

Published in final edited form as:

Comput Vis Image Underst. 2008 October 1; 112(1): 67–80. doi:10.1016/j.cviu.2008.06.005.

Limited view CT reconstruction and segmentation via constrained metric labeling

Vikas Singh¹, Lopamudra Mukherjee², Petru M. Dinu³, Jinhui Xu², and Kenneth R. Hoffmann⁴

¹Biostatistics & Medical Informatics and Computer Sciences, UW-Madison.

²Computer Science and Engineering, SUNY-Buffalo.

³Toshiba Medical Research Institute USA, Vernon Hills, Illinois.

⁴Toshiba Stroke Research Center, Dept. of Neurosurgery, SUNY-Buffalo.

Abstract

This paper proposes a new discrete optimization framework for tomographic reconstruction and segmentation of CT volumes when only a few projection views are available. The problem has important clinical applications in coronary angiographic imaging. We first show that the limited view reconstruction and segmentation problem can be formulated as a “constrained” version of the metric labeling problem. This lays the groundwork for a linear programming framework that brings metric labeling classification and classical algebraic tomographic reconstruction (ART) together in a unified model. If the imaged volume is known to be comprised of a finite set of attenuation coefficients (a realistic assumption), given a regular limited view reconstruction, we view it as a task of voxels reassignment subject to maximally maintaining consistency with the input reconstruction and the objective of ART simultaneously. The approach can reliably reconstruct (or segment) volumes with several multiple contrast objects. We present evaluations using experiments on cone beam computed tomography.

1 Introduction

Tomographic reconstruction from only a limited number of projection images is an important problem in medical imaging. For example, in coronary interventions, the vasculature undergoes simultaneous motion and deformation due to respiration and coronary contraction – thus, a complete rotational acquisition of a ‘stationary’ object is practically infeasible. As a result, 3D volumetric data in coronary studies is difficult (Note: reconstruction of the vasculature from two views may be performed by indicating the segments of interest followed by a triangulation like procedure). On the other hand, when the imaged object is stationary and a complete rotational set of projection images are available, it is well known that the corresponding reconstruction task can be precisely formulated as an inverse problem. This problem is very well studied and many algorithms with proven convergence properties exist, see [1] and references therein for more information. However, when imaging an object with periodic motion such as the coronary vasculature, the ‘limited’ number of views are typically the ones that correspond to the heart in approximately the same *phase* (e.g., diastole or systole) as the

© 2008 Elsevier Inc. All rights reserved.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

imaging system (gantry) rotates around it, see Fig. 1 (b). Here, the computed tomography (CT) reconstruction problem formulated as a linear system of equations is under-defined. In the frequency domain, this can be interpreted as certain regions of the Fourier space which are insufficiently characterized. The key problem in limited view CT reconstruction is to reconstruct a 3D volume that corresponds to what may have been a complete rotational acquisition, although the given angular spacing of the projection image set is quite far from the Nyquist criterion. In addition to coronary vessel reconstruction, there are instances of the limited view problem in breast imaging [2] (as a result of the maximum radiation dose that can be given to the breast tissue) as well as in dental radiography [3,4].

We illustrate the problem of artifacts and blurring using an example in Fig. 2 where Fig. 2 (a) shows a single projection view and Figs. 2 (b) and 2(c) show a 2D slice from a reconstruction volume obtained from 360 (fully sampled) and 15 images respectively. Note the artifacts (radial lines from the center to the image boundary) in Fig. 2(c) compared with the ‘clean’ reconstruction in Fig. 2(b). Our primary objective is to recover Fig. 2(b) from Fig. 2(c) reliably and efficiently. This will enable us to reliably approximate what might have been the volumetric reconstruction for a complete rotational acquisition (i.e., fully sampled), even when only a limited number of views of the object are available.

The remainder of this paper is organized as follows. In §1.1, we review related literature on this problem. We introduce the main ideas of this paper in §2 – first, by looking at limited view CT reconstruction as a case of metric labeling with additional requirements in §2.1 and then elaborating on the form our additional requirements take in §2.2. We formally establish some hardness results for the resultant optimization problem in §2.3. We introduce the optimization model in §3.1, which is followed by a discussion on how to obtain integral solutions such that (1) we do not move far away from the global optimum and (2) our additional requirements still remain satisfied. We present experimental results in §4, and finally highlight some avenues for future research and conclude in §5. A preliminary version of this paper appeared in [5].

1.1 Previous work

Volumetric reconstruction using only a limited number of projection views has been an active topic of research because of its relevance and far reaching practical implications. A common feature of some of the earlier approaches was the extension or modification of standard CT reconstruction algorithms to the limited view case, this was done by using the iterative reconstruction-reprojection algorithm recursively while trying to *estimate the data in the missing views* [6,7]. This strategy has had limited success; as Anderson pointed out in [8] – no additional information is introduced during the estimation process and hence the convergence of these iterative approaches is not well defined. Subsequent efforts have focused on the design of algorithms specifically for the limited view case using a wide variety of tools – from modified algebraic reconstruction [2] to Fourier analysis [9]. Several researchers have reported on satisfactory practical performance of Expectation Maximization (EM) algorithms [10], but such algorithms usually have difficulty escaping local minima suggesting that we may obtain sub-optimal results; see the discussion in [11]. Here, we will focus on approaches that are combinatorial in nature and relevant as background for the subsequent sections.

A key feature of tomographic reconstruction is the calculation of line integrals of an unknown attenuation function (imaged volume) along projection lines originating from an x-ray source. If the attenuation distribution (to-be-estimated function) is considered to be discrete, each continuous integral can be substituted by a summation of discrete values (see Fig. 3), the reconstruction problem then naturally transforms into a discrete optimization problem. Recognizing this connection, several researchers have proposed mathematical programming as well as certain greedy frameworks to solve this problem. A noteworthy example is [12] where Fishburn *et al.* formulated the problem of determining a subset S given line integrals as

a linear programming problem. Later, this approach was extended to the image reconstruction domain by Gritzmann *et al.* [13] who introduced several objective functions to obtain “maximal” consistent reconstructions. A recent work [14] used the linear programming approach of [12] together with the objective function of [13] and added additional “smoothness” terms in the objective to obtain tomographic reconstruction of 3D volumes. In a subsequent work, the authors proposed a quadratic objective function [15] and reported improvements over their approach in [14]. A common limitation of many existing algorithms is that they are only applicable for *binary* images or volumes – the implicit assumption is that the imaged volume only draws from two distinct attenuation coefficients, namely air (*i.e.*, background) and one among tissues, contrast agent like iodine or bone as foreground. This assumption is only occasionally satisfied making the practical applicability of such techniques rather limited. Here, we remove this restriction; the algorithm allows multiple contrast objects in the imaged volume, they can all be reconstructed and segmented simultaneously.

The limited view problem has also been encountered in other medical imaging modalities (outside of CT reconstruction). For example, in the magnetic resonance (MR) imaging community, a number of recent papers have tackled the problem of dealing with artifacts and noise (in parallel MR image reconstruction) within an optimization framework. The artifacts in the reconstruction arise in an effort to reduce the number of acquired data samples (and consequently the patient scan time). Several interesting approaches based on a least-squares type framework [16], a formulation amenable to the use of efficient graph-cuts methods incorporating edge-preserving priors [17], and methods using L_1 norm regularization [18] have been proposed. For more on the background of the problem in MRI and a review of existing techniques, we refer the reader to [19].

1.2 Brief Introduction to metric labeling

We will briefly digress here to familiarize the reader with *metric labeling*, an interesting combinatorial problem introduced independently by Kleinberg and Tardos [20] and Boykov, Veksler, and Zabih [21]. In this classification framework, the goal is to assign a small set of *representative* labels, T , to a larger set of items, P , where $|T| \ll |P|$. It is elegant in that it effectively captures two competing influences – (1) the *separation cost* conveys the idea that if two items are alike, it is desirable to classify them together by assigning them the same label; (2) the *assignment cost* on the other hand depends only on the individual label assignments we make for each item. For instance, in an image processing application, it makes sense to assign a pair of neighboring pixels with similar intensity values to the same cluster (or class). At the same time, we must prefer assigning a ‘light-blue’ pixel to ‘blue’ (as a representative label) rather than denoting it by ‘black’. While the problem is **NP**-hard even when the separation cost is metric, Kleinberg and Tardos [20] proposed an elegant randomized-rounding-based approximation algorithm⁵. Independently, Boykov, Veksler, and Zabih [21] pointed out the relationship of graph-cuts to metric labeling and proposed an efficient max-flow-based algorithm that relies on computing local minima of energy functions repeatedly (polynomial number of steps) for image segmentation. The models are based on Markov Random Fields [23], and allow the design of very efficient optimization schemes that lead to provably good solutions. The limited view CT reconstruction problem benefits from these ideas, as we discuss in the subsequent sections.

2 Theory and Main Ideas

Our linear programming formulation for the limited view tomographic reconstruction and segmentation problem builds upon Kleinberg and Tardos’ metric labeling model.

⁵In [22], Hochbaum showed that polynomial time algorithms were possible for convex assignment cost and linear smoothness cost.

2.1 Limited view reconstruction as metric labeling

Let us start with the volume obtained from a standard simultaneous algebraic reconstruction procedure (SART) as a first step, we denote this volume as V where $V(i, j, k)$ or $V(p)$ denotes the intensity value at voxel $p \in V$ whose coordinates are (i, j, k) . Recall that V contains severe streak and blurring artifacts as illustrated in Fig. 2(c). Our primary goal is to recover an artifact free reconstruction.

A simple visual analysis of the images in Figs. 2(b) and (c) indicates that the distribution of the intensity values of Fig. 2(b) has only two well concentrated regions, the first centered around the foreground intensity values (vessel cross section) and background values comprising the second region – we know that the imaged volume has only two distinct attenuation coefficients. Generalizing this further, if the imaged volume consists of bone, air, and a contrast agent, we must see three distinct contrast values in the reconstruction, one for each unique material in the imaged configuration. The distribution of contrast values in Fig. 2(c), on the other hand, is far more “spread out” as a result of blurring and streak artifacts. The key observation here is the following – when imaging the human body, in most cases, one knows in advance the *expected* materials (and their physical properties) and consequently the contrast values in the reconstruction given information about the imaging system. The x-ray attenuation coefficients of materials that may be imaged are available and the exact subset can be determined based on the particular anatomy. Therefore, given the characteristics of the imaging system, the specific contrast values can be calculated. Hence, it makes intuitive sense to view the problem of going from Fig. 2(c) to (b) as a reassignment task.

The loss of texture information—The discussion above raises one possible concern: that the fully sampled (360-view) reconstruction has a natural “texture”, which we may not be able to recover by reassigning the pixels to three (or four) distinct contrast values. This could be problematic if such texture adds value to the diagnostic process (e.g., if one were analyzing bone density). However, what is of interest in many applications (such as the present one) is a geometric reconstruction of the vasculature (e.g., a mesh representation), so it may be used for quantitative evaluations by a clinician or for other biomechanical experiments. For example, even when a textured 360-view reconstruction is available, quantitative analysis of the vessel (e.g., tortuosity calculation, vessel sizing) is performed after “binarizing” [24]. Therefore, discrete reconstructed volumes are useful and the reassignment formulation turns out to be a desirable characteristic in this application. We will focus on the reassignment task next.

Our objective will be to reassign the voxels to one among a set of distinct contrast values. This set of labels (or contrast values) should ideally be given as an input based on the physics of the imaging system and the imaged anatomy; however, if unavailable, may be estimated from the input volume⁶. Assigning each voxel to its closest label greedily is unlikely to work well, so we must put similar intensity-valued voxels in the same partition (or segment); by the same argument, pixel pairs that are dissimilar in the input image volume must be assigned to dissimilar labels. To reflect this logic⁷, the cost consists of two parts, data assignment cost and smoothness cost [25,21]:

$$\mathcal{E} = \mathcal{E}_d + \mathcal{E}_s. \quad (1)$$

⁶In general, we may perform a simple preprocessing step using k -means clustering (where k is chosen based on the anatomy being imaged).

⁷This strategy is standard in segmentation schemes based on Markov Random Fields in vision. We discuss this briefly and then move to our formulation which allows incorporation of CT specific special constraints.

The first term *i.e.*, \mathcal{E}_d evaluates the cost of assigning a voxel to a distinct label. The second term in the objective function, \mathcal{E}_s is useful to preserve neighborhood continuity. These two terms of the energy functional are effectively captured by the objective function of metric labeling. The likelihood of a voxel being assigned to a particular label is given by the data term, where as pairwise relationships among objects are preserved through the smoothness term.

Consider a binary function, $h : V \rightarrow T$, mapping voxels to elements in T , where T is the set of labels; h yields the mapping, X , defined as follows.

$$X(p, l) = \begin{cases} 1 & \text{if } p \in V \text{ is assigned label } l \in T; \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

For the sake of presentation, let us assume that X is known. If $T(l)$ is the intensity value of label l , then if p is assigned to l , the data term is

$$\mathcal{E}_d(p, l) = g_d(V(p), T(l)) = \|V(p) - T(l)\|_1 \quad (3)$$

Denoting $g_d(V(p), \cdot)$ as $g_d(p, \cdot)$, the data energy penalty for p is

$$\mathcal{E}_d(p) = \sum_{l=1}^{|T|} g_d(p, T(l)) \cdot X(p, l). \quad (4)$$

We can extend this for all voxels in V as

$$\mathcal{E}_d = \sum_{p=1}^{|V|} \sum_{l=1}^{|T|} g_d(p, T(l)) \cdot X(p, l). \quad (5)$$

Let us now move to the smoothness term of the energy. If neighboring pixels p and q are assigned labels $f(p)$ and $f(q)$, the penalty imposed must depend on the extent of similarity of p and q as well as $f(p)$ and $f(q)$. If p and q are very similar, we assign a high weight $w(\cdot)$ to their corresponding edge $e_{pq} \in E$. This encourages the optimization process to assign them to similar labels to avoid incurring a penalty of $g_s(f(p), f(q)) \cdot w(e_{pq})$, where $w(e_{pq})$ is the weight of e_{pq} (similarity of p and q).

