Distribution Agreement

In presenting this thesis as a partial fulfillment of the requirements for a degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis in whole or in part in all forms of media, now or hereafter now, including display on the World Wide Web. I understand that I may select some access restrictions as part of the online submission of this thesis. I retain all ownership rights to the copyright of the thesis. I also retain the right to use in future works (such as articles or books) all or part of this thesis.

Ziyi Yin April 9, 2019

Edge Detection and Enriched Subspaces

by

Ziyi Yin

James Nagy

Adviser

Department of Mathematics

Victoria Powers

Committee Member

Yuanzhe Xi

Committee Member

Bree Ettinger

Committee Member

Davide Fossati

Committee Member

2019

Edge Detection and Enriched Subspaces

by

Ziyi Yin

James Nagy

Adviser

Abstract of

a thesis submitted to the Faculty of Emory College of Arts and Sciences

of Emory University in partial fulfillment

of the requirements of the degree of

Bachelor of Sciences with Honors

Department of Mathematics

2019

Abstract

Edge Detection and Enriched Subspaces

By Ziyi Yin

We describe an image reconstruction algorithm that reconstructs the image using approximate image basis obtained from the image segmentation algorithm. Unlike similar approaches which reconstruct the image directly, our algorithm highly improves the accuracy of reconstruction in most cases especially with only a limited range of X-ray projection angles. We first use the image segmentation method to detect the edge of the object, then use the segmentation result as the approximate image basis to reconstruct the image. Moreover, our algorithm can be applied to different kinds of objects, such as a single object, separated objects and overlay objects, and the accuracy of the reconstruction image is always improved.

Edge Detection and Enriched Subspaces

by

Ziyi Yin

James Nagy

Adviser

A thesis submitted to the Faculty of Emory College of Arts and Sciences

of Emory University in partial fulfillment

of the requirements of the degree of

Bachelors of Science with Honors

Department of Mathematics

2019

Acknowledgements

First, I would like to thank my adviser, Dr. James Nagy, for offering me this important opportunity to work with him and introducing me to the field of image sciences, and generally, computational mathematics. I understand that, as the chair of Department of Mathematics, he is very busy dealing with a variety of things in the department everyday and traveling around the world to attend conferences. However, he still guides me throughout the whole process with every effort and gives me many enlightening ideas for my work. Moreover, he also helps me make my determination to pursue an advanced degree in the field after graduating from college. I would also like to thank Drs. Victoria Powers, Yuanzhe Xi, Bree Ettinger and Davide Fossati for taking time to attend my honor thesis defense as committee members. Next, I would like to thank Emory honors program to support me along the way. Finally, I would thank my parents for supporting my studies here at Emory University and letting me choose my career path afterwards.

Contents

List of Figures

Chapter 1

Introduction

Image processing is a scientific method to convert an image into digital form on which operations are performed. The field of image technology is important in many real-life applications such as military [4] [6], computerized photography [7] [17], medical/biological imaging [14] [15] and also other industrial applications [1][2].

In this thesis, two of the most significant fields of image processing are to be mentioned, which are image reconstruction and image segmentation. The image reconstruction problem, which is often the most important part, requires solving an inverse problem,

$$\boldsymbol{Ax} = \boldsymbol{b}$$

where $\boldsymbol{A}$ is the projection matrix related to the models how the X-ray

beams travel through the object; $\boldsymbol{b}$ is the projection data which represents the energy loss of the X-ray beams while they pass through the object, which is also called *sinogram*; $\boldsymbol{x}$ is the information about the shape of the object.

In real applications, there are many difficulties of reconstructing the image from the projection data. First, measuring the projection data $\boldsymbol{b}$ accurately is always very difficult, which means that $\boldsymbol{b}$ of the system is usually perturbed with some inevitable error $\boldsymbol{\epsilon}$. Also, the range of projection angles may not be a full range of angles (from 0 to 180), but a limited range of angles (e.g. from 0 to 45). As a result, even though the perturbation on $\boldsymbol{b}$ is comparatively small, the error of the solution $\boldsymbol{x}$ from the inverse problem may be very high due to the ill-conditioned property of $\boldsymbol{A}$. Therefore, the aim of image processing algorithms is to solve the inverse problem in order to obtain the object information from the projection data, with high accuracy and efficiency.

To understand more about the image processing problem, more background material about inverse problems is mentioned in the following Chapter 2. The image reconstruction process is discussed in Chapter 3, and the image segmentation process is discussed in Chapter 4. The idea of enriched Krylov subspace methods is discussed in Chapter 5, and how segmentation can be used to construct a good set of basis vectors for the enrichment process. Numerical experiments are given in Chapter 6, and concluding

remarks in Chapter 7.

Chapter 2

Background

To solve the inverse problem in Chapter 1

$$\boldsymbol{Ax = b}$$

We first need to understand the computing difficulties when attempting to solve inverse problems with some numerical linear algebra concepts. In a numerical point of view, we could also consider solving the problem as

$$\min_{\boldsymbol{x}} \{||\boldsymbol{Ax - b}||_{\boldsymbol{2}}\}$$

2.1 Computational Difficulties

A naive approach to solve $\boldsymbol{x}$ in this inverse problem, is to multiply the $\boldsymbol{A}^{-1}$ on both sides, so that a naive solution $\boldsymbol{x}$ is yielded by

$$\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b}$$

However, there are a lot of disadvantages of doing this:
1. To obtain the inverse of matrix $\boldsymbol{A}$ is very computationally expensive. In many cases, $\boldsymbol{A}$ has a certain sparse structure which can be exploited for numerical computing, but $\boldsymbol{A}^{-1}$ will be very dense. Therefore, explicitly computing the whole matrix of $\boldsymbol{A}^{-1}$ and doing matrix multiplications with $\boldsymbol{A}^{-1}$ becomes very expensive.
2. Moreover, the solution from $\boldsymbol{A}^{-1}$ may not be numerically accurate for ill-posed problems even though $\boldsymbol{A}^{-1}$ is computed accurately. This is because $\boldsymbol{A}$ is highly ill-conditioned.

To understand how the ill-conditioning of $\boldsymbol{A}$ can produce inaccurate solutions, consider the singular value decomposition (SVD) of $\boldsymbol{A}$,

$$\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^{\boldsymbol{T}} = \sum_{i=1}^{n} \sigma_i \boldsymbol{u_i} \boldsymbol{v_i^T}$$

where $\boldsymbol{A}$ is an $m \times n$ matrix, $\boldsymbol{U}$ is an $m \times m$ matrix, $\boldsymbol{\Sigma}$ is an $m \times n$ matrix and $\boldsymbol{V}$ is an $n \times n$ matrix. Also, $\boldsymbol{U}$ and $\boldsymbol{V}$ are both orthogonal matrices, which means that $\boldsymbol{U^T U} = \boldsymbol{U U^T} = \boldsymbol{I_{m\times m}}$ and $\boldsymbol{V^T V} = \boldsymbol{V V^T} = \boldsymbol{I_{n\times n}}$.

And $\boldsymbol{\Sigma}$ is a diagonal matrix with diagonal entries as $\sigma_1, \sigma_2, ..., \sigma_n$, which are called singular values and satisfy

$$\sigma_1 \geq \sigma_2 \geq ... \geq \sigma_n \geq 0$$

In the SVD sense, an ill-conditioned matrix $\boldsymbol{A}$ indicates that

1. The singular values σ_i decay very fast and cluster at 0.
2. The right singular vectors $\boldsymbol{v_i}$ oscillate way more intensely for small singular values [3][11].

Suppose that we compute $\boldsymbol{U, \Sigma, V}$ accurately for matrix $\boldsymbol{A}$, then a true solution $\boldsymbol{x}$ can be yielded by

$$\boldsymbol{x_{true}} = \boldsymbol{A^{-1}b} = \boldsymbol{(U\Sigma V^T)^{-1}b} = \sum_{i=1}^{n} \frac{\boldsymbol{u_i^T b}}{\sigma_i} \boldsymbol{v_i}$$

However, if there is a small perturbation $\boldsymbol{\epsilon}$ in our projection data vector $\boldsymbol{b}$ which changes $\boldsymbol{b}$ to $\tilde{\boldsymbol{b}} = \boldsymbol{b} + \boldsymbol{\epsilon}$, then the solution $\boldsymbol{x}$ will be also perturbed to $\tilde{\boldsymbol{x}}$, where

$$\tilde{\boldsymbol{x}} = \boldsymbol{A^{-1}(b+\epsilon)} = \boldsymbol{(U\Sigma V^T)^{-1}(b+\epsilon)} = \sum_{i=1}^{n} \frac{\boldsymbol{u_i^T(b+\epsilon)}}{\sigma_i} \boldsymbol{v_i}$$

$$= \sum_{i=1}^{n} \frac{\boldsymbol{u_i^T b}}{\sigma_i} \boldsymbol{v_i} + \sum_{i=1}^{n} \frac{\boldsymbol{u_i^T \epsilon}}{\sigma_i} \boldsymbol{v_i} = \boldsymbol{x_{true} + error}$$

which yields the error term as

$$\boldsymbol{error} = \sum_{i=1}^{n} \frac{\boldsymbol{u_i^T \epsilon}}{\sigma_i} \boldsymbol{v_i}$$

Here, even assuming the perturbation $\boldsymbol{\epsilon}$ is very small with respect to vector $\boldsymbol{b}$, since the small singular values σ_i are almost zero when i is large, and the reciprocals of these multiply extremely oscillating vectors $\boldsymbol{v_i}$, the error term could be terribly large.

Due to the disadvantages of the naive approach of computing $\boldsymbol{A^{-1}b}$, it is rarely used in numerical inverse problems. More methods including matrix decomposition or iterative approaches are derived to avoid the numerical instability, which are used more widely.

2.2 Regularization Methods

There are many remedies to avoid the ill-conditioned property of matrix $\boldsymbol{A}$. The basic idea is to use a filter vector $\boldsymbol{\phi}$ specifically for the terms with small singular values in order to decrease the error term.
The general solution form is

$$\boldsymbol{x_{sol}} = \sum_{i=1}^{n} \phi_i \frac{\boldsymbol{u_i^T \tilde{b}}}{\sigma_i} \boldsymbol{v_i}$$

2.2.1 Truncated SVD

In the truncated singular value decomposition (TSVD) method, the regularization is implemented by replacing inversion of the small singular values in $\boldsymbol{A}$ by 0. Therefore, denoting the user-defined tolerance as τ, the TSVD solution is given by

$$\boldsymbol{x_{TSVD}} = \sum_{i=1}^{k} \frac{\boldsymbol{u_i^T \tilde{b}}}{\sigma_i} \boldsymbol{v_i}$$

where $\sigma_k \geq \tau > \sigma_{k+1}$.

The filter factors ϕ_i for the TSVD are

$$\phi_i = \begin{cases} 1 \text{ if } i \leq k \\ 0 \text{ if } i > k \end{cases}$$

In this way, we avoid the numerical instability of the terms with small singular values, which helps control the error term so that it does not grow too large.

However, there are some disadvantages of using the TSVD method.

1. Overall, computing the SVD can be very computationally expensive when matrix $\boldsymbol{A}$ is large. Even though it avoids oscillations at the end of the singular values, it is not an efficient algorithm compared to other approaches, such as iterative methods.

2. Since the filter vector is so sharp that it simply cuts off the last $n - k$

components, an important limitation for this method is where to truncate the SVD. Choosing a tolerance which is too big would end up with losing too many singular values that contain important information of matrix $\boldsymbol{A}$, and can also cause error. Choosing a tolerance which is too small would end up with some small singular values in the denominator of the filtered inverse solution and thus would still lead to a large error. Therefore, there are some other methods with smoother filters to avoid this kind of problem.

2.2.2 Tikhonov Regularization

For Tikhonov regularization, we shift the problem from solving

$$\min_{\boldsymbol{x}}\{||\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b}||_2\}$$

to

$$\min_{\boldsymbol{x}}\{||\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b}||_2^2+\lambda^2||\boldsymbol{x}||_2^2\}$$

which is the same as

$$\min_{\boldsymbol{x}}\left\{\left|\left|\begin{bmatrix}\boldsymbol{A}\\ \lambda\boldsymbol{I}\end{bmatrix}\boldsymbol{x}-\begin{bmatrix}\boldsymbol{b}\\ \boldsymbol{0}\end{bmatrix}\right|\right|_2^2\right\}$$

which is also the same as using filter factors

$$\phi_i=\frac{\sigma_i^2}{\sigma_i^2+\lambda^2}$$

where λ controls the "smoothness" of the regularized solution.

Given that the original problem to solve $\min_{\boldsymbol{x}}\{||\boldsymbol{Ax}-\boldsymbol{b}||_2\}$ is ill-conditioned, the basic idea of the Tikhonov method is to add a penalty term $\lambda^2||\boldsymbol{x}||_2^2$ to keep $||\boldsymbol{x}||_2$ small. We could also see this as a shift of the normal equations (see [3][11] for more information of normal equations) from

$$\boldsymbol{A^T Ax} = \boldsymbol{A^T b}$$

to

$$(\boldsymbol{A^T A} + \lambda^2 \boldsymbol{I})\boldsymbol{x} = \boldsymbol{A^T b}$$

where $\boldsymbol{A^T A}$ and $\boldsymbol{A^T A} + \lambda^2 \boldsymbol{I}$ are both $n \times n$ square matrices.

For the original problem, the eigenvalues of $\boldsymbol{A^T A}$ are the square of singular values of $\boldsymbol{A}$, as σ_i^2, which decay very fast to 0. However, for the regularized problem, the eigenvalues of $\boldsymbol{A^T A} + \lambda^2 \boldsymbol{I}$ are $\sigma_i^2 + \lambda^2$, which will be bounded away from 0 if $\lambda > 0$. Therefore, the Tikhonov regularized problem is more well-conditioned than the original problem, which means that it is more numerically stable.

Similarly to the TSVD method, choosing the correct λ for the Tikhonov method is also very important to solve the problem. Especially, when k is chosen such that $\sigma_k = \lambda$, the sharp filter TSVD can be seen as an approximation to the smooth filter of Tikhonov regularization.

2.2.3 Modified Tikhonov Regularization

In the classical Tikhonov regularization method, the aim is to keep $||\boldsymbol{x}||_2$ small. However, this may not always be desirable. Sometimes, we already have a temporary guess for the solution, denoted as $\boldsymbol{x_0}$ and we do not want the final solution too far from our guess $\boldsymbol{x_0}$. Then, we can change the penalty term $\lambda^2||\boldsymbol{x}||_2^2$ to $\lambda^2||\boldsymbol{x}-\boldsymbol{x_0}||_2^2$, which yields the method as

$$\min_{\boldsymbol{x}}\{||\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b}||_2^2+\lambda^2||\boldsymbol{x}-\boldsymbol{x_0}||_2^2\}$$

which is the same as solving the corresponding normal equation

$$(\boldsymbol{A^T}\boldsymbol{A}+\lambda^2\boldsymbol{I})\boldsymbol{x}=\boldsymbol{A^T}\boldsymbol{b}+\lambda^2\boldsymbol{x_0}$$

In some special cases, the guess for the solution is more complicated and there can also be other constraints on our solution $\boldsymbol{x}$. Then a matrix $\boldsymbol{L}$ is needed for the penalty term as $\lambda^2||\boldsymbol{L}\boldsymbol{x}||_2^2$, which yields the modified Tikhonov method as

$$\min_{\boldsymbol{x}}\{||\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b}||_2^2+\lambda^2||\boldsymbol{L}\boldsymbol{x}||_2^2\}$$

which is the same as

$$\min_{\boldsymbol{x}}\left\{\left|\left|\begin{bmatrix}\boldsymbol{A}\\ \lambda\boldsymbol{L}\end{bmatrix}\boldsymbol{x}-\begin{bmatrix}\boldsymbol{b}\\ \boldsymbol{0}\end{bmatrix}\right|\right|_2^2\right\}$$

If $\boldsymbol{L} = \boldsymbol{I}$, then this reduces to the classical Tikhonov regularization. This modified Tikhonov method is very important for the image reconstruction idea in Chapter 3 and Chapter 5.

We provide some background material on the basic image reconstruction problem in Chapter 3, and for image segmentation in Chapter 4. The idea of enriched Krylov subspace methods is discussed in Chapter 5, and how segmentation can be used to construct a good set of basis vectors for the enrichment process. Numerical experiments are given in Chapter 6, and concluding remarks in Chapter 7.

