Multidimensional Data Recovery via Iterative Regularization based on Higher Order Singular Value Decomposition

Seyedeh Faegheh Yeganli

Submitted to the
Institute of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Electrical and Electronic Engineering

Eastern Mediterranean University
March 2018
Gazimağusa, North Cyprus

Approval of the Institute of Graduate Studies and Research

Assoc. Prof. Dr. Ali Hakan Ulusoy
Acting Director

I certify that this thesis satisfies the requirements as a thesis for the degree of Doctor of Philosophy in Electrical and Electronic Engineering.

Prof. Dr. Hasan Demirel
Chair, Department of Electrical and Electronic Engineering

We certify that we have read this thesis and that in our opinion it is fully adequate in scope and quality as a thesis of the degree of Doctor of Philosophy in Electrical and Electronic Engineering.

Prof. Dr. Hasan Demirel
Co-Supervisor

Prof. Dr. Runyi Yu
Supervisor

Examining Committee

1. Prof. Dr. Hasan Demirel
2. Prof. Dr. Aytül Erçil
3. Prof. Dr. Fikret S. Gürgen
4. Prof. Dr. Erhan A. İnce
5. Prof. Dr. Hüseyin Özkaramanlı
6. Prof. Dr. Runyi Yu
7. Assoc. Prof. Dr. Önsen Toygar

ABSTRACT

With the recent advances of networking, sensors, and storage technologies, many multidimentional datasets are being generated in various fields. These datasets are often incomplete or contaminated during the acquisition process. Recovering the missing or noise-free data from degraded observations thus becomes crucial to obtaining precise information to refer to. The aim of this thesis is towards restoration of multidimensional (tensor) data. Specifically, we consider three problems: 1) tensor inpainting, 2) magnetic resonance image denoising and, 3) hyperspectral image denoising.

The theory of tensors has become popular in dealing with multidimensional data, due to the capability of tensors in exploiting additional structure in comparison with matrix based alternatives. The most commonly used decomposition of multidimensional data to date is higher order singular value decomposition (HOSVD). The HOSVD is an efficient way for eliciting intrinsic structure of multidimensional data. It offers a simple, adaptive and natural way to exploit sparsity among all dimensions of multidimensional data. The HOSVD decomposes a particular tensor data into the product of a sparse tensor and a few orthogonal matrices, each of which captures the subspace information corresponding to one dimension. In this work, we solve the restoration problems by employing the HOSVD transform and by exploiting the sparsity of the multidimensional signals. We enforce the sparsity using iterative regularization technique, which is shown to be very effective for our problems.

The first contribution of this work is employing the iterative regularization scheme for tensor inpainting. The rationale of this approach is based on an enhanced sparse representation in HOSVD domain and it uses the iterative regularization procedure for inpainting. Improved performances of this algorithm are demonstrated in our experiments on three dimensional tensors, taken from multi-channel (color) images, video sequences, and magnetic resonance images. The evaluation is made quantitatively in terms of peak signal-to-noise ratio and structural similarity index, and qualitatively via visual comparisons.

Despite the success of magnetic resonance imaging techniques in many applications, acquisition noise is still a limiting factor for the quality and hence the usefulness of the techniques. Our second contribution is improving the application of the iterative higher order singular value decomposition framework to denoising the magnetic resonance images. The proposed algorithm forms a single tensor from the noisy data. This tensor undergoes an HOSVD, where its sparse representation coefficients are calculated with respect to a set of directional orthogonal basis matrices. Denoising is achieved by iteratively applying soft thresholding on the calculated sparse representation coefficients. The proposed algorithm is further enhanced with a post-process of Wiener filtering. The performance of the proposed method is evaluated using synthetic and real magnetic resonance images. Validation results and quantitative comparisons with the state-of-the-art in magnetic resonance image denoising clearly demonstrate the advantages of the proposed method.

The hyperspectral data cube is considered as a three-order tensor that is able to jointly treat both the spatial and spectral dimensions. Noise in hyperspectral image can degrade the visual quality and limit the applicability of computerized analysis processes. Hence, toward the third contribution we consider the denoising of the hyperspectral images to improve the performance of the subsequent applications. In this work, not only we use the proposed iterative higher order singular value decomposition, but also we go one step further and propose a new iterative denoising method which utilizes the advantages of the patch-based HOSVD sparse model and the iterative regularization technique. The experiments with both synthetic noisy data and real hyperspectral data reveal that the proposed iterative algorithm improves the hyperspectral data quality in terms of both quality metrics and visual inspection. The subsequent classification results further validate the effectiveness of the proposed hyperspectral noise reduction algorithm.

In conclusion, extensive experiments on synthetic and real world datasets have shown the competitive performance of the proposed algorithms for inpainting, magnetic resonance image denoising, and hyperspectral image denoising over existing state-of-the-art ones.

Keywords: Denoising, higher order singular value decomposition, hyperspectral, iterative regularization, MR images, patch-based, soft thresholding, sparse representation, Tucker decomposition.

ÖZ

Son zamanlarda ağ oluşturma, sensörler, ve depolama teknolojileri alanında yaşanan ilerleme ve gelişmeler ile birlikte farklı alanlarda birçok çok-boyutlu veri seti oluşturulmuştur. Bu veri setleri çoğunlukla veri toplama süreci boyunca eksik kalmakta veya bozulmaktadır. Bu neden ile parazit içermeyen verilerin eksilen gözlemleme verilerinden geriye alınması, atıfta bulunmak üzere dakik bilgilerin elde edilmesi açısından oldukça önemlidir. Bu tez çalışmasının amacı çok-boyutlu (tensör) verilerin restorasyonu ve yeniden yapılandırılmasından ibarettir. Bu tez çalışmasında 3 problem özellikle dikkate alınmaktadır: 1) Tensör içboyaması, 2) manyetik yankılama görüntülerinin parazit temizlemesi, 3) aşırı spektral görüntülerin parazit temizlemesi.

Tensör teorisi, tensörlerin matris bazlı alternatifler ile karşılaştırıldığında ek yapıları işleme kapasitesi nedeniyle çok-boyutlu verilerin işlenmesi için popüler hale gelmiştir. Günümüzde en yaygın olarak kullanılan çok-boyutlu veri ayrıştırma yöntemi üst seviye tekil değer ayrıştırmasıdır (ÜSTDA). ÜSTDA, çok-boyutlu verilerin içsel yapısının ortaya çıkarılması için etkin bir yöntemdir. Bu yöntem, çok-boyutlu verilerin tüm boyutları arasında seyrekliğin işletilmesi için sade, uyarlanabilir ve doğal bir yöntem sunmaktadır. ÜSTDA belirli bir tensör verisini seyrek bir tensör ile her biri tek bir boyuta ait altuzay bilgilerini yansıtan birkaç ortogonal matrisin çarpımı halinde ayrıştırmaktadır. Bu çalışmada ÜSTDA dönüşümünden yararlanılarak ve çok-boyutlu sinyallerin seyrekliği işletilerek çözünürlük problemleri çözülmektedir. Seyreklik, problemlerimiz açısından çok etkili olduğu gösterilen tekrarlayan düzenleme tekniği kullanılarak uygulanmaktadır.

Bu çalışmanın temel katkısı tensör içboyaması için tekrarlayan düzenleme planının kullanımından ibarettir. Bu yöntemin mantığı, ÜSTDA etkinlik alanında gelişmiş bir seyrek temsile dayalı olup içboyama için tekrarlayan düzenleme yöntemini kullanmaktadır. Bu algoritmanın geliştirilmiş performansı, çok kanallı (renkli) görüntüler, video dizinleri, ve manyetik yankılama görüntülerinden alınan üç boyutlu tensörler üzerindeki deneylerimizde gösterilmiştir. Değerlendirmeler nicel olarak en yüksek sinyal gürültü oranı ve yapısal benzerlik indeksi açısından ve nitel olarak ise görsel kıyaslama aracılığı ile gerçekleştirilmiştir.

Manyetik yankılama görüntülerinin birçok uygulamadaki başarısına rağmen, veri toplama parazitleri kalite açısından halen sınırlandırıcı bir faktör olup dolayısıyla bu tekniklerin yararlı olmamasına neden olmaktadır. Bu tez çalışmasının ikinci katkısı ise manyetik yankılama görüntülerinin parazitlerden arındırılmaları amacıyla tekrarlayan üst seviye tekil değer ayrıştırma çerçevesinin uygulamasının geliştirilmesinden ibarettir. Önerilen algoritma parazitli veriden tek bir tensör oluşturmaktadır. Bu tensör, seyrek temsil katsayılarının bir takım yönlü ortogonal taban matrisi bakımından hesaplandığı bir ÜSTDA'ya maruz kalmaktadır. Parazit temizleme işlemi, hesaplanan seyrek temsil katsayıları üzerinde yumuşak eşiklemenin tekrarlanarak uygulanması yolu ile elde edilmektedir. Önerilen algoritma daha sonra Wiener filtreleme ile sağlanan ileri bir işlem ile geliştirilmektedir. Önerilen yöntemin performansı sentetik ve gerçek manyetik yankılama görüntüleri kullanılarak değerlendirilmektedir. Onaylama sonuçları ve manyetik yankılama görüntülerinin parazit temizleme alanındaki en son gelişmeler ile yapılan nicel karşılaştırmalar önerilen yöntemin avantajlarını açıkça ortaya çıkarmaktadır.

