ON THE BAYESIAN RECONSTRUCTION METHOD FOR RANDOMLY ORIENTED PARTICLES IN CRYO-EM

Sami S. Brandt, Katrine Hommelhoff Jensen and François Lauze

Department of Computer Science, University of Copenhagen, Denmark.

ABSTRACT

In this work, we address the problem of reconstructing the 3D structure of an object from a set of transmission electron microscopy (TEM) images, taken at unknown, random directions around the object. We use the expectation maximisation (EM) algorithm for finding the maximum a posteriori (MAP) estimates for the 3D structure from the marginal posterior, where the view orientations are integrated out. In comparison to previous work related to this single particle reconstruction application, we have made the following novel contributions. (1) We use Monte Carlo integration to approximate the expected complete data log posterior to reduce the computational complexity; (2) we use a uniform prior in the space of rotations instead of the space of rotation angles; (3) we use the positivity constraint for the reconstructed density that is both a physical constraint as well as it acts as a natural sparsity prior; (4) on the M-step we use a large scale, subspace trust-region method based on the interior-reflective Newton method for efficient computation of the reconstruction. We experimented the approach on cryo-electron microscopy (cryo-EM) protein images. The results are promising and show that the 3D structure can be robustly recovered with the proposed method in spite of the very low signal-to-noise ratio (SNR).

1. INTRODUCTION

Single particle cryo-EM is a form of TEM aimed at determining the 3D structure of proteins from images of randomly oriented copies of the macromolecular complex, each of which represents a certain view and eventually a certain conformation. Whereas the strength of single particle cryo-EM is its unique ability to capture the macromolecules in their natural state, it suffers from very low SNR in the projection images. The single particle reconstruction problem principally constitutes two main problems: (1) to recover the underlying imaging geometry [1], i.e., the relative position and orientation of each image, and (2) to solve the ill-posed reconstruction problem, while dealing with low SNR, unevenly distributed projection directions as well as particle shape variation.

Several works have been proposed developing towards a statistically sound reconstruction via maximum likelihood (ML) or maximum a posteriori (MAP) approaches, see [2] for an overview. To deal with the complexity of the likelihood or posterior distributions, the expectation maximisation (EM) algorithm [3] has been applied in various ways [4, 5]. EM methods were first introduced in cryo-EM by Sigworth [6] for clustering 2D projection images, from which class averages are computed and used for calculating an initial reconstruction. A 3D method [6] uses an EM-like approach to recover projections and conformational classes of 3D particles. However, their final numerical alternation procedure does not maximise the marginal likelihood, as an EM-ML would, but it maximises an approximate complete data likelihood instead. The EM-MAP approach [7] in the Fourier domain uses a smoothness prior for the particle. In addition, MAP estimation methods, not based on EM, have been proposed such as one based on sparsity [8]; and another computing the marginal posterior with a Monte Carlo estimate and a Gaussian smoothness prior for the reconstruction [9].

We propose an EM algorithm for finding the MAP estimates for the 3D orientation for the projection images and the 3D object density. The algorithm incrementally improves the 3D model based on the projection images and current model estimates by marginalising over the projection directions, referred to as the missing data, w.r.t. the complete data distribution, conditioned to the observed data and the previous 3D model and the noise variance estimate.

2. PROJECTION MODEL

Let us first define how the volume density $f$ projects onto the observed image $m_i$. Formally, the projection is represented as

$$m_i(x, y) = (B_{\eta_i} \circ A_{\theta_i}) f(x, y, z),$$

(1)

where $f \in \mathcal{C}_c^0(\mathbb{R}^3)$, $f(x, y, z) \geq 0 \forall (x, y, z) \in \mathbb{R}^3$ and $m \in \mathcal{C}_c^0(\mathbb{R}^2)$, where $\mathcal{C}_c^0(\mathbb{R}^k)$ denotes the class of compact, continuous functions in $\mathbb{R}^k$, $k = 2, 3$, and $A_{\theta_i}$ is a linear projection operator $f \mapsto m_i^\ast$, $m_i^\ast \in \mathcal{C}_c^0(\mathbb{R}^2)$, and $B_{\eta_i}$ is a linear operator $m_i \mapsto m_i$ on the image plane, where the orientation $\theta_i \in \Theta$ and transformation $\eta_i \in H$ are parameter vectors. The projection $A_{\theta_i}$ onto the image plane is defined as

$$m_i^\ast(x, y) = \int (R_i f)(x, y, z) dz$$

(2)
numerical implementation, we also need the adjoint operators \( B_i \) the finite dimensional operator between the measured images where

\[
m_i(x, y) = B_i^{-1} m_i^*(x, y),
\]