Intensity Similarity—Recall that for a pair of intensities p and q the edge weight or similarity

(using RBF kernel) may be given as $\exp\left(-\frac{\|p - q\|^2}{2\sigma^2}\right)$ where σ controls the smoothness. In the Potts model, it is expressed as $K \cdot \delta(f(p) \neq f(q))$, where $\delta(\cdot)$ is 1 if $f(p) \neq f(q)$ and K is a either a constant or the penalty for a discontinuity at (p, q) . Similarly, the function, $g_s(f(p), f(q))$ can be related to the difference between $f(p)$ and $f(q)$ or be a constant value. In this application, assigning a “bone” voxel to a “tissue” label is as incorrect as another misclassification, say to “iodine”. Hence, we consider $g_s(f(p), f(q)) = M$, to penalize all misclassifications equally.

We consider the smoothness energy for p and q to be given as

$$\mathcal{E}_s(p, q) = \begin{cases} M \cdot w(e_{pq}) & \text{if } p, q \text{ have different labels;} \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

We can rewrite (6) using X as

$$\mathcal{E}_s(p, q) = \frac{1}{2} \left(M \cdot w(e_{pq}) \sum_{l=1}^{|T|} \|X(p, l) - X(q, l)\| \right). \quad (7)$$

Also, each voxel must only be assigned to one label,

$$\sum_{l=1}^{|T|} X(p, l) = 1 \quad \forall p \in V. \quad (8)$$

Using the ideas in [20], we have the following (preliminary) integer programming problem where $X(p, l)$ (or x_{pl}) is 1 or 0 depending on whether or not the voxel p is assigned to label l (similar to h in (2))

$$\begin{aligned} \min \quad & \sum_{p=1}^{|V|} \sum_{l=1}^{|T|} g_d(p, T(l)) \cdot x_{pl} + \frac{M}{2} \sum_{e_{pq}} \sum_{l=1}^{|T|} w(e_{pq}) \|x_{pl} - x_{ql}\|, \\ \text{s.t.} \quad & \sum_{l=1}^{|T|} x_{pl} = 1, \\ & x_{pl} \in \{0, 1\}. \end{aligned} \quad (9)$$

2.2 Line Sum Constraints

While the model above accurately captures the requirement of reassigning voxels to a discrete set of contrast values subject to data and smoothness terms, it ignores the available projection view data completely in the assignment process. As a result, the final reassigned volume may not be consistent with the projection views. Also, if the artifacts are conspicuous enough they will be classified as one of the foreground labels. To address this difficulty, we must determine a volume that not only preserves the desired qualities above but also maintains complete consistency with available projection data. In other words, the reassignment must *agree* with the *witnesses* (projection images of the volume). We address this by introducing ‘line-sum’ constraints, the key principle guiding tomographic reconstruction. Why imposing line-sum as a witness is preferable to solving for it directly is a question we address at the end of §2.3.

Let V , $|V| = n^3$ denote the unknown volume. Let $m = n^2k$ projection lines come out of an x-ray source, pass through the volume and cast a projection on k image planes. Then, the ‘line sum’ in tomographic reconstruction is

$$A_i x = b_i \quad \forall_i \in \{1, \dots, n^2k\}, \quad (10)$$

where the to-be-determined x is a function of the attenuation coefficients of the imaged distribution and b is the vector of *observed* projections, *i.e.*, a pixel on the projection image. The entry $A_{ij} \in \mathbb{R}_+$ indicates the distance traveled by line i through voxel j (see Fig. 3), the contribution of j to the i^{th} integrand. Here, the distribution is continuous. But we seek to assign the voxels of V to discrete intensity levels, given as T . Using $X \in \mathbb{R}^{|V| \times |T|}$ as a variable matrix denoting to-be-computed reassignments where $X(p, l) \in \{0, 1\}$ maps voxel p to label l , we can denote the *discretized* effective attenuation coefficient as XT where $X \in \mathbb{R}^{|V| \times |T|}$ and $T \in \mathbb{R}^{|T| \times 1}$. For a row in X , the product with the assignment matrix on the right gives the contrast value the voxel assumes *after reassignment*. The equivalent form is then

$$A_i X T \approx b_i, \forall i \in \{1, \dots, n^2 k\}. \quad (11)$$

Note that after reassignment to discrete values, the projection of the volume may not completely agree with the given projections. In fact, such a reassignment will most likely not even exist. Nonetheless, a good solution should be close to the projection images. Let ϵ denote the allowable variation between the projection line sum and the given pixel value. We can rewrite (11) as

$$A_i X T \geq b_i - \epsilon, \quad A_i X T \leq b_i + \epsilon, \quad \forall i, \quad (12)$$

where ϵ can either be a user input or be determined by some preprocessing. In summary, we would like to perform a MRF-type reassignment under the requirement that the final volume be consistent with the set of witnesses (projection images).

2.3 Hardness results

In this section, we show that determining a voxel \rightarrow label assignment satisfying the line-sum constraints is hard. Recall that a problem W is **NP**-hard if $\forall W' \in \mathbf{NP}, W' \leq_p W$, where \leq_p denotes polynomial time reducibility. If it is also the case that W is in **NP**, then W is called **NP**-complete, where **NP** denotes the class of non-deterministic polynomial time solvable problems.

We consider the simplest case where we must determine an assignment for only a single line-sum constraint. Obviously, the result holds for multiple lines also. For any line i , let the coefficient row, A_i , be denoted as $F = \{F_1, \dots, F_m\} \in \mathbb{Z}_+$, note that the integrality condition is wlog since the coefficients can be scaled. Define a set $\bar{F}, |\bar{F}| = |F| \cdot |T|$ given by the product of terms in F and T , picking one from each set. It is simple to see that finding an assignment satisfying the line-sum constraint is a form of the following simplified decision problem.

Problem 1 (Feasibility problem (FPr))—Given n coefficients $\bar{F} = \{\bar{F}_1, \bar{F}_2, \dots, \bar{F}_{|\bar{F}|}\}$ with $\bar{F}_i \geq 0, \bar{F}_i \in [0, 1]$, a set $T = \{T(1), T(2), \dots, T(c)\}$ for some constant c , and a non-negative scalar value b , determine the existence of a vector $X \in T^n$ such that $b - \epsilon \leq \bar{F} X \leq b + \epsilon$.

The above problem does not ask for an explicit voxel \rightarrow label assignment, but only the number of times each term in \bar{F} can be considered to satisfy the inequality. We state and then establish the following result.

Theorem 1: *The Feasibility problem is **NP**-complete.*

Proof: Consider the following well-known **NP**-complete problem [26].

Problem 2 (Subset Sum (SS))—Given a set $D = \{d_1, d_2, \dots, d_n\}$ where $d_i \geq 0$ and $B > 0$, determine whether there is a subset $S', S' \subset S$ such that $\sum_{d_i \in S'} d_i = B$.

The reduction is: $SS \leq_p \text{FPr}$. First, we rewrite SS as follows. Let X be a vector in $\{1, 0\}^n$. Then, the SS problem asks for the existence of a X such that $DX = B$. For every instance of SS, we can construct an equivalent instance of FPr by setting $c = 2$ and $\epsilon = 0$. If we can solve FPr (in polynomial time), we will also have obtained a solution to SS. Hence, FPr is **NP**-hard. Given a feasible solution to FPr, it can be easily verified whether the equality $b - \epsilon \leq \bar{F} X \leq b + \epsilon$ is satisfied. Thus, FPr is in **NP**. The theorem follows.

The usefulness of ART solution as an initialization: Before proceeding, let us briefly investigate why the ART solution is useful as a starting point, i.e., the utility of viewing our problem as a reassignment task instead of solving for the label assignments for pixels directly. It is easy to see that determining such label assignments directly would require solving the line-sum constraints (in a discrete setting). The statement of the theorem above tells us that this is difficult. Of course, we may still avoid using the ART solution: the alternative is to solve (12) in a continuous (rather than discrete) setting. But comparing the continuous version of (12) and (10) shows that we will get results similar to the ART reconstruction of (10). It is challenging to design a principled rounding strategy for this solution. Using the algebraic reconstruction as a starting point (i.e., initialization), despite being artifact prone, proves to be helpful if we exploit the projection data as additional information.

3 Proposed Algorithm

3.1 Model—Combining (9) with the line sum constraints we have the following IP model:

$$\begin{aligned}
 \min \quad & \sum_{p=1}^{|V|} \sum_{l=1}^{|T|} g_d(p, T(l))x_{pl} + \frac{M}{2} \sum_{e_{pq}} \sum_{l=1}^{|T|} w(e_{pq}) \|x_{pl} - x_{ql}\| \\
 \text{s.t.} \quad & \sum_{l=1}^{|T|} x_{pl} = 1, \\
 & b_i - \epsilon \leq \sum_{p=1}^{|V|} \sum_{l=1}^{|T|} A_{ip} x_{pl} T(l) \leq b_i + \epsilon \quad \forall i, \\
 & x_{pl} \in \{0, 1\}.
 \end{aligned} \tag{13}$$

where $T(l)$ denotes the l^{th} label of T . In the above model, the variables x_{pl} are considered binary, though these variables can be relaxed to be in $[0, 1]$; hence the above model can be formulated as a linear program. The non-linear absolute term in the objective can be addressed with an additional variable z_e , one for each edge. Some additional constraints are needed to complete the linearization. For completeness, the transformed model follows.

$$\begin{aligned}
 \min \quad & \sum_{p=1}^{|V|} \sum_{l=1}^{|T|} g_d(p, T(l))x_{pl} + \frac{M}{2} \sum_{e_{pq}} \sum_{l=1}^{|T|} w(e_{pq})z_e \\
 \text{s.t.} \quad & \sum_{l \in T} x_{pl} = 1 \quad \forall p \in V, \\
 & z_e \geq x_{pl} - x_{ql} \quad \forall p, q \in V, \forall l \in \{1, \dots, |T|\}, \\
 & z_e \geq x_{ql} - x_{pl} \quad \forall p, q \in V, \forall l \in \{1, \dots, |T|\}, \\
 & b_i - \epsilon \leq \sum_{p=1}^{|V|} \sum_{l=1}^{|T|} A_{ip} x_{pl} T(l) \leq b_i + \epsilon \quad \forall i, \\
 & x_{pl}, z_e \in [0, 1].
 \end{aligned} \tag{14}$$

Theorem 2: *The number of constraints in (14) is $O(n^2k + |V|)$, the number of variables are $O(|V|)$. The linear program in (14) can be solved optimally in polynomial time.*

3.2 Rounding—In this section, we discuss how to obtain an integral solution, X , from a fractional one, x^* . Let $x^* = [x_{pl}^*]$ be an optimal solution to (14), $x_{pl}^* \in [0, 1]$. Our interest is in the following two issues.

Approximation: Not move far away from the optimal solution of (14).

Feasibility: The rounded solution still respects the line-sum constraints.

The first point refers to the so-called ‘gap’ – the ratio of the rounded integral and fractional optimized values. The second aspect, i.e., feasibility, in the present problem is challenging. Observe that in most LPs, *any* rounded solution is feasible. For example, consider constraints

of the form $\sum_{i=1}^n x_i \geq 0$. Clearly, any 0–1 solution will trivially satisfy this requirement, hence the focus is primarily placed on *provably* good approximations. However, the situation here is more complicated since the line sum constraints have lower *and* upper bounds; the inequalities with “ \leq ” are the *packing* constraints and “ \geq ” denotes *covering* constraints. Simultaneously satisfying both while not moving too far away from the optimal value makes this difficult. We will focus on rounding primarily in the following context.

1. Deterministic: where both the upper and lower bounds of the line-sum constraints are violated by at most a known factor in the *worst* case.
2. Randomized: where the upper (lower) bound is satisfied with high probability, the lower (upper) bound is violated by at most a constant factor.

3.2.1 Rounding for two label case: Let the label set be T , the indices of the two labels be a_1 and a_2 , and $T(a_1)$ and $T(a_2)$ denote their intensity values. Wlog, assume $T(a_1) > T(a_2) > 0$, x^* denotes the optimal LP solution and X is the integral solution.

3.2.1.1 Deterministic rounding: The approach is as follows.

1. If $x_{pa1}^* \geq x_{pa2}^*$ set $X_{pa1} = 1$ and $X_{pa2} = 0$.
2. If $x_{pa1}^* < x_{pa2}^*$ set $X_{pa1} = 0$ and $X_{pa2} = 1$.