Chapter 3

Image Reconstruction

3.1 Math Modeling

In the field of image reconstruction, X-rays are projected through the object and then projection data is collected. Notice that the loss of X-ray energy follows the Beer-Lambert Law [20],

$$T = \frac{\phi_e^t}{\phi_e^i} = e^{-\tau}$$

and

$$\tau = \sum_{i=1}^{N} \tau_i = \sum_{i=1}^{N} \sigma_i \int_0^l \mu_i(z) dz$$

where

ϕ_e^t is the radiant flux transmitted by the material sample,

ϕ_e^i is the radiant flux received by the material sample,

τ is the optical depth of the material sample,

σ_i is the attenuation cross section of the attenuating species i in the material sample,

μ_i is the attenuation coefficient of species i in the material sample,

l is the path length of the X-ray beam through the material sample.

X-ray

↘ ↘ ↘ ↘ ↘ ↘ ↘ ↘

μ_1	μ_2	μ_3	μ_4
μ_5	μ_6	μ_7	μ_8
μ_9	μ_{10}	μ_{11}	μ_{12}
μ_{13}	μ_{14}	μ_{15}	μ_{16}

Therefore, assuming that in each pixel of the object, the attenuation coefficient is distributed evenly, and we use μ_i to denote the attenuation coefficient for the *i-th* pixel in the object, then considering one beam of X-ray of index k we obtain that

$$-\ln T_k = \sum_{i=1}^{N} \mu_i d_{ik} = (\boldsymbol{d_k}, \boldsymbol{\mu})$$

where d_{ik} denotes the length of the path through which the beam of light goes through pixel i, and the parentheses are used to denote the inner product.

By adding more X-rays in each projection angle and rotating the X-rays gradually, a set of projection data with a number of beams in different angles is collected.

Therefore, the image reconstruction problem can be modeled in a linear algebra sense.

$$\boldsymbol{Ax} = \boldsymbol{b}$$

where

$\boldsymbol{A}$ is projection matrix with entries of d_{ik}, which only contains information about the X-ray projection methods,

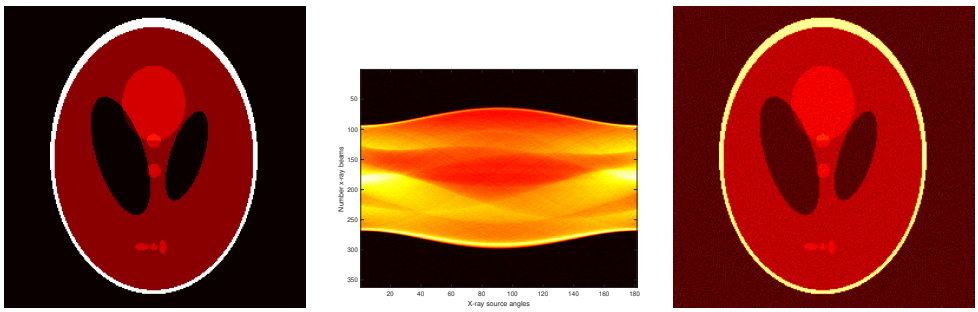
$\boldsymbol{b}$ is the projection data collected from multiple beams in different angles, which is called a *sinogram*,

$\boldsymbol{x}$ is the coefficient vector with entries of attenuation coefficient u_i, which contains the object information.

Therefore, to compute the object $\boldsymbol{x}$, an inverse problem needs to be solved. Also, due to the ill-conditioned property of the projection matrix $\boldsymbol{A}$, regularization methods mentioned in Chapter 2 need to be used to solve the problem more accurately. More image modeling background can be found in [12].

3.2 Reconstruction Example

Figure 3.1 shows an example for the image reconstruction process.
In this example, the true object is generated by `phantom` in MATLAB, along with the associated projection data for an x-ray projection problem, where projections were done at angles $0, 1, 2, \ldots, 180$ degrees. The reconstructed image was computed using a hybrid version of LSQR, called HyBR [8], which is an iterative method to solve the least squares problem with Tikhonov regularization mentioned in Chapter 2; see also [16, 10]. We can see that with full projection angles, the image reconstruction result given by the `IRhybrid_lsqr` method is a very good approximation to the true object.

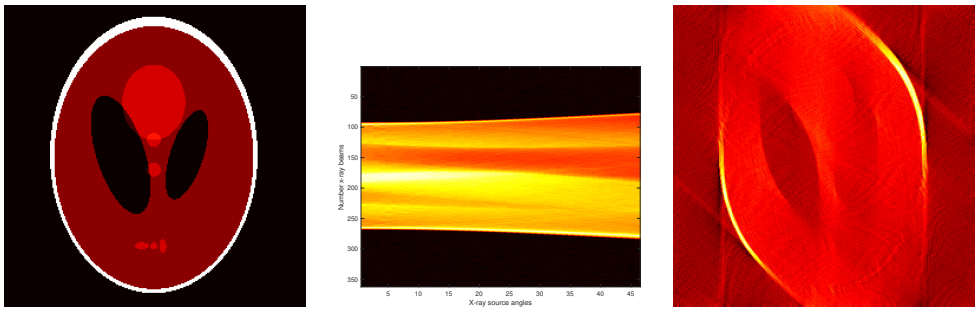


True image | Projection data | Reconstructed image

Figure 3.1: Example of an image reconstruction problem, where projection data is taken for angles $0, 1, 2, \ldots, 180$ degrees, and the `IRhybrid_lsqr` in [10] is used to compute the reconstructed image.

However, there are some cases in which the standard iterative approach does not work well. In the example of Figure 3.2, the projection is done by

a limited range of angles of $0, 1, 2, \ldots, 45$. Therefore, the width of the projection data decreases from 180 to 45. With the limited information, the reconstruction image is distorted a lot compared to the original object. More importantly, the reconstruction object in the middle is almost mixed with the background and it is somewhat hard to separate the object from the background in the upper left and lower right corners.



True image Projection data Reconstructed image

Figure 3.2: Example of an image reconstruction problem, where projection data is taken for angles $0, 1, 2, \ldots, 45$ degrees, and the `IRhybrid_lsqr` in [10] is used to compute the reconstructed image.

These two examples of image reconstructions indicate that even though Tikhonov regularization method is implemented to avoid the ill-positioned property of the projection matrix $\boldsymbol{A}$, due to the limited angle projection, the reconstructed image is still not good enough. Therefore, improving the reconstruction result with only limited projection angles becomes a very important issue. In the following chapters, an image segmentation method in Chapter 4 is used to detect the edge of the object and an enriched

subspace method in Chapter 5 is used as a crucial idea to include an image basis as constraints so that a more accurate image can be reconstructed.

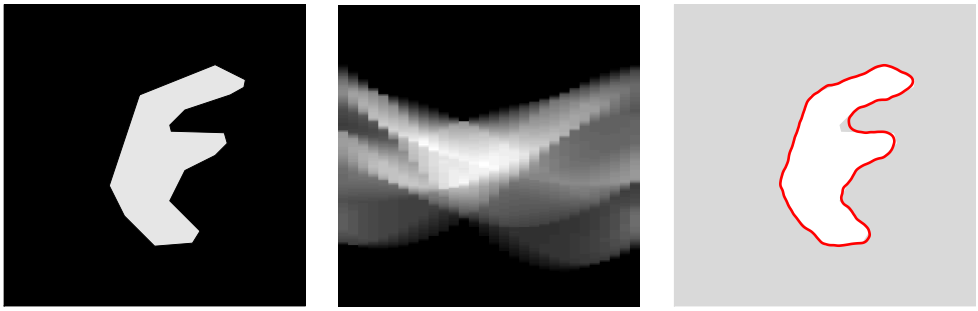
Chapter 4

Image Segmentation

Image segmentation is to consider the object in another way. Instead of seeing the object from the view of intensity values in each pixel, a parametric curve is used to represent the edge of the object. And then we consider the intensity is 1 inside the curve, 0 outside the curve. Therefore, a parametric curve corresponds to a particular object.

4.1 Segmentation Example

Figure 4.1 shows an example of image segmentation given by the method from [9]. The left image of Figure 4.1 shows the true image. Instead of marking intensity values inside and outside correspondingly, the object is described by 18 vertices in a clockwise order. The intensity of the region inside the curve is marked by 1 - white, and the intensity of the region



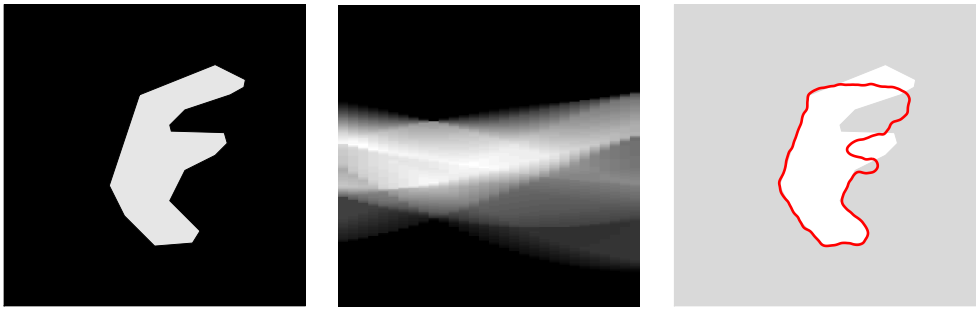
True image Projection data Image segmentation result

Figure 4.1: Example of an image segmentation problem, where projection data is taken for angles $0, 12, 24, \ldots, 180$ degrees, and the iterative method in [9] is used to compute image segmentation result.

outside the curve is marked by 0 - black. The image in the middle of Figure 4.1 shows the projection data of the object, with 15 columns corresponding to 15 projection angles. The right image of Figure 4.1 shows the result of image segmentation, as a red parametric curve circumscribing the object. Notice that the number of vertices may be different for the true curve and the result curve. There are 18 vertices on the original curve but there are 500 vertices on the result curve from the segmentation. Despite the different of number of vertices, the two objects described by the parametric curve are almost the same.

In this case, the projection range is from 0 to 180 degrees, in 12 degrees increment, which collects information all around the object. It is obvious to see that the parametric curve result is a very good approximation to the object, as it almost follows the edge of the object. However, the image

segmentation process may also be subject to limited angles sometimes, which is shown in Figure 4.2.



True image Projection data Image segmentation result

Figure 4.2: Example of an image segmentation problem, where projection data is taken for angles $0, 6, 12, \ldots, 90$ degrees, and the iterative method in [9] is used to compute image segmentation result.

Figure 4.2 shows an example of projecting the same object in a limited range of angles from $0, 6, 12, \ldots, 90$. In this case, even though the true object is the same as in the Figure 4.1, the segmentation result is not a good approximation compared to the full range of angles in Figure 4.1. Although the lower part of the curve almost follows the edge of the object, the right part of the curve is somewhat distorted and fails to follows the shape of the object, due to the limited range of projection angles.

In the following sections, the method in [9] will be discussed to show the importance of image segmentation before reconstructing the image.

4.2 Segmentation Method

In this section, an iterative method of image segmentation is given by [9].

Initialize curve $\boldsymbol{c}$ as a circle;

while *not converged* **do**

compute $\hat{p}_k j$ for $\boldsymbol{c}$ and $\hat{\mu} = 1$ - forward model;

compute μ for $\hat{p}_k j$ and s_{kj} - update attenuation coefficients;

update $\boldsymbol{c}$ - curve deformation;

end

Algorithm 1: Image Segmentation Method

Finally, the method aims to minimize the external energy, as

$$E_{ext}(\mu, \boldsymbol{c}) = \sum_{j=1}^{J} \sum_{k=1}^{K} (s_{kj} - \mu \hat{p}_{kj})^2$$

where

k denotes the projection angle,

j denotes x-ray in each projection angle,

μ denotes attenuation coefficient,

p_{kj} denotes predicted sinogram,

$\hat{p}_{kj}$ denotes projection of the curve c for attenuation $\mu = 1$,

s_{kj} denotes the true sinogram.

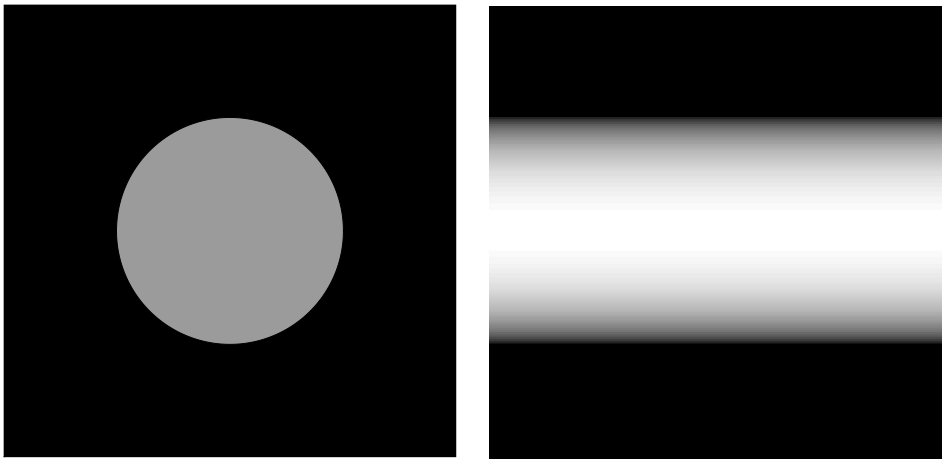
4.2.1 Step 1: Forward Model

The part of forward model aims to get the temporary sinogram denoted as $\hat{p}_{kj}$ from current curve assuming the attenuation coefficient $\mu = 1$. The basic idea of the model is to project two adjacent vertices to the detector pixels and then update the sinogram entry corresponding to the line segment. The forward model algorithm is shown below.

close the curve;
for $k = 1, 2, \ldots, K$ **do**
  for $n = 1, 2, \ldots, N$ **do**
    compute $l(n, k)$;
    compute $l(n + 1, k)$;
    compare $l(n, k)$ and $l(n + 1, k)$ (determine sign);
    add signed contribution of $\overline{c_n c_{n+1}}$;
  end
end

Algorithm 2: Forward Model

Figure 4.3 shows an example of using the forward model on the initial guess as a circle. As the projection angle changes from 0 to 180, the projection data is exactly the same. The forward model is used to project the current curve to obtain the current projection data. In comparison to the real projection data, the curve information can be efficiently updated.



Initial Guess Current Sinogram

Figure 4.3: Example of the forward model, where the current curve is the initial guess as a circle, which is shown on the left. The current sinogram is shown on the right.

4.2.2 Step 2: Attenuation Update

After obtaining the current sinogram $\hat{p}_{kj}$ from the current curve by the forward model, the method updates the attenuation coefficient μ for the object by minimizing the external energy

$$E_{ext}(\mu) = \sum_{j=1}^{J} \sum_{k=1}^{K} (s_{kj} - \mu \hat{p}_{kj})^2$$

By taking the derivative of E with respect to μ

$$\frac{dE_{ext}(\mu)}{d\mu} = \sum_{j=1}^{J} \sum_{k=1}^{K} (s_{kj} - \mu \hat{p}_{kj}) \hat{p}_{kj} = 0$$

The optimal μ is given by

$$\mu = \frac{\sum_{j=1}^{J}\sum_{k=1}^{K} s_{kj}\hat{p}_{kj}}{\sum_{j=1}^{J}\sum_{k=1}^{K}(\hat{p}_{kj})^2}$$

which corresponds to a least squares fit of $\mu\hat{p}_{kj}$ to s_{kj}.

After updating the attenuation, the projection is given by

$$p_{kj} = \mu\hat{p}_{kj}$$

4.2.3 Step 3: Curve Update

After obtaining the current sinogram of the curve, the curve can be updated by minimize the difference of the true sinogram and the current sinogram. The curve update contains two parts - curve deformation and curve regularization.

We denote δ_n as a signed displacement for the vertex c_n, as

$$c_n^{new} = c_n + \delta_n \boldsymbol{j_n}$$

where $\boldsymbol{j_n}$ is a unit vector perpendicular to the gradient of the curve at c_n. Then, the form of external energy is given by a quadratic function of δ_n

$$E_{ext}(\delta_n) = E_0 + \sum_{k=1}^{K}[(s_{kl(n,k)} - p_{kl(n,k)} + \mu\delta_n)^2 - (s_{kl(n,k)} - p_{kl(n,k)})^2]$$

where E_0 is the energy before displacing the point c_n.

Then by taking the derivative of E with respect to δ_n

$$\frac{dE_{ext}}{d\delta_n} = 2\mu \sum_{k=1}^{K} (s_{kl(n,k)} - p_{kl(n,k)} + \mu\delta_n) = 0$$

the optimal δ_n is given by

$$\delta_n = \frac{1}{\mu K} \sum_{k=1}^{K} (p_{kl(n,k)} - s_{kl(n,k)})$$

Therefore, every vertex on the curve is deformed by

$$c_n^{new} = c_n + \delta_n j_n$$

In order to maintain stability of the curve, the curve is regularized by

$$\Delta c_n = \alpha(c_{n-1} - 2c_n + c_{n+1}) + \beta(-c_{n-2} + 4c_{n-1} - 6c_n + 4c_{n+1} - c_{n+1})$$

where α controls the elasticity term (second derivative)-to control the curve length, and β controls the rigidity term (fourth derivative)-to control the curve bending. See also [13] for more details about curve regularization.

