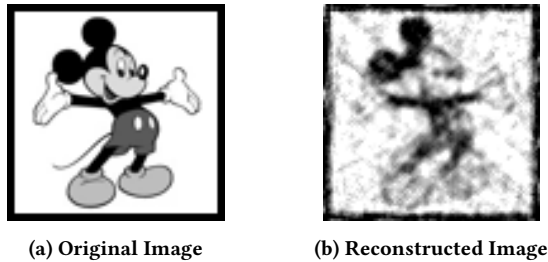**(a) Original Image** **(b) Reconstructed Image**

Figure 15: 10% Noise and unknown shifts of maximum amplitude ±2, RMSE-5.65%

4 DISCUSSION AND CONCLUSION

From the results presented here, we conclude that the algorithm described in this paper is capable of estimating the original underlying structure to a high degree of accuracy, without any prior knowledge about the angles of projections and any prior structural information such as templates. Further, we have also seen that our method is robust to the presence of outliers, projections with extremely high amounts of noise and projections with random unknown shifts. Our method does not make any assumption on the distribution of the projection angles. In fact, in some situations in cryo-EM, certain orientations are more favorable than others and this is where the assumptions made by other algorithms would fail. The estimation of orientations is based on sound mathematical concepts which rely on HLCC at the foundation. The next step, the alternating gradient descent also does not make any assumption and descends onto the correct structure despite it being a non-convex optimization problem. This is due to an initial estimate provided to us by the moment based approach.

Future work: There are two major directions for future work.

- There is an interesting insight that we have observed through these experiments: Despite the promising results of compressed sensing in general image reconstruction from compressive measurements [3] including in tomography [4], it does not achieve an accurate reconstruction in this case where there is uncertainty in the sensing matrix, i.e. in the projection angles. The sparsity-based optimization framework is also very sensitive to the initial estimate in such scenarios, and often does not converge at all. On the other hand, the FBP algorithm, which is used in the gradient descent approach is quite robust to the initial estimate and gives us very accurate reconstructions, despite the inevitable errors (even if they are small in value) in the estimates of the orientations and the shifts. This is even more surprising considering the fact that the FBP does not exploit the powerful sparsity prior that is a common feature of compressed sensing algorithms. A detailed theoretical or analytical study of this surprising observation is a major avenue for future research.

- The other important direction is the validation of our algorithm on actual Cryo-EM datasets, insect tomography datasets, and CT reconstructions with patient motion.

REFERENCES

[1] S. Basu and Y. Bresler. 2000. Feasibility of tomography with unknown view angles. *IEEE Transactions on Image Processing* 9, 6 (6 2000), 1107–1122. DOI: http://dx.doi.org/10.1109/83.846252

[2] S. Basu and Y. Bresler. 2000. Uniqueness of tomography with unknown view angles. *IEEE Transactions on Image Processing* 9, 6 (6 2000), 1094–1106. DOI: http://dx.doi.org/10.1109/83.846251

[3] E. J. Candes and M. B. Wakin. 2008. An Introduction To Compressive Sampling. *IEEE Signal Processing Magazine* 25, 2 (2008), 21–30.

[4] K. Choi, J. Wang, L. Zhu, T-S Suh, S. Boyd S, and L. Xing. 2010. Compressed sensing based cone-beam computed tomography reconstruction with a first-order method. *Medical Physics* 37, 9 (2010), 5113–5125.

[5] R. Coifman, Y. Shkolnisky, F. Sigworth, and A. Singer. 2008. Graph laplacian tomography from unknown random projections. *IEEE Transactions on Image Processing* 17, 10 (2008), 1891âĂŞ1899.

[6] E. Malhotra et al. 2016. Tomographic reconstruction from projections with unknown view angles exploiting moment-based relationships. In *2016 IEEE International Conference on Image Processing (ICIP)*. IEEE, 1759–1763. DOI: http://dx.doi.org/10.1109/ICIP.2016.7532660

[7] Yi Fang, Mengtian Sun, S.V.N. Vishwanathan, and Karthik Ramani. 2011. sLLE: Spherical locally linear embedding with applications to tomography. In *CVPR 2011*. IEEE, 1129–1136. DOI: http://dx.doi.org/10.1109/CVPR.2011.5995563

[8] L. A. Feldkamp, L. C. Davis, and J. W. Kress. 1984. Practical cone-beam algorithm. *Journal of the Optical Society of America A* 1, 6 (6 1984), 612. DOI: http://dx.doi.org/10.1364/JOSAA.1.000612

[9] J. Fessler. 2000. Statistical image reconstruction methods for transmission tomography. In *Handbook of Medical Imaging: Medical Image Processing and Analysis*. Vol. 2. SPIE, 1–70.

[10] Joachim Frank. 2006. *Three-dimensional electron microscopy of macromolecular assemblies: visualization of biological molecules in their native state.* Oxford University Press.

[11] Jorge F. Genise and Gerardo Cladera. 1995. Application of computerized tomography to study insect traces. *Ichnos* 4, 1 (4 1995), 77–81. DOI: http://dx.doi.org/10.1080/10420949509380116

[12] C. Huang and H. D. Tagare. 2016. Robust w-Estimators for Cryo-EM Class Means. *IEEE Transactions on Image Processing* 25, 2 (Feb 2016), 893–906.

[13] P. Huber. 2009. *Robust Statistics.* John Wiley Sons.

[14] Seung-Jean Kim, K. Koh, M. Lustig, Stephen Boyd, and Dimitry Gorinevsky. 2007. An Interior-Point Method for Large-Scale -Regularized Least Squares. *IEEE Journal of Selected Topics in Signal Processing* 1, 4 (12 2007), 606–617. DOI: http://dx.doi.org/10.1109/JSTSP.2007.910971

[15] Alex Krizhevsky. 2009. *Learning Multiple Layers of Features from Tiny Images.* Technical Report.

[16] S. Lloyd. 1982. Least squares quantization in PCM. *IEEE Transactions on Information Theory* 28, 2 (3 1982), 129–137. DOI: http://dx.doi.org/10.1109/TIT.1982.1056489

[17] D.D. Muresan and T.W. Parks. Adaptive principal components and image denoising. In *Proceedings 2003 International Conference on Image Processing (Cat. No.03CH37429)*, Vol. 1. IEEE, 101–4. DOI: http://dx.doi.org/10.1109/ICIP.2003.1246908

[18] F. Natterer. 2001. *The Mathematics of Computerized Tomography.* Society for Industrial and Applied Mathematics. DOI: http://dx.doi.org/10.1137/1.9780898719284

[19] Fred J. Sigworth. 2016. Principles of cryo-EM single-particle image processing. *Microscopy* 65, 1 (2 2016), 57–67. DOI: http://dx.doi.org/10.1093/jmicro/dfv370

[20] A. Singer and H.-T. Wu. 2013. Two-Dimensional Tomography from Noisy Projections Taken at Unknown Random Directions. *SIAM Journal on Imaging Sciences* 6, 1 (1 2013), 136–175. DOI: http://dx.doi.org/10.1137/090764657

[21] Hengyong Yu and Ge Wang. 2007. Data Consistency Based Rigid Motion Artifact Reduction in Fan-Beam CT. *IEEE TRANSACTIONS ON MEDICAL IMAGING* 26, 2 (2007). DOI: http://dx.doi.org/10.1109/TMI.2006.889717