Lemma 1: The deterministic rounding produces an integral solution X s.t.

$$\frac{2}{y+1}(b_i - \epsilon) \leq A_i X T \leq 2(b_i + \epsilon) \quad \forall_i, T(a_1) = y T(a_2). \quad (15)$$

Proof: Since x^* is feasible, we have $x_{pa1}^* + x_{pa2}^* = 1$. Hence, $x_{pa1}^* \geq 0.5$ implies $x_{pa2}^* < 0.5$. Consider the (worst) case when the upper bound is violated. This happens when all X_{pa1} ’s are set to 1 (using x_{pa1}^*); simultaneously, all X_{pa2} ’s are set to 0 (using x_{pa2}^*). Clearly, the lower bound is

trivially satisfied. Also, $x_{pa1}^* \geq \frac{1}{2} X_{pa1}$. For each p , the contribution to the line sum *before* rounding was $A_{ip} T(a_1) x_{pa1}^* + A_{ip} T(a_2) x_{pa2}^*$, after rounding the contribution is $A_{ip} T(a_1)$. Since $T(a_1) > T(a_2)$, $A_{ip} T(a_1) \geq A_{ip} T(a_1) x_{pa1}^* + A_{ip} T(a_2) x_{pa2}^*$. Since $T(a_1) = y T(a_2)$ where $y > 1$ and N denotes the number of voxels, we have the following.

$$\begin{aligned} \sum_{p=1}^N A_{ip} T(a_1) X_{pa1} &= \sum_{p=1}^N A_{ip} T(a_1) (x_{pa1}^* + x_{pa2}^*) \\ &= \left(1 + \frac{y-1}{y+1}\right) \sum_{p=1}^N (A_{ip} T(a_1) x_{pa1}^* + A_{ip} T(a_2) x_{pa2}^*) \\ &\leq \left(1 + \frac{y-1}{y+1}\right) (b_i + \epsilon) \\ &\leq 2(b_i + \epsilon) \end{aligned} \quad (16)$$

The detailed derivation is available in §6 (Appendix). The violation on the lower bound can be shown similarly, by considering the (worst) case where the lower bound is violated (when all p are assigned to the smaller label $T(a_2)$).

3.2.1.2 Randomized rounding: When there is a constant factor violation in the upper bound,

the lower bound (in the worst case) may be violated by a larger ratio of $\frac{2}{y+1}$. To improve these results, in this section, we propose a randomized rounding schema. The main result of this approach is

- The lower bound is satisfied with high probability.
- The upper bound is violated (in worst case) by at most a factor of 2.
- The obtained integral solution is ‘good’, *i.e.*, close to the optimal solution.

The key in our rounding procedure is a non-linear rounding scheme, a mechanism recently used to obtain improved approximation ratios for a number of problems [27,28,29,30]. In using such a scheme, the design of an appropriate non-linear rounding function is important. We first introduce a few technical results, see [31].

Theorem 3 (Hoeffding): *If X_1, \dots, X_n are mutually independent 0/1 random variables, such that $\Pr[X_i = 1] = p_i$ and $\Pr[X_i = 0] = 1 - p_i$, and $\psi = \sum_{i=1}^n w_i X_i$, $0 \leq w_i \leq 1$, then for any $\lambda > 0$*

$$\Pr[\psi \geq \mathbb{E}(\psi) + \lambda] \leq \exp\left(\frac{-2\lambda^2}{n^2}\right), \Pr[\psi \leq \mathbb{E}(\psi) - \lambda] \leq \exp\left(\frac{-2\lambda^2}{n^2}\right) \quad (17)$$

A direct consequence of the Hoeffding Inequality when $\psi \ll n$ is the Anluin Valiant Inequality as follows.

Theorem 4 (Anluin Valiant [32]): *If X_1, \dots, X_n are mutually independent variables with $0 \leq X_i \leq 1$, $\Pr[X_i = 1] = p_i$ and $\Pr[X_i = 0] = 1 - p_i$, and $\psi = \sum_{i=1}^n X_i$, then for any $\beta > 0$*

$$\Pr[\psi \geq (1+\beta)\mathbb{E}(\psi)] \leq \exp\left(\frac{-\beta^2\mathbb{E}(\psi)}{2(1+\beta/3)}\right) \quad (18)$$

$$\Pr[\psi \leq (1-\beta)\mathbb{E}(\psi)] \leq \exp\left(\frac{-\beta^2\mathbb{E}(\psi)}{2}\right) \quad (19)$$

We now establish a simple result invoking the Anluin Valiant Inequality for use later. Note that (18) was used to show an upper bound result in [33].

Lemma 2: *Let X_1, \dots, X_n be mutually independent 0/1 random variables, with $\Pr[X_i = 1] = p_i$ and $\Pr[X_i = 0] = 1 - p_i$ for all i ; let $\psi(X) = \sum_{i=1}^n w_i X_i$, $0 \leq w_i \leq 1$. If ψ_{lb} be the lower bound on $\mathbb{E}(\psi(X))$, then for any $\beta > 0$*

$$\Pr[\psi(X) \leq (1-\beta)\psi_{lb}] \leq \exp\left(\frac{-\beta^2\psi_{lb}}{2}\right). \quad (20)$$

Proof: Let $f = \frac{\psi_{lb}}{\mathbb{E}(\psi(X))}$. Let Y_1, \dots, Y_n be random variables in $[0, 1]$ such that Y_i takes the value $f w_i X_i$ with probability 1. Clearly Y_i are independent variables with $1 \geq \mathbb{E}(Y_i) \geq 0$. Therefore, the random variable $\psi(Y) = Y_1 + \dots + Y_n$ is such that $\mathbb{E}(\psi(Y)) = \psi_{lb}$ and $\psi(Y) \leq \psi(X)$. It follows that

$$\Pr[\psi(X) \leq (1 - \beta)\psi_{lb}] \leq \Pr[\psi(Y) \leq (1 - \beta)\psi_{lb}] \leq \exp\left(\frac{-\beta^2 \psi_{lb}}{2}\right). \quad (21)$$

3.2.1.3 Rounding Approach: We start with x^* . Choose a $\delta \geq 1$ (the precise value of δ will be specified later). Let X_p be a random variable, $X_p = 1$ indicates that the voxel p maps to the higher-valued label, $T(a_1)$, p maps to the lower-valued label $T(a_2)$ if $X_p = 0$. Initially, voxels do not have a label assignment. We iterate through the following steps until all voxels are labeled.

1. Pick a label index $l \in \{a_1, a_2\}$ at random and a real number $\alpha \in [0, 1]$.
2. For each voxel p which is not yet assigned, assign it the label $T(a_1)$ (making $X_p = 1$) if $l = a_1$ and $\alpha \leq \delta x_{pa_1}^*$.
3. Else, assign it the label $T(a_2)$ (making $X_p = 0$) if $l = a_2$ $\alpha \leq 1 - \delta x_{pa_1}^*$.

From the above, it is clear that we have $\Pr[X_p=1]=\delta x_{pa_1}^*$ and $\Pr[X_p=0]=1 - \delta x_{pa_1}^*$.

3.2.1.4 Constraint Satisfiability: The integral solution X_{pa_1} can be thought of as an assignment to a random variable of the same name, which in turn, is the assignment to random variable X_p . It is easy to see that $\mathbb{E}(X_{pa_1})=\mathbb{E}(X_p)=\delta x_{pa_1}^*$.

Lemma 3: Let s be a value that satisfies

$$\min(b_i - \epsilon - \sum_{j=1}^N A_{ij} T(a_2)) = s \ln(4m) \text{ and } \delta = \frac{s+1 + \sqrt{2s+1}}{s}, \text{ then}$$

$$\Pr\left[\sum_{j=1}^N \sum_{l=1}^2 A_{ij} X_{jl} \leq b_i - \epsilon\right] \leq \frac{1}{4} \forall i \in \{1, \dots, m\}, \quad (22)$$

where $m \leq n^2 k$ is the number of lines.

Proof: The line-sum (AXT) constraints in (15) can be restated (considering only lower bound) for each line i as

$$b_i - \epsilon \leq A_{i1} T(a_1) x_{1a_1} + A_{i1} T(a_2) x_{1a_2} + \dots + A_{iN} T(a_1) x_{Na_1} + A_{iN} T(a_2) x_{Na_2},$$

with N entries (voxels) in row i of A . We assume wlog that $A_{ij} T(l) \in [0, 1]$ (Note: while the input may not always satisfy this condition, we can enforce this by normalization). Replacing x_{ja_2} by $1 - x_{ja_1}$ for all $j \in \{1, \dots, N\}$, we have

$$b_i - \epsilon - \underbrace{\sum_{j=1}^N A_{ij} T(a_2)}_{\bar{b}_i} \leq \underbrace{A_{i1}(T(a_1) - T(a_2))}_{\bar{A}_{i1}} x_{1a_1} + \dots + \underbrace{A_{iN}(T(a_1) - T(a_2))}_{\bar{A}_{iN}} x_{Na_1},$$

which gives

$$\bar{b}_i \leq \sum_{j=1}^N \bar{A}_{ij} x_{ja_1} \quad \forall i \in \{1, \dots, m\}. \quad (23)$$

This also implies that the fractional solution to the LP (*i.e.*, x^*) obeys the previous condition, since it is a feasible solution to the LP. Therefore, $\forall i \in \{1, \dots, m\}$

$$\bar{b}_i \leq \sum_{j=1}^N \bar{A}_{ij} x_{ja_1}^*. \quad (24)$$

First, we compute the expectation of the summation term, for all i

$$\mathbb{E}\left(\sum_{j=1}^N \bar{A}_{ij} X_{ja_1}\right) = \sum_{j=1}^N \bar{A}_{ij} \mathbb{E}(X_j) = \sum_{j=1}^N \bar{A}_{ij} \delta x_{ja_1}^* \geq \delta \bar{b}_i \quad (25)$$

Now, we must choose an appropriate value of β such that we can bound the probability of

“undesirable events” away from 1. Setting $\beta = 1 - \frac{1}{\delta}$,

$$\begin{aligned} \Pr\left[\sum_{j=1}^N \bar{A}_{ij} X_{ja_1} \leq \bar{b}_i\right] &= \Pr\left[\sum_{j=1}^N \bar{A}_{ij} X_{ja_1} \leq (1 - \beta)\delta \bar{b}_i\right] \\ &\leq \exp\left(-\frac{\beta^2 \delta \bar{b}_i}{2}\right) \\ &= \exp\left(-\frac{(1-\delta)^2 \bar{b}_i}{2\delta}\right) \\ &\leq \frac{1}{4m} \end{aligned} \quad (26)$$

where (26) follows due to Lemma 2. Therefore, the probability of *any* AXT lower bound violation is

$$\Pr[AXT > b] \leq m \left(\frac{1}{4m}\right) = \frac{1}{4}. \quad (27)$$

For the upper bound case, the worst case violation is at most 2 by invoking Lemma 1.

3.2.2 Approximation: In order to prove the desired approximation ratio, we emulate the proof of Kleinberg and Tardos [20] for the Uniform Metric Labeling problem (ULP) that did not include line-sum constraints. Using the expected value of our random variables as a modification, we restate the lemmas. We refer the reader to [20] for their proofs. In the

following, an edge $e = (p, q)$ being ‘separated’ in a rounding iteration implies that p and q were unassigned before the iteration and one of p and q was assigned to a label during the iteration.

- The probability that an unassigned voxel, p , is assigned a label a_1 in a given iteration is $\frac{x_{pa_1}^* \delta}{2}$. and that an unassigned voxel, p , is assigned a label a_2 is $\frac{1 + \delta + x_{pa_2}^* \delta}{2}$
- The probability that an unassigned voxel, p , is assigned any label in a given phase is $\frac{1}{2}$.
- Over all phases, the probability that an unassigned voxel, p , is assigned label a_1 is $x_{pa_1}^* \delta$ and label a_2 is $1 - \delta + x_{pa_2}^* \delta$.
- For an edge $e = (p, q)$, the probability that e is ‘separated’ by an iteration is $\frac{2\delta z_e}{2} = \sum_{l \in T} |x_{pl}^* \delta - x_{ql}^* \delta|$ for any $l \in \{a_1, a_2\}$.
- The probability that p and q have different labels (over all iterations) is $2\delta z_e$.

Theorem 5: If OPT is the optimal LP solution to (14)–(15) and W_{exp} is the solution obtained from the randomized rounding procedure, then with high probability

$$W_{exp} \leq 2\delta OPT. \quad (28)$$

By Markov’s Inequality, the probability that the obtained solution is larger than 2δ times the optimal is

$$\Pr[W_{exp} \geq 2\delta OPT] \leq \frac{1}{2\delta} \quad (29)$$

Finally, we summarize the results of our randomized rounding approach.

Corollary 2: Let X be the 0/1 integral solution obtained by rounding a fractional solution x^* . The probability that X does not satisfy

$$W_{exp} \leq 2\delta OPT, \quad b - \epsilon \leq AX \leq 2(b + \epsilon)$$

is at most $\frac{1}{2\delta} + \frac{1}{4}$, where δ is as in Lemma 3.

3.2.3 Arbitrary Label case—The results in the above sections relied on n independent 0/1 random variables whose expected value was obtained as a function of the LP solution, x^* . For the two label case, for a voxel p , we have $\Pr[X_p=1]=f(x_{pa_1}^*)$ and $\Pr[X_p=0]=f(x_{pa_2}^*)$ for label indices a_1, a_2 , where $f(\cdot)$ is an appropriate function (in the previous case, $f(\cdot)$ was $\delta x_{pa_1}^*$). However, for the arbitrary label case, this conversion no longer works since X_p takes one among $|T|$ values with some probability. While probabilistic rounding similar to the two-label case works in practice as we discuss in §4, we do not derive concentration bounds (equivalent to the two-label case) in this paper.

