A MULTI-RESOLUTION APPROACH TO COMPLEXITY REDUCTION IN TOMOGRAPHIC RECONSTRUCTION

Boxiao Ma, Nour Zalmai, and Hans-Andrea Loeliger

ETH Zurich, Dept. of Information Technology & Electrical Engineering
{ma, zalmai, loeliger}@isi.ee.ethz.ch

ABSTRACT

Most of the algorithms for tomographic reconstruction face the same problem: high computational complexity. In order to tackle this problem, this paper proposes a general multi-resolution approach that enables a flexible choice of reconstruction focus and thus saves computational power in reconstructions. The approach is demonstrated in this paper based on a reconstruction algorithm using a (improper) Markov random field prior with sparsifying NUV terms (normal with unknown variance), where the unknown variances are learned by approximate EM (expectation maximization). The experimental and practical results show that both for simulated and real-world objects the proposed framework yields satisfying results with much lower computational cost.

Index Terms— Tomographic reconstruction, multi-resolution, NUV, computational complexity

1. INTRODUCTION

Tomographic reconstruction is a multidimensional inverse problem where the goal is to obtain an estimate of a 2D or 3D object from noisy projections. It has been a challenge for decades due to the huge dimension of the corresponding mathematical objects and different physical artifacts. For example, the reconstruction of a 3D cube discretized with 1000 elements on each dimension involves 10^9 variables and can easily require a memory of several hundred Gigabytes. Despite its many deficiencies, the filtered back projection (FBP) is nowadays still the most widely used method [1]. The main reason for that is the high computational cost of more sophisticated reconstruction algorithms.

Therefore, the complexity reduction of those reconstruction algorithms is crucial for their practical use. Several attempts to accelerate the computations by parallelization using GPUs [2–4] have been proposed. Although the reduction in computation time is enormous, it does not lower the massive memory requirement, which is still a bottleneck for standard computers. Besides, those parallelizations are usually designed specifically for certain algorithms and are thus not portable. Inspired by the fact that a big part of the reconstructed object in tomography is just emptiness or not of actual interest, we propose a general multi-resolution approach to reduce the computational complexity by using a multi-resolution grid and a multi-stage reconstruction. The multi-resolution approach enables a flexible distribution of computational power and reduces both computational time and required memory. Moreover, the results have arguably good quality in the area of interest. A number of tomographic reconstructions, which previously could not achieve a satisfactory resolution on ordinary computers, can now be done with the multi-resolution approach.

A great variety of tomographic reconstruction algorithms exists [5–11] that outperform the FBP approach. Some of them try to exploit the sparsity in the gradient and use regularization involving the total variation [8, 9]. Several authors have also proposed algorithms using Markov random field (MRF) [10,11], and [11] also demonstrates a message passing method. In this paper we use an algorithm using a MRF prior with sparsifying NUV terms. The reconstruction algorithm is inherited from [12, 13] but extended with a new ability to handle boundary conditions. The unknown parameters (variances) of the MRF prior are learned by an approximate EM algorithm. The pertinent computations boil down to iterative scalar Gaussian message passing, which scales linearly with the number of voxels.

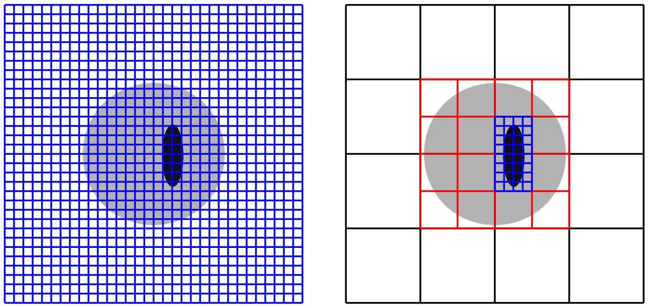
The paper is structured as follows: the general multi-resolution approach is proposed and illustrated in Section 2. In Section 3, we introduce the specific reconstruction algorithm based on [12]. The experimental and practical results are presented in Section 4.