Therefore, the curve regularization is given by

$$c_n = c_n^{new} - \Delta c_n^{new}$$

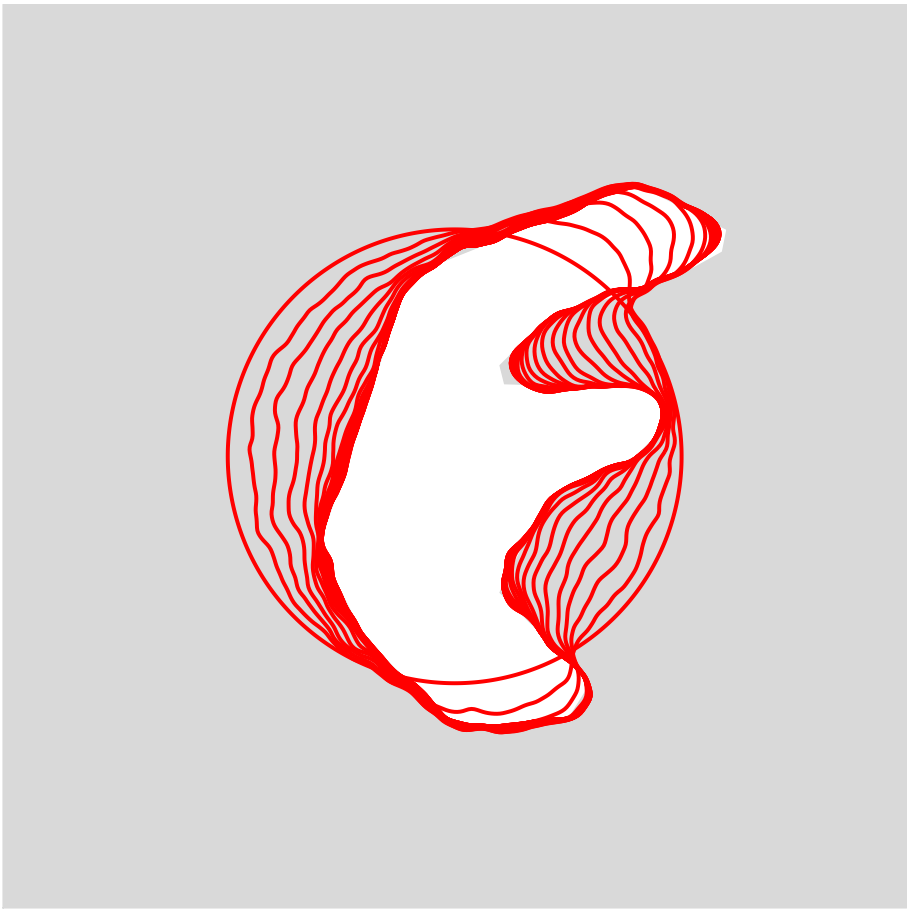


Figure 4.4: Example of the curve update process in segmentation process

Figure 4.4 shows an example of the curve update process. The image segmentation process starts from an initial guess as a circle. During each iteration, each point on the curve moves according to the difference of the true sinogram and current sinogram. The whole curve keeps moving along with the shape of the white object in the center. Finally, the curve successfully follows the shape of the edge of the object.

After updating the curve, the method goes back to step 1 to repeat the iteration. More details about stopping rule and the accuracy of the algorithm will be discussed in Chapter 6.

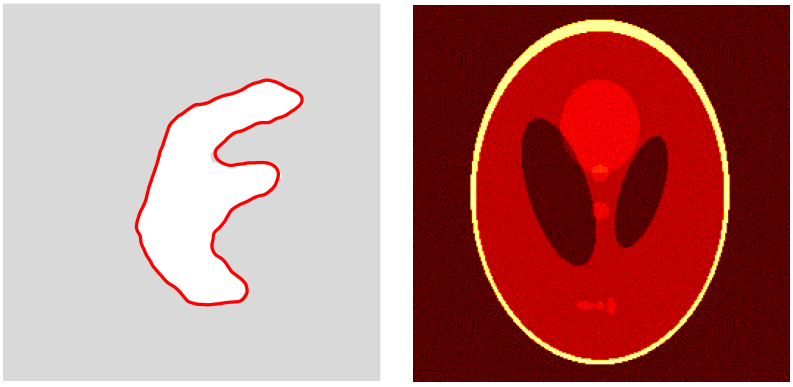
4.3 Comparison of Reconstruction and Segmentation

Image reconstruction and image segmentation are two different kinds of process to obtain information of the object. Image reconstruction represents the object as intensity values in each pixel. The reconstruction is processed by forming the matrices and updating the attenuation coefficient in each pixel accordingly.

However, image segmentation is a process which represents the object as a parametric curve. The object is given by the region inside the curve, which takes the same attenuation coefficient uniformly.

Therefore, the image segmentation process focuses primarily on a single

object which takes the same attenuation coefficient uniformly. The segmentation method aims to detect the edge of the object and clearly separates the object from the background, as is shown in Figure 4.5. However, the image reconstruction process focuses primarily on computing the attenuation coefficient in every pixel in the full range. Therefore, the reconstruction process does a good job of detecting the part inside the edge of object. But it is sometimes difficult for the reconstruction process to separate the object from the background, as is shown in Figure 4.5. Therefore, the basic idea of enriched subspace, discussed in the following chapters, is to segment the image before reconstruction, so that we can detect the edge of the object and feed this information into the reconstruction algorithm to obtain a more accurate final image. More numerical experiments will be mentioned in Chapter 6.



Segmentation example Reconstruction example

Figure 4.5: Example of some results from image reconstruction and image segmentation

Chapter 5

Enriched Krylov Subspace Methods

In this chapter, the image segmentation method and the image reconstruction method are to be combined to solve the image processing problem. First, it is necessary to understand the background of Krylov subspace methods.

5.1 Krylov Subspace Method

A Krylov subspace method is an iterative method to solve the inverse problem.

$$\boldsymbol{Ax} = \boldsymbol{b}$$

In linear algebra, the order-r Krylov subspace generated by n-by-n matrix $\boldsymbol{A}$ and a vector $\boldsymbol{b}$ of dimension n is given by

$$K_r(\boldsymbol{A}, \boldsymbol{b}) = span\{\boldsymbol{b}, \boldsymbol{Ab}, \boldsymbol{A^2b}, \ldots, \boldsymbol{A^{r-1}b}\}$$

The Krylov subspace method aims to find the solution of $\boldsymbol{Ax} = \boldsymbol{b}$ in the span of the Krylov subspace. In order to get a good approximation to the exact solution of $\boldsymbol{Ax} = \boldsymbol{b}$, in each iteration of the Krylov subspace method, the euclidean norm of residual is minimized, with the solution $\boldsymbol{x_m}$ inside the m-order Krylov subspace $K_m(\boldsymbol{A}, \boldsymbol{b})$.

$$\min_{\boldsymbol{x_m}} ||\boldsymbol{Ax_m} - \boldsymbol{b}||_2$$

subject to

$$\boldsymbol{x_m} \in span\{\boldsymbol{b}, \boldsymbol{Ab}, \boldsymbol{A^2b}, \ldots, \boldsymbol{A^{m-1}b}\}$$

Notice that because every subspace is contained in the next subspace, the residual does not increase. Also, assume $\boldsymbol{A}$ is an n-by-n matrix, then (assuming exact arithmetic) after n iterations, the Krylov subspace method will get the exact solution. More details can be found in [18]. Some preconditioning schemes to improve efficiency can be found in [19].

5.2 Enriched Subspace

From the Krylov subspace method, the final solution is restricted in a particular r-order Krylov subspace $K_r(\boldsymbol{A}, \boldsymbol{b})$. However, there may be sometimes that more restrictions can be added to the object if more information about the object is provided. Therefore, those restrictions can be added in the form of enriched basis of the Krylov subspace.

If some parts of the object are already provided, then the information of those parts can be added directly to the subspace in the form of basis vectors.

Assume that there are k parts in the object that are already provided to describe the object, then if we set 1 inside the part and 0 outside the part, k vectors can be formed as

$$\boldsymbol{u_1}, \boldsymbol{u_2}, \ldots, \boldsymbol{u_k}$$

By adding those k vectors into the subspace, the final solution from the Krylov subspace method will contain the information of the image basis.Therefore, instead of starting from an empty subspace in the Krylov subspace method, the enriched Krylov subspace starts from the subspace of $span\{\boldsymbol{u_1}, \boldsymbol{u_2}, \ldots, \boldsymbol{u_k}\}$.

Then, the enriched Krylov subspace is given by

$$\hat{K}_r(\boldsymbol{A}, \boldsymbol{b}) = span\{\boldsymbol{u_1}, \boldsymbol{u_2}, \ldots, \boldsymbol{u_k}, \boldsymbol{b}, \boldsymbol{Ab}, \boldsymbol{A^2b}, \ldots, \boldsymbol{A^{r-1}b}\}$$

Since there are more restrictions on the object, the final solution from the enriched Krylov subspace method would ideally perform better than the general Krylov subspace method. More details about enriched Krylov subspace method can be found in [5].

Here is an example. In Chapter 3 we showed an example of a limited angle reconstruction; recall the results in Figure 3.2. We first consider the best possible basis, which consists of segmenting the true image. Each basis image is obtained by assigning a value of 1 inside the region of each segmented piece, and 0 outside the region; Figure 5.1 displays these basis images.

If we use this ideal basis in the enriched Krylov method (which is implemented in the IR Tools code `IRenrich` [10]), only one iteration is needed to compute an outstanding reconstruction. However, if we continue the iterations, finally the parts other than the ideal basis will dominate the solution, which makes the solution almost the same as the solution given by `IRhybrid_lsqr`. This is shown in Figure 5.2, along with the solution computed using `IRhybrid_lsqr`.

We also show a plot of the relative errors at each iteration in Figure 5.4. Note that with the ideal basis, the enriched Krylov subspace computes an

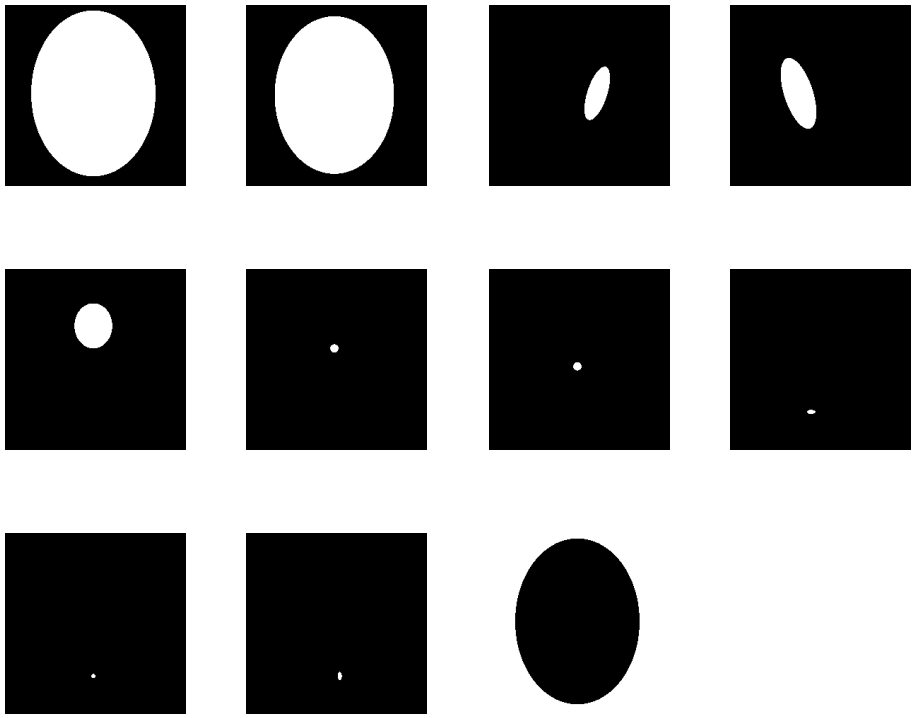
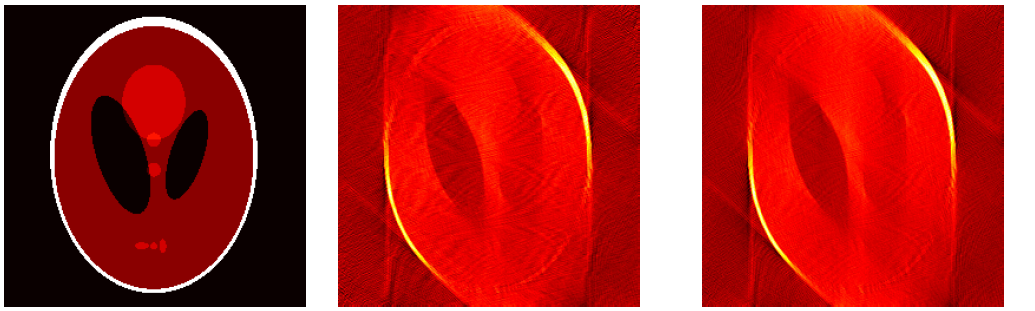


Figure 5.1: Segmenting the true image in Figure 3.2, we obtain this set of basis images for each segmented piece.



True image | `IRenrich` Reconstruction | `IRhybrid_lsqr` Reconstruction

Figure 5.2: Example of an image reconstruction problem using the enriched Krylov subspace method, with an ideal basis obtained by segmenting the true image.

excellent solution after one iteration, but if we continue to iterate, the solution eventually converges to the same solutions as computed by the `IRhybrid_lsqr` methods. More experiments will be shown in the next chapter when the basis is only an approximation.

5.3 Alternative Approach - Modified CGLS

Instead of adding basis to the Krylov subspace, the general conjugate gradient least square (CGLS) method can be modified in order to add the basis information into the approach.

Instead of solving the general least squares problem as

$$\min_x \{||\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}||_2\}$$

we first denote a matrix $\boldsymbol{W}$, which contains all of the basis vectors, as

$$\boldsymbol{W} = [\boldsymbol{u_1}, \boldsymbol{u_2}, \ldots, \boldsymbol{u_k}]$$

We want the final result to contain the basis information, which means that we want the numerical solution $\boldsymbol{x}$ to be a linear combination of all possible basis vectors, which ideally would be

$$\boldsymbol{x} = c_1\boldsymbol{u_1} + c_2\boldsymbol{u_2} + \ldots + c_k\boldsymbol{u_k} = \boldsymbol{W}\boldsymbol{c}$$

where $\boldsymbol{c}$ denotes a column vector with the weights of all the basis vectors.

Therefore, recall from the idea of modified Tikhonov regularization in Chapter 2, a penalty term of $||\boldsymbol{x} - \boldsymbol{W}\boldsymbol{c}||$ can be added to restrict the solution x, which finally yields the modified CGLS method of minimizing

$$
\begin{aligned}
&||\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}||_2^2 + \lambda^2||\boldsymbol{x} - \boldsymbol{W}\boldsymbol{c}||_2^2 \\
&= \left|\left| \begin{bmatrix} \boldsymbol{b} \\ \boldsymbol{0} \end{bmatrix} - \begin{bmatrix} \boldsymbol{A} & \boldsymbol{0} \\ -\lambda\boldsymbol{I} & \lambda\boldsymbol{W} \end{bmatrix} \begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{c} \end{bmatrix} \right|\right|_2^2 \\
&= ||\hat{\boldsymbol{b}} - \hat{\boldsymbol{A}}\hat{\boldsymbol{x}}||_{\boldsymbol{2}}^{\boldsymbol{2}}
\end{aligned}
$$

where

$$
\hat{\boldsymbol{b}} = \begin{bmatrix} \boldsymbol{b} \\ \boldsymbol{0} \end{bmatrix}
$$

$$
\hat{\boldsymbol{A}} = \begin{bmatrix} \boldsymbol{A} & \boldsymbol{0} \\ -\lambda\boldsymbol{I} & \lambda\boldsymbol{W} \end{bmatrix}
$$

$$
\hat{\boldsymbol{x}} = \begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{c} \end{bmatrix}
$$

In this way, we transform the penalized function to a larger least squares problem. Finally we will use the CGLS method to solve the least squares problem and get the solution for $\hat{\boldsymbol{x}}$. Then after taking the first n entries, we will get the solution for $\boldsymbol{x}$. More details about CGLS method can be found

in [3].

True image | `IRcgls` Reconstruction | `IRhybrid_lsqr` Reconstruction

Figure 5.3: Example of an image reconstruction problem using `IRcgls` method in [10], with an ideal basis obtained by segmenting the true image.