4 Experimental Results

The implementation of the algorithm was done in C++ using CImg [34] and CPLEX (as a linear program solver) and was run on a 1.7GHz computer with 2GB RAM. In this section, we present performance evaluations of the algorithm on simulations, phantom data, as well as animal acquisitions. We will provide a few examples from each experimental category.

Our primary objective was to analyze if streak artifacts and blurring, the distinguishing characteristics of limited view reconstruction, can be sufficiently mitigated. Our second objective was to determine how a “denoised” image compares with one obtained via reconstruction with complete 360-view (i.e., fully sampled) information.

4.1 Head Phantom simulations

We used a modified Shepp-Logan head phantom image for the first set of 2D simulations. The distribution was projected onto 360 views (1° apart). The projections were then used for a simultaneous ART reconstruction to serve as ground-truth data (see Fig. 4). We emulated the limited view case by picking $k = \{10, 15, 20\}$ equally spaced projections from the complete rotational set. Again, a reconstruction was performed for each case. We can see in Fig. 5 (top row), as the number of views (k) decreases, the quality of the reconstructed image progressively becomes worse and the effect of streak and blurring artifacts becomes more pronounced. The task of our algorithm was then to recover the original image – given the reconstruction V , the limited projection views corresponding to V (i.e., line-sum constraints in §2.2), and the number of labels desired ($|T|$) in the final solution.

As a preprocessing step, we used a k -means clustering algorithm on the set of intensity values in V to find $|T| = 3$ disjoint clusters and their corresponding centers. The centers served as the labels in T . The grid-graph for V had 128×128 vertices, we used four-neighborhood adjacency to introduce edges between neighboring pixels. The edge weights were assigned using the RBF kernel. Having initialized the required parameters, the algorithm set up the linear program matrix, solved the model and obtained a solution vector (i.e., x^*). Then, we probabilistically round the values to 0 or 1 to obtain an integral solution. We observed that in all cases the fractional solution was good (see Fig. 6), we will discuss this in more detail shortly. The results (V') are shown in Fig. 5 (bottom row) for limited view cases corresponding to $k = \{10, 15, 20\}$. In all three cases, the algorithm is able to suppress the streak artifacts and recognize the regions of interest from the background. Observe that in the limited view reconstruction corresponding to $k = 10$ (top-row), the black ellipses are hardly noticeable due to blurring, our algorithm is still able to distinguish it from the background (bottom-row). As can be expected, when we move from $k = 10$ to $k = 20$, we notice fewer outliers in the background and the quality of the images determined by our algorithm improve.

Using the fully sampled (360-view) reconstruction, we calculated the number of mislabeled pixels, this was done by manually segmenting the ground truth and establishing cluster-to-cluster correspondence between V and V' . The misclassification for $k = \{10, 15, 20\}$ was $\{9\%, 7\%, 5\%\}$, with the majority of such pixels lying on the object boundary. We saw similar error values when repeating the experiments for other sets of projection data. The running time of the algorithm was ~ 30 seconds. Finally, as we mentioned above, the quality of the solutions was very good w.r.t. the integrality gap. In Fig. 6, we show the distribution of the entries of x^* , the values are highly concentrated towards 0 and 1 (only 0.6% of the entries of x^* are in $[0.1, 0.9]$), note that the y-axis is on a logarithmic scale. This indicates that the loss in optimality due to rounding was quite small in practice.

We also repeated the head phantom experiments with 4 labels and 20 projection views. For this purpose, the bright big ellipse in the interior was assigned a different contrast value

compared to its smaller circles. The original image is shown in Figure 7 (a), the solution from our algorithm is shown in 7 (b). The solution is only slightly worse than the 3 label case with a misclassification rate of 7%. We also ran an implementation of the graph-cuts framework [21], the same label values used in our algorithm were used as input to the graph-cuts implementation. Since the graph cut formulation does not naturally allow side constraints (i.e., line-sum constraints), the eventual segmentation is based primarily on the input. As a result, some of the artifacts are segmented as valid foreground regions, see Fig. 7 (c). The misclassification error was 17%. In our experiments, we tried increasing the smoothness values gradually but that did not significantly reduce the artifacts, instead merged the smaller circles. In summary, the line-sum constraints are useful for enforcing consistency with the projection data during reconstruction/segmentation.

4.2 Vessel Phantom reconstruction

Our second set of evaluations focused on limited-view reconstructions of an aluminum vessel phantom. The acquisition was performed on a Toshiba Infinix imaging system at 90 kVp with a 9" field of view, see Fig. 8. A copper filter was used to avoid additional beam hardening artifacts. A simultaneous ART reconstruction of the phantom was performed using the complete rotational acquisition to serve as ground-truth data. The resolution was $64^2 \times 64$. To emulate the limited view case, we picked $k = 20$ equally spaced projections and performed a reconstruction. Similar to §4.1, we calculated the label values, T , using k -means clustering, and then determined the spatial and grouping cues which were assigned as the edge weights and the data terms in the model. Note that since $|T| = 2$, we have a special constrained case of two-way graph cut (and not multi-way cut); two-way graph cuts (without line-sum constraints) can be solved optimally in polynomial time.

We illustrate the results for the vessel reconstruction in Fig. 9. The first column (a) show 2D slices corresponding to a complete rotational reconstruction. The second column (b) show the same 2D slice for the limited view case, we can clearly see that the image exhibits streak and blurring artifacts. What makes the task challenging is that in several regions the objects of interest have similar contrast as the 'structured' artifacts. Column (c) shows the solution obtained by our algorithm, the streak and blurring artifacts have been eliminated and regions of interest have been identified. We repeated such experiments for many different sets of projection images but that did not have an effect on the solution quality.

In Fig. 10, we show the volume visualization results from two views for the 360-view reconstruction and V'_{20} determined by our algorithm. The illustrations show that the two reconstructions are comparable though the full-view reconstruction has "cleaner" boundaries. To evaluate the accuracy of our limited view reconstruction, we calculated the misclassification error between V' and the 360-view reconstruction after manual cropping and threshold segmentation. These results are encouraging – the percentage of misclassified voxels were less than 1% (i.e., background pixels labeled as foreground as well as foreground pixels labeled as background), the vessel regions correspond to $\sim 1\%$ of the volume. Since the input images were blurred, the errors were primarily in the first category. The running time of the algorithm was ~ 3 minutes.

4.3 Aneurysm Phantom

Our evaluations in this section focus on reconstruction of an aneurysm phantom⁸ from a limited number of views. Several of the image acquisition parameters and preprocessing were similar to what was discussed in §4.2 in addition to injecting iodine as a contrast material. Some of the images from the sequence are shown in Fig. 11.

⁸An aneurysm is a bulge of the vessel walls due to disease. The analysis of vessel/aneurysm topologies is important in stroke research.

We illustrate our results for 20-view reconstruction in Fig. 12. Again, the first column (a) show two 2D slices corresponding to a complete rotational reconstruction. The second column (b) show the same 2D slice for the limited view case which is severely affected by artifacts. The reader will notice that the blurring artifacts here are slightly worse than those shown in §4.2. Column (c) shows the solution obtained by our algorithm, the streak and blurring artifacts have been eliminated and we are able to reliably determine the regions of interest, that agrees well with the complete reconstruction. The three sets of slices, one in each row, correspond to the choice of slicing plane (sagittal, axial, and coronal). The running time of the technique was ~ 4 minutes. The resolution was $64^2 \times 64$. We also note that choosing different sets of projection images had minimal impact on the results overall.

In Fig. 13, we show the volume visualization results from two views for the 360-view reconstruction and V'_{20} determined by our algorithm. Similar to §4.2, the two reconstructions are comparable though the full-view reconstruction looks better qualitatively. The misclassification error between V' and the 360-view reconstruction was less than 2%, mostly around the boundaries and the thinner (2 – 3 voxels) vessels. Because of blurring, the misclassifications around the boundaries (background labeled as foreground) comprised most of the errors, these may decrease upon a careful choice of edge-weights. However, extremely thin vessels will still be difficult as they were almost completely blurred out in the 20-view reconstruction.

4.4 Rabbit Head images

The images in this section are taken from an animal study, the volume comprised of air, soft-tissue and bone, the three distinct contrast materials corresponding to these materials were calculated. We omit the discussion of other clinical aspects of the image acquisition. A reconstruction corresponding to 20 views was used as an input to our algorithm. The resolution was $64^2 \times 64$. We first illustrate the results using representative 2D slices. In Figure 14, we show three slices (sagittal, coronal, and axial) from the full reconstruction, limited view reconstruction, and our algorithm. As in the previous results, the limited view reconstruction has artifacts that are addressed by our technique. The results match well with those from a full reconstruction.

An informative visualization of a three-contrast volume is difficult to show, hence, we show some images of the visualization in two contrast mode in Fig. 15 (a)–(b), and three-contrast mode in Fig. 15 (c)–(d) (where the third contrast i.e., soft-tissue, is set to be ~80% transparent). We see that the results for a complete reconstruction in Fig. 15 (b), (d) are comparable to the one obtained from our algorithm Fig. 15 (a), (c). We do not report on misclassification errors in this case because we were unable to obtain ground-truth data. As in the experiments reported in the previous sections, we noticed that choosing different sets of projection images had little impact on the results overall.

4.5 Dog Head images

The images in this section correspond to another animal study. Similar to the rabbit acquisition, the volume comprised of air, soft-tissue and bone, the three distinct contrast materials corresponding to these materials were calculated. A reconstruction corresponding to 30 views was used as an input to our algorithm. The resolution was $64^2 \times 64$.

We show the limited view reconstruction of the head section using the projection images. Fig. 16 (a)–(b), and three-contrast mode in Fig. 16 (c)–(d) (where the third contrast i.e., soft-tissue, is set to be ~80% transparent). Similar to §4.4, we see that the results for a complete reconstruction in Fig. 16 (b), (d) are comparable to the one obtained from our algorithm Fig. 16 (a), (c).

Finally, we note that for all these data-sets the loss due to rounding was quite small. In general, the distribution of entries in the solution vector for all datasets was similar to the plot shown in Fig. 6. A large percentage of entries were concentrated around 0 or 1 with fewer than 1% of the entries in x^* having values in the $[0.1, 0.9]$ range. A summary of misclassification errors is provided below.

5 Conclusions and Future Work

In this paper, we have proposed an optimization framework based on a constrained version of metric labeling to provide reliable CT reconstruction and segmentation from only a limited number of projection views. The algorithm allows simultaneous reconstruction and segmentation of multiple contrast values and brings ideas from algebraic reconstruction together with combinatorial optimization approaches for energy minimization in a unified model. We report on promising numerical results on simulations, as well as phantom images and animal studies. The algorithm is easy to implement and runs in 3–4 minutes on $64^2 \times 64$ volume and ~ 30 secs for $2D$ images. An interesting direction of future research will be to see if purely combinatorial algorithms can be designed for the problem. While recent work has explored the question of which form of energy functions in vision can be minimized in a graph-cuts framework [35,36], it will be very useful if such frameworks can also be extended to tackle additional constraints that arise from “side knowledge”. The application of randomized rounding for energy minimization in vision is also a relatively unexplored area, which should be investigated in future work.

Acknowledgments

This work was supported in part by NSF CCF-0546509, IIS-0713489, NIH grant HL 52567 and Toshiba Medical Systems Corporation. The first author was also supported by the Department of Biostatistics and Medical Informatics, University of Wisconsin-Madison and UW Institute for Clinical and Translational Research. The authors thank the reviewers for suggestions that led to improvements in the organization and presentation of this paper.

6 Appendix

6.1 Derivation of (16)

$$\begin{aligned}
 \sum_{p=1}^N A_{ip} T(a_1) X_{pa_1} &= \sum_{p=1}^N A_{ip} T(a_1) (x_{pa_1}^* + x_{pa_2}^*) \\
 &= \sum_{p=1}^N (A_{ip} T(a_1) x_{pa_1}^* + A_{ip} T(a_2) x_{pa_2}^* + A_{ip} (T(a_1) - T(a_2)) x_{pa_2}^*) \\
 &= \sum_{p=1}^N (A_{ip} T(a_1) x_{pa_1}^* + A_{ip} T(a_2) x_{pa_2}^* + \frac{y-1}{y+1} (y+1) A_{ip} T(a_2) x_{pa_2}^*) \\
 &= \sum_{p=1}^N (A_{ip} T(a_1) x_{pa_1}^* + A_{ip} T(a_2) x_{pa_2}^* \\
 &\quad + \frac{y-1}{y+1} (A_{ip} T(a_1) x_{pa_2}^* + A_{ip} T(a_2) x_{pa_2}^*)) \\
 &\leq \sum_{p=1}^N (A_{ip} T(a_1) x_{pa_1}^* + A_{ip} T(a_2) x_{pa_2}^*)
 \end{aligned} \tag{30}$$

$$+\frac{y-1}{y+1}(A_{ip}T(a_1)x_{pa_1}^*+A_{ip}T(a_2)x_{pa_2}^*) \quad (31)$$

$$=(1+\frac{y-1}{y+1})\sum_{p=1}^N(A_{ip}T(a_1)x_{pa_1}^*+A_{ip}T(a_2)x_{pa_2}^*) \quad (32)$$

$$\leq(1+\frac{y-1}{y+1})(b_i+\epsilon) \quad (33)$$

$$\leq 2(b_i+\epsilon) \quad (34)$$

where the first inequality is valid because $x_{pa_1}^* \geq x_{pa_2}^*$.