2. MULTI-RESOLUTION APPROACH

The multi-resolution approach enables a computationally efficient method of reconstruction in tomography. The basic idea is to reconstruct the area of interest with high resolution and the remaining areas with lower resolution.

2.1. Notation

We assume that the voxels (or pixels) of the object to be reconstructed are arranged in a 3 (or 2)-dimensional rectangular grid. Each voxel is associated with a unique spatial index $s_\ell = (i, j, k)$, $\ell \in \{1, \dots, L\}$ and a random variable X_{s_ℓ}



Uniform-resolution grid Multi-resolution grid

Fig. 1. Different grids for a simple object. Left is the uniform-resolution grid, right is the multi-resolution grid.

denoting the voxel value, where L is the total number of voxels. The measurements $\mathbf{y} \in \mathbb{R}^N$ are obtained by projecting the object to the detector. The object $\mathbf{x} = (x_1, \ldots, x_L)^T$ and the measurements $\mathbf{y}$ are linked by the projection matrix $A \in \mathbb{R}^{N \times L}$ with $\mathbf{y} = A\mathbf{x}$. The projection matrix is generated by methods such as distance-driven projection [14]. The goal is to reconstruct the object $\mathbf{x}$ from the measurements $\mathbf{y}$.

We denote by $C = \{c_p : c_p \in \mathbb{R},\ p \in \{1, \ldots, P\}\}$ the set of boundary conditions for boundary voxels. Let Θ with size P be the set of all boundary voxel-condition pairs (ℓ, p). Note that one boundary pixel can have multiple corresponding boundary conditions. Let Δ be the set of nearest-neighbor pairs: for any two voxels at locations $s_\ell = (i, j, k)$ and $s_{\ell'} = (i', j', k')$, $(\ell, \ell') \in \Delta$ if and only if both $|i - i'| + |j - j'| + |k - k'| = 1$ and $\ell < \ell'$.

2.2. An Example

Consider the toy example in Fig. 1, where the gray circle with a black ellipse inside are the object to be reconstructed and the ellipse is the actual area of interest.

With the multi-resolution approach the reconstruction is performed as follows: the object is initially reconstructed with a very coarse resolution, as the black grid shown in the right part of Fig. 1. Obviously, in the 4 central pixels we find something which is different from the background and thus consider them as the current area of interest, then we use a higher resolution (the red grid) to reconstruct this area and find more details in the 2 central right pixels. Finally, we update the area of interest to these two pixels and reconstruct them with the highest resolution (the blue grid). The final estimate of the object $\hat{\mathbf{x}}$ can be obtained by merging the reconstructions from all three steps.

It can be seen from Fig. 1 that the pixel number in the multi-resolution grid ($16 + 16 + 32 = 64$) is much less than in the uniform-resolution grid ($32 * 32 = 1024$), where the whole space was filled with the resolution of the blue grid. We are going to see that in this way the computational complexity can be substantially reduced while preserving the reconstruction quality in the area of interest.

Algorithm 1 Successive Cancellation Algorithm

Input: $\mathbf{y}_0$
Output: Object estimate $\hat{\mathbf{x}}$
Generate A_0 and C_0
Reconstruct $\hat{\mathbf{x}}_\mathbf{0}$ from $\mathbf{y}_0$, C_0 and A_0
$i = 1$
while unsatisfied with the reconstruction **do**
    Choose a new area of interest S_i and a new resolution
    $\mathbf{y}_i = \mathbf{y}_{i-1} - A_{i-1}\hat{\mathbf{x}}_{i-1}^{S_i^c}$
    Generate C_i according to $\hat{\mathbf{x}}_{i-1}$
    Generate A_i for S_i
    Reconstruct $\hat{\mathbf{x}}_i$ with $\mathbf{y}_i$, C_i, and A_i
    $i = i + 1$
end while
Merge the results $\hat{\mathbf{x}}_0 \ldots \hat{\mathbf{x}}_i$ into $\hat{\mathbf{x}}$

2.3. The Successive Cancellation Algorithm

The general procedure of the multi-resolution approach is stated in Algorithm 1. The inputs are the measurements, the output is the final object estimate.