Figure 5.3 shows an example of using the modified CGLS method to reconstruct the phantom image. If we use the ideal basis in the modified CGLS method, which is implemented in `IRcgls` in [10], we will get a very good approximation which is shown in the middle of Figure 5.3.

We also show a plot of the relative errors at each iteration in Figure 5.4. Note that with the ideal basis, the error term of `IRcgls` keeps decreasing and finally reaches an acceptable tolerance. In comparison to the solution given by `IRenrich` and `IRhybrid_lsqr`, `IRcgls` gives a very good approximation to the object. Therefore, `IRcgls` is always used instead of `IRenrich` with the aim of adding basis information.

In all the examples in this chapter, the true basis is used as an enriched subspace to reconstruct the image. However, the true basis is never

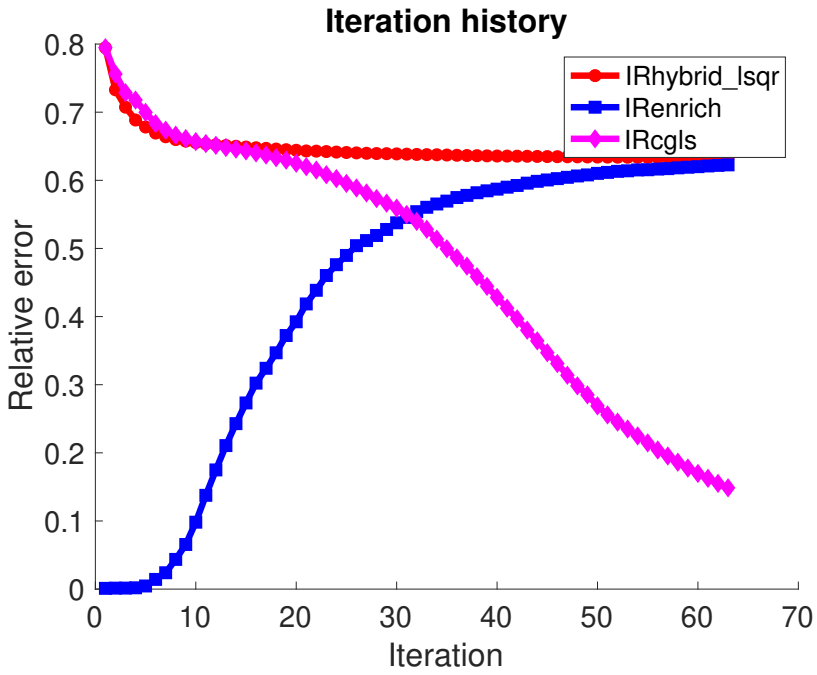


Figure 5.4: This plot shows the iteration history. Specifically, it shows the relative errors of the computed reconstruction at each iteration, using `IRenrich`, `IRhybrid_lsqr` and `IRcgls` method.

provided in the real case. Therefore, in order to improve the accuracy of image reconstruction result, instead of directly reconstructing the image, image segmentation can be first performed to get an approximate basis of the image. Then, this approximate basis will be used through the reconstruction process. More numerical experiments about these methods with an approximate basis will be shown in the next chapter.

Chapter 6

Numerical Experiments

6.1 Segmentation Experiments

First, in order to understand how accurate a solution is compared to the object, two factors are considered in the segmentation case, which are residual and error.

Error is defined as the difference of the pixels covered by the approximate curve and the pixels covered by the true curve. It can be considered as the difference from $\boldsymbol{x}$ to $\boldsymbol{x_{true}}$. It is calculated by assigning 1 inside the curve and 0 outside the curve for both curves and take the difference of two vectors. To minimize the error, we use

$$\min_{\boldsymbol{x}}\{||\boldsymbol{x} - \boldsymbol{x_{true}}||_2\}$$

Residual is defined as the difference of the true sinogram and the result sinogram, which represents the difference of the projection data. It can be considered as the difference from $\boldsymbol{Ax}$ to $\boldsymbol{b}$. To minimize the residual, we try

$$\min_{\boldsymbol{x}}\{||\boldsymbol{Ax}-\boldsymbol{b}||_2\}$$

Actually in the real case, since we never know the true object, it is impossible for us to detect the error of our solution. Therefore, residual term is usually used as a measurement of the error term. However, in the following sections, we will see that sometimes the residual is not a good measurement of error, which means that minimizing the residual does not minimize the error term.

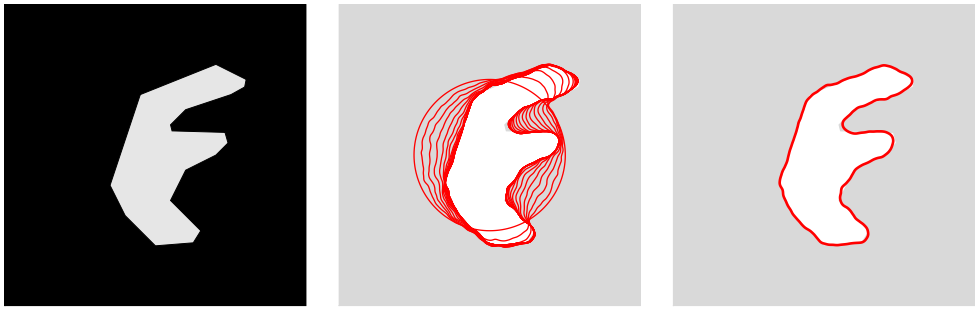
In the following subsections, we are going to implement the image segmentation process on different kinds of objects and different ranges of projection angles.

6.1.1 Single Object

Full Range of Angle: 0-180

Figure 6.1 shows an example of an image segmentation result from a single object in a full range of projection angles. We can clearly see that with the full range of angles, the segmentation process does a very good job to circumscribe the true object on the left. Also, Figure 6.2 shows the error

plot along with the iterations. Notice that we plot each point through 25 iterations. We can see that both the error and the residual decrease to an acceptable range.



True image Curve Evolution Segmentation Result

Figure 6.1: Example of an image segmentation from a single object in full range of projection angles.

In this example, we can clearly see that for a single object, if we use full range of angles, the segmentation result is very good.

Half Range of Angle: 0-90

Figure 6.3 shows an example of an image segmentation result from the same single object using only a half range of projection angles. We can clearly see that with half range of angles, the segmentation result is somewhat distorted on the right side. Also, Figure 6.4 shows the error plot along with the iterations. We can see that the error term is somewhat high at the end, but the residual still decreases to a comparatively small term at

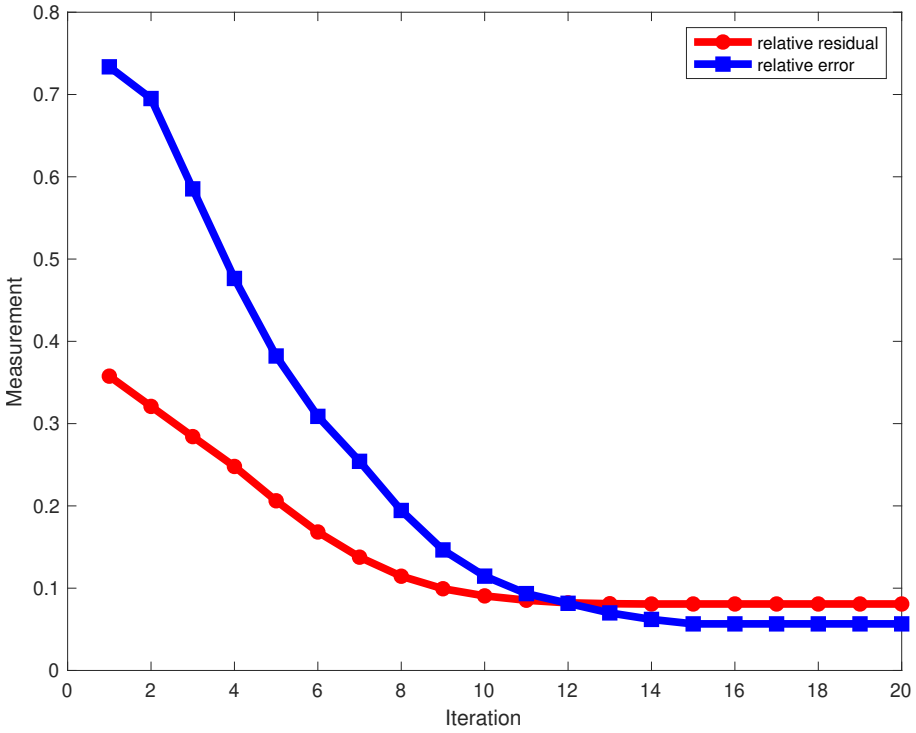


Figure 6.2: Error of an image segmentation from a single object in full range of projection angles.

the end. Therefore, in this case, the residual term fails to be a good measurement for the error term.

True image Curve Evolution Segmentation Result

Figure 6.3: Example of an image segmentation from a single object in half range of projection angles.

In this example, we can see that for a single object, if we use half range of angles, the segmentation result is not very good.

Quarter Range of Angle: 0-45

Figure 6.5 shows an example of an image segmentation result from the same single object using only a quarter range of projection angles. We can clearly see that with this limited range of angles, the segmentation result is totally distorted on the right side and upper side of the object. Also, Figure 6.6 shows the error plot along with the iterations. We can see that the error term is very high at the end, but the residual still decreases to a comparatively small term at the end. Therefore, in this case, the residual term still fails to be a good measurement for the error term.

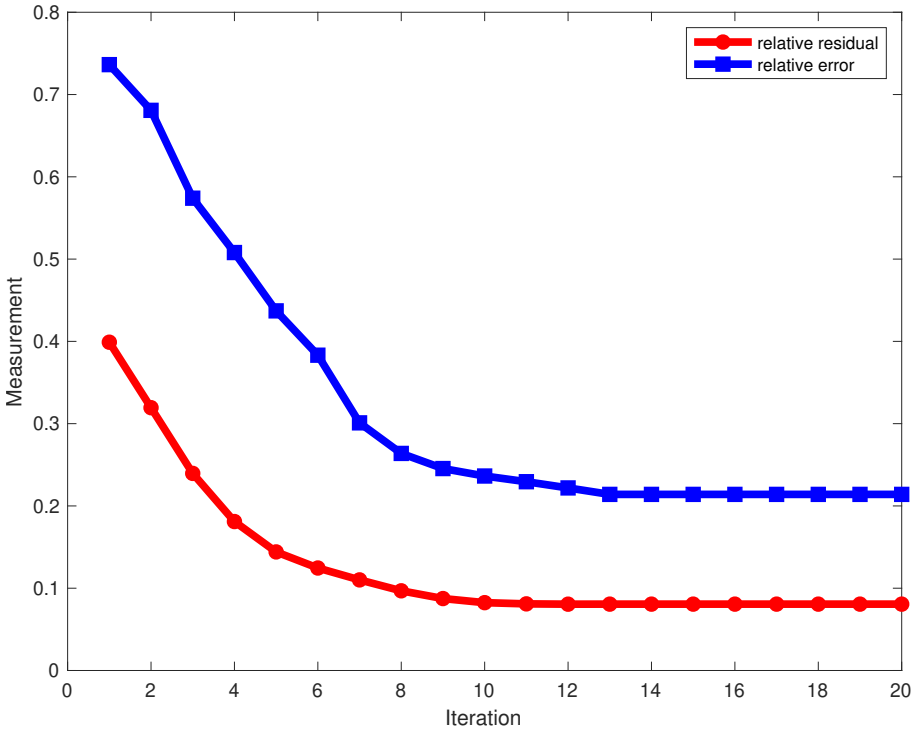
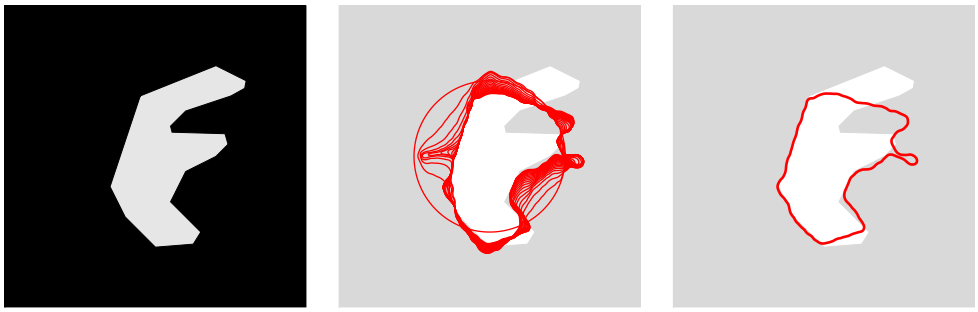


Figure 6.4: Error of an image segmentation from a single object in half range of projection angles.



True image | Curve Evolution | Segmentation Result

Figure 6.5: Example of an image segmentation from separated object in quarter range of projection angles.

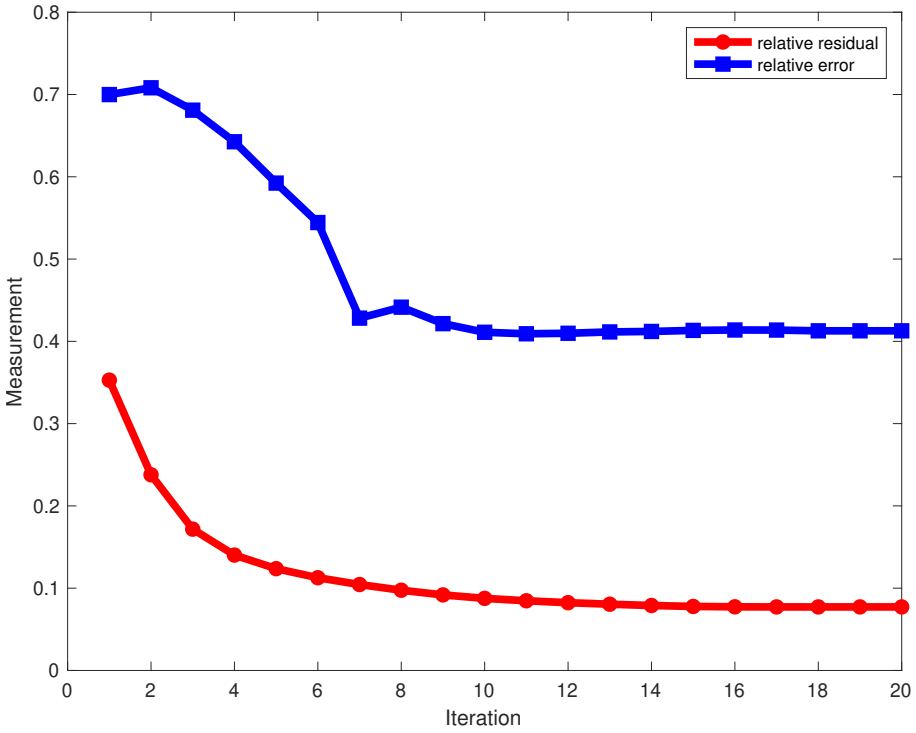


Figure 6.6: Error of an image segmentation from a single object in quarter range of projection angles.

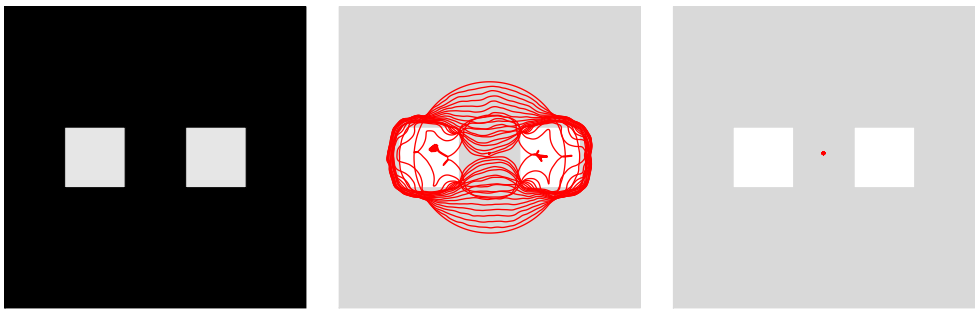
In this example, we can see that for a single object, if we use only one quarter of the range of angles, the segmentation result is comparatively bad. We will see more details about the segmentation result in the next section along with image reconstructions.

6.1.2 Separated Objects

Original Method

In this section, a full range of angles are projected onto two separated objects. First, the image segmentation method in [9] is used to do the segmentation, which is shown in Figure 6.7. We can clearly see that the image segmentation result (which is only a small red dot in the middle of the image) is far off from the true image. More importantly, the plot in the middle of Figure 6.7 shows that the iteration goes through a semi-convergence process. The curve first goes inside toward the edge of the objects, but finally keeps going inside and fails to follow the edge of the object. This is because the image segmentation method is primarily designed for only one object, which corresponds to only one curve. If there are two separated objects, some parts of the curve will converge, which means that the forward model of the segmentation method fails to measure the objects any more.