(3)

Discretising the volume to finite number \((N^3)\) of voxels, and continuous images to finite number \((N^2)\) of pixels, we may write the view projection formula as a matrix equation

\[
m_i^* = A_{\theta_i} f = PR_i^2 f,
\]

(4)

where \( f \) is a vector of voxel values, \( m_i \) a vector of pixel values of the image \( i \); \( R_i^2 \) is the rotation-interpolation matrix, which includes the rotation of the volume grid and linear interpolation of the density values; \( P \) represents the density projection by integration over the z-dimension (see Fig. 1(b)). Similarly, we write the mapping on the image plane as

\[
m_i^* = B_{\eta_i} m_i = WT_i R_i^2 m_i,
\]

(5)

where \( R_i^2 \) is another rotation-interpolation matrix, \( T_i \) is a translation matrix, and \( W \) a 2D windowing matrix. The total discretised model, with all images \( i = 1, 2, \ldots, I \) stacked into \( m \), can be represented as

\[
A_{\theta} f = B_{\eta} m,
\]

(6)

The finite dimensional operator between the images and the volume density is represented by the \((N^2 I) \times N^3\) matrix \( A_{\theta} \), such that it decomposes as \( A_{\theta} = (A_{\theta_1}, A_{\theta_2}, \ldots, A_{\theta_I}) \), where \( A_{\theta_i} \) is the projection for image \( i \). Similarly, \( B_{\eta} \) is the finite dimensional operator between the measured images and the rotation translation aligned images, represented by the \((N^2 I) \times (M^2 I)\) matrix which decomposes as the block diagonal matrix \( B_{\eta} = \text{blockdiag}(B_{\eta_1}, B_{\eta_2}, \ldots, B_{\eta_I}) \). In the numerical implementation, we also need the adjoint operators \( A_{\theta}^T \) and \( B_{\eta}^T \), which include the back-projections and adjoint of the linear interpolation, see [10].

3. STATISTICAL MODEL

Since (6) is ill-posed, we find the regularised solution, i.e., the MAP solution, which maximises the posterior

\[
p(f, \theta, \eta | m) \propto p(m|f, \theta, \eta)p(f, \theta, \eta),
\]

(7)

with subject to \( f \geq 0 \), where \( p(m|f, \theta, \eta) \) is the likelihood and \( p(f, \theta, \eta) \) the prior. Assuming i.i.d. Gaussian noise in the measurements with a known variance \( \sigma^2 \), an independent Gaussian prior for the voxel values, a uniform prior in the space of rotations, and non-informative priors on \( \theta \) and \( \eta \), the map solution is equivalent to maximising the log posterior

\[
l(f, \theta, \eta) = -\frac{1}{2\sigma^2} \| A_{\theta} f - B_{\eta} m \|^2 + C,
\]

(8)

where \( A_{\theta}'(\theta, \sqrt{X}) \), \( B_{\eta}' = (B_{\eta}, 0) \), \( \lambda \) is the regularisation parameter depending on the variance of the noise and \( f \), and the constant \( C \) is independent of \( f, \theta \) and \( \eta \).

4. EXPECTATION MAXIMISATION ALGORITHM

Finding the MAP solution to (7) is difficult due to the very high dimensionality of the problem. Considering the projection directions \( \theta \) and the 2D transformation parameters \( \eta \) as the missing data, we may apply EM to find the MAP estimate. E-step We construct the expected log posterior of the parameters conditioned to the complete data \( z = \{f, \theta, \eta\} \), or,

\[
\chi(f, \hat{f}^{(m-1)}) = E_{z|m, \hat{f}^{(m-1)}} \{ \log l(f, \theta, \eta) \}
\]

(9)

where the expectation is taken over the distribution of the complete data given the measurements and the previous estimate \( \hat{f}^{(m-1)} \) for the reconstruction.