The area of interest S_i is a set of voxels that form a rectangular part of the object with a certain resolution. In Algorithm 1, the resolution always becomes increasingly higher as the algorithm proceeds. The new measurements $\mathbf{y}_i$ for each step are obtained by successive cancellation where $\hat{\mathbf{x}}_i^{S_i^c}$ is a vector same as $\hat{\mathbf{x}}_i$ except that the elements included in S_i are set to 0. The set of boundary condition C_i for each step is obtained by directly taking the previous reconstructed voxel values, which lie on the border of S_i. Concerning the new projection matrix A_i, although most of the projection methods usually assume that the rotation axis is at the center of the object, it is straightforward to obtain the projection matrix for an arbitrary part of an object using the linearity of the projection.

The approach is applicable to any reconstruction method that can handle boundary conditions. In the next section, we develop this approach for a specific reconstruction algorithm.

3. SPECIFIC RECONSTRUCTION APPROACH

We now work out the proposed multi-resolution approach by extending the reconstruction algorithm from [12] so that it can handle boundary conditions. We thus only introduce some basic definitions and the differences to [12].

3.1. Prior Model

The interesting part of the statistical model is the (improper) prior. Here we penalize not only the differences between neighboring voxels but also between voxels and boundary conditions. That is to say, except the $U_{\ell,\ell'}$ we define the slack variables $U_{\ell,p}$ for each boundary voxel-condition pairs $(\ell, p) \in \Theta$ as

$$U_{\ell,p} = X_{s_\ell} \tag{1}$$

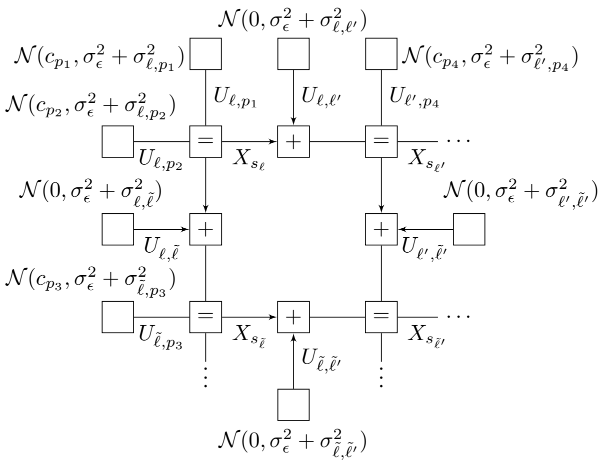


Fig. 2. Factor graph representation of the prior model with nearest neighbor pairs $\{(\ell,\ell'),(\ell,\tilde{\ell}),(\ell',\tilde{\ell}'),(\tilde{\ell},\tilde{\ell}')\} \subset \Delta$ and voxel-boundary pairs $\{(\ell,p_1),(\ell,p_2),(\tilde{\ell},p_3),(\ell',p_4)\} \subset \Theta$.

with $U_{\ell,p} \sim \mathcal{N}(c_p,\sigma_\epsilon^2+\sigma_{\ell,p}^2)$. Rewriting the relations $U_{\ell,\ell'} = X_{s_\ell} - X_{s_{\ell'}}$ and (1) as $\mathbf{U} = D\mathbf{X}$ with a pertinent matrix D, we define our (improper) prior model as

$$\tilde{p}\,(\mathbf{x};\;\boldsymbol{\sigma}^2\,) = \int \delta(\mathbf{u}-D\mathbf{x})\,p(\mathbf{u}|\boldsymbol{\sigma}^2)\,\mathrm{d}\mathbf{u} \tag{2}$$

$$= \prod_{(\ell,\ell')\in\Delta} \frac{1}{2\pi(\sigma_\epsilon^2+\sigma_{\ell,\ell'}^2)} \exp\left(-\frac{(x_\ell - x_{\ell'})^2}{2(\sigma_\epsilon^2+\sigma_{\ell,\ell'}^2)}\right)$$