Figure 6.8 shows the error and residual plot during the iteration. We can see that the algorithm goes through the semi-convergence process. Both the



True image Evolution Segmentation Result

Figure 6.7: Example of an image segmentation from a single object in half range of projection angles.

error and residual term first go down to an acceptable tolerance, then keeps going up. In this case, we can see that the variation of the residual term successfully follows the variation of the error term. Even though in the real case, the error term is never really detected, it is very effective to use residual term as a measurement of error. Therefore, we can modify the stopping rule and stop the algorithm as soon as residual starts to increase.

Modified Method

Figure 6.9 shows the segmentation result from the same object while we stop the iteration when the residual term reaches the minimum. From the evolution plot, we can see that the curve starts from a circle and keeps going inside, but stops before the curve converges. Even though the segmentation result is not good enough to completely follow the shape of those two objects, it is much better than the original method and it could

be good enough to form a basis which will be shown in the later reconstruction part.

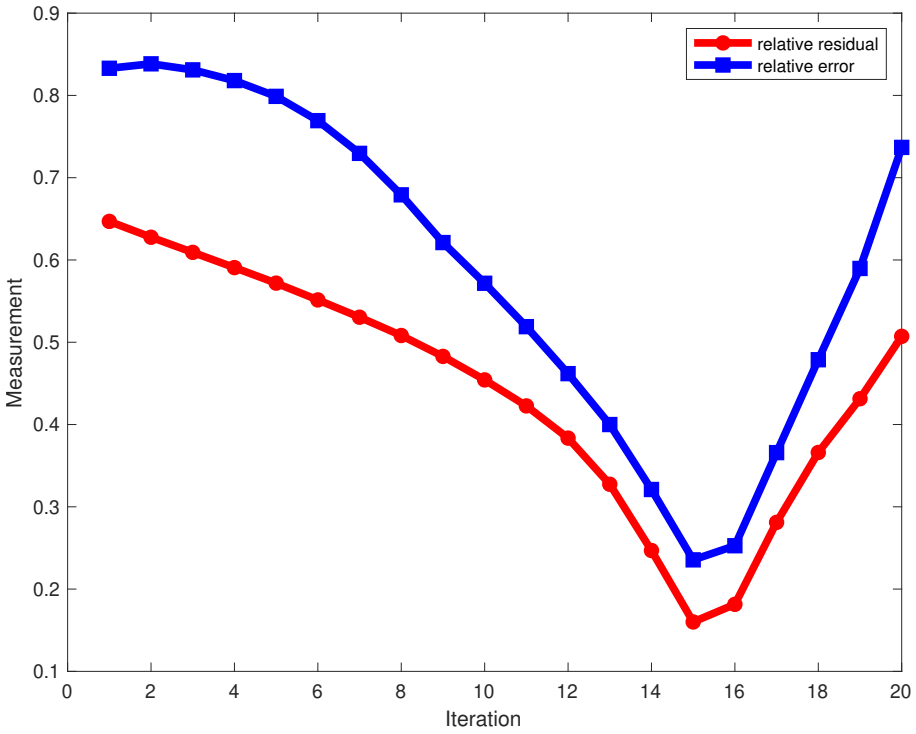
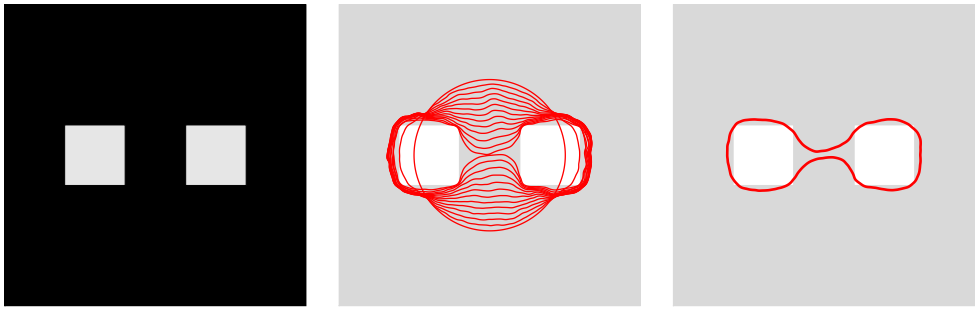


Figure 6.8: Error of an image segmentation from a single object in full range of projection angles.

Figure 6.10 shows the error and residual plot during the iteration. The only difference from the previous subsection is that we stop the iteration ahead of time. We can see that the error term at last falls into an acceptable tolerance. In this case, the residual term successfully becomes a good measurement of the error term. Even though error cannot be detected, the stopping rule based on the residual term successfully improves the

algorithm of two separated object. We will see more experiments of using the image segmentation scheme to form the basis for image reconstruction in the following sections.



True image Curve Evolution Segmentation Result

Figure 6.9: Example of an image segmentation from separated object in full range of projection angles.

From these experiments, we can see that the image segmentation result may be somewhat distorted due to the limited range of projection angles. Also, the segmentation process does not work very well for separated objects even though the new stopping rule is implemented. However, the segmentation result will provide some information as the basis of the image, which can be used in the later image reconstruction process. More experiments in the next section will combine the segmentation and reconstruction process together to improve the accuracy of reconstruction result.

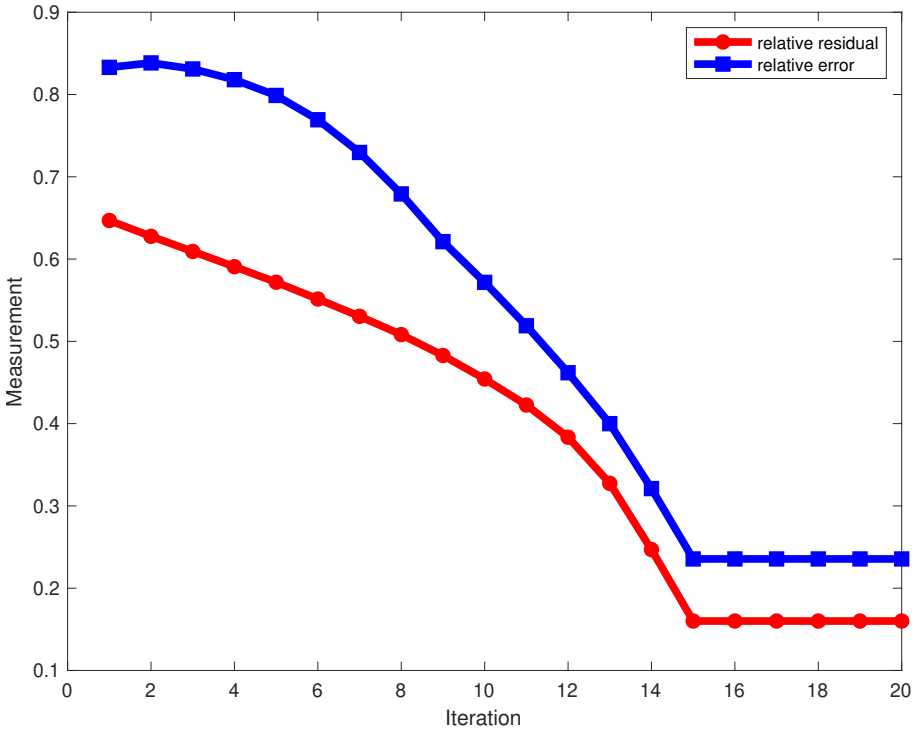


Figure 6.10: Error of an image segmentation from a single object in full range of projection angles. Modified stopping rule is used.

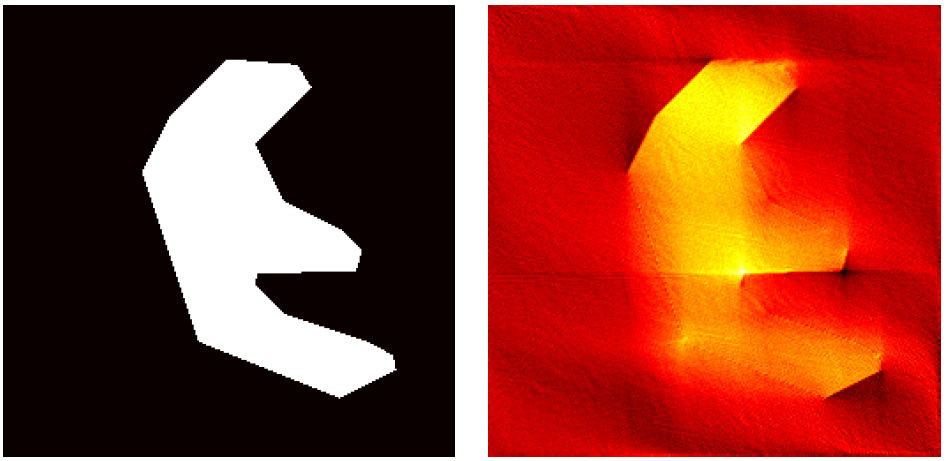
6.2 Reconstruction Experiment

In this section, we will follow the most important idea in this thesis: first segmentation, then reconstruction. We are going to compare the result of image reconstruction in a couple of cases: no basis provided (direct reconstruction), true basis provided and approximate basis provided by segmentation result. We are going to see how accurate the reconstruction result can be if the basis is provided by image segmentation. Notice that because the difficulty of image reconstruction primarily exists in the case of limited range of projection angles, we are going to focus on the case of projection angles ranging from 0 to 90, or even more extremely, from 0 to 45.

6.2.1 Single Object

Half Range of Angles: 0-90

Direct Reconstruction Figure 6.11 shows an example of direct image reconstruction from a single object using only a half range of angles. The left of Figure 6.11 shows the true image, and the reconstructed image given by `IRhybrid_lsqr` is shown on the right. We can see that the image reconstruction does not work well for this case. Much of the edge of the image is blurry, and it is very hard to detect the object from the blurry background.



True image `IRhybrid_lsqr` Reconstruction

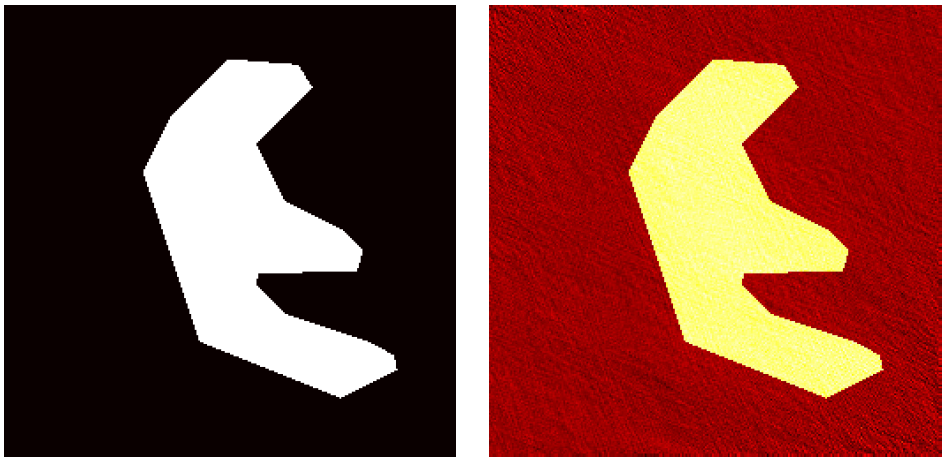
Figure 6.11: Example of a direct image reconstruction from a single object in half range of projection angles.

True Basis Provided Here, we are going to use a basis of segmenting the true image, which is shown in Figure 6.12. Notice the basis not only contains the object, but also contains the outside background.

True basis image

Figure 6.12: The true basis image of the single object in Figure 6.11

Figure 6.13 shows the reconstruction result when using the true basis of the object given by `IRcgls` with parameter $\lambda = 2$. We can clearly see that the



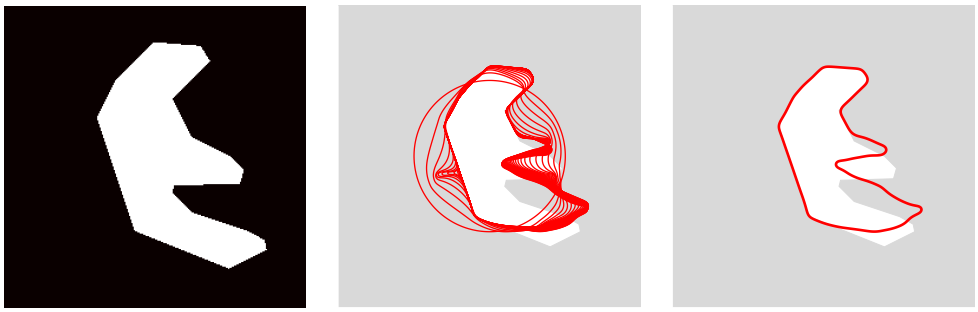
True image IRcgls Reconstruction

Figure 6.13: Example of IRcgls image reconstruction from a single object using true image basis in half range of projection angles.

reconstruction result is much more accurate than the direct reconstruction result, and it is almost exactly the shape of the true object. However, in the real case, the true information of the object is never provided so we have no access to this basis information. What we want to see is how the reconstruction process goes when the basis information is only an approximation given by the image segmentation result. In the case of limited projection angles, it is very necessary to provide the basis information in order to improve the accuracy of the reconstructed image.

Approximate Basis Provided First, the image segmentation method in [9] is used to segment the image into two parts - inside and outside, by a parametric curve. Notice that the range of angles is still limited to 0-90 for

the segmentation case. Then, the segmentation result and the curve evolution process is shown in Figure 6.14. Using this image segmentation, we form an approximate basis which is shown in Figure 6.15.



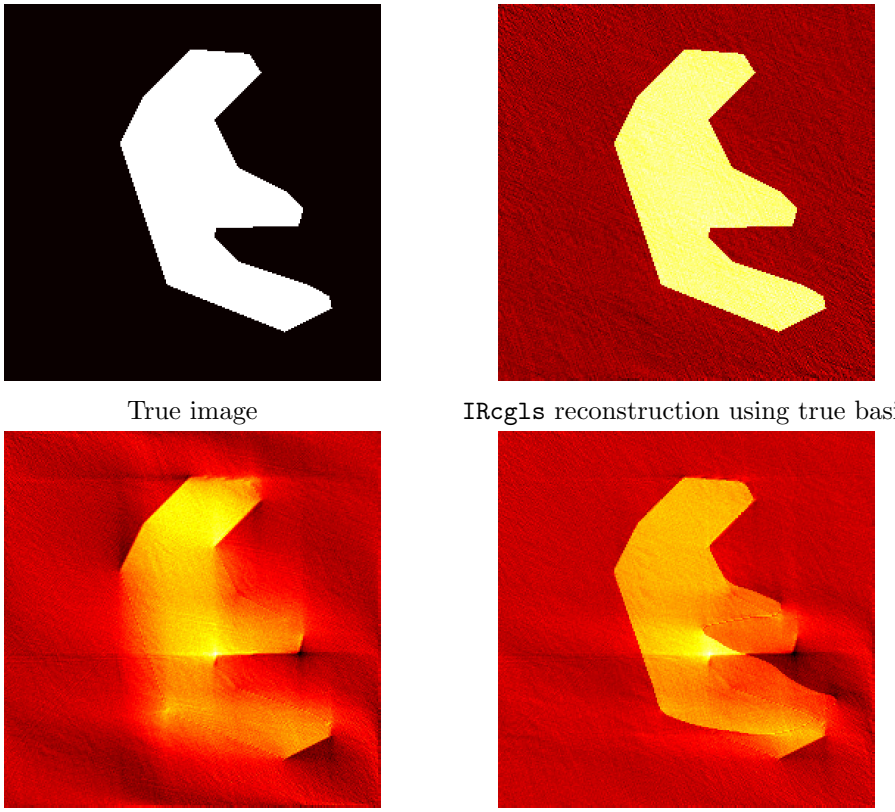
True image | Curve Evolution | Segmentation result

Figure 6.14: Image segmentation process from a single object in half range of projection angles to form an approximate basis.

Basis from segmentation

Figure 6.15: The basis formed by image segmentation from a single object in half range of projection angles.

We then use the basis from the segmentation result and the `IRcgls` method is implemented to reconstruct the image. Figure 6.16 shows the true image and three kinds of reconstruction results, which are the `IRcgls` reconstruction using the true basis, the direct `IRhybrid_lsqr`



True image

`IRcgls` reconstruction using true basis

`IRhybrid_lsqr` reconstruction

`IRcgls` reconstruction using approximate basis

Figure 6.16: Comparison of reconstruction result from a single object in half range of projection angles.

reconstruction and the IRcgls reconstruction using the approximate basis obtained from the segmentation algorithm. We can see that even though the IRcgls reconstruction result using the approximate basis is not as good as the result from true basis, it is significantly better than the direct IRhybrid_lsqr reconstruction result. In particular, the left part of the reconstructed image has been improved a lot. This example indicates that our scheme, first segmentation then reconstruction, highly improves the reconstruction result for a single object in the case of limited angle projections.