Let \( A_{\theta_j}' = (A_{\theta_j}, \sqrt{X_j}I_j) \) be the modelled projection corresponding to a discretised direction \( \theta_j \in \Theta_j \), where \( \Theta_j \cap \Theta_{j'} = \emptyset \) for all \( j \neq j' \). Let \( \Theta = \cup_{j=1}^{J} \Theta_j \). Similarly, let \( B_{\eta_k}' = (B_{\eta_k}, 0) \), and consider discretised 2D rotations and translations \( \eta_k \in H_{kl} \), where \( H_{kl} \in H_{kl}' = \emptyset \) for all \( k \neq k' \) and \( \cup_{k=1}^{K} \cup_{l=1}^{L} H_{kl} = H \). If the parameters \( \theta^i, \eta^i \) for each image \( i \) were available, the Kronecker delta \( \delta(\theta^j - \theta_j, \eta^j - \eta_k) \) would give unity value when \( \theta^i = \theta_j, \eta^i = \eta_k \) and zero otherwise. The expected log posterior conditioned to the complete data is thus

\[
\chi(f, \hat{f}^{(m-1)}) \equiv - E_{z|m, \hat{f}^{(m-1)}} \left\{ \sum_i \delta(\theta^i - \theta_j, \eta^i - \eta_k) || A_{\theta_j}' f - B_{\eta_k}' m_i ||^2 \right\}
\]

\[- \sum_{i, j, k, l} P(\theta^i \in \Theta_j, \eta^i \in H_{kl}|m, \hat{f}^{(m-1)}) || A_{\theta_j}' f - B_{\eta_k}' m_i ||^2 \]

where constants depending only on the noise variance have been dropped, and the posterior probability that \( \theta^i = \theta_j, \eta^i = \eta_k \),

\fig{1}{0.8}{(a) Illustration of the single particle projection geometry; (b) illustration of the forward projection model containing rotation, interpolation and projection of the volume.}
given the measurements and the previous model \( \hat{f}^{(m-1)} \), is

\[
P(\theta' \in \Theta_j, \eta' \in H_k; |\hat{f}^{(m-1)}|) = \int_{\Theta} \int_{H} \exp(-\frac{1}{2\sigma^2} ||A_{\theta_j} f - B_{\eta_k} m_i||^2) \det J r(\theta') |d\theta' d\eta'|
\]

\[
\approx \sum_{i,j,k,l} \exp(-\frac{1}{2\sigma^2} ||A_{\theta_j} f - B_{\eta_k} m_i||^2) \det J r(\theta_j)
\]

where \( \det J r(\theta') \) arises the uniform prior for point \( r \) on the manifold of rotations and the substitution rule for integrals.

The weights \( w_{ijkl} \) form a discrete probability distribution each fixed \( i \). Since \( \chi \) involves summing over all discretised orientations \( j \), 2D rotations \( k \) and translations \( l \), which is costly on the M-step, we draw \( L \) random samples from the distribution defined by the weights for each view (cf. [11]), and define the approximation for \( \chi \), as

\[
\tilde{\chi}(f, \hat{f}^{(m-1)}) = -\frac{1}{2\sigma^2} \sum_{i,j,k,l} \tilde{w}_{ijkl} ||A_{\theta_j} f - B_{\eta_k} m_i||^2
\]

(10)

where \( \tilde{w}_{ijkl} \) is the empirical probability distribution based on the \( L \) orientation samples for the view \( i \).