$$\cdot \prod_{(\ell,p)\in\Theta} \frac{1}{2\pi(\sigma_\epsilon^2+\sigma_{\ell,p}^2)} \exp\left(-\frac{(x_{\ell_p} - c_p)^2}{2(\sigma_\epsilon^2+\sigma_{\ell,p}^2)}\right). \tag{3}$$

where σ_ϵ^2 is a parameter of the algorithm and $\sigma_{\ell,\ell'}^2$ and $\sigma_{\ell,p}^2$ are unknown and will be estimated. From now on we use the notation $(\ell,\ell'(p))$ to denote the pairs (ℓ,ℓ') or (ℓ,p).

The structure of this prior is illustrated in Fig. 2. It is sparsifying in the following sense: at any local maximum of the likelihood, many estimated variance parameters $\sigma_{\ell,\ell'(p)}^2$ are likely to be zero. The corresponding differences $U_{\ell,\ell'(p)}$ are then regularized to be small (but not necessarily zero), thus prompting smooth areas in the reconstruction. On the other hand, when the estimated $\sigma_{\ell,\ell'(p)}^2$ are nonzero, the cost for arbitrarily large jumps is very small, which allows for sharp edges in the reconstruction [15, 16]. The sparsity level can be adjusted by tuning σ_ϵ^2.

3.2. Estimation and Scalar Gaussian Message Passing

As mentioned, we wish to estimate the unknown variances $\boldsymbol{\sigma}^2$ by maximizing the "likelihood". The actual computations boil down to iterative scalar Gaussian message passing. With a similar derivation as in [12], the maximization problem splits for each $\sigma_{\ell,\ell'}^2$ and $\sigma_{\ell,p}^2$:

$$\hat{\sigma}_{\ell,\ell'(p)}^2 = \max\left(0, \mathbb{E}\left[\|U_{\ell,\ell'(p)}\|^2\right] - \sigma_\epsilon^2\right). \tag{4}$$

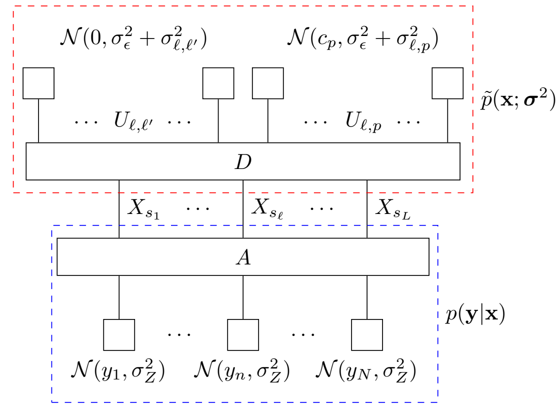


Fig. 3. Factor graph representation of $p(\mathbf{u},\mathbf{x},\mathbf{y}|\boldsymbol{\sigma}^2)$. Top dashed box: $\tilde{p}(\mathbf{x};\boldsymbol{\sigma}^2)$. Bottom dashed box: $p(\mathbf{y}|\mathbf{x})$.

The factor graph in Fig. 3 represents the joint distribution $p(\mathbf{u},\mathbf{x},\mathbf{y}|\boldsymbol{\sigma}^2)$. We use iterative scalar Gaussian message passing [17] in the factor graph of Fig. 3 to compute an approximate estimate of $\mathbb{E}\left[\|U_{\ell,\ell'(p)}\|^2\right]$, which turns out to work well in practice. If the algorithm converges, the estimated means are correct, but the variances are not [18, 19]. The details of this computation can be obtained through extending the algorithm in [12] by introducing boundary conditions. Specifically, we do some extra computations in steps 2 and 3 of the message passing procedure. In step 2 we compute

$$\overrightarrow{w}_{X_{s_\ell}^{(\ell,p)}} := (\sigma_\epsilon^2+\sigma_{\ell,p}^2)^{-1} \tag{5}$$

$$\overrightarrow{\xi}_{X_{s_\ell}^{(\ell,p)}} := c_p(\sigma_\epsilon^2+\sigma_{\ell,p}^2)^{-1}. \tag{6}$$