Figure 6.17 shows the corresponding error plot of the three reconstruction methods, which are thedirect IRhybrid_lsqr reconstruction, the IRcgls reconstruction using the approximate basis obtained from segmentation algorithm and the IRcgls reconstruction using the true basis. Notice that the direct IRhybrid_lsqr reconstruction is the original approach to solve the inverse problem with no basis information and the IRcgls reconstruction using the true basis is only an ideal situation. The most important curve on the plot is IRcgls_seg which reconstructs the image using the approximate basis. We can clearly see that although the error term for the magenta curve in the middle is not very close to the best reconstruction result from the true basis, it is much better than the direct IRhybrid_lsqr method. Therefore, our scheme, first segmentation then reconstruction, truly improves the reconstruction result for a single object in half range of angles.

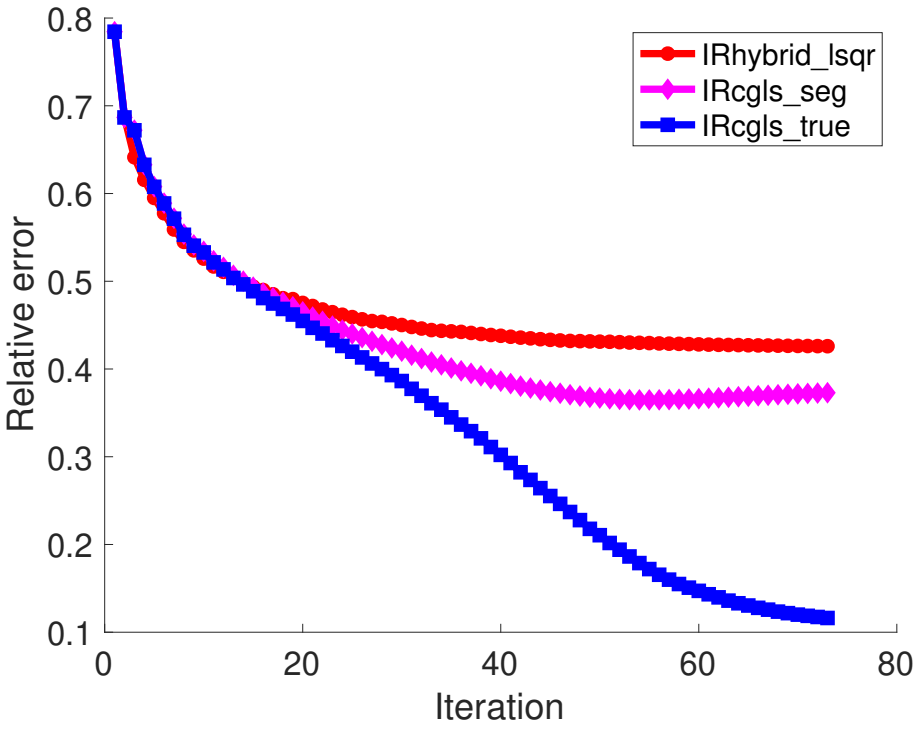
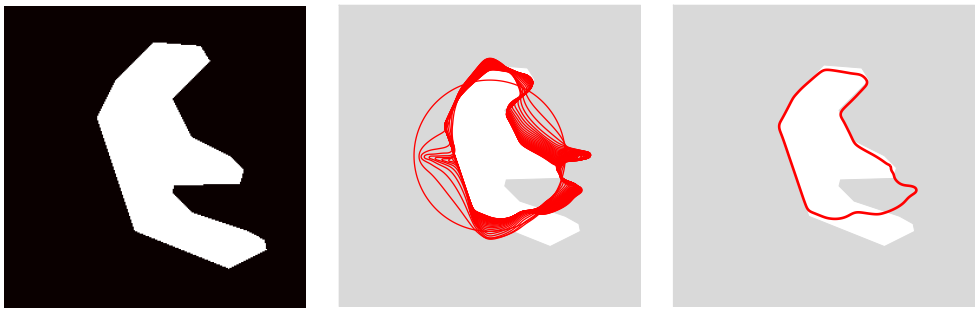


Figure 6.17: This plot shows the iteration history of image reconstruction from a single object in half range of projection angles. Specifically, it shows the relative errors of the computed reconstruction at each iteration, using `IRhybrid_lsqr`, `IRcgls` from segmentation basis and `IRcgls` from true basis

Quarter Range of Angles: 0-45

Now, assume that we have only a quarter of range of angles, which is 0-45, and we still aim to do image reconstruction for the same object. In this limited range of angles, it is very difficult to reconstruct or segment the image accurately. Following the scheme, first segmentation then reconstruction, we first use the limited range of angles from 0 to 45 to segment the image. Figure 6.18 shows the image segmentation process and Figure 6.19 shows the resulting basis images. Now, we hope to use this basis information to reconstruct the image with the `IRcgls` method.



True image | Curve Evolution | Segmentation result

Figure 6.18: Image segmentation process from a single object in quarter range of projection angles to form an approximate basis.

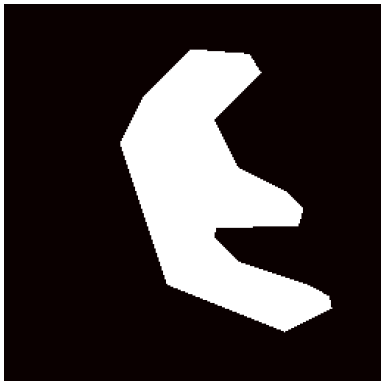
Figure 6.20 shows the true image and three kinds of reconstruction results, which are the `IRcgls` reconstruction using the true basis, the direct `IRhybrid_lsqr` reconstruction and the `IRcgls` reconstruction using the approximate basis obtained from segmentation algorithm. We can see that in this very limited angle case, the direct `IRhybrid_lsqr` reconstruction

Basis from segmentation

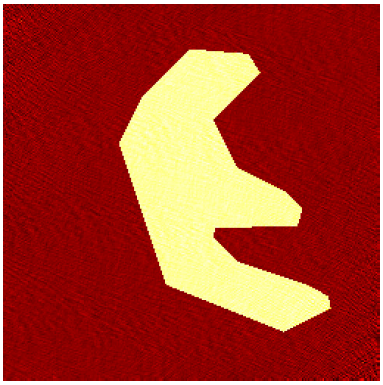
Figure 6.19: The basis formed by image segmentation from a single object in quarter range of projection angles.

with no basis information completely fails to detect the object from the background, which is shown in the lower left image. Using the true basis information, `IRcgls` always gives a very good reconstruction compared to the true image. In this case, although there is some part missing in the lower part of the object, the `IRcgls` reconstruction using the approximate basis in the lower right image works much better than the direct `IRhybrid_lsqr` reconstruction. In the reconstructed image, the object is clearly separated from the background. This example indicates that our scheme, first segmentation then reconstruction, still highly improves the reconstruction result.

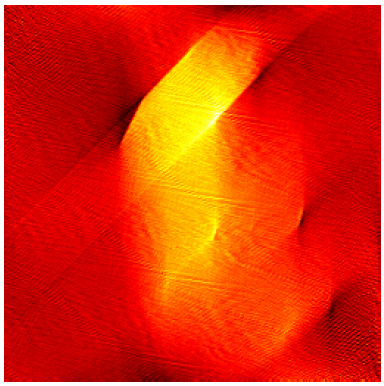
Figure 6.21 shows the corresponding error plot of the three reconstruction methods, which are the direct `IRhybrid_lsqr` reconstruction, the `IRcgls` reconstruction using the approximate basis obtained from segmentation algorithm and `IRcgls` reconstruction using the true basis. In this case, the magenta curve goes through a semi-convergence period and finally reaches



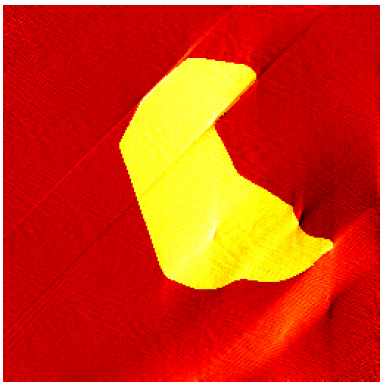
True image



`IRcgls` reconstruction using true basis



`IRhybrid_lsqr` reconstruction



`IRcgls` reconstruction using approximate basis

Figure 6.20: Comparison of reconstruction result from a single object in quarter range of projection angles.

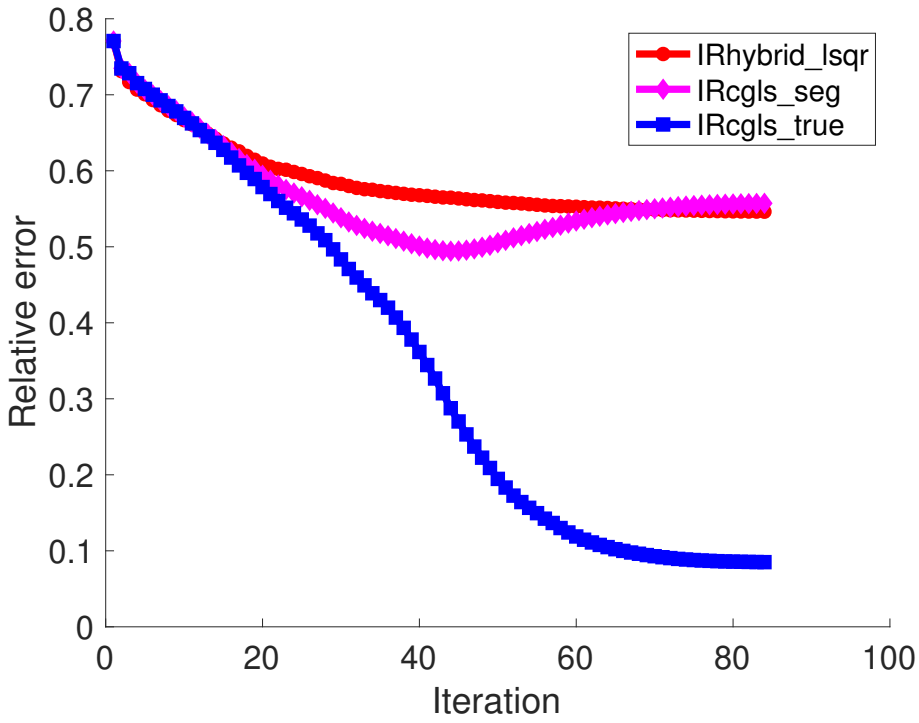


Figure 6.21: This plot shows the iteration history of image reconstruction from a single object in quarter range of projection angles. Specifically, it shows the relative errors of the computed reconstruction at each iteration, using `IRhybrid_lsqr`, `IRcgls` from segmentation basis and `IRcgls` from true basis

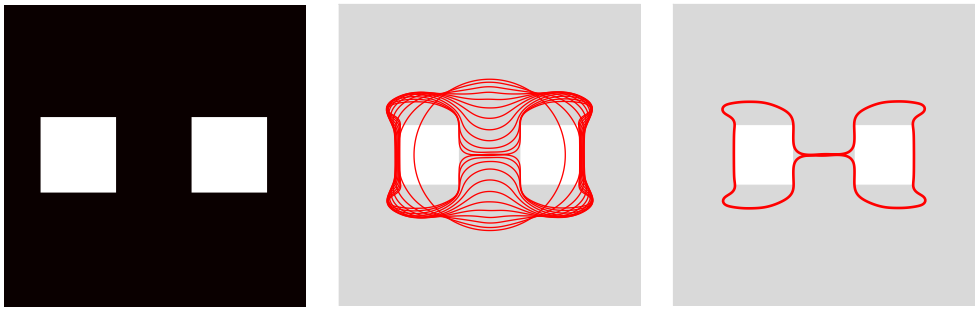
the red curve, which means that finally the reconstruction by `IRhybrid_lsqr` and `IRcgls` using the approximate basis reach the same relative error. Actually, it does not mean `IRcgls` with the approximate basis does not improve the accuracy. There may be some reasons for this. First, in the reconstructed image from `IRcgls` using the approximate basis, there are some added parts on the right of the object, which enlarges the error. Also, the missing part on the lower part of the object enlarges the error as well. Finally, both two types of errors are very high. However, the `IRcgls` reconstruction using the approximate basis still detects the object despite some distortion.

6.2.2 Separated Object

Half Range of Angles: 0-90

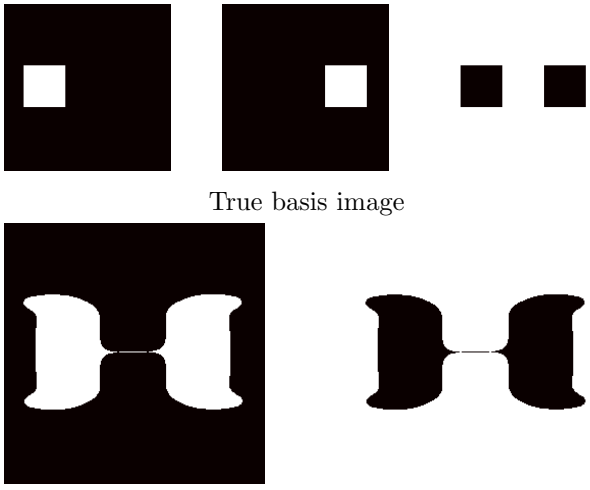
Figure 6.22 shows the image segmentation process of two separated objects, and Figure 6.23 shows the basis images obtained from image segmentation. Figure 6.23 also mentions the three true basis images shown on the top, which are the left square, the right square and the remaining background. Notice that the segmentation goes through a semi-convergence process; recall the results in Figure 6.8. In this segmentation process, the iteration is stopped as soon as the residual term starts to increase. Now, we hope to use this basis information to reconstruct the image by the `IRcgls` method.

Figure 6.24 shows the true image and three kinds of reconstruction results,



True image Curve Evolution Segmentation result

Figure 6.22: Image segmentation process from separated object in half range of projection angles to form an approximate basis.



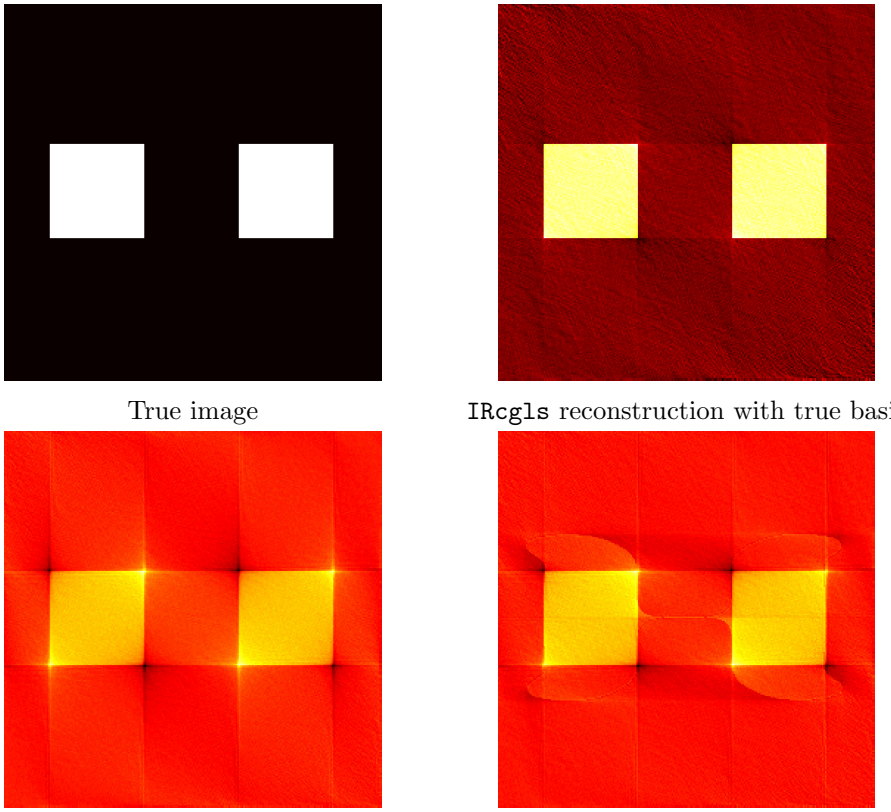
True basis image

Basis from segmentation

Figure 6.23: The true basis and the basis formed by image segmentation from separated objects in half range of projection angles.

which are the IRcgls reconstruction using the true basis, the direct IRhybrid_lsqr reconstruction and IRcgls reconstruction using the approximate basis obtained from segmentation algorithm. We can see that on the upper right corner, using the true basis information, IRcgls always gives a very good reconstruction which is almost exactly the same as the true image shown in the upper left corner. In the lower right corner, we can see that the IRcgls reconstruction using the approximate basis is somewhat influenced by errors from segmentation process, which has some blurred parts around the squares. However, the squares in the IRcgls reconstruction using the approximate basis are somewhat brighter than the squares in the direct IRhybrid_lsqr reconstruction, which means that the difference in the attenuation coefficients of the object and background are larger, indicating that the IRcgls reconstruction using the approximate basis still works more successfully on distinguishing the object from the background.