Aşırı spektral veri küpü, hem uzaysal hem de spektral boyutları ortaklaşa bir şekilde işleme kabiliyeti olan üç dereceli bir tensör olarak dikkate alınmaktadır. Aşırı spektral bir görüntüdeki parazit, görsel kaliteyi düşürüp bilgisayarlı analiz işlemlerinin uygulanabilirliğini sınırlandırabilmektedir. Dolayısıyla üçüncü katkı olarak bu tez çalışmasında müteakip uygulamaların performansının geliştirilmesi amacıyla aşırı spektral görüntülerin parazitten arındırılması dikkate alınmıştır. Bu tez çalışmasında yalnızca önerilen tekrarlayan üst seviye tekil değer ayrıştırması dikkate alınmamış olup aynı zamanda bir adım ileriye giderek yama bazlı ÜSTDA seyrek model ile tekrarlayan düzenleme tekniğinin avantajlarını kullanan yeni bir tekrarlayan parazit temizleme yöntemi önerilmiştir. Hem sentetik parazitli veriler hem de gerçek aşırı spektral veriler üzerinde gerçekleştirilen deneylerin sonuçları önerilen tekrarlayan algoritmanın aşırı spektral veri kalitesini hem kalite ölçüleri hem de görsel inceleme açısından geliştirdiğini ortaya çıkarmaktadır. Müteakip sınıflandırma sonuçları daha sonra önerilen aşırı spektral parazit azaltma algoritmasının etkinliğini onaylamaktadır.

Sonuç olarak sentetik ve gerçek veri setleri üzerinde gerçekleştirilen kapsamlı deneyler, içboyama, manyetik yankılama görüntülerinin parazit temizlemesi ve aşırı spektral görüntülerin parazit temizlemesi için önerilen algoritmaların ilgili alanlarda mevcut olan en son gelişmeler ile karşılaştırıldığında rekabet edebilir performansa sahip olduğunu göstermiştir.

Anahtar Kelimeler: Parazit temizleme, üst seviye tekil değer ayrıştırması, aşırı spektral, tekrarlayan düzenleme, MR görüntüleri, yama bazlı, yumuşak eşikleme, seyrek temsil, Tucker ayrıştırması.

DEDICATION

It is dedicated to:

My Beloved Parents, My Lovely Sisters and My Brother Doğus

ACKNOWLEDGMENT

My journey to achieve my doctoral degree was fruitful because of the different roles played by many people worth acknowledging. My sincere gratitude is first to my supervisors, Prof. Dr. Runyi Yu and Prof. Dr. Hasan Demirel, who both not only enriched my knowledge but encouraged me and scrutinized my research. I must mention my gratitude for their quality advices which enabled me to explore new ideas at the various stages of the way. With their positive attitude, I was spurred to remain confident and inspired to fulfill my big dream.

I am sincerely thankful for the role played by the academic and administrative staff of the Electronic Engineering Department at the Eastern Mediterranean University, for providing me with enabling environment and cooperation, first as a student and eventually as a research assistant.

In no small measure, my love and respect goes to my amiable parents and siblings for their support, patience and care. Without your understanding, my studies would not have been successful.

Last but not the least, to my valuable friends Azadeh Didari, Pouya Bolourchi, Mohammad Harastani, Masoud Moradi, and Lara Mohammad Shhab, I cherish your well-wishes and enjoyable quality time we shared. I will not leave out all my loving friends whose names are too numerous to mention here, I thank you all for how you have decorated my life in different shades.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

1D	One Dimensional
2D	Two Dimensional
3D	Three Dimensional
4D	Four Dimensional
δ	Step Size
ϵ	Tolerance
$\mathbf{\Lambda}$	Sample Set
ρ	Missing ratio
σ	Standard Deviation
σ_e	Estimated Standard Deviation
σ_v	Stabilized Standard Deviation
τ	Regularization Parameter
a	A Scalar
$\boldsymbol{u}$	A Vector
A_i	The ith Factor Matrix
$\boldsymbol{D}_\tau(\cdot)$	A Soft Shrinkage Operator
M	A Matrix
$\boldsymbol{G}$	Core Tensor
$\boldsymbol{P}$	A Linear Operator
$\boldsymbol{T}$	A Tensor
U_i	The ith Orthogonal Factor Matrix

$T_{(i)}$	Mode-i Matricized Version of Tensor $\boldsymbol{T}$
$\circ$	Outer Product
$\times_i$	Mode-i Product of Tensor and Matrix
$\langle \cdot, \cdot \rangle$	The Euclidean Inner Product
$\| \cdot \|_1$	l_1 Norm
$\| \cdot \|_F$	Frobenius Norm
AWGN	Additive White Gaussian Noise
ADMM	Alternating Direction Method of Multipliers
BM3D	Block Matching and 3D Filtering
BM4D	Block Matching and 4D Filtering
CP	(C)NADECOM/(P)ARAFAC
FaLRTC	Fast Low Rank Tensor Completion
FBCP	Fully Bayesian CP Factorization
ICA	Independent Component Analysis
HaLRTC	Highly Accurate Low Rank Tensor Completion
HOSVD	Higher Order Singular Value Decomposition
HSI	Hyperspectral Imaging
LRTC	Low Rank Tensor Completion
MAD	Mean Absolute Difference
MRI	Magnetic Resonance Imaging
NLM	Non-Local Means
NTF	Non-Negative Tensor Factorization
PARAFAC	Parallel Factor

PCA	Principle Component Analysis
PES-TV	Projections onto Epigraph Sets of the TV Function
PSNR	Peak Signal to Noise Ratio
RSE	Relative Square Error
SCA	Sparse Competent Analysis
sgm	Signum Function
SiLRTC	Simple Low Rank Tensor Completion
SNR	Signal to Noise Ratio
SSIM	Structural Similarity Index
STDC	Simultaneous Tensor Decomposition and Completion
SVD	Singular Value Decomposition
SVM	Support Vector Machine
VSt	Variance Stabilization
WTucker	Weighted Tucker

Chapter 1

INTRODUCTION

1.1 Introduction

In recent years, with advances in networking, sensors, and storage technologies bigger and bigger data are being emerged in a wide range of fields in science, including medical imaging, mobile internet, and cloud computing. In order to succeed in this big data era, it becomes crucial to be able to extract useful information from data for efficient processing. Most big data are multidimensional and they can often be represented as multidimensional arrays, which are often referred to as tensors [1, 2].

Since the nineteenth century, tensors which are higher order generalization of matrices and vectors, have played a major part in applied and theoretical frameworks for mathematicians and physicists respectively. In physics, they offer a convenient language for expressing certain natural laws. A famous example is Einstein's theory of general relativity, whose fundamental equations are expressed in terms of tensors. Despite the successful developments of tensor models and analysis in physics and mathematics, the tensor models have undergone a lack of interest in many other scientific and engineering communities. This is spurred in part by pioneering works in psychometrics which applied tensor-based techniques for data analysis purposes [3, 4]. Later in the twentieth century, blind source separation techniques were developed by researchers to

exploit the tensor structure of higher-order cumulants [5, 6], and many works employing tensor models have surfaced in the chemometrics literature [7, 8]. Nowadays, the ever expanding list of applications of tensors embraces problems in signal processing, computer vision, telecommunications, dynamical system modeling and identification, biomedical engineering, and data mining [9, 10, 11, 12, 13, 14, 15].

Nevertheless, the theory of tensors in signal processing should be of great interest following the recent development of multicomponent modern sensor modalities, especially in imagery. Indeed, the data generated from these sensors are fundamentally multidimensional (tensor) objects. Once multidimensional arrays (tensors) have been employed to store multidimensional data, the application will dictate what type of manipulation or post-processing of the data is required. In one and two dimensional (vector and matrix) cases, methodologies have been well established, including sparse component analysis (SCA), principle component analysis(PCA), independent component analysis (ICA), and nonnegative matrix factorization (NMF) [16, 17, 8]. However, the existing framework of vector and matrix algebra is insufficient. In other words, hidden components within multidimensional data cannot be extracted by using the classical matrix and vector techniques. On the other hand, we can utilize hidden components within multidimensional data and keep the multidimensional data as a whole entity if the analysis tools account for the intrinsic multidimensional patterns present, motivating the development of multilinear techniques. That is to say, multilinear algebra is adapted to multidimensional data, as it involves two tensor decompositions [19, 20].

Similar to the popularity of matrix decompositions, tensor decompositions play significant roles in processing, and analysing the tensor data. The two leading frameworks of tensor decomposition to reveal the algebraic structure in the data, are the Tucker decomposition and CANDECOMP/PARAFAC (CP) [2, 21].