In step 3, the original formulas are modified as

$$\overrightarrow{w}_{X_{s_\ell}} := \sum_{\ell:(\ell,i)\in\Delta\cup\Theta} \overrightarrow{w}_{X_{s_\ell}^{(\ell,i)}} + \sum_{\ell':(\ell',\ell)\in\Delta} \overrightarrow{w}_{X_{s_\ell}^{(\ell',\ell)}} \tag{7}$$

$$\overrightarrow{\xi}_{X_{s_\ell}} := \sum_{\ell:(\ell,i)\in\Delta\cup\Theta} \overrightarrow{\xi}_{X_{s_\ell}^{(\ell,i)}} + \sum_{\ell':(\ell',\ell)\in\Delta} \overrightarrow{\xi}_{X_{s_\ell}^{(\ell',\ell)}}. \tag{8}$$

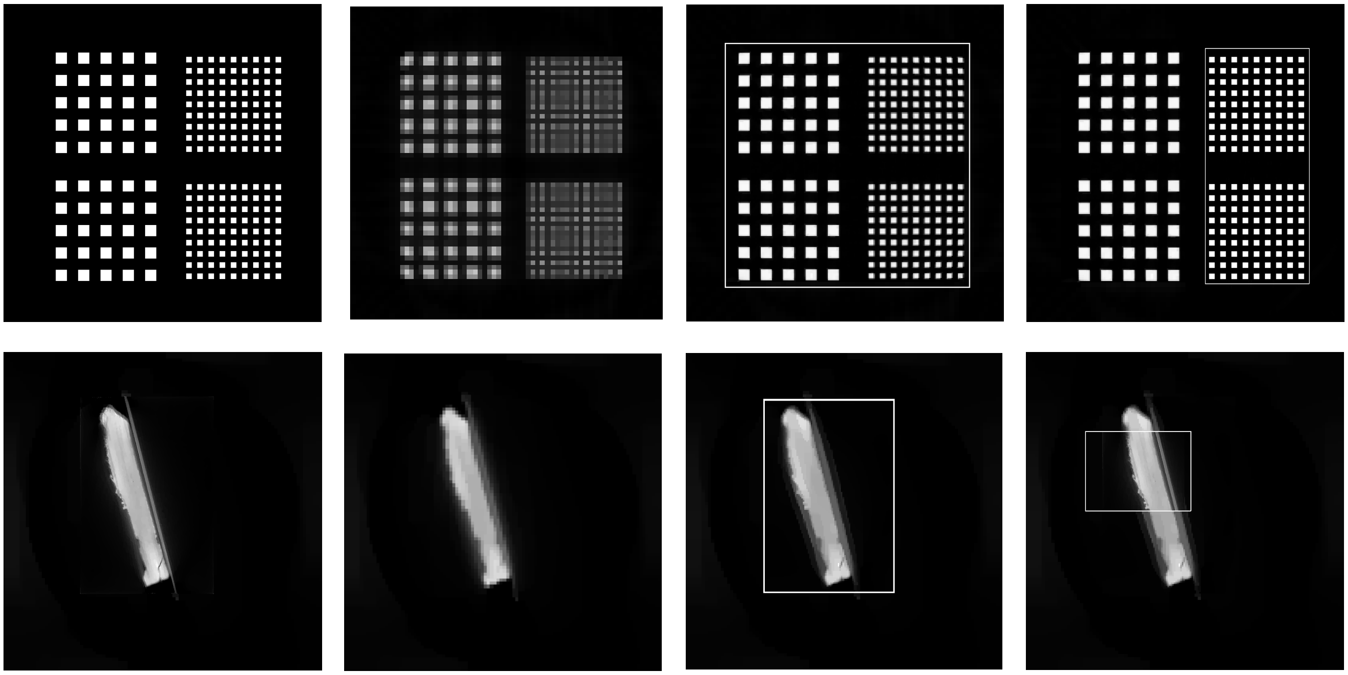
After the algorithm has converged, $\mathbb{E}\left[\|U_{\ell,\ell'(p)}\|^2\right]$ can be calculated from the messages in the factor graph, and $\hat{\sigma}_{\ell,\ell'(p)}^2$ can then be obtained from (4). Finally, messages are used to compute the posterior means