Figure 6.25 shows the corresponding error plot of the three reconstruction methods, which are the direct IRhybrid_lsqr reconstruction, the IRcgls reconstruction using the approximate basis obtained from segmentation algorithm and the IRcgls reconstruction using the true basis. In this case, the magenta curve is obviously lower than the red curve, which means that, finally in terms of error, the reconstruction given by IRcgls using the approximate basis is somewhat better than the reconstruction given by IRhybrid_lsqr. Overall, due to the very limited range of angles, errors of



True image

`IRcgls` reconstruction with true basis

`IRhybrid_lsqr` reconstruction

`IRcgls` reconstruction with segmentation basis

Figure 6.24: Comparison of reconstruction result from a separated objects in half range of projection angles.

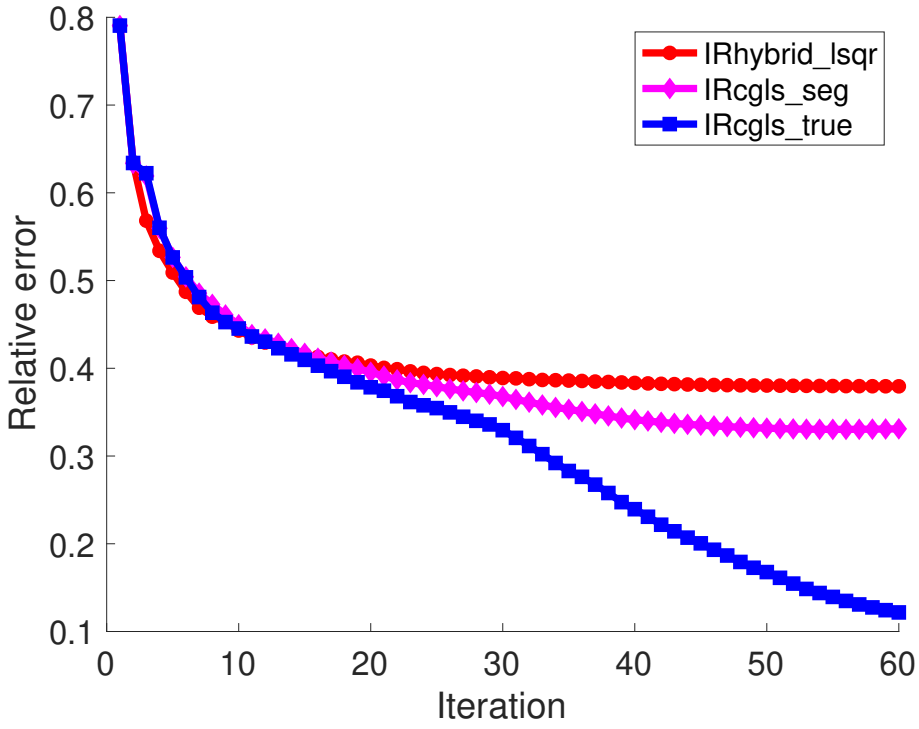
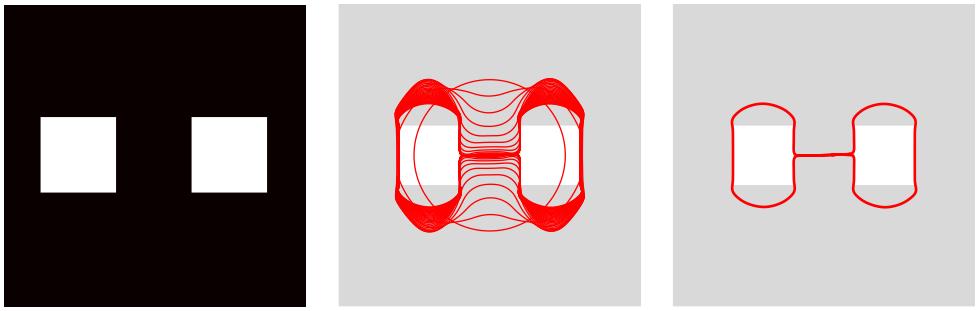


Figure 6.25: This plot shows the iteration history of image reconstruction from separated objects in half range of projection angles. Specifically, it shows the relative errors of the computed reconstruction at each iteration, using `IRhybrid_lsqr`, `IRcgls` from segmentation basis and `IRcgls` from true basis

both reconstructed images are somewhat high. However, the `IRcgls` reconstruction using the approximate basis still detects the object out of the background and improves the accuracy of image reconstruction result.

Quarter Range of Angles: 0-45

Figure 6.26 shows the image segmentation process of two separated objects, and Figure 6.27 shows the image basis obtained from image segmentation. We can also see that since the image segmentation process always starts with an initial guess as a circle, the segmentation works generally better for symmetric objects (e.g. in Figure 6.26 and Figure 6.22) than non-symmetric objects (e.g. in Figure 6.14). Now, we hope to use this basis information to reconstruct the image with the `IRcgls` method.



True image | Curve Evolution | Segmentation result

Figure 6.26: Image segmentation process from separated object in quarter range of projection angles to form an approximate basis.

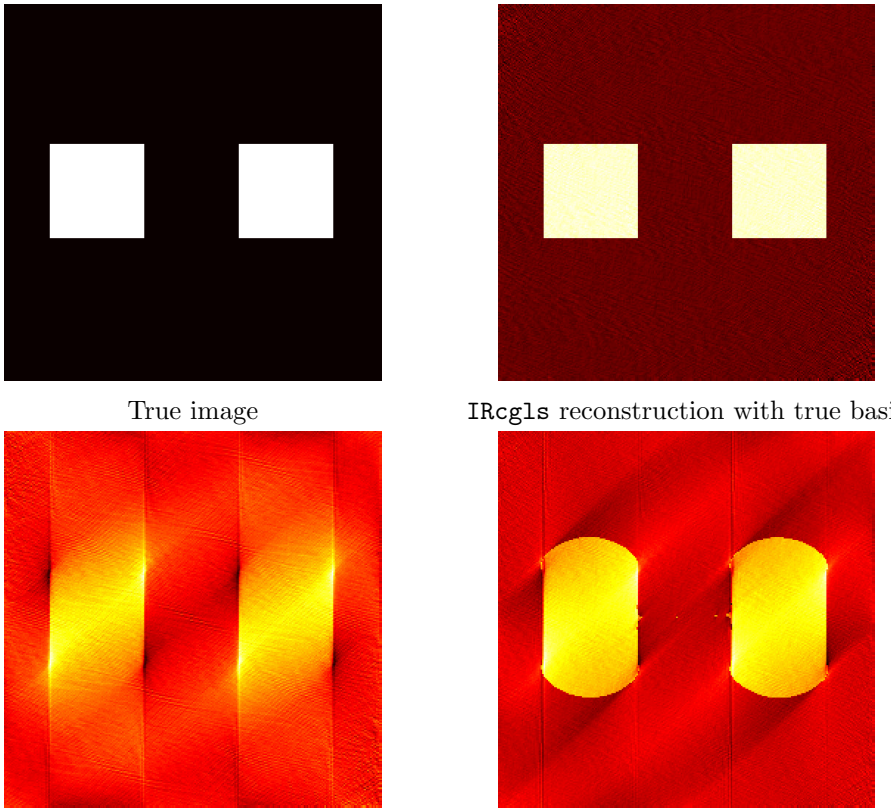
Figure 6.28 shows the true image and three kinds of reconstruction results, which are the `IRcgls` reconstruction using the true basis, the direct

Basis from segmentation

Figure 6.27: The true basis and the basis formed by image segmentation from separated objects in quarter range of projection angles.

`IRhybrid_lsqr` reconstruction and the `IRcgls` reconstruction using the approximate basis obtained from the segmentation algorithm. We can see that on the upper right corner, using true basis information, the `IRcgls` always gives a very good reconstruction which is almost exactly the same as the true image shown in the upper left corner. The reconstruction given by the direct `IRhybrid_lsqr` is shown in the lower left corner, which is very blurred and mixed with the background. In the lower right corner, we can see that the `IRcgls` reconstruction using the approximate basis is largely influenced by the error from segmentation process, which is almost exactly the same as the shape of image segmentation result. However, the difference of the yellow part and the red part is very clear, which indicates that the `IRcgls` reconstruction using the approximate basis again successfully detects the edge of the object and distinguishes it from the background.

Figure 6.29 shows the corresponding error plot of the three reconstruction methods, which are the direct `IRhybrid_lsqr` reconstruction, the `IRcgls`



True image

`IRcgls` reconstruction with true basis

`IRhybrid_lsqr` reconstruction

`IRcgls` reconstruction with segmentation basis

Figure 6.28: Comparison of reconstruction result from a separated objects in quarter range of projection angles.

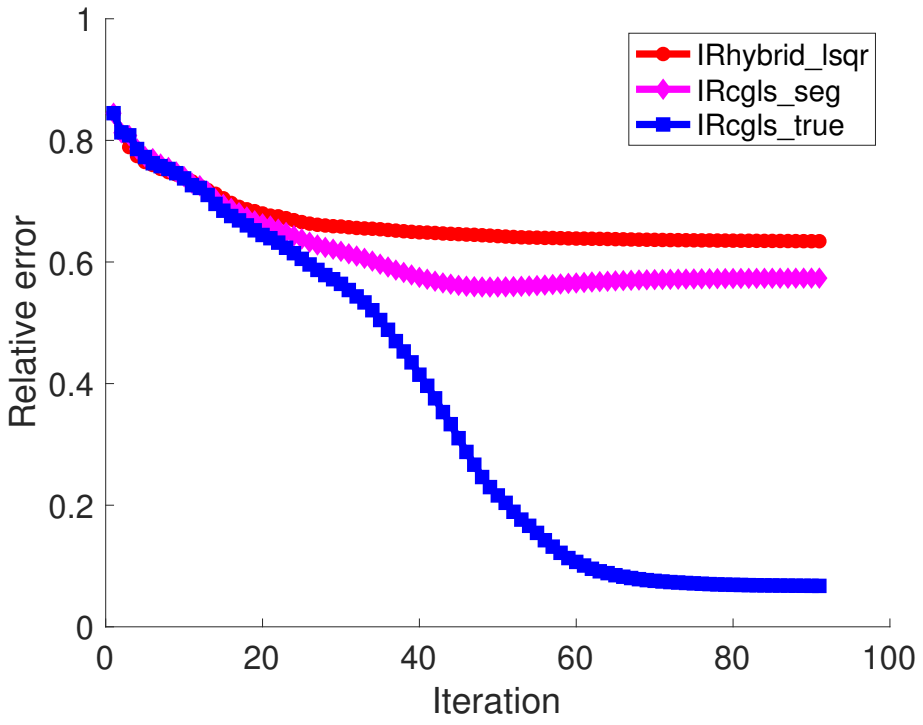


Figure 6.29: This plot shows the iteration history of image reconstruction from separated objects in quarter range of projection angles. Specifically, it shows the relative errors of the computed reconstruction at each iteration, using `IRhybrid_lsqr`, `IRcgls` from segmentation basis and `IRcgls` from true basis

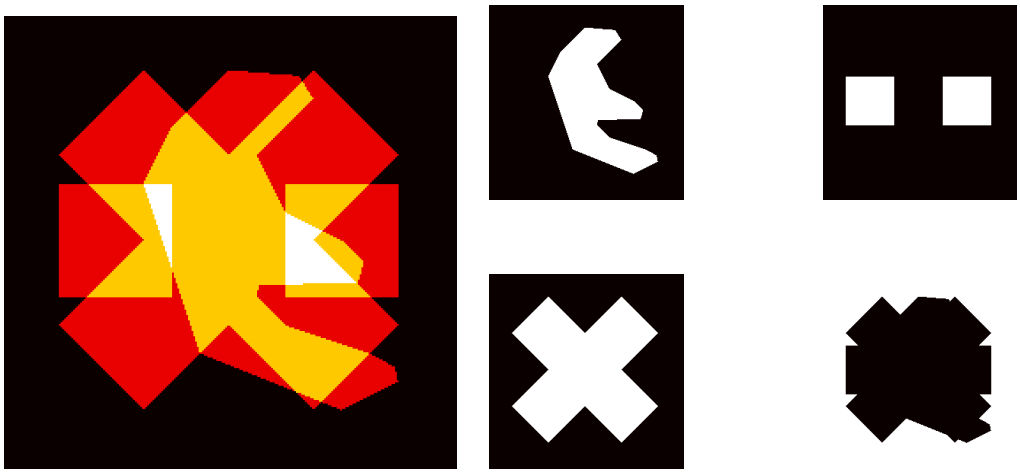
reconstruction using the approximate basis obtained from segmentation algorithm and the `IRcgls` reconstruction using the true basis. In this case, the magenta curve is also lower than the red curve, which means that, finally in terms of error, the reconstruction given by `IRcgls` using the approximate basis is better than the reconstruction given by `IRhybrid_lsqr`. However, it is very difficult to reconstruct the image accurately given such a limited range of angles. Errors in both reconstructed images are very high. However, the `IRcgls` using the approximate basis still detects the object out of the background and on some level improves the accuracy of image reconstruction result.

6.2.3 Overlay Object

First, an overlay object is shown in Figure 6.30 which is a combination of several basis images shown on the right of Figure 6.30. We want to see how our scheme, first segmentation then reconstruction, works for this kind of overlay object. We also consider two cases of limited range of projection angles, half range and quarter range.

Half Range of Angles: 0-90

First, the image segmentation scheme is implemented for the whole object, and the resulting basis is shown in Figure 6.31. We can see that the segmentation successfully segments the whole image from the background but has no indication of the area inside the outer edge of the object. In this



True image True image basis

Figure 6.30: An example of overlay object with the true basis image.

case, since the some parts of the object overlay each other, the segmentation basis is very different from the true basis. Next, we will see how reconstruction works with the segmentation basis.

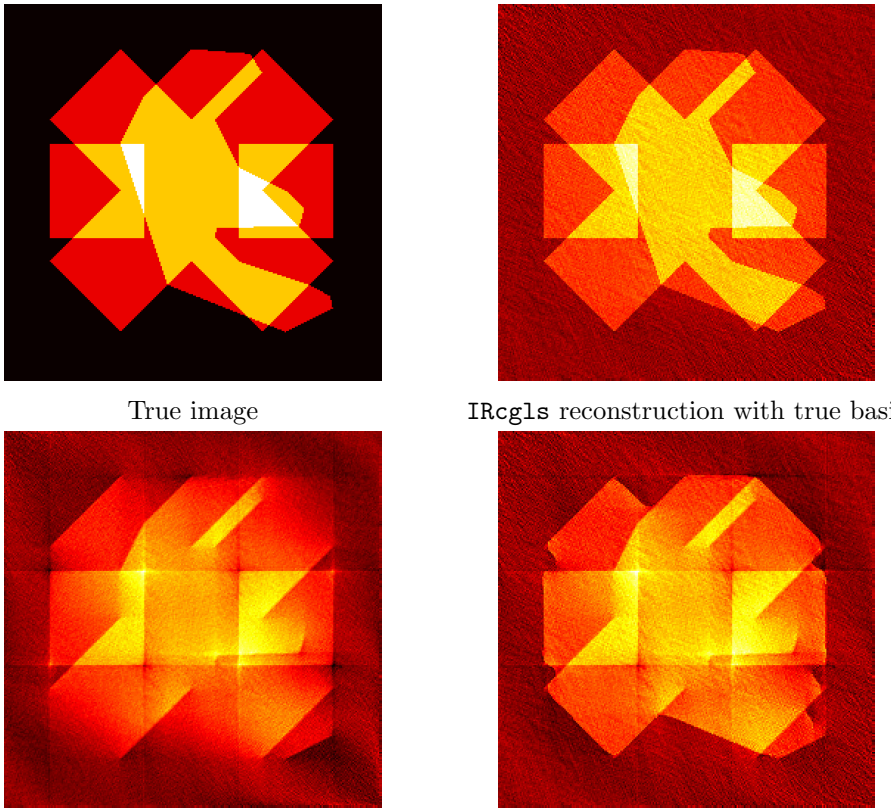
Figure 6.32 shows the true image and three different reconstruction results, which are the `IRcgls` reconstruction using the true basis, the direct `IRhybrid_lsqr` reconstruction and the `IRcgls` reconstruction using the approximate basis obtained from the segmentation algorithm. We can see that on the upper right corner, using the true basis information, the `IRcgls` always gives a very good reconstruction which is almost exactly the same as the true image shown in the upper left corner. The reconstruction given by the direct `IRhybrid_lsqr` is shown in the lower left corner, which is very blurred but gives somewhat clear intensity levels as yellow, orange, red,

Segmentation basis

Figure 6.31: Segmentation basis of overlay object with half range of angles.

showing the overlay levels. In the lower right corner, we can see that the `IRcgls` reconstruction using the approximate basis successfully detects the outer edge of the object and again distinguishes it from the background. In terms of image visualization, it is almost as good as the result given by the `IRcgls` reconstruction using the true image basis.