Tucker decomposition and the CP decomposition are the most well known tensor decomposition, originally introduced in Psychometrics and later found many applications in diverse areas [2, 4, 18, 20, 21, 22, 23, 24, 25]. From theoretical view, one may consider decompositions as higher order extension of matrix singular value decomposition (SVD). Specifically, CP decomposition decomposes the given tensor as an approximation of rank-1 tensors by taking their weighted sum. In contrast, in Tucker decomposition a tensor can be decomposed into a tensor of dense but smaller size which is represented by the core tensor, while its factor matrices span the subspace occupied by fibers of the tensor data. CP decomposition can be viewed as a special case of Tucker decomposition with super-diagonal core tensor. Compared to CP, a Tucker decomposition has more flexibility due to the core tensor which allows for interactive bases, and has a better generalization ability than CP decomposition for different type of data. However, Tucker decomposition may not decompose the tensor uniquely as CP does. In order to obtain meaningful and unique representation by the Tucker decomposition, orthogonality, sparsity, and non-negativity constraints are often imposed on the factors yielding non-negative tensor factorization (NTF) and sparse non-negative tucker decomposition. Tucker decomposition with orthogonality constraints on the factors, is also known as higher order singular value decomposition (HOSVD) or multilinear

SVD. The HOSVD can be computed by unfolding the tensor in each mode and calculating the singular vectors corresponding to each mode [2, 21, 26, 27].

1.2 Thesis Objective

Most often, in multidimensional data analysis applications, the data is incomplete or contaminated during the acquisition. Therefore, recovering information of the missing or noise free data may tend to pose a problem which is easy to describe, yet unfortunately difficult to solve. This has necessitated the availability of high data quality as essential for several practical applications. To bridge the gap between reality and need, it is necessary to be able to estimate uncorrupted information which remains hidden within the available, possibly corrupted and/or incomplete data.

Data restoration is often formulated as an inverse problem. The objective of restoration is to consider the estimation of the unknown true data $\boldsymbol{T}$ from an observed data defined by

$$\boldsymbol{X} = \boldsymbol{P}(\boldsymbol{T}) + \eta \tag{1.1}$$

where η is the noise accumulated in the acquisition process, and $\boldsymbol{P}$ is a linear operator, a projection in inpainting, or the identity in denoising. However, in many practical problems, ill-conditioning of $\boldsymbol{P}$ precludes the possibility to use an inverse filtering to the measurements. In this manner, prior knowledge of the data is required to regularize the restoration problem. A recently developed approach is to regularize the inverse problem with the sparse-promoting in some transform domain [28, 29, 30, 31, 32].

Recently, a family of powerful approaches referred to as sparse vector and robust low

rank matrix modelling have been introduced. Although, the ℓ_0-norm, and the rank minimization have been proven to be strong global constraints and good measures of sparsity, the optimization problem involving the ℓ_0-norm or the rank minimization is NP-hard in general due to their discrete nature. The ℓ_1-norm and the nuclear norm (also known as the trace norm) are widely used to approximate the ℓ_0-norm and the rank of a matrix, and the resulting problems are convex optimization problems. However, sparse and low-rank modelling typically treats the multidimensional data in the form of vectors or matrices, which are however inadequate in representing the multidimensional data. Meanwhile, the extension of the sparse vector and matrix modelling techniques within the tensorial framework has led to significant performance improvements. This is due to the fact that the proposed approaches can better preserve and employ information about the structure of the multidimensional space the data lie in [20, 33, 34, 35]. To this end, to investigate the possibilities of preserving the multidimensional data structure, this thesis concerns itself with applications of multilinear algebra in the field of reconstruction of multidimensional data.

This thesis aims to solve the restoration problem by employing the HOSVD transform model. We enforce the sparsity using some iterative regularization technique, which is shown to be quite effective for the restoration problem. The proposed restoration method encourages the sparsity of HOSVD by thresholding of the HOSVD core. This thresholding, well adopted since the objective multidimensional data takes the compressible HOSVD model, renders an approximate solution of an ℓ_1 regularized least square problem.

1.3 Problem Statement

Many real world signals are sparse (or approximately sparse) in a certain transform domain[16]. Inspired by a recent application of the matrix singular value decomposition in employing the sparsity, the HOSVD is exploited in the current investigation. The HOSVD is an efficient way for eliciting intrinsic structure of multidimensional data. It offers a simple, adaptive and natural way to exploit sparsity among all dimensions of multidimensional data. The HOSVD decomposes a particular tensor data into the product of a sparse tensor and a few orthogonal matrices, each of which captures the subspace information corresponding to one dimension.

In this thesis, we shall solve the multidimensional data restoration problem. We employ the tensor transform model and exploit the sparsity of the multidimensional signals. Specifically, in this work three problems are considered: 1) tensor inpainting, 2) magnetic resonance image (MRI) denoising, and 3) hyperspectral image (HSI) denoising.

1.4 Thesis Contributions

The main contributions made in this thesis are listed below:

1. In the first contribution of this work we employ the iterative regularization scheme for tensor inpainting. The rationale of this approach is based on an enhanced sparse representation in HOSVD domain and it uses the iterative regularization procedure for inpainting. Improved performances of this algorithm are demonstrated in our experiments on three dimensional (3D) tensors, taken from multichannel (color) images, video sequences, and magnetic resonance images.

2. Despite the success of magnetic resonance imaging techniques in many applications, acquisition noise is still a limiting factor for the quality and hence the usefulness of the techniques. The main goal of the second contribution is investigating and improving the application of the iterative higher order singular value decomposition framework to denote the MRIs. The proposed algorithm forms a single tensor from the noisy data. This tensor undergoes an HOSVD, where its sparse representation coefficients are calculated with respect to a set of orthogonal basis matrices. Denoising is achieved by iteratively applying soft thresholding on the calculated sparse representation coefficients. The proposed algorithm is further enhanced with a post-process of Wiener filtering.

3. The hyperspectral data cube is considered as a three-order tensor that is able to jointly treat both the spatial and spectral modes. Noise in HSI can degrade the visual quality and limit the applicability of computerized analysis processes. Hence as the third contribution we consider the denoising of the HSIs to improve the performance of the subsequent applications. In this work, we not only use the proposed iterative higher order singular value decomposition, but go one step further and propose the new iterative denoising method which utilizes the advantages of the patch-based HOSVD sparse model and the iterative regularization technique.

1.5 Thesis Outline

This thesis is organized as follows: Chapter 1 describes the thesis introduction, aims, and outline of the research. Chapter 2 presents some notations, and an introduction to tensor definition and tensor decompositions. Chapter 3 introduces an iterative method

for tensor inpainting based on higher order singular value decomposition. Noise removal from MR images via iterative regularization based on higher order singular value decomposition is studied in Chapter 4. Furthermore, Chapter 5 considers reducing the noise from HSIs via patch-based higher order singular value decomposition sparse model. In Chapter 6, conclusions and possible future works are presented.

Chapter 2

HIGHER ORDER SINGULAR VALUE DECOMPOSITION SHRINKAGE AND ITS RELATION WITH ITERATIVE REGULARIZATION

2.1 Introduction

There are various ways to define tensor depending on the context, in which it is used. In most of the applications, the tensor is employed as a structure for sorting and organizing data. Therefore, this chapter aligns on providing a foundation to understand the subsequent chapters. In this chapter, we define the algebraic operation and the notation used throughout the thesis. The fundamental properties of two of the important tensor decompositions, the Tucker and the CANDECOM/PARAFAC, are also presented and discussed.

2.2 Notation

The term tensor refers to a multidimensional array, also known as multiway or multimode array. The modes of the tensor are referred to by the number of its dimensions. They are also called orders or ways. A vector is a tensor of order one, and is usually described by one dimensional (1D). A matrix is described by a two-dimensional (2D) array is a tensor of order two. We use a d dimensional array to describe a d order tensor. Throughout this thesis, the distinguish between multidimensional tensors, matrices, vectors and scalars will be shown by their representation: d order tensors are

bold-face upper case letters $(\boldsymbol{T}, \boldsymbol{S}, \dots)$, matrices are upper case letters $(M, U, \dots)$, vectors are bold-face lower case letters $(\boldsymbol{u}, \boldsymbol{v}, \dots)$, and scalars are lower case letters $(a, b, \dots)$. For instance, Let $\boldsymbol{T}$ be a dth order tensor with entries $t_{i_1,i_2,\dots,i_d}$.

Fibers are the higher order analogue of matrix rows and columns. The fiber of a tensor can be obtained by fixing all tensor indices except one, i.e, $\boldsymbol{T}_{:,i_2,\dots,i_d}$, $\boldsymbol{T}_{i_1,:,\dots,i_d}$, $\dots$, $\boldsymbol{T}_{i_1,i_2,\dots,i_{d-1},:}$ respectively. Fibers are always assumed to be column vector. Tensor slices are two-dimensional sections of a tensor, which can be computed by fixing all tensor indices except two, i.e, $\boldsymbol{T}_{:,:,i_3,\dots,i_d}$,$\boldsymbol{T}_{i_1,:,:,i_4,\dots,i_d}$,$\dots$,$\boldsymbol{T}_{i_1,i_2,\dots,i_{d-2},:,:}$. Figure 2.1 shows tensor fibers (column, row, and tube) and slices (horizontal, lateral, and frontal) of a 3rd order tensor [2, 21].