$$m_{X_{s_\ell}} = (\overrightarrow{\xi}_{X_{s_\ell}} + \overleftarrow{\xi}_{X_{s_\ell}})(\overrightarrow{w}_{X_{s_\ell}} + \overleftarrow{w}_{X_{s_\ell}})^{-1} \tag{9}$$

which form our image estimate $\hat{\mathbf{x}} = (m_{X_{s_1}},\ldots,m_{X_{s_L}})^T$.

4. RESULTS

Fig. 4 shows some experimental results using two 2D test objects: first, a simulated phantom and second, a cross-section of a (real) fossil. The multi-resolution reconstruction is done in 3 stages, with a coarse-resolution, an intermediate-resolution and a fine-resolution. The parameters of the reconstruction algorithm (σ_ϵ^2 and σ_Z^2) decrease gradually as the algorithm proceeds. At each stage, we used 20 iterations of message passing and 10 EM updates.



(a) Uniform-resolution (b) Coarse-resolution (c) Intermediate-resolution (d) Final-resolution

Fig. 4. Reconstruction results in different stages of the proposed approach. The areas of interest are framed with white boxes. Top: a simulated phantom. Bottom: a cross-section of a fossil.

In Fig. 4 the leftmost column shows the reconstructions with uniform-resolutions 1024×1024 and 800×800 for the simulated phantom and the fossil respectively. The other columns present the results of the 3 different stages of the multi-resolution approach. The fine resolution used in the last stage is same as that used in the uniform-resolution reconstruction.

Assuming that in Fig. 4 the smaller squares (top row) are the final area of interest of the first object and the small protuberances, which are the fossilized bones, on the left side of the stone (bottom row) are the final area of interest of the second object, we can see from Fig. 4 that the multi-resolution approach provides satisfying reconstruction results. All the small squares are correctly recognized despite being still quite blurred after the first step and the small details of the fossil are not washed out either. For the final area of interest of the first phantom, the squared error between the reconstructions and the ground truth for uniform-resolution reconstruction and multi-resolution reconstruction are 0.0007 and 0.0013 respectively. For the final area of interest of the second phantom, the squared error between the uniform- and multi-resolution reconstruction is 0.0017. The dynamic range of the gray scale in the representation is $[0, 1]$.

The decrease in computational complexity is significant. Compared to the reconstruction with a high uniform-resolution grid, of which the results are shown in the first column of Fig. 4, the multi-resolution approach reduces the computation time by 75.4% and 82.0% and the required memory by 80.5% and 89.4% for the first and second objects respectively. The detailed computation time and memory consumption are given in Tables 1 and 2

Table 1: Memory consumption

	Uniform	Multi
Squares	12.2GB	2.2GB
Fossil	8.5GB	0.9GB

Table 2: Time consumption

	Uniform	Multi
Squares	366s	90s
Fossil	185s	35s

The reduction in complexity is even more significant for 3D reconstructions, since the computational complexity of the applied reconstruction approach scales linearly with the number of pixels (voxels).

5. CONCLUSION

We proposed a general multi-resolution approach for tomography and worked it out for a reconstruction method using a graphical model with NUV terms. The multi-resolution approach enables the selection of an area of interest and the flexible allocation of computational power. The specific reconstruction algorithm uses approximate EM with computations amounting to iterative scalar Gaussian message passing. According to our results, the reconstruction in the final area of interest is essentially as good as the reconstruction using a uniform high-resolution grid, while the complexity is reduced substantially in 2D and much more in 3D.

6. REFERENCES

[1] G. T. Herman, *Fundamentals of Computerized Tomography: Image Reconstruction From Projections*, Springer-Verlag, 2009.