References

1. Jiang M, Wang G. Convergence of the simultaneous algebraic reconstruction technique (sart). IEEE Transactions on Image Processing 2003;12(8)2
2. Zhang Y, Chan H-P, Sahiner B, Wei J, Goodsitt MM, Hadjiiski LM, Ge J, Zhou C. A comparative study of limited-angle cone-beam reconstruction methods for breast tomosynthesis. Medical Physics 2006;33(10):3781–3795. [PubMed: 17089843]2, 3
3. Rantala M, Vänskä S, Järvenpä S, Kalke M, Lassas M, Moberg J, Siltanen S. Wavelet-based reconstruction for limited angle x-ray tomography. IEEE Transactions on Medical Imaging 2006;25(2):210–217. [PubMed: 16468455]2
4. Kolehmainen V, Vanne A, Siltanen S, Järvenpää S, Kaipio JP, Lassas M, Kalke M. Parallelized bayesian inversion for three-dimensional dental x-ray imaging. IEEE Transactions on Medical Imaging 2006;25(2):218–228. [PubMed: 16468456]2
5. Singh, V.; Dinu, PM.; Mukherjee, L.; Xu, J.; Hoffmann, KR. Limited view ct reconstruction via constrained metric labeling; Proc. of IEEE Computer Society Workshop on Mathematical Methods in Biomedical Image Analysis; 2007. 3
6. Nassi M, Brody WR, Medoff BP, Macovski A. Iterative reconstruction–reprojection: an algorithm for limited data cardiac-computed tomography. IEEE Transactions on Biomedical Engineering 1982;29:333–341. [PubMed: 7084960]3
7. Kim JH, Kwak KY, Park SB, Cho ZH. Projection space iteration reconstruction-projection. IEEE Transactions on Medical Imaging 1985;4:139–143. [PubMed: 18243964]3
8. Anderson AH. Algebraic reconstruction in CT from limited views. IEEE Transactions on Medical Imaging 1989;8(1):50–55. [PubMed: 18230499]3
9. Velikina, J.; Leng, S.; Chen, G-H. Limited view angle tomographic image reconstruction via total variation minimization; Proc. SPIE; 2007. 4
10. Fessler, J. PhD thesis. Stanford University; 1990. Object-based 3-D reconstruction of arterial trees from a few projections. 4
11. Rangarajan, A. Tutorial on the EM algorithm. 1998. <http://www.cise.ufl.edu/~anand/pdf/mixem.pdf>. 4
12. Fishburn P, Schwander P, Shepp L, Vanderbei RJ. The discrete radon transform and its approximate inversion via linear programming. Discrete Applied Mathematics 1997;75(1):39–61.4

13. Gritzmann P, de Vries S, Wiegelmann M. Approximating binary images from discrete x-rays. *SIAM Journal on Optimization* 2000;11(2):522–546.4
14. Weber S, Schüle T, Schnörr C, Hornegger J. A linear programming approach to limited angle 3d reconstruction from dsa projections. *Methods of Information in Medicine* 2004;43(4):320–326. [PubMed: 15472741]4
15. Schule T, Schnörr C, Weber S, Hornegger J. Discrete tomography by convex-concave regularization and d.c. programming. *Discrete Applied Mathematics* 2005;151(1–3):229–243.4
16. Pruessmann KP, Weiger M, Scheidegger M, Boesiger P. SENSE: sensitivity encoding for fast mri. *Magnetic Resonance in Medicine* 1999;42:952–962. [PubMed: 10542355]4
17. Raj A, Singh G, Zabih R, Kressler B, Wang Y, Schuff N, Weiner M. Bayesian parallel imaging with edge-preserving priors. *Magnetic Resonance in Medicine* 2007;57:8–21. [PubMed: 17195165]4
18. Lustig M, Donoho DL, Pauly JM. Sparse MRI: The application of compressed sensing for rapid mr imaging. *Magnetic Resonance in Medicine* 2007;58(6):1182–1195. [PubMed: 17969013]4
19. Raj, A. PhD thesis. Cornell University; 2005 May. Improvements in Magnetic Resonance Imaging using Information Redundancy. 4
20. Kleinberg J, Tardos E. Approximation algorithms for classification problems with pairwise relationships: metric labeling and markov random fields. *Journal of the ACM* 2002;49(5):616–639.5, 9, 17
21. Boykov Y, Veksler O, Zabih R. Fast approximate energy minimization via graph cuts. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 2001;23(11):1222–1239.5, 7, 20
22. Hochbaum DS. An efficient algorithm for image segmentation, markov random fields and related problems. *Journal of the ACM* 2001;48(2):686–701.5
23. Li, SZ. *Markov Random Field Modeling in Computer Vision*. Springer-Verlag; 1995. 5
24. Hoffmann KR, Nazareth DP, Miskolczi L, Gopal A, Wang Z, Rudin S, Bednarek DR. Vessel size measurements in angiograms: A comparison of techniques. *Medical Physics* 2002;29(7):1622–1633. [PubMed: 12148745]7
25. Ishikawa, H.; Geiger, D. Segmentation by grouping junctions; *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*; 1998. p. 125-131.7
26. Kleinberg, J.; Tardos, E. *Algorithm Design*. Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc.; 2005. chapter 8. 11
27. Goemans MX, Williamson David P. New 34-approximation algorithms for the maximum satisfiability problem. *SIAM Journal on Discrete Mathematics* 1994;7(4):656–666.14
28. Kolliopoulos SG, Young NE. Approximation algorithms for covering/packing integer programs. *Journal of Computer and System Sciences* 2005;71(4):495–505.14
29. Srivastav, A.; Stangier, P. Integer multicommodity flows with reduced demands; *Proc. of European Symposium on Algorithms*; 1993. p. 360-371.14
30. Bertsimas D, Vohra R. Rounding algorithms for covering problems. *Mathematical Programming* 1998;80(1):63–89.14
31. Mitzenmacher, M.; Upfal, E. *Probability and Computing: Randomized Algorithms and Probabilistic Analysis*. New York, NY, USA: Cambridge University Press; 2005. 14
32. Angluin D, Valiant LG. Fast probabilistic algorithms for hamiltonian circuits and matchings. *Journal of Computer and System Sciences* 1979;18(2):155–193.14
33. Fohlin, H.; Kliemann, L.; Srivastav, A. Technical Report 0611. Institut für Informatik, Christian-Albrechts-Universität zu Kiel; 2006. Randomized algorithms for mixed matching and covering in hypergraphs in 3d seed reconstruction in brachytherapy. 14
34. Tschumperle, D. The C++ Template Image Processing Library (CImg). <http://cimg.sourceforge.net>. 18
35. Kolmogorov V, Zabih R. What energy functions can be minimized via graph cuts? *IEEE Transactions on Pattern Analysis and Machine Intelligence* 2004;26(2):147–159. [PubMed: 15376891]28
36. Freedman D, Drineas P. Energy minimization via graph cuts: settling what is possible. *IEEE Conference on Computer Vision and Pattern Recognition* 2005:939–946.28

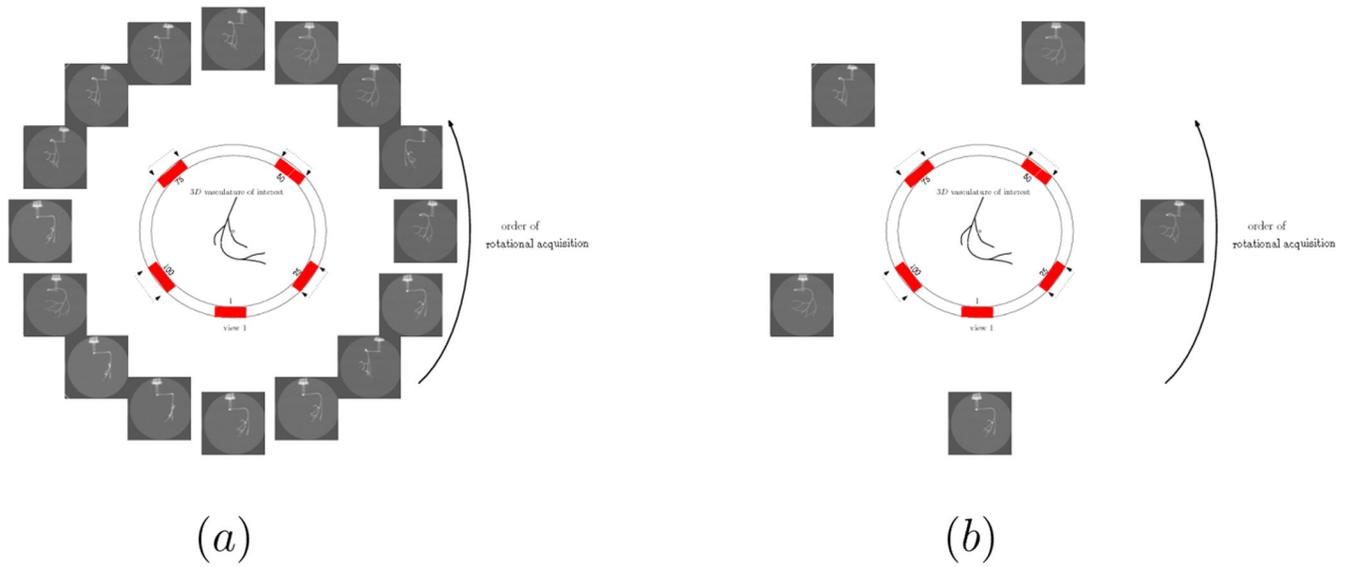


Fig. 1. (a) A complete rotational (i.e., fully sampled) acquisition of a vessel phantom. (b) In a respiratory cycle, the ‘limited views’ correspond to the vessel structure in the same respiratory phase (position and shape) when the patient is breathing normally during acquisition. Typically the imaging system rotates around the vasculature during acquisition, and the same phase images are determined using cardiac gating (e.g., with ECG).

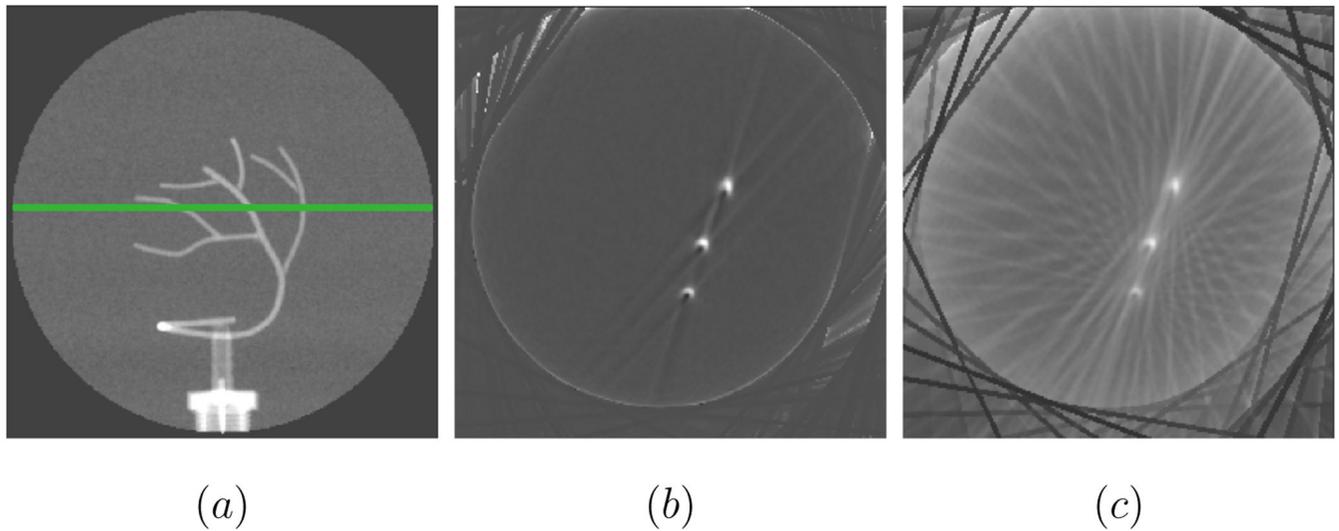


Fig. 2. (a) A single projection image of a vessel phantom, phantoms are physical devices that may have properties similar to a specific human anatomy and are used in radiation related procedures to evaluate performance. (b) A $2D$ cross sectional slice (perpendicular to the view direction and the horizontal green line in (a)) of the reconstruction volume using 360 views at 1° separation. (c) The same $2D$ cross sectional slice as in (b) but the reconstruction was performed using 15 views only.