Figure 6.33 shows the corresponding error plot of the three reconstruction methods, which are the direct `IRhybrid_lsqr` reconstruction, `IRcgls` reconstruction using the approximate basis obtained from segmentation algorithm and `IRcgls` reconstruction using the true basis. In this case, the magenta curve is also a little bit lower than the red curve, which means that, finally in terms of error, the reconstruction image given by `IRcgls` using the approximate basis is better than the reconstruction given by the direct `IRhybrid_lsqr` reconstruction. In this case, `IRcgls` with



True image

`IRcgls` reconstruction with true basis

`IRhybrid_lsqr` reconstruction

`IRcgls` reconstruction with segmentation basis

Figure 6.32: Comparison of reconstruction result from overlay object in half range of projection angles.

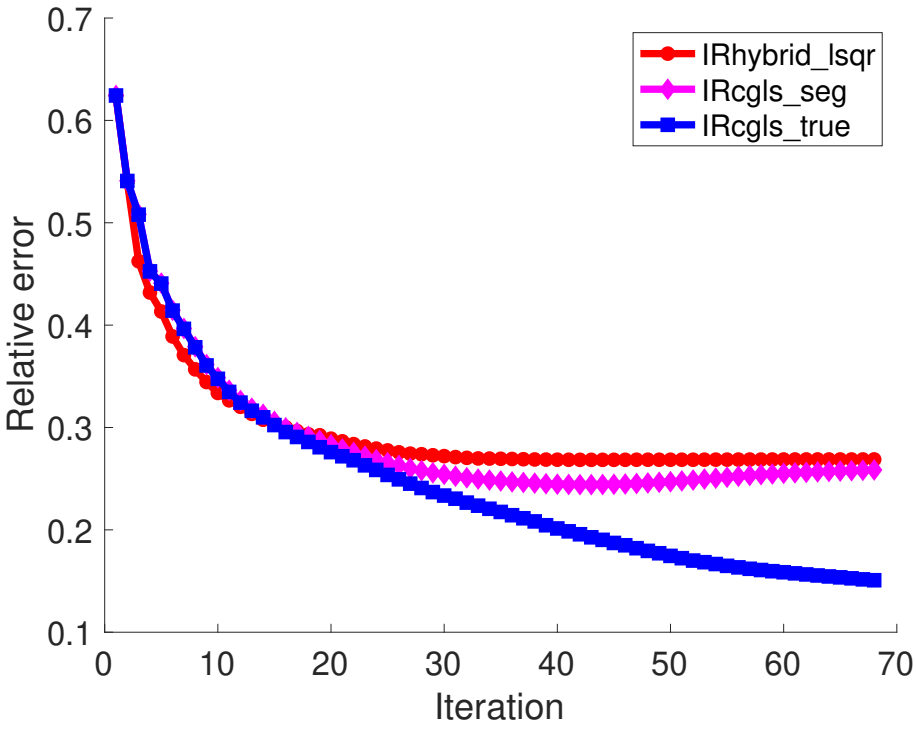


Figure 6.33: This plot shows the iteration history of image reconstruction from overlay object in half range of projection angles. Specifically, it shows the relative errors of the computed reconstruction at each iteration, using `IRhybrid_lsqr`, `IRcgls` from segmentation basis and `IRcgls` from true basis

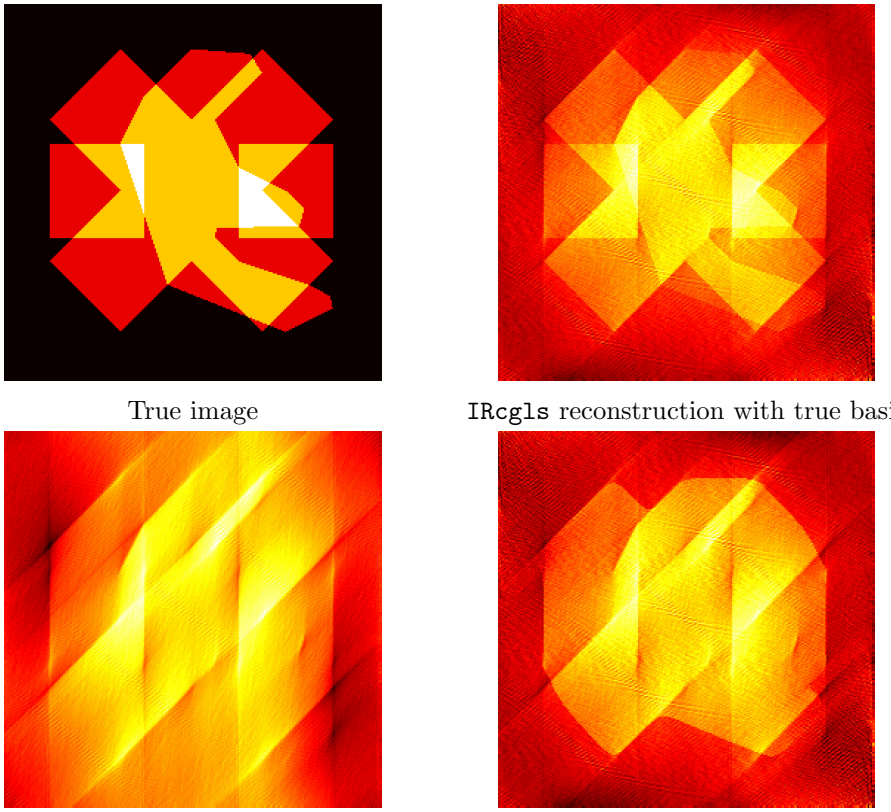
segmentation basis still improves the accuracy of the image reconstruction result.

Quarter Range of Angles: 0-45

Now we decrease the range of projection angles from half range to quarter range, and investigate the reconstruction performance in such a limited range of projection angles. The image segmentation scheme is implemented to the whole object, and the resulting approximate basis images are shown in Figure 6.34. We can see that the edge of the segmentation result is very elastic, which means that in such a limited range of projection angles, the segmentation result is not very accurate as expected. Similarly, the segmentation successfully segments the whole image from the background but has no indication of the area inside the outer edge of the object.

Segmentation basis

Figure 6.34: Segmentation basis of overlay object in quarter range of angles.



True image

`IRcgls` reconstruction with true basis

`IRhybrid_lsqr` reconstruction

`IRcgls` reconstruction with segmentation basis

Figure 6.35: Comparison of reconstruction result from overlay object in quarter range of projection angles.

Figure 6.35 shows the true image and three kinds of reconstruction results, which are the `IRcgls` reconstruction using the true basis, the direct `IRhybrid_lsqr` reconstruction and the `IRcgls` reconstruction using the approximate basis obtained from segmentation algorithm. We can see that on the upper right corner, using the true basis information, the solution given by `IRcgls` is slightly blurred but is still very similar to the true object. The reconstruction given by direct `IRhybrid_lsqr` is shown in the lower left corner, which is very poor. It is very difficult to tell the shape of the object from this blurry image. In the lower right corner, we can see that even though the `IRcgls` reconstruction using the approximate basis is not good enough to see what is inside the outer edge of object, it still roughly tells the shape of the whole object. In such a limited range of angles from 0 to 45, it is overall very difficult to reconstruct the image with high accuracy.

Figure 6.36 shows the corresponding error plot of the three reconstruction methods, which are the direct `IRhybrid_lsqr` reconstruction, the `IRcgls` reconstruction using the approximate basis obtained from segmentation algorithm and the `IRcgls` reconstruction using the true basis. Surprisingly, the red curve blows up at the end of the iteration, which is corresponding to the final poor reconstructed image from `IRhybrid_lsqr`. In this case, the magenta curve is close enough to the blue curve, which means that the solution provided by the approximate basis is somewhat close to the solution provided by the true basis. In this case, the `IRcgls` reconstruction using the approximate basis highly improves the accuracy of image

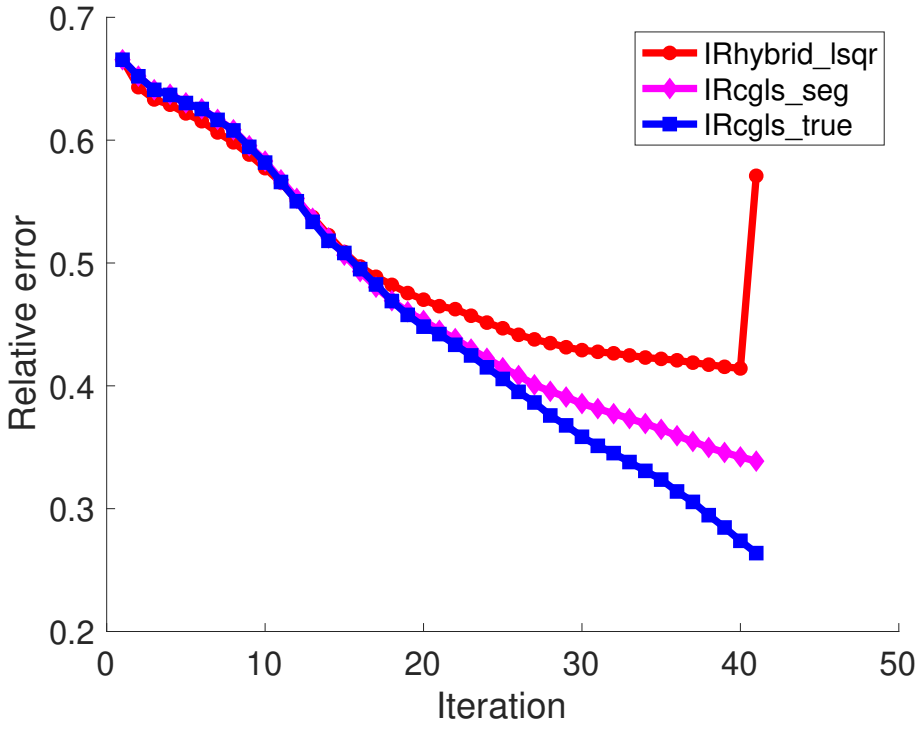


Figure 6.36: This plot shows the iteration history of image reconstruction from overlay object in quarter range of projection angles. Specifically, it shows the relative errors of the computed reconstruction at each iteration, using `IRhybrid_lsqr`, `IRcgls` from segmentation basis and `IRcgls` from true basis

reconstruction result.

Chapter 7

Concluding Remarks

In this thesis, the image segmentation scheme is discussed to detect the edge of the object. Then, we investigated using the segmentation results to form an enriched subspace to improve the performance of image reconstruction. Extensive numerical experiments are presented in Chapter 6 to see the performance of the algorithms.

- We first tried image segmentation with different ranges of projection angles for different kinds of objects. We find that image segmentation performs best on cases of single object and aims to detect the outer edge of the object. We conclude that for specifically limited range of projection angles, the image segmentation result is somewhat distorted. Also, the segmentation process fails to converge for separated objects. A new stopping rule is needed to improve the accuracy.

- We then tried image reconstruction with different ranges of angles. We can see that when the range of angles is limited, the result from reconstruction is highly perturbed and it is very difficult even to detect the shape of the object from the background, which corresponds to the solution given by the iterative solver `IRhybrid_lsqr`.

- Next, we tried to add basis information into the subspace of the image in order to improve the efficiency. We can see that the iterative solver `IRcgls` always works well given the true basis of the object. However, the true basis of the object is never given. Therefore, the segmentation result can be used to obtain an approximate basis for the image reconstruction. From the experiments, we can see that approximate basis successfully improves the accuracy of the reconstructed image in most cases. Even though in some cases the error term is not reduced too much, the solution by `IRcgls` reconstruction using the approximate basis can be considered better in a view of visualization.

The effective implementation of the scheme, first segmentation then reconstruction, demonstrates that this scheme can be used in more future studies. More segmentation methods could be generated to detect the inside area of the object in order to improve the basis. Also, the ways to add the segmentation basis to the subspace can be generated to improve

the efficiency and accuracy of the reconstruction algorithm.

Bibliography

[1] Maurice S Beck et al. *Process tomography: principles, techniques and applications*. Butterworth-Heinemann, 2012.

[2] ZM Bi and Lihui Wang. Advances in 3d data acquisition and processing for industrial applications. *Robotics and Computer-Integrated Manufacturing*, 26(5):403–413, 2010.

[3] Ake Bjorck. *Numerical Methods in Matrix Computations*. Springer, Cham, 2015.

[4] X Briottet, Y Boucher, A Dimmeler, A Malaplate, A Cini, Marco Diani, HHPT Bekman, P Schwering, T Skauli, I Kasen, et al. Military applications of hyperspectral imagery. In *Targets and backgrounds XII: Characterization and representation*, volume 6239, page 62390B. International Society for Optics and Photonics, 2006.

[5] Daniela Calvetti, Lothar Reichel, and Abdallah Shuibi. Enriched krylov subspace methods for ill-posed problems. *Linear algebra and its applications*, 362:257–273, 2003.

[6] Antonin Chambolle, Ronald A De Vore, Nam-Yong Lee, and Bradley J Lucier. Nonlinear wavelet image processing: variational problems, compression, and noise removal through wavelet shrinkage. *IEEE Transactions on Image Processing*, 7(3):319–335, 1998.

[7] Robin L Chazdon and CB Field. Photographic estimation of photosynthetically active radiation: evaluation of a computerized technique. *Oecologia*, 73(4):525–532, 1987.

[8] Julianne Chung, James G Nagy, and DIANNE P O'Leary. A weighted gcv method for lanczos hybrid regularization. *Electronic Transactions on Numerical Analysis*, 28(Electronic Transactions on Numerical Analysis), 2008.

[9] Vedrana Andersen Dahl, Anders Bjorholm Dahl, and Per Christian Hansen. Computing segmentations directly from x-ray projection data via parametric deformable curves. *Meas. Sci. Technol.*, 29, https://doi.org/10.1088/1361-6501/aa950e, 2018.

[10] Silvia Gazzola, Per Christian Hansen, and James G. Nagy. Ir tools: A matlab package of iterative regularization methods and large-scale test problems. *Numerical Algorithms*, https://doi.org/10.1007/s11075-018-0570-7, 2008.

[11] Gene H. Golub and Charles F. Van Loan. *Matrix Computations*. Johns Hopkins University Press, Baltimore, third edition, 1996.

[12] Per Christian Hansen, James G. Nagy, and Dianne P. O'Leary. *Deblurring Images: Matrices, Spectra, and Filtering.* SIAM: Fundamentals of Algorithms, 2006.

[13] Michael Kass, Andrew Witkin, and Demetri Terzopoulos. Snakes: Active contour models. *International journal of computer vision*, 1(4):321–331, 1988.

[14] Thomas Martin Lehmann, Claudia Gonner, and Klaus Spitzer. Survey: Interpolation methods in medical image processing. *IEEE transactions on medical imaging*, 18(11):1049–1075, 1999.

[15] Matthew J McAuliffe, Francois M Lalonde, Delia McGarry, William Gandler, Karl Csaky, and Benes L Trus. Medical image processing, analysis and visualization in clinical research. In *Proceedings 14th IEEE Symposium on Computer-Based Medical Systems. CBMS 2001*, pages 381–386. IEEE, 2001.

[16] Christopher C. Paige and Michael A. Saunders. Lsqr: An algorithm for sparse linear equations and sparse least squares. *ACM: Transactions on Mathematical Software*, 8:43–71, 1982.

[17] Eli Peli, Thomas R Hedges, and Bernard Schwartz. Computerized enhancement of retinal nerve fiber layer. *Acta ophthalmologica*, 64(2):113–122, 1986.

[18] Yousef Saad. Krylov subspace methods for solving large unsymmetric linear systems. *Mathematics of computation*, 37(155):105–126, 1981.

[19] Arvind K Saibaba and Peter K Kitanidis. Efficient methods for large-scale linear inversion using a geostatistical approach. *Water Resources Research*, 48(5), 2012.

[20] DF Swinehart. The beer-lambert law. *Journal of chemical education*, 39(7):333, 1962.