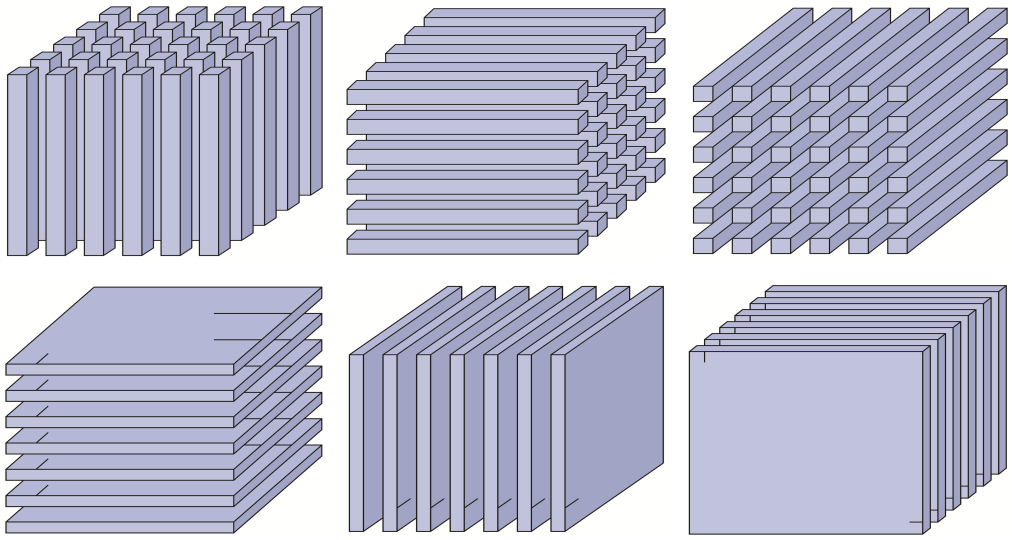


Figure 2.1: An example of tensor fibers (column, row, and tube; respectively) (row 1) and slices (horizontal, lateral, and frontal sides; respectively) (row 2).

In order to work with tensors, it is often convenient to transform them to matrices, a process known as matricisation; also known as unfolding or flattening. A tensor

of order d can be unfolded in d various ways. The mode-i unfolding of a tensor $\boldsymbol{T}$ is symbolized by the matrix $T_{(i)}$. The matricisation reorders the entries of the tensor $\boldsymbol{T}$ into a matrix from a given mode. The dimension of $T_{(i)}$ is $\mathbf{R}^{n_i \times J}$, where $J = n_1, \ldots, n_{i-1}, n_{i+1}, \ldots, n_d$. Figure 2.2 shows the way of flattering a 3rd order tensor $\boldsymbol{T} \in \mathbf{R}^{3\times 2\times 2}$ into three modes unfolding matrices, i.e. $T_{(1)}$, $T_{(2)}$, and $T_{(3)}$ [2, 21]. In mathematical terms, the tensor element $(i_1, \ldots, i_{n-1}, i_n, i_{n+1}, \ldots, i_d)$ maps to matrix element (i_n, j), given by the expression:

$$j = 1 + \sum_{l=1, l\neq n}^{d} (i_l - 1)J_l \quad with \quad J_l = \prod_{m=1, m\neq n}^{l-1} n_m \tag{2.1}$$

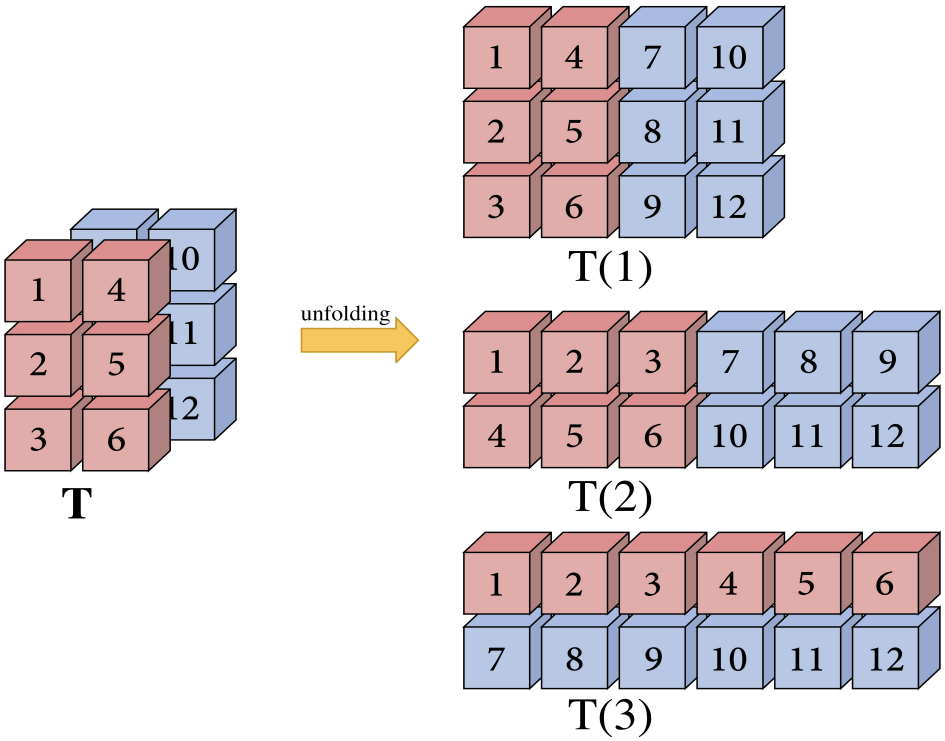


Figure 2.2: An example of Tensor unfolding.

Tensors can be multiplied by a matrix, the product between a tensor and a matrix can be expressed by employing the concept of mode-i product. The mode-i matrix product between a tensor $\boldsymbol{T} \in \mathbf{R}^{n_1 \times n_2 \times \cdots \times n_d}$ and a matrix $U \in \mathbf{R}^{r_i \times n_i}$, denoted by

$\boldsymbol{T} \times_i U \in \mathbf{R}^{n_1 \times \cdots \times r_i \times \cdots \times n_d}$, returns a new tensor of size $n_1 \times \cdots \times r_i \times \cdots \times n_d$; the operator $\times_i$ denotes the mode-i product of tensor with the matrix. The scalar or inner product of two given tensors $\boldsymbol{X}, \boldsymbol{T}$ is the sum of the products of their entries which is defined as: $\langle \boldsymbol{X}, \boldsymbol{T} \rangle = \sum_{i_1=1}^{n_1} \sum_{i_2=1}^{n_2} \cdots \sum_{i_d=1}^{n_d} (x_{i_1,i_2,\ldots,i_d}) \times (t_{i_1,i_2,\ldots,i_d})$. Based on the definition of the inner product the Frobenius norm of dth order which is the square root of the sum of the squared absolute values of the elements of the tensor, is denoted as [2, 21]

$$\|\boldsymbol{T}\|_F = \Big(\sum_{i_1,i_2,\ldots,i_d} (t_{i_1,i_2,\ldots,i_d})^2 \Big)^{1/2} \tag{2.2}$$

2.3 Tensor Decompositions

In recent decays, conception of matrix decomposition has extensively contributed to development of science and engineering. As a matter of fact, advanced linear algebra provided the researchers with numerical solutions which decompose a matrix into more representative matrices. Several techniques such as QR, LU and SVD have been successfully practiced in engineering applications. More recently, multilinear algebra has opened a new horizon to the field by introducing tensors. Although tensor decomposition has not yet been evolved as much as matrix decomposition has, this concept has attracted the attention of researcher as a compelling alternative to matrix decomposition. Particularly, two major tensor decomposition trends namely Tucker decomposition and CANDECOM/PARAFAC are known as well-developed notions in modern science [2, 21, 36, 37]. Before continuing, the traditional computation of SVD will be reviewed and then discussed regarding the Tensor decompositions.

2.3.1 Matrix Singular Value Decomposition (SVD)

A basic concept in linear algebra is the singular valued decomposition (SVD). The SVD of a matrix $M \in \mathbf{R}^{n_1 \times n_2}$ is a well-known, rank revealing factorization. Consider the singular value decomposition of a matrix M of rank R [36, 37]:

$$M = USV^T \tag{2.3}$$

where $U \in \mathbf{R}^{n_1 \times n_1}$, and $U \in \mathbf{R}^{n_2 \times n_2}$ are orthogonal matrices (i.e., $UU^T = I$, $VV^T = I$), S is a non-negative diagonal matrix. The diagonal elements of S, denoted as s_i are called the singular values, and are arranged in decreasing order. The columns of U and V, denoted by $\boldsymbol{u_i}$ and $\boldsymbol{v_i}$ are called as left and right singular vectors of M respectively. The number of nonzero singular values is equal to the rank of the matrix, i.e. to the number of linearly independent columns or rows of the matrix. The SVD enables us to write the M as the sum of rank-1 terms:

$$M = s_1 u_1 {v_1}^T + \cdots + s_R u_R {v_R}^T \tag{2.4}$$

Note that, the rank $R = \mathrm{rank}(M)$ can also be defined as the minimum number of rank-1 terms whose sum is equal to M. The SVD of a matrix is unique (assuming all the singular values are distinct) only because it has the orthogonality constraints and the ordered singular values.

The so called nuclear norm of a matrix M which appears in several problems related to low rank matrix completion or rank minimization problems, is defined in terms of singular values [36, 37]:

$$\|M\|_* = \sum_{i=1}^{R} s_i \tag{2.5}$$

Bearing in mind the above discussion, we shall present in the following two principle

tensor decompositions that can be considered as two different generalizations of the SVD.

2.3.2 CANDECOMP/PARAFAC

The CP decomposition of a tensor is one of the important notations of tensor decomposition, which leads to the definition of CP rank. The idea of the CP decomposition was first proposed by Hitchcock (1927) [38, 39]. However, the CP decomposition became more popular mostly in the psychometrics community with the works of Carrol and Chang (1970)[40], who introduced this decomposition as the canonical decomposition (CANDECOM), and parallel factors (PARAFAC) by Harshman (1970) [41]. This decomposition is currently known as (CP), as the abbreviation of both names. The CP decomposition is based on the fact that a given tensor can be written as the sum of several rank-1 tensors. Given a 3rd order tensor $\boldsymbol{T} \in \mathbf{R}^{n_1 \times n_2 \times n_3}$, the CP decomposition can be approximated as a sum of R rank-1 tensors,

$$\boldsymbol{T} \cong \sum_{r=1}^{R} (\boldsymbol{a_r} \circ \boldsymbol{b_r} \circ \boldsymbol{c}_r) \tag{2.6}$$

where $\boldsymbol{a}_r \in \mathbf{R}^{n_1}$, $\boldsymbol{b}_r \in \mathbf{R}^{n_2}$, and $\boldsymbol{c}_r \in \mathbf{R}^{n_3}$, and $\circ$ is the tensor outer product. In this model R is called the rank of $\boldsymbol{T}$. Since, the CP of a tensor expose its rank, it is said to be a rank revealing decomposition or rank retaining decomposition. The computation of the CP decomposition requires knowing the amount R of rank-1 tensors to involve in the sum (equation (2.6)). Most algorithms calculate many alternative decompositions to find the one that fits the best. When denoted in element-wise form, equation (2.6) can be rewritten as follows:

$$\begin{aligned} &t_{i_1,i_2,i_3} \cong \textstyle\sum_{r=1}^{R} (a_{i_1 r} b_{i_2 r} c_{i_3 r}), \\ &i_1 = 1, 2, \ldots, n_1, \quad i_2 = 1, 2, \ldots, n_2, \quad i_3 = 1, 2, \ldots, n_3 \end{aligned} \tag{2.7}$$

A key feature of this decomposition, which is often considered as its prime advantage over matrix decompositions, is its uniqueness up to permutation and scaling under the usually fulfilled conditions [42]. A more general framework for uniqueness has been recently presented in [43, 44]. Figure 2.3 illustrates the graphical representation of the CP decomposition.

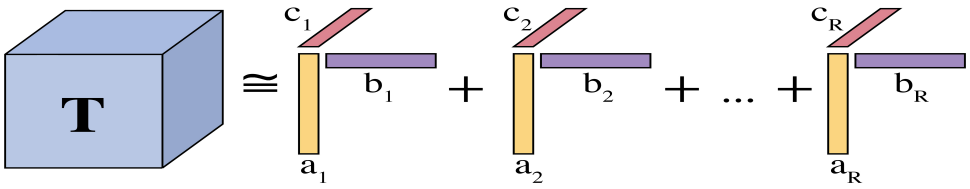


Figure 2.3: CP decomposition of a 3rd order tensor.

2.3.3 Tucker Decomposition

It is noted that, the CP decomposition can be seen about a tensor generalization of matrix SVD. There is another generalization of the matrix SVD which is named Tucker decomposition. In 1963, the Tucker decomposition of a higher order tensor is introduced by Tucker that reveals the multilinear rank [3]. As with the CP decomposition, tensor models based on the Tucker decomposition also find many practical applications [2, 21].

Tucker decomposition states that each dth order tensor $\boldsymbol{T} \in \mathbf{R}^{n_1 \times n_2 \times \cdots \times n_d}$ can be expressed in terms of one core tensor $\boldsymbol{S}$ and d factor matrices as follows:

$$\boldsymbol{T} = \boldsymbol{S} \times_1 A_1 \times_2 A_2 \cdots \times_d A_d \tag{2.8}$$

where $A_i \in \mathbf{R}^{n_i \times r_i}$ is factor matrix. The tensor $\boldsymbol{S}$ is called as core tensor. Note that the

core tensor does not always need to have the same dimensions as $\boldsymbol{T}$. Without further restrictions, infinitely many Tucker decomposition of a tensor exist.

Unlike the CP decomposition, the Tucker decomposition is not unique. However, in the special case where the core tensor has nonzero elements only on the superdiagonal, the Tucker decomposition is reduced uniquely to the CP under some mild condition. To obtain meaningful and unique representation by the Tucker decomposition, orthogonality, sparsity, and nonnegativity constraints are often imposed on hidden factors and the core tensor of the Tucker decomposition. Recently, the Tucker model has been further developed by a setting constraint of the orthogonality on the factor matrices and on the slices of the core tensors, which is called as an orthogonal Tucker decomposition. The purpose of these developments is to generalize the SVD matrix decomposition to tensor analysis, so it can also be defined as higher order singular value decomposition [2, 21]. In the HOSVD of a tensor $\boldsymbol{T}$, factor matrices $(A_1, A_2, \ldots, A_d)$, must be orthogonal and will be from now on represented with letters $(U_1, U_2, \ldots, U_d)$. The HOSVD mode can be written as:

$$\boldsymbol{T} = \boldsymbol{S} \times_1 U_1 \times_2 U_2 \cdots \times_d U_d \tag{2.9}$$

where $U_i \in \mathbf{R}^{n_i \times r_i}$ is an orthogonal matrix which captures structural information of $T_{(i)}$, and $\boldsymbol{S} \in \mathbf{R}^{r_1 \times r_2 \times \cdots \times r_d}$ is the so-called core tensor which describes contributions of $\boldsymbol{T}$ in the structural of all U_i's. The values $(r_1, r_2, \ldots, r_i)$ correspond to the ranks of the different matrix unfolding of $\boldsymbol{T}$ along the different modes. Furthermore, the core tensor $\boldsymbol{S}$ possesses the following properties:

1. In multilinear algebra as well as in matrix algebra the Frobenious norm is invari-

ant, i.e., $\|\boldsymbol{T}\|_F = \|\boldsymbol{S}\|_F$.

2. The subtensors $\boldsymbol{S}_{n_i=\alpha}$ obtained by fixing the ith index to α, have the properties of:

   (a) all-orthogonality: two subtensors $\boldsymbol{S}_{n_i=\alpha}$ and $\boldsymbol{S}_{n_i=\beta}$ are orthogonal for all possible value of i, α and β subject to $\alpha \neq \beta$, i.e., $\langle \boldsymbol{S}_{n_i=\alpha}, \boldsymbol{S}_{n_i=\beta} \rangle = 0$ when $\alpha \neq \beta$;

   (b) Ordering: The subtensors are ordered decreasingly according to their Frobenius norms as $\|\boldsymbol{S}_{n_i=1}\|_F \geq \|\boldsymbol{S}_{n_i=2}\|_F \geq \cdots \geq \|\boldsymbol{S}_{n_i=d}\|_F \geq 0$.

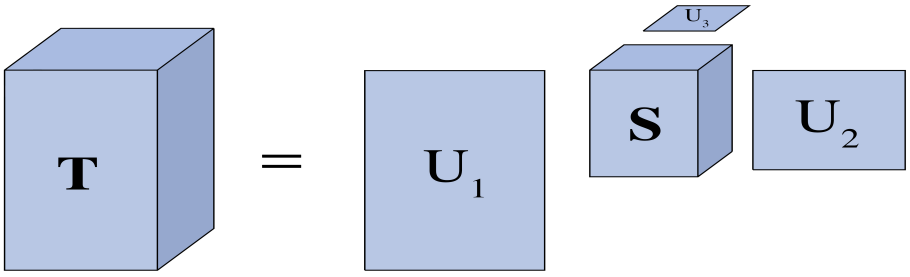


Figure 2.4: HOSVD decomposition of a 3th order tensor.

The HOSVD can also be represented as an expansion of mutually orthogonal rank-1 tensors,

$$\boldsymbol{T} = \sum_{i_1=1}^{r_1} \sum_{i_2=1}^{r_2} \cdots \sum_{i_d=1}^{r_d} s_{i_1,i_2,\ldots,i_d} (\boldsymbol{u}_1^{i_1} \circ \boldsymbol{u}_2^{i_2} \cdots \circ \boldsymbol{u}_d^{i_d}) \tag{2.10}$$

where $s_{i_1,i_2,\ldots,i_n}$ are the entries of the core tensor which show the level of interaction between the different components, $\boldsymbol{u}_1^{i_1}, \boldsymbol{u}_2^{i_2}, \ldots, \boldsymbol{u}_d^{i_d}$ are vectors, and $\circ$ represents vector outer product. In Figure 2.4 the graphical representation of the HOSVD decomposi-

tion is given. An interesting characteristic of the HOSVD is that when employed for tensor of order two, the HOSVD reduces to the traditional SVD.

For tensors, there is not a unique concept of rank. Although the tensor rank is defined as the matrix rank, a particularity of the rank of a tensor is that there is not a straightforward way of calculation; it is an NP hard problem.

2.4 HOSVD Shrinkage and its Relation with Iterative Regularization

In this work, we enforce the sparsity using iterative regularization technique, which is shown to be very effective for our inverse problems. Here, the corrupted data is filtered by using the HOSVD sparse model and iterative regularization technique.

Suppose that $\boldsymbol{X}^k (k = 1, 2, \ldots, K)$ stands for the reconstructed result in the kth iteration and K is the number of iterations. Moreover, the first step can be put into the iterative framework by choosing initial values. Let $\boldsymbol{Y}^0$ be an initial tensor with the given measurement $\boldsymbol{X} = \boldsymbol{P}(\boldsymbol{T}) + \eta$. Now, we are proposing the following iterative regularization procedure:

1. The tensor $\boldsymbol{Y}$ at $k-1$ iteration is transformed into the HOSVD as equation (2.9), to obtain the HOSVD coefficients:

$$\boldsymbol{Y}^{k-1} = \boldsymbol{S} \times_1 U_1 \times_2 U_2 \cdots \times_d U_d \tag{2.11}$$

2. A soft shrinkage operator $\boldsymbol{D}_\tau(\boldsymbol{S})$ (with parameter τ) is applied on every coefficients of the core tensor. That is, the shrinkage operator is defined via element operation as follow:

$$\boldsymbol{D}_\tau(\boldsymbol{S}) = ([s_{i_1,i_2,\ldots,i_d}]_\tau) \tag{2.12}$$

in which the soft thresholding is given as

$$\left[s_{i_1,i_2,\ldots,i_d}\right]_\tau = \begin{cases} s_{i_1,i_2,\ldots,i_d} - \operatorname{sgn}(s_{i_1,i_2,\ldots,i_d})\tau, & \text{if } |s_{i_1,i_2,\ldots,i_d}| > \tau \\ 0, & \text{otherwise.} \end{cases} \tag{2.13}$$

3. Then, the kth approximate is obtained by

$$\boldsymbol{X}^k = \boldsymbol{D}_\tau(\boldsymbol{S}) \times_1 U_1 \times_2 U_2 \cdots \times_d U_d \tag{2.14}$$

And the tensor on which HOSVD is performed in the next iteration is updated by

$$\boldsymbol{Y}^k = \boldsymbol{Y}^{k-1} - \delta\, \boldsymbol{P}(\boldsymbol{X}^k - \boldsymbol{T}) \tag{2.15}$$

The parameter $\delta > 0$ is the scalar step size at each iteration k. Thus, the proposed iterative method for tensor data restoration can be formed as Algorithm 1.

Algorithm 1 The Iterative Regularization Algorithm for Tensor Data Restoration

1: **INPUT:** step size δ, threshold τ, and the maximum iteration count K.
2: **Initialize** $\boldsymbol{Y}^0$ by $\boldsymbol{Y}$.
3: **Repeat**
Compute HOSVD of $\boldsymbol{Y}^k = \boldsymbol{S} \times_1 U_1 \times_2 U_2 \cdots \times_d U_d$.
Perform soft thresholding $\boldsymbol{D}_\tau(\boldsymbol{S}) = ([s_{i_1,i_2,\ldots,i_d}]_\tau)$.
Construct $\boldsymbol{X}^k = \boldsymbol{D}_\tau(\boldsymbol{S}) \times_1 U_1 \times_2 U_2 \cdots \times_d U_d$.
Update $\boldsymbol{Y}^k = \boldsymbol{Y}^{k-1} - \delta\, \boldsymbol{P}(\boldsymbol{X}^k - \boldsymbol{T})$.
Until the maximum iteration count K is met.
$\hat{\boldsymbol{X}} \leftarrow \boldsymbol{X}^k$
4: **OUTPUT:** The recovered result $\hat{\boldsymbol{X}}$.

Chapter 3

AN ITERATIVE METHOD FOR TENSOR INPAINTING BASED ON HIGHER ORDER SINGULAR VALUE DECOMPOSITION

3.1 Introduction

In this chapter, we consider the problem of tensor (i.e., multidimensional array) inpainting. By using HOSVD, we propose an iterative algorithm that performs soft-thresholding on entries of the core tensor and then reconstructs via the orthogonal matrices. An inpainted tensor is obtained at the end of the iteration. Simulations conducted over color images, video frames, and MR images validate that the proposed algorithm is competitive with state-of-the-art completion algorithms.

3.1.1 Related Work

Inpainting usually refers to the process of reconstructing missing elements or removing unwanted objects of an image or a video. The key idea is to identify and utilize relationships between the missing elements and the known ones [45, 46]. Approaches to inpainting can be categorized into local [45, 47, 48] and non-local approaches [49, 50, 51, 52].

Recently, sparse vector recovery and low rank matrix completion have been widely used in image inpainting like the other field of image analysis and computer vision [28, 29, 37, 53, 54]. The notion of rank minimizing has been naturally exploited in non-

local approaches; and the so-called nuclear norm is used as an alternative to the standard rank in order to render tractable optimization problems for matrix completion [36, 37, 53] and for image restoration [53, 54]. Approximate solutions can be obtained via iterative soft singular value thresholding.

In many applications, one needs to deal with multi-way data represented by tensors. During the last several years, there has been a rapidly growing interest in the recovery of unknown low-rank tensors [55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66]. Recent research has addressed tensor completion as a matrix completion problem [56, 58, 59, 60, 61, 62]. This is customarily done by recasting the tensor completion into minimizing rank(s) of some unfolded matrices. Accordingly, these approaches are essentially matrix-based, thus do not effectively exploit the structure of a tensor. Another line of research in low-rank tensor completion is based on the tensor decompositions [63, 67, 68, 69, 70].

Recall that with HOSVD, each d-dimensional tensor $\boldsymbol{T}$ can be expressed in terms of one core tensor (sparse tensor) $\boldsymbol{S}$ and d factorizing matrices U_i of appropriate sizes as in equation (2.9), where all matrices are orthogonal and carry directional information of $\boldsymbol{T}$, and the entries of $\boldsymbol{S}$ represent relative contributions of these matrices. It is noted that, unlike the case with matrix singular value decomposition, the core tensor $\boldsymbol{S}$ is not necessarily diagonal and can have negative-valued entries as well. Approaches using the HOSVD for tensor completion include the works in [57, 60, 61, 64, 66]. Note that they all involve modifying the orthogonal factorization matrices. As a result, the

directional structure of the tensor is also altered.

3.1.2 Tensor Completion Methods

Based on unfolding tensors, Liu *et al.* [56] proposed a low rank tensor completion algorithm for tensor inpainting. This algorithm is considered as the baseline of the tensor completion methods. It minimizes the nuclear norms of matrices obtained from unfolding along all modes. In a subsequent work, Liu *et al.* [60] developed three algorithms, namely simple low rank tensor completion (SiLRTC), fast low rank tensor completion (FaLRTC), and highly accurate low rank tensor completion (HaLRTC). The SiLRTC uses a block coordinate descent method which employs a relaxation technique to separate the dependent relationships, and guaranteed to achieve a globally optimal solution. The FaLRTC algorithm, which uses a smoothing scheme to convert the original non-smooth problem into a smooth one, is developed to improve the convergence speed. Finally, the HaLRTC algorithm employs the alternating direction method of multipliers (ADMMs) to the low rank tensor completion problems. It has been shown than HaLRTC can achieve higher recovery accuracy, but it is slower than FaLRTC. With the unfolding scheme, low rank tensor completion is developed by using Douglas Rachford splitting technique in [58]. In these methods, the rank or rank approximation is used as a powerful tool to capture the global information. It is noted that these approaches to completion are essentially matrix approaches rather than tensor ones, which do not effectively exploit the structure of a tensor.

Along the line of tensor decomposition, Filipovic and Jukic [66] proposed a weighted Tucker model (WTucker) for low rank tensor completion based on pre-specified mul-

tilinear rank, in which the rank of the tensor can be over- or underestimated. Chen et al. [57] presented a completion method named simultaneous tensor decomposition and completion (STDC) based on Tucker decomposition which employs nuclear norm minimization for the factor matrices. Furthermore, in STDC method a graph Laplacian term is utilized for characterizing the underlying structure between factor matrices. STDC can be considered as a tensor extension of low-rank and smooth matrix completion.

On the other hand, Zhao *et al.* [63] proposed a tensor completion method based on Bayesian tensor factorization model with mixture prior assumption for the factor matrices under CP framework (FBCP). The FBCP algorithm is characterized as a tuning parameter-free approach which can effectively avoid parameter selections. The significant advantage of FBCP is that it can automatically estimate the tensor rank. In contrast, other exiting CP algorithms require the tensor rank to be manually specified.

3.2 Proposed Algorithm

In this section, we describe the process of the proposed algorithm in detail [71].

The completion problem studied in this thesis can be stated as follows: consider a 3rd order tensor $\boldsymbol{T} \in R^{n_1 \times n_2 \times n_3}$. Assume that only a portion of its elements are known, and the values of the rest elements are unknown (missing or altered). Let $\boldsymbol{\Lambda} \subset \bigotimes_{i=1}^{3} \{1, 2, 3\}$ denote the set of all locations of corresponding observed elements. Our problem is to complete tensor $\boldsymbol{T}$ by determining the values of these missing elements.

We now employ an iterative core tensor thresholding (CTT) algorithm to address this problem. At each iteration, we first perform the HOSVD; the core tensor obtained is then modified via soft thresholding. While preserving the structural information and acknowledging contribution of large entries (in terms of magnitude) in the core tensor, we further obtain an approximate solution of the inpainting problem by synthesizing the HOSVD using the modified core tensor.

As in section (2.4), we note that $\boldsymbol{P}_{\boldsymbol{\Lambda}}\colon R^{n_1\times n_2\times n_3} \longrightarrow R^{n_1\times n_2\times n_3}$ is the orthogonal projection on $\boldsymbol{\Lambda}$, i.e., the (i_1, i_2, i_3)th element of $P_{\boldsymbol{\Lambda}}(\boldsymbol{X})$ is equal to x_{i_1,i_2,i_3} if $(i_1, i_2, i_3) \in \boldsymbol{\Lambda}$, or is equal to 0 otherwise, and $\boldsymbol{Y}^0$ is initialized by the measurements $\boldsymbol{X}$ with missing value elements. Then the iteration scheme can be expressed as:

$$\begin{aligned} \boldsymbol{X}^k &= \boldsymbol{D}_\tau(\boldsymbol{Y}^{k-1}) \\ \boldsymbol{Y}^k &= \boldsymbol{Y}^{k-1} - \delta\, \boldsymbol{P}_{\boldsymbol{\Lambda}}(\boldsymbol{X}^k - \boldsymbol{T}) \end{aligned} \tag{3.1}$$

The pseudo-code and the flowchart of the iterative regularization algorithm for tensor inpainting are given in Algorithm 2 and Figure 3.1, respectively.

Algorithm 2 The CTT algorithm for Tensor Inpainting

1: **INPUT:** Sample set $\boldsymbol{\Lambda}$, observed entries $\boldsymbol{P}_{\boldsymbol{\Lambda}}(\boldsymbol{T})$, step size δ, threshold τ, tolerance ϵ, and iteration number K.
2: **Initialize** $\boldsymbol{Y}^0$ by $\boldsymbol{P}_{\boldsymbol{\Lambda}}(\boldsymbol{T})$.
3: **for k= 0, 2, ..., K do**
Compute HOSVD of $\boldsymbol{Y}^k = \boldsymbol{S} \times_1 U_1 \times_2 U_2 \times_3 U_3$.
Calculate $\boldsymbol{D}_\tau(\boldsymbol{S})$ by applying soft thresholding on $\boldsymbol{S}$.
Construct $\boldsymbol{X}^k = \boldsymbol{D}_\tau(\boldsymbol{S}) \times_1 U_1 \times_2 U_2 \times_3 U_3$.
If $\|\boldsymbol{P}_{\boldsymbol{\Lambda}}(\boldsymbol{X}^k - \boldsymbol{T})\|_F / \|\boldsymbol{P}_{\boldsymbol{\Lambda}}(\boldsymbol{T})\|_F \leq \epsilon$ then break.
Update $\boldsymbol{Y}^k = \boldsymbol{Y}^{k-1} - \delta\, \boldsymbol{P}_{\boldsymbol{\Lambda}}(\boldsymbol{X}^k - \boldsymbol{T})$.
Until the maximum iteration count K is met.
$\hat{\boldsymbol{X}} \leftarrow \boldsymbol{X}^k$
4: **OUTPUT:** The inpainted result $\hat{\boldsymbol{X}}$.

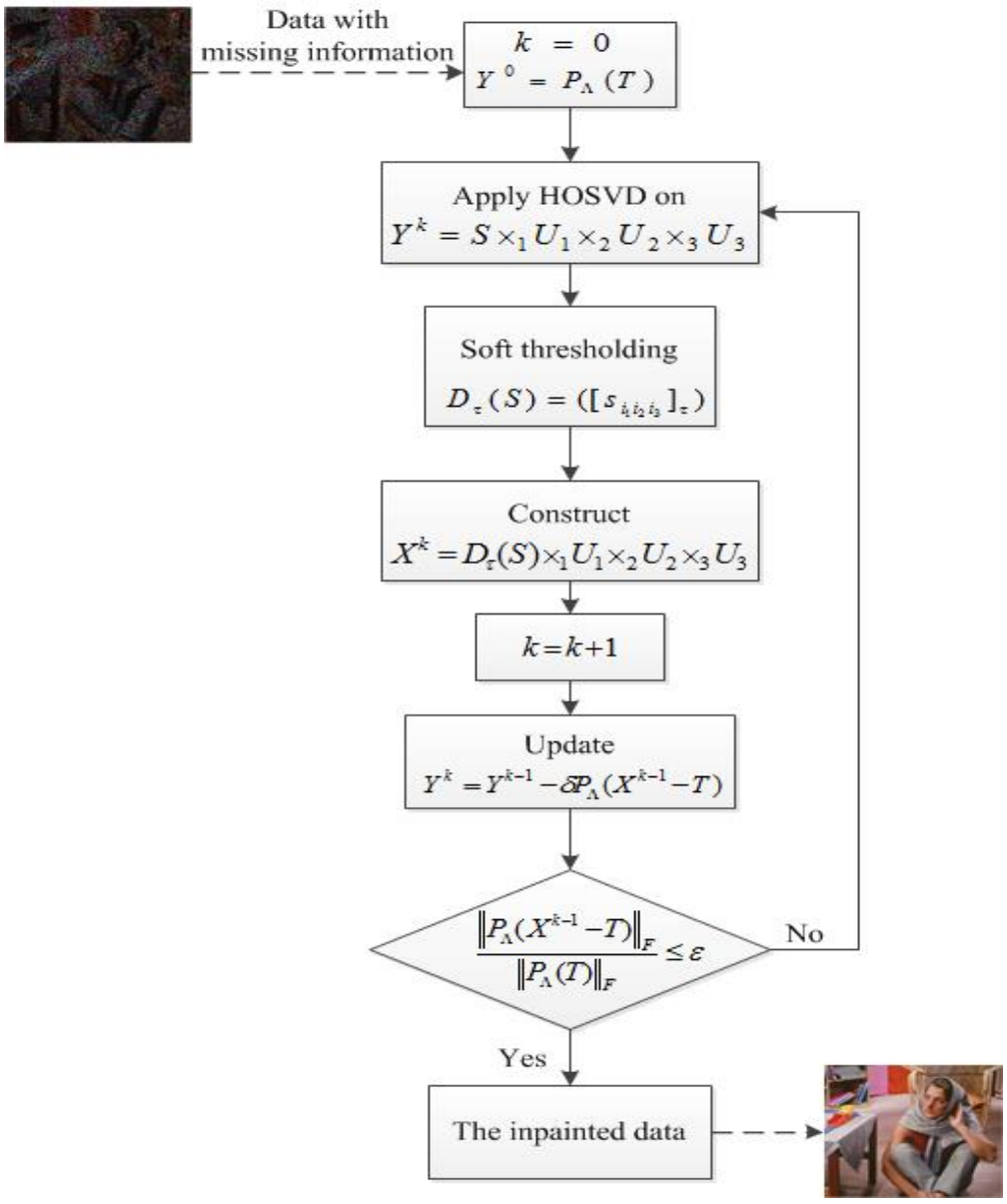


Figure 3.1: Flowchart for tensor inpainting.

In the proposed CTT approach, instead of unfolding or separating the slices of the tensor, the tensor is projected onto the HOSVD transform. In this approach the iterative soft thresholding is performed to regularize the result of the inpainting, leading to a sparse representation of the core tensor. This sparsity is imposed by applying soft thresholding on the representation coefficients, which were calculated with respect to a transform basis. This corresponds to the minimization of the ℓ_1 norm of the core

tensor coefficients.

The main step of the proposed approach is to modify the elements of the core tensor via thresholding. By soft thresholding coefficients of the core tensor and keeping those of orthogonal matrices, we achieve minimization of tensor $\boldsymbol{T}$ through $\boldsymbol{S}$ and preserving its directional structure at the same time. As a result, the coherence between the slices (i.e., the information along all dimensions) is used. This implies that the proposed approach takes advantage of global property of the observed data properly.

The iteration scheme may be regarded as an extension of the singular value thresholding (SVT) algorithm for matrix completion [37]. Note that this extension is not straightforward as the core tensor $\boldsymbol{S}$ now is not diagonal and it may contain negative elements. Nevertheless, we can still consider the following optimization problem in the framework of sparse regularization:

$$\hat{\boldsymbol{X}} = \arg\min_{\boldsymbol{X}} \frac{1}{2}\|\boldsymbol{X} - \boldsymbol{T}\|_F^2 + \tau\psi(\boldsymbol{X}) \tag{3.2}$$

where $\hat{\boldsymbol{X}}$ is the approximation variable, τ is the regularization parameter, and $\psi(\boldsymbol{X})$ is a sparse regularization term. Recall that the Frobenius norm is invariant under HOSVD (i.e., $\|\boldsymbol{X}\|_F = \|\boldsymbol{S}\|_F$), and it is consistent with the ℓ_1 norm [34, 35, 68, 72]. Since the core tensor $\boldsymbol{S}$ is a sparse representation of $\boldsymbol{X}$ in the HOSVD domain [34, 35, 68, 72], the ℓ_1 norm of $\boldsymbol{S}$ can thus be used for regularization. Accordingly, the proposed CTT algorithm would lead to a tensor with fewer rank-1 terms and yield an approximate solution to the sparse problem given below [71]:

$$\min_{\boldsymbol{X}} \tfrac{1}{2}\|\boldsymbol{X}-\boldsymbol{T}\|_F^2 + \tau\|\boldsymbol{X}\|_1$$
$$\text{such that } P_{\boldsymbol{\Lambda}}(\boldsymbol{X}) = P_{\boldsymbol{\Lambda}}(\boldsymbol{T}). \tag{3.3}$$

3.3 Experimental Results

In this section, the performance of the proposed CTT algorithm is evaluated and compared with some state-of-the-art inpainting algorithms. These are FaLRTC [65], HaLRTC [65], WTucker [66], STDC [57], and FBCP [63]. The FaLRTC, HaLRTC, WTucker, and STDC algorithms are included in the comparison as it is based on the Tucker decomposition; the FBCP algorithm is included as the state-of-the-art method based on CP decomposition.

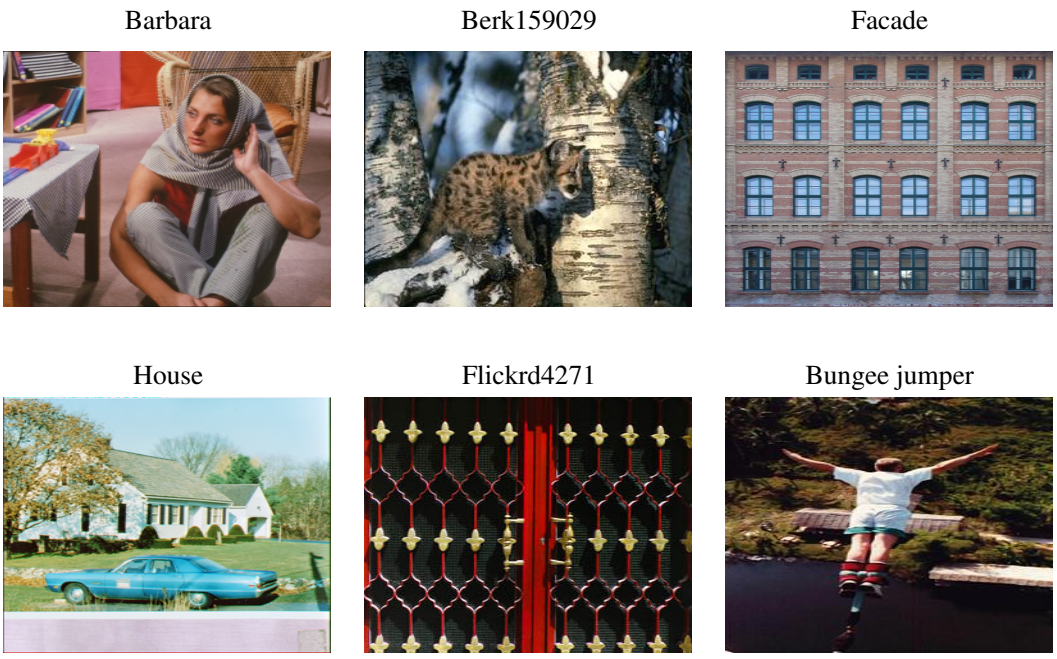


Figure 3.2: Test images.

Simulations are conducted over color images, video sequences, and MR images. Performance evaluation is carried out in terms of two error based measures: 1) the peak-

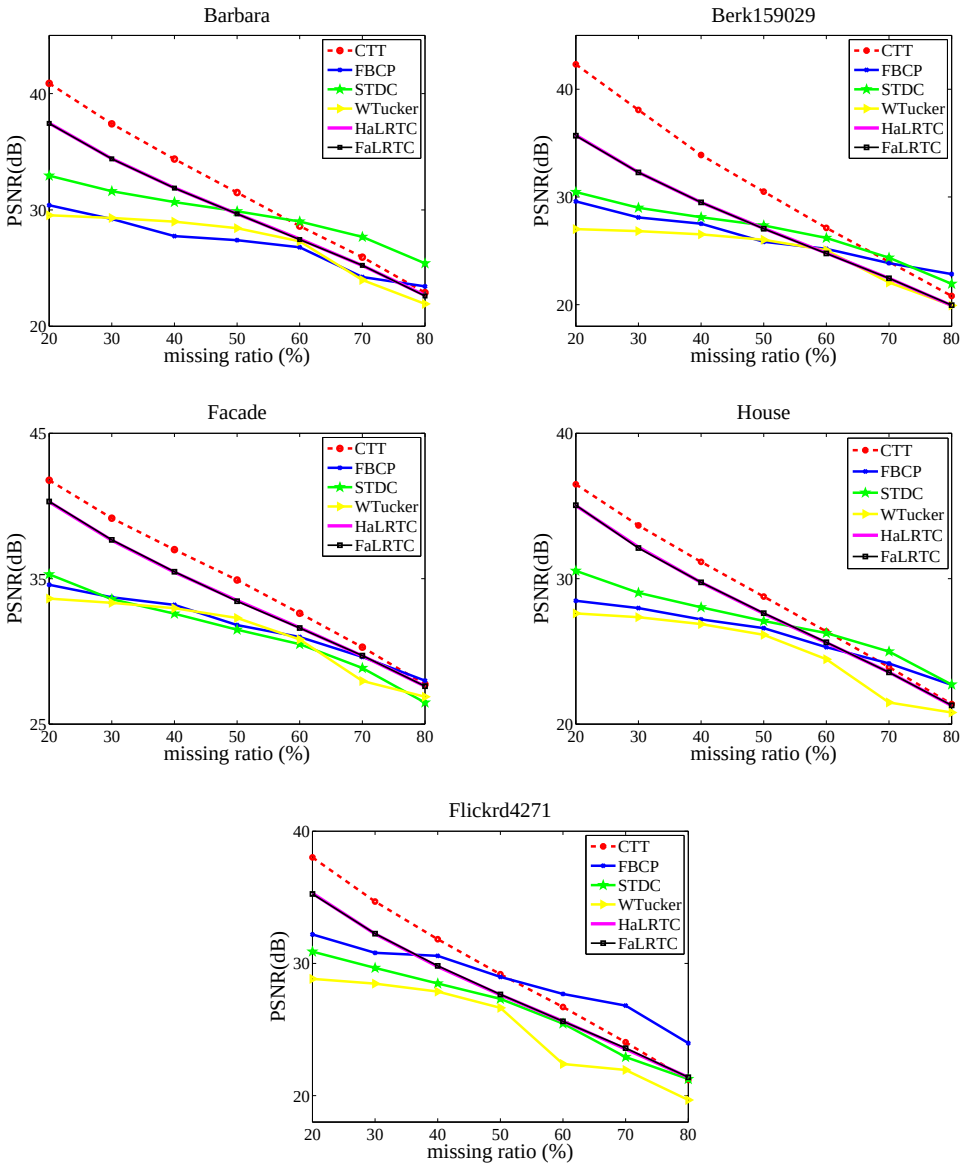


Figure 3.3: PSNR (dB) comparison of different algorithms on five color images at various missing ratios.

signal-to-noise-ratio (PSNR) [see Appendix A] and 2) relative squared error (RSE) [see Appendix B] which are widely used in tensor completion and visual data inpainting task. The third one is a correlation-based measure, referred to as the structural similarity index (SSIM) [see Appendix C]. In the experiments, the parameters are set as follows:

$$\tau = \gamma \times \sqrt[3]{\frac{m}{n_1 \times n_2 \times n_3}} \times (\sqrt[3]{n_1} + \sqrt[3]{n_2} + \sqrt[3]{n_3}) \tag{3.4}$$

where m is the total number of the available elements, with $1 \leq \gamma \leq 20$, $\delta = 0.4 \times n_1 \times n_2 \times n_3/m$, $\epsilon = 10^{-4}$, and the maximum iteration count is $K = 500$.

3.3.1 Color Image Inpainting

In this experiment, the aforementioned algorithms are run on the following color images of 256×256: Barbara, Berk159029, Facade, House