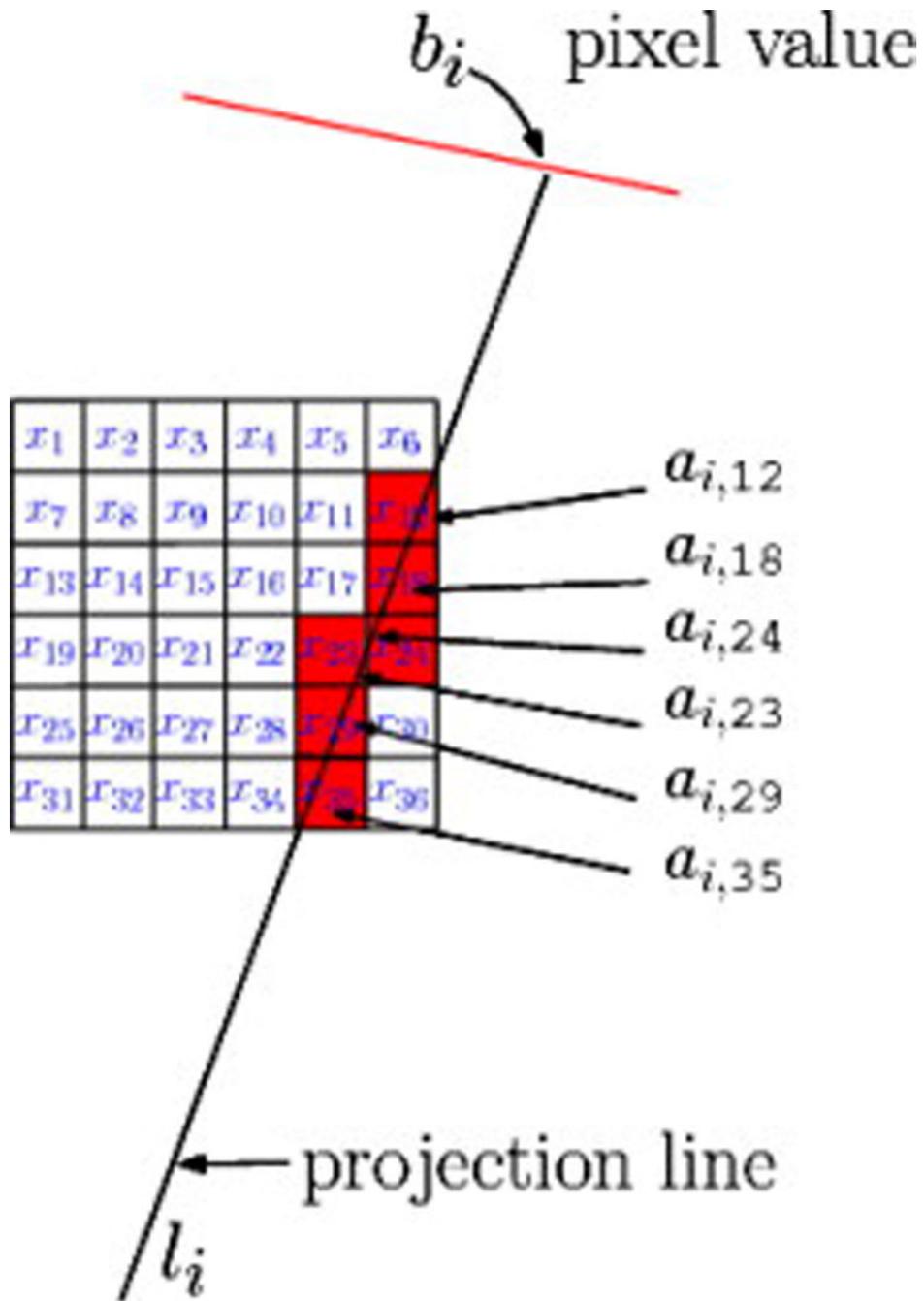
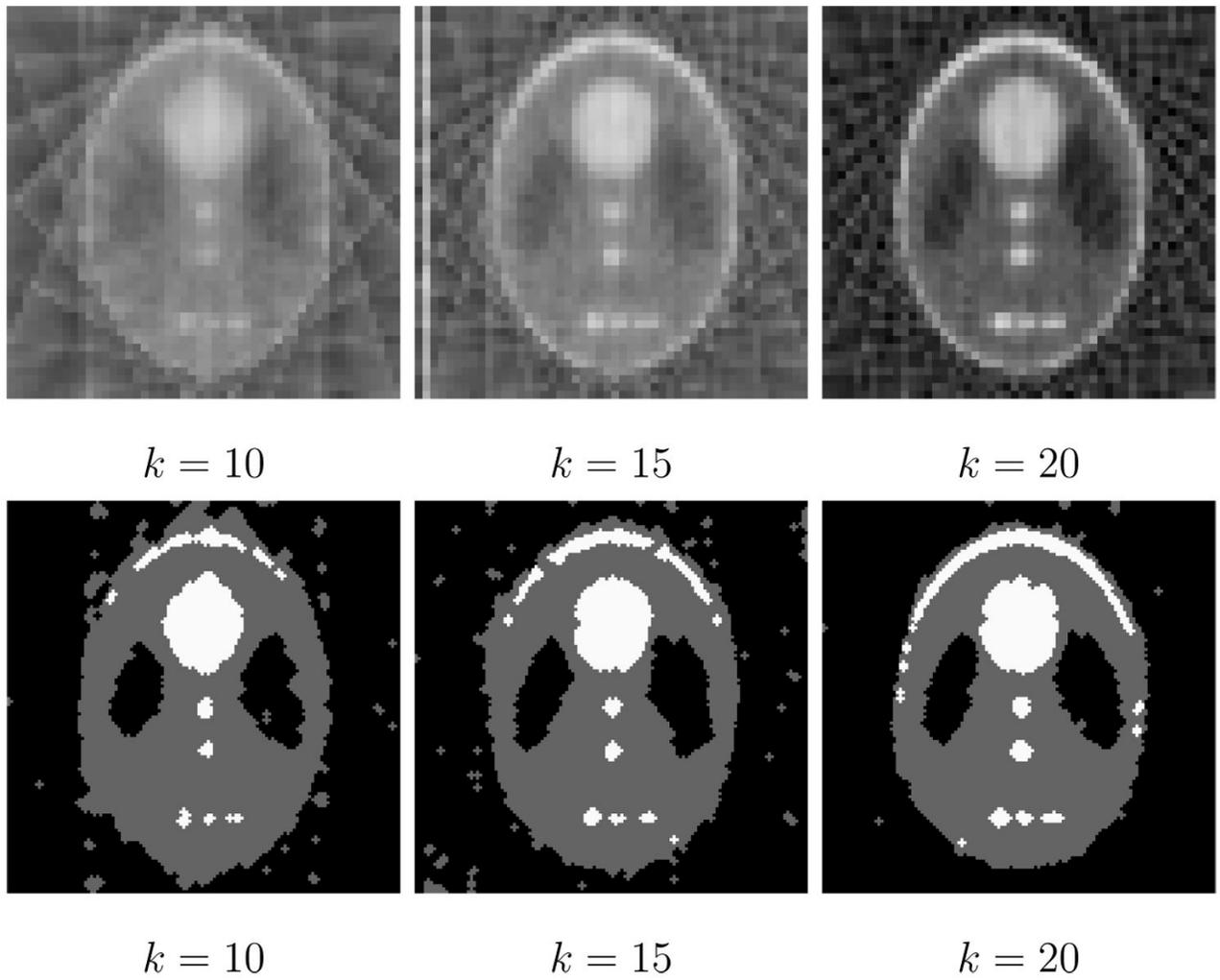


Fig. 3. The case for a single projection line, the unknown distribution is $x = [x_i]$ and a_{ij} denotes the coefficients (distance traveled by line i through voxel j). The relationship is then given by $a_{35}x_{35} + a_{29}x_{29} + a_{23}x_{23} + a_{24}x_{24} + a_{18}x_{18} + a_{12}x_{12} = b_i$.



Fig. 4.
Shepp-Logan reconstruction using 360 views.

**Fig. 5.**

The top row corresponds to limited view reconstructions $(V_{10}, V_{15}$ and $V_{20})$ for $k = \{10, 15, 20\}$ views respectively. The second row corresponds to the solutions $(V'_{10}, V'_{15}$ and $V'_{20})$ obtained from our algorithm.

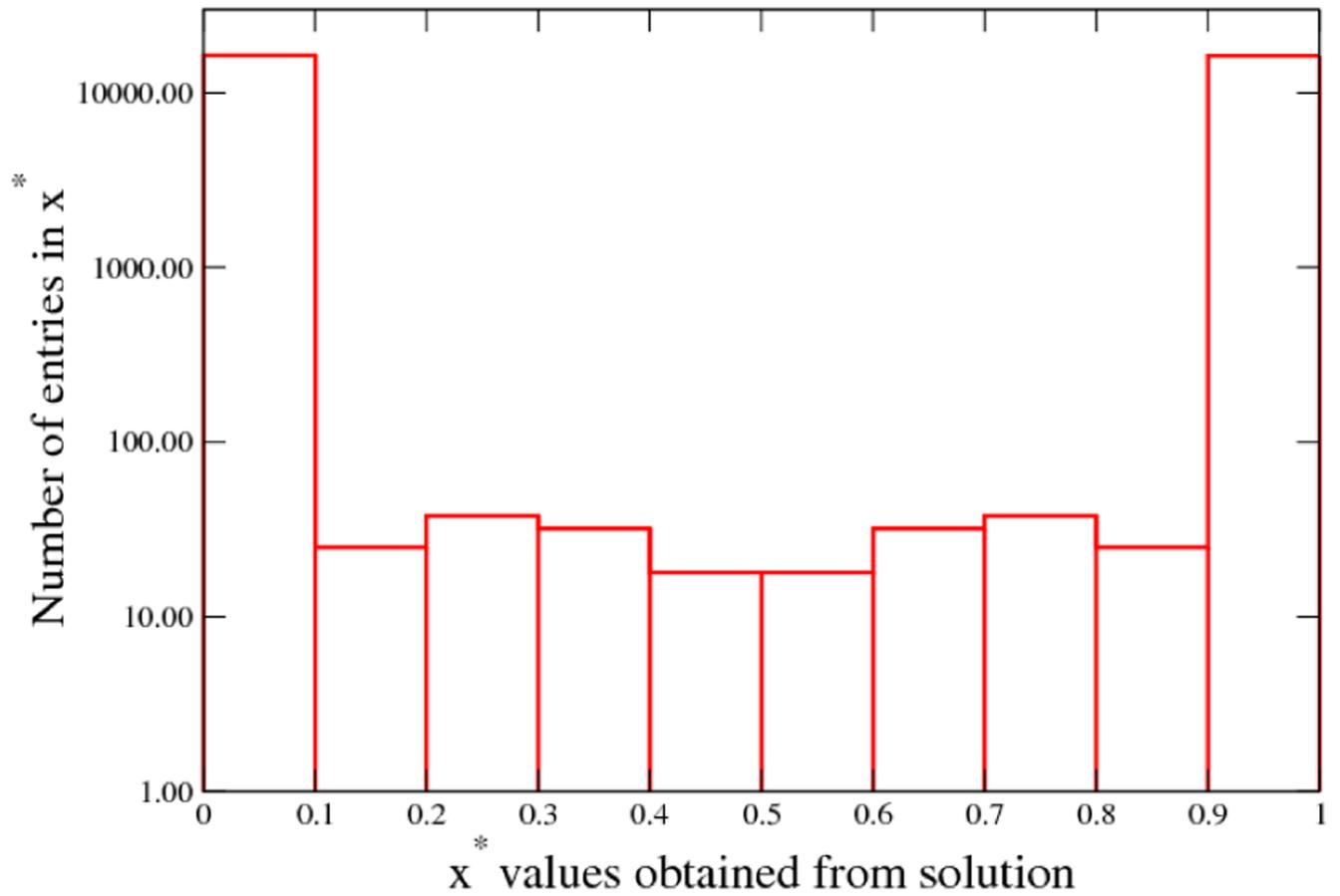


Fig. 6. A histogram of the values of x^* for $k = 20$, y-axis is logarithmic scale.

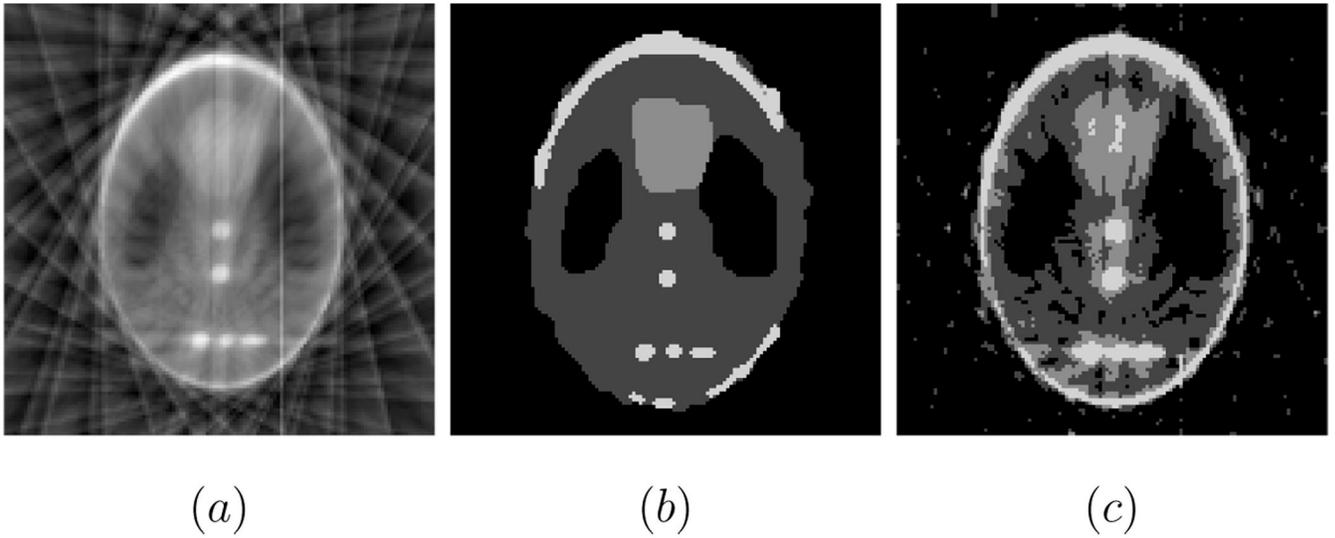


Fig. 7. Results for (a) original image after reconstruction for $k = 20$, (b) segmentation/reconstruction using our algorithm and (c) segmentation without line-sum constraints (i.e., Graph Cuts).