**M-step** We maximise the (approximated) expected posterior distribution of the parameters \( f \) conditioned to the complete data, i.e., solve for the estimate

\[
\hat{f}^{(m)} = \arg \max_{f: f \geq 0} \chi(f, \hat{f}^{(m-1)}).
\]

(11)

This is equivalent to

\[
\arg \min_{f: f \geq 0} \sum_{i,j,k,l} \tilde{w}_{ijkl} ||A_{\theta_j} f - B_{\eta_k} m_i||^2
\]

\[
= \arg \min_{f: f \geq 0} \sum_{j} W_j \sum_{i,k,l} \tilde{w}_{ijkl} ||A_{\theta_j} f - B_{\eta_k} m_i||^2
\]

(12)

where \( W_j = \sum_{i',k',l'} \tilde{w}_{ij'kl'} \) and \( W_{ijkl} = \sum_{i',k',l'} \tilde{w}_{ijkl} \).

Let \( \bar{m} = \sum_{i,j,k,l} W_{ijkl} B_{\eta_k} m_i \) be the weighted mean image into direction \( j \) computed over the image population, rotations and translations. Solving (12) is then equal to solving

\[
\arg \min_{f: f \geq 0} ||A' f - \bar{m}'||^2
\]

(13)

where

\[
A' = \begin{pmatrix} W_1 A_{\theta_1} \\ \vdots \\ W_J A_{\theta_J} \end{pmatrix}, \quad \bar{m}' = \begin{pmatrix} \bar{m}^1 \\ \vdots \\ \bar{m}^J \end{pmatrix}.
\]

We thus solve the next reconstruction estimate from (13). The whole algorithm is summarised in Algorithm 1.

**Algorithm 1** Expectation Maximisation reconstruction.

1. E-step: Compute \( K \) orientation samples for each view and construct the approximate weights \( \tilde{w}_{ijkl} \) from the samples to form \( \tilde{\chi}(f, \hat{f}^{(m-1)}) \) according to (10).
2. M-step: Maximise \( \tilde{\chi}(f, \hat{f}^{(m-1)}) \) by solving (13).
3. Iterate from Step 1 until convergence.

**Fig. 2.** (a) 3D ribosome model from the EMAN database; (b) simulated projections of the 3D phantom model; (c) clustered class averages of real projections.

**5. EXPERIMENTS**

To evaluate our reconstruction method, we used a priorly computed 3D protein model \((58 \times 58 \times 58)\) as the ground truth (Fig. 2a). We first created a set of 100 simulated \(58 \times 58\) projection images of the phantom with uniformly distributed rotations (Fig. 2b). Then we added Gaussian noise onto the projections, with the inverse SNR from \(20\%\) to \(160\%\), to experiment the sensitivity of the algorithm, see Fig. 3a–e. We set the regularisation parameter to \( \lambda = C \sigma_{\text{noise}}^2 / \sigma_f^2 \), with the correction coefficient \( C = 1.1 \) to avoid overfitting. As the initial guess we used the phantom perturbed with Gaussian noise with the inverse SNR of \(80\%\) (Fig. 4a). The results are in Fig. 3f–j which show that our method is able to robustly recover the 3D structure. To additionally experiment the sensitivity to the initial guess, we perturbed the phantom by adding i.i.d. Gaussian noise onto it with the inverse SNR from \(80\%\) to \(1280\%\), see Fig. 4a–e, and reconstructed the projections with the SNR of \(20\%\) (Fig. 3b). The results (Fig. 4f–j) indicate particular robustness to bad initialisation.

The 3D phantom had originally been reconstructed from the set of 13 class averages (Fig. 2c) by using the standard projection matching. We had the class average images available, thus we could compare our method to the phantom by re-reconstructing the class averages. To find the proper regularisation level for this data, we estimated the noise level by computing the standard deviation of the background pixel values and \( \sigma_f \) was set to the standard deviation of the phantom. Due to the small number of images, we used early stopping regularisation in addition to the Gaussian prior. The results after the first iteration (Fig. 5) show more detail than the original phantom, while the regularisation level is appropriate.

1EMAN: http://blake.bcm.edu/emanwiki/EMAN2
6. CONCLUSION

We have studied the single particle reconstruction problem in cryo-EM. We use the EM algorithm for finding the MAP structure estimates of the marginal posterior, where the projection directions have been integrated out. This approach allows for efficient computation and robustness against very high levels of measurement noise and bad initialisation. The method can also be easily extended for the generalised Tikhonov regularisation with the positivity prior that facilitates statistical reconstruction with a large class of priors.

7. REFERENCES