[2] G. Yan, J. Tian, S. Zhu, Y. Dai, and C. Qin, "Fast cone-beam CT image reconstruction using GPU hardware," *Journal of X-ray Science and Technology*, vol. 16, no. 4, pp. 225–234, 2008.

[3] P. B. Noël, A. M. Walczak, J. Xu, J. J. Corso, K. R. Hoffmann, and S. Schafer, "GPU-based cone beam computed tomography," *Computer methods and programs in biomedicine*, vol. 98, no. 3, pp. 271–277, 2010.

[4] W. J. Palenstijn, K.J. Batenburg, and J. Sijbers, "Performance improvements for iterative electron tomography reconstruction using graphics processing units (GPUs)," *Journal of structural biology*, vol. 176, no. 2, pp. 250–253, 2011.

[5] K. Choi, J. Wang, L. Zhu, T. Suh, S. Boyd, and L. Xing, "Compressed sensing based cone-beam computed tomography reconstruction with a first-order method," *Medical physics*, vol. 37, no. 9, pp. 5113–5125, 2010.

[6] J. A. Fessler, "Penalized weighted least-squares image reconstruction for positron emission tomography," *IEEE Trans. on Medical Imaging*, vol. 13, no. 2, pp. 290–300, 1994.

[7] S. Maeda, W. Fukuda, A. Kanemura, and S. Ishii, "Maximum a posteriori X-ray computed tomography using graph cuts," in *Journal of Physics: Conference Series*. IOP Publishing, 2010, vol. 233.

[8] E. Y. Sidky and X. Pan, "Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization," *Physics in Medicine and Biology*, vol. 53, no. 17, pp. 4777, 2008.

[9] M. Yan and L. A. Vese, "Expectation maximization and total variation-based model for computed tomography reconstruction from undersampled data," in *Proc. SPIE 7961, Medical Imaging*. Int. Society for Optics and Photonics, 2011.

[10] N. Dobigeon, A. O. Hero, and J.-Y. Tourneret, "Hierarchical Bayesian sparse image reconstruction with application to MRFM," *IEEE Trans. on Image Processing*, vol. 18, no. 9, pp. 2059–2070, 2009.

[11] M. Borgerding, P. Schniter, J. Vila, and S. Rangan, "Generalized approximate message passing for cosparse analysis compressive sensing," in *IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, 2015, pp. 3756–3760.

[12] N. Zalmai, C. Luneau, C. Stritt, and H.-A. Loeliger, "Tomographic reconstruction using a new voxel-domain prior and Gaussian message passing," in *Proc. 24th European Signal Processing Conf. (EUSIPCO)*, 2016, pp. 2295–2299.

[13] B. Ma, N. Zalmai, R. Torfason, C. Stritt, and H.-A. Loeliger, "Color image segmentation using iterative edge cutting, NUV-EM, and Gaussian message passing," in *Proc. 5th IEEE Global Conf. on Signal and Information Processing (GlobalSIP)*, 2017.

[14] B. De Man and S. Basu, "Distance-driven projection and backprojection in three dimensions," *Physics in Medicine and Biology*, vol. 49, no. 11, pp. 2463, 2004.

[15] M. E. Tipping, "Sparse Bayesian learning and the relevance vector machine," *Journal of Machine Learning Research*, vol. 1, pp. 211–244, 2001.

[16] H.-A. Loeliger, L. Bruderer, H. Malmberg, F. Wadehn, and N. Zalmai, "On sparsity by NUV-EM, Gaussian message passing, and Kalman smoothing," *Information Theory and Applications Workshop (ITA)*, Feb 2016.

[17] H.-A. Loeliger, "An introduction to factor graphs," *IEEE Signal Processing Magazine*, vol. 21, no. 1, pp. 28–41, 2004.

[18] J. S. Yedidia, W. T. Freeman, and Y. Weiss, "Generalized belief propagation," in *NIPS*, 2000, vol. 13, pp. 689–695.

[19] Y. Weiss and W. T. Freeman, "Correctness of belief propagation in Gaussian graphical models of arbitrary topology," *Neural Computation*, vol. 13, no. 10, pp. 2173–2200, 2001.