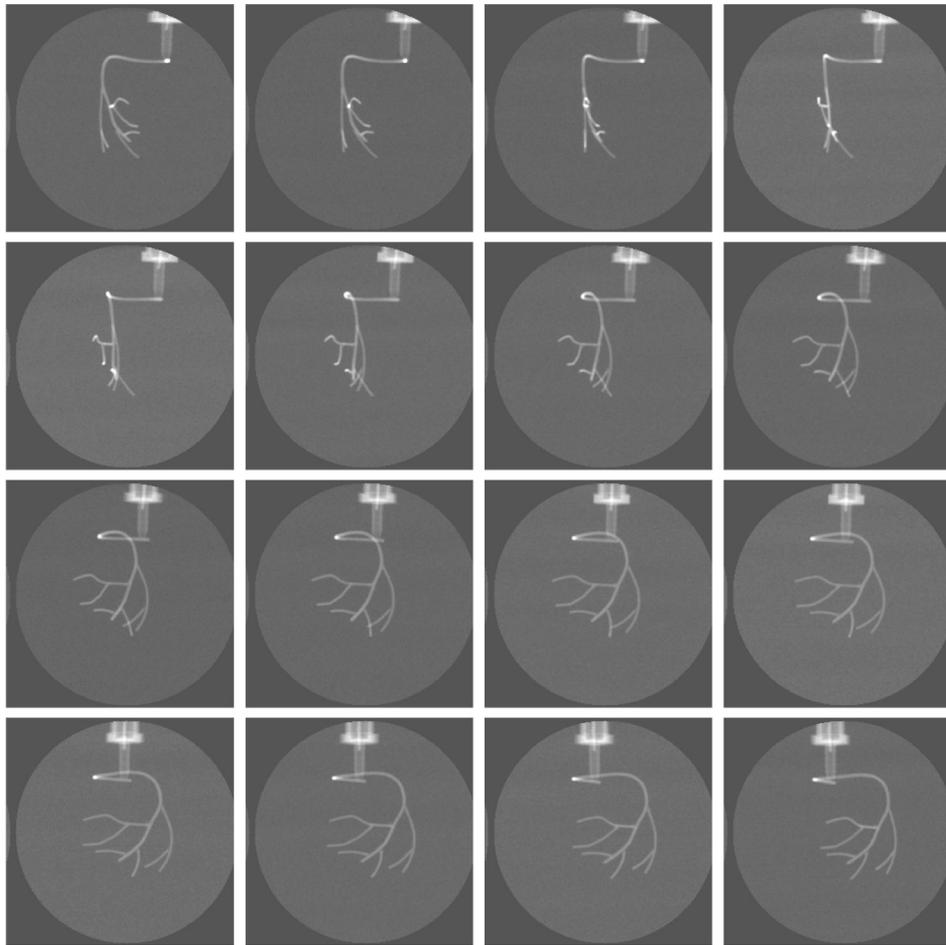
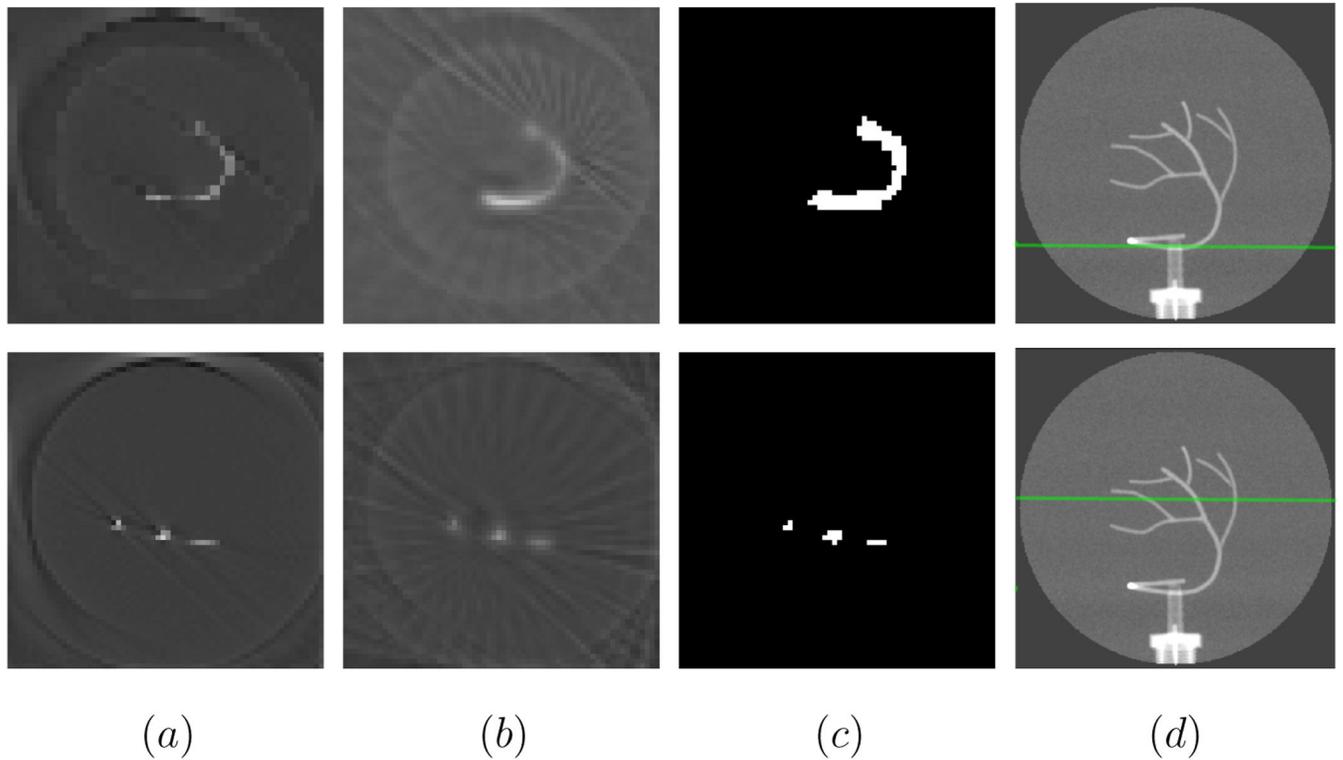


Fig. 8.
Rotational image sequence of vessel phantom.

**Fig. 9.**

Two $2D$ slices for a full rotational reconstruction in column (a), corresponding $2D$ slices taken from a limited view reconstruction with $k = 20$ in (b), the results from our algorithm for the $2D$ slices in (c). The slices in this figure correspond to a plane orthogonal to the imaging plane in (d), the slice intersects it at the horizontal green line shown in (d). This is true for slices shown in other figures also.

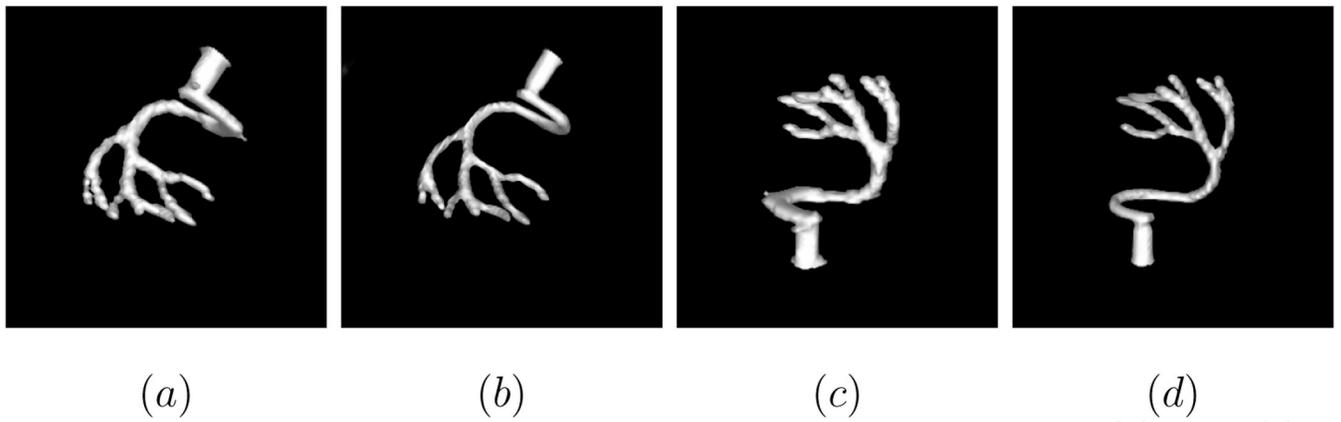


Fig. 10. Volume visualization of vessel phantom reconstruction, (a) and (c) correspond to V'_{20} from two different views, (b) and (d) correspond to the 360-view reconstruction.

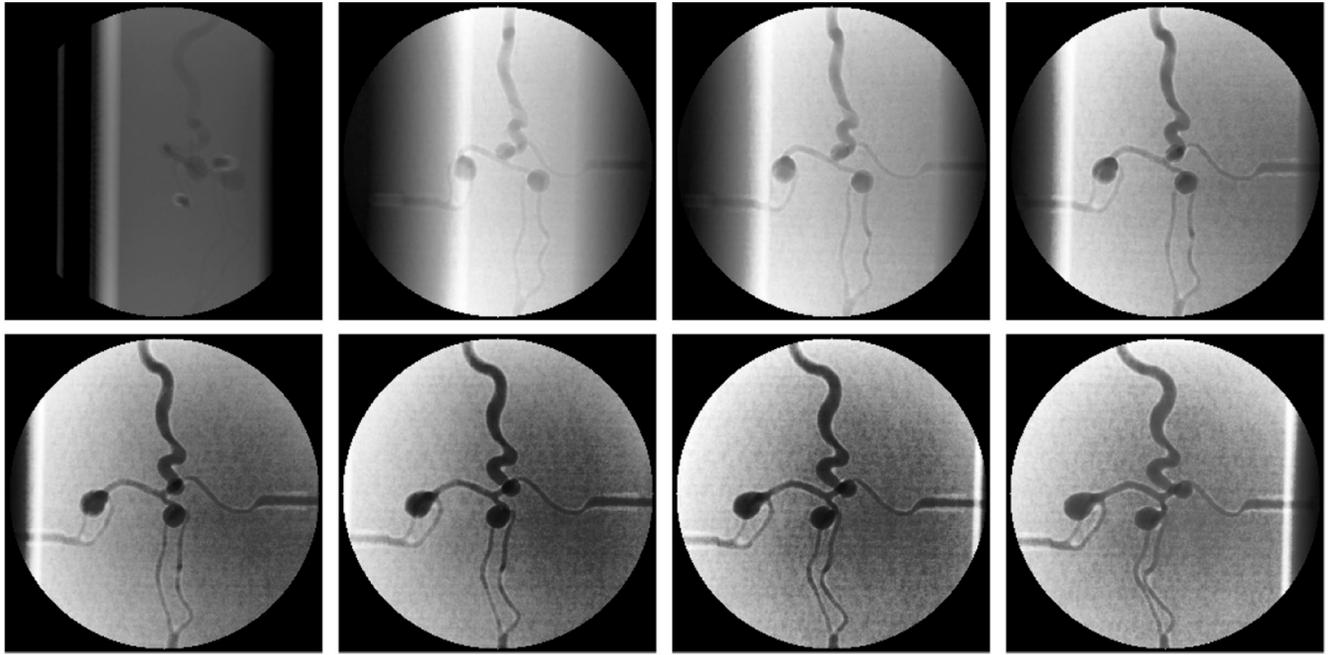


Fig. 11.
Rotational image sequence of an aneurysm phantom.

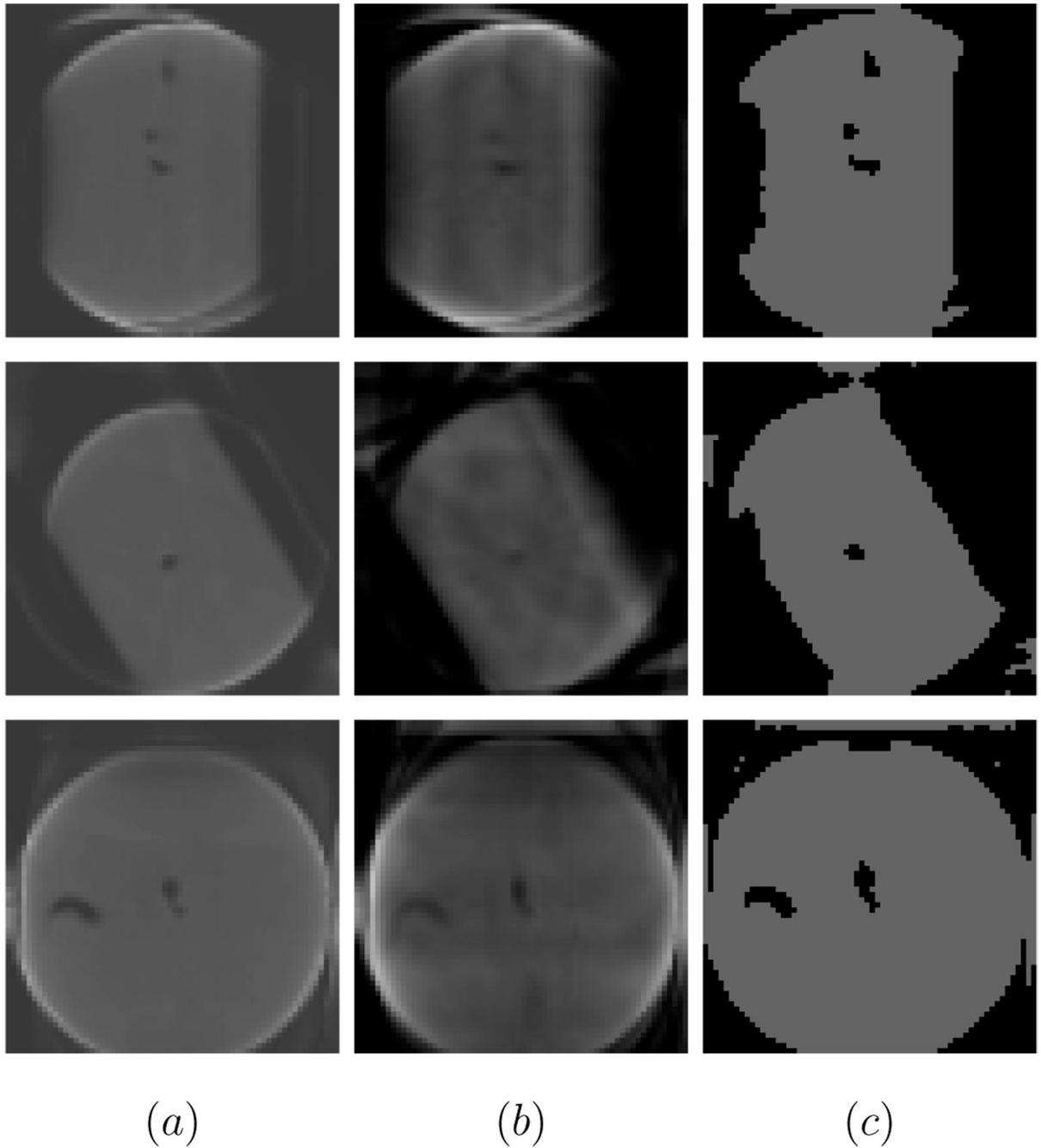


Fig. 12. Sagittal, Coronal, and axial slices of the vessel aneurysm volume. Slices corresponding to the complete reconstruction in column one, the limited view (20 view) reconstruction in column two, and for our reassignment algorithm with line-sum constraints in column three.

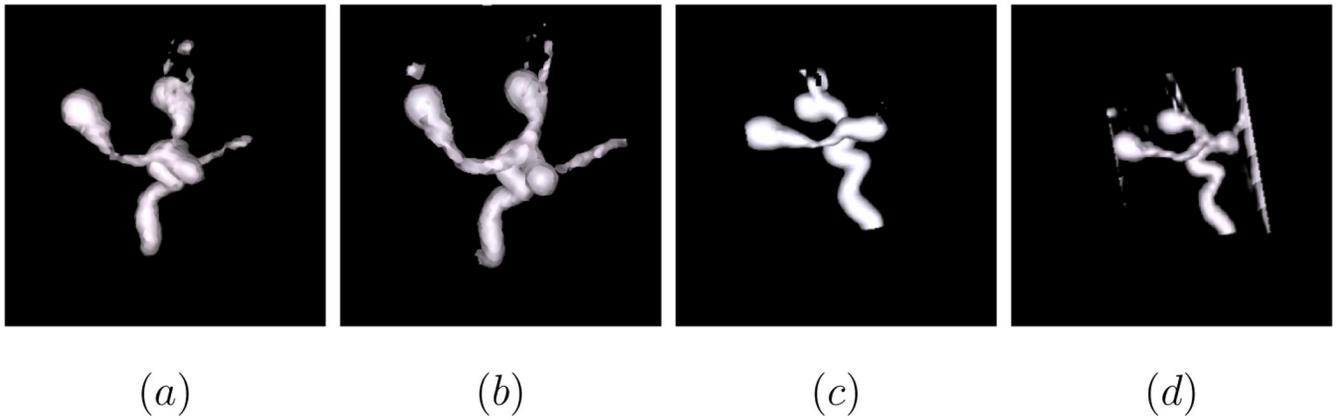


Fig. 13. Volume visualization of aneurysm reconstruction, (a) and (c) correspond to V'_{20} from two different views, (b) and (d) correspond to the 360-view reconstruction of the vessel aneurysm.

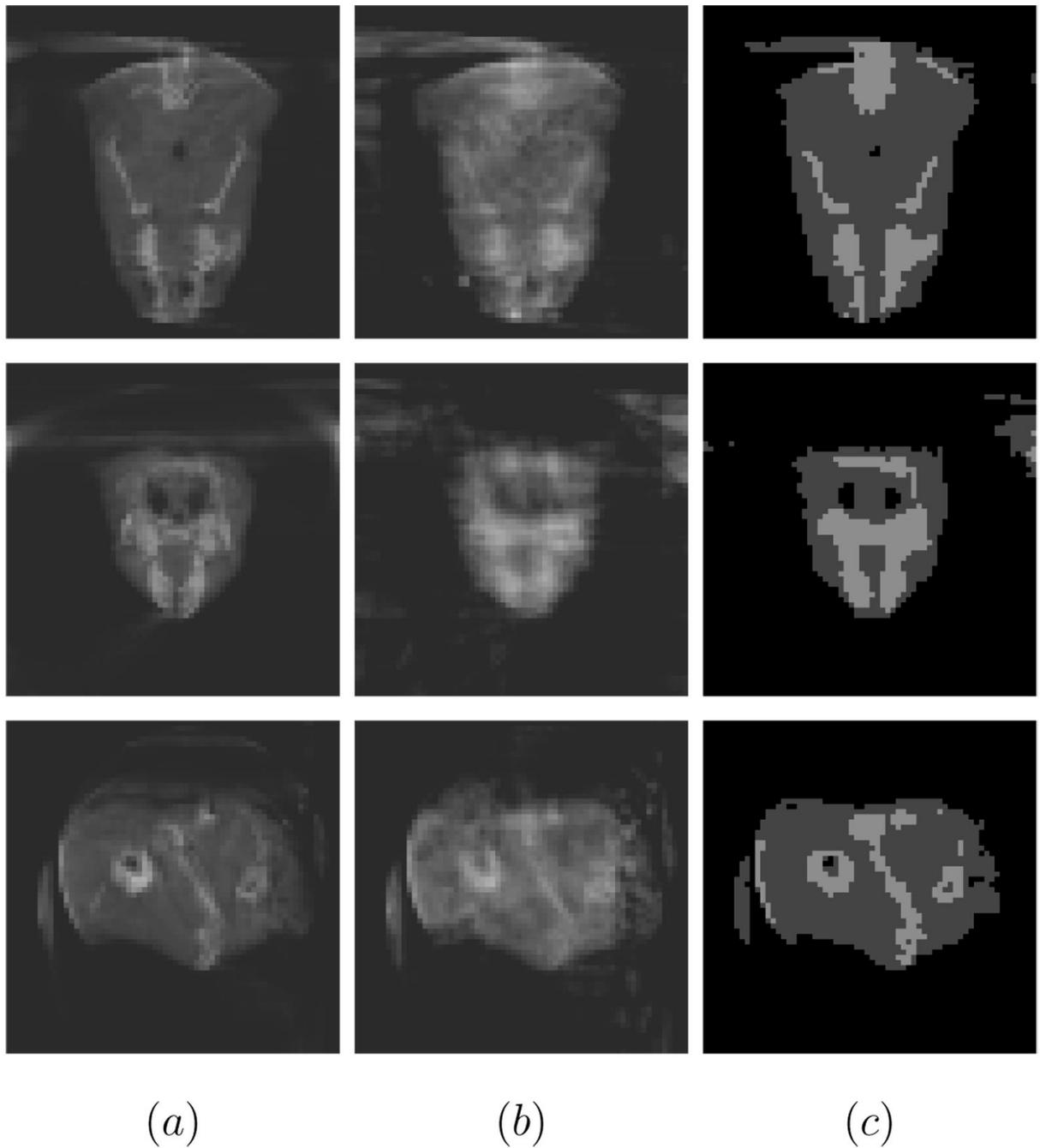


Fig. 14. Sagittal, coronal, and axial slices of the rabbit head volume. Slices corresponding to the complete reconstruction in column (a), the limited view (20 view) reconstruction in column (b), and for our reassignment algorithm with line-sum constraints in column (c).

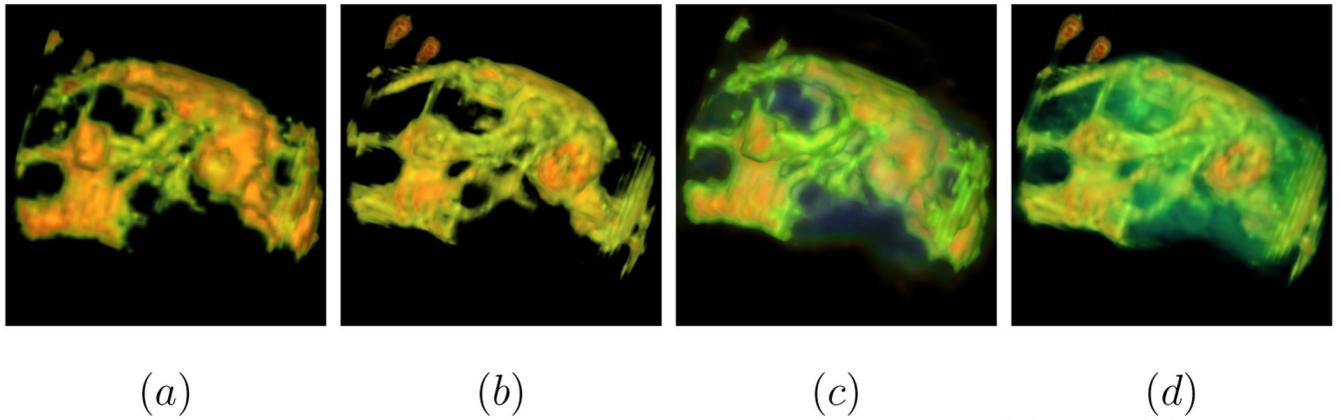
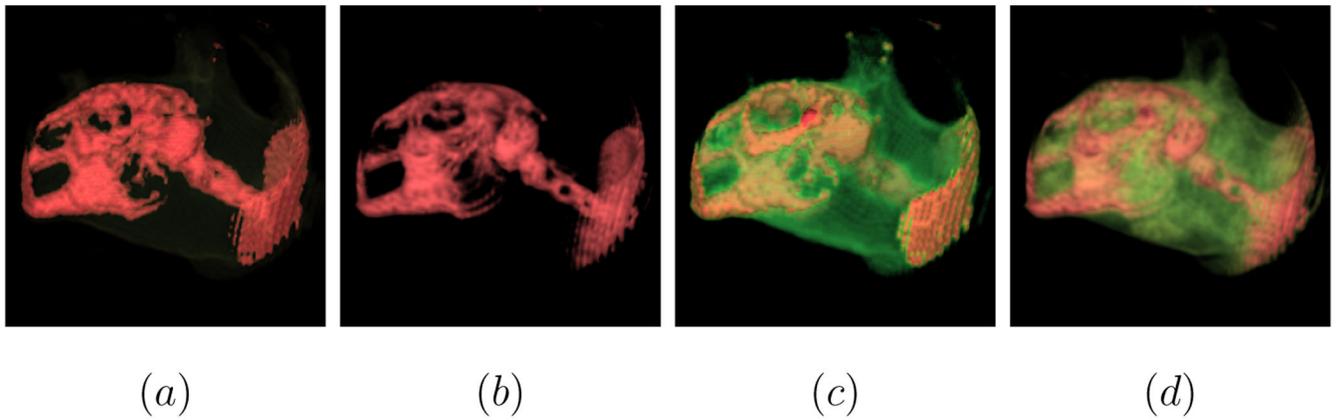


Fig. 15.

Volume visualization of rabbit head reconstruction, (a) and (c) correspond to V'_{20} , (b) and (d) correspond to the 360-view reconstruction of the rabbit skull reconstruction. Note that (a)–(b) correspond to only two contrast values (bone and air) whereas in (c)–(d), the third contrast value (soft-tissue) is set to be $\sim 80\%$ transparent.

**Fig. 16.**

Volume visualization of dog head reconstruction, (a) and (c) correspond to V'_{30} , (b) and (d) correspond to the 360-view reconstruction of the dog head reconstruction. Note that (a)–(b) correspond to only two contrast values (bone and air) whereas in (c)–(d) we can also see the soft-tissue (partly transparent).

Table 1
Summarization of the misclassification errors on each data set

Data Set	views (k)	contrast values	misclassification
head phantom	10	3	9%
head phantom	15	3	7%
head phantom	20	3	5%
head phantom	20	4	7%
vessel phantom	20	2	1%
aneurysm phantom	20	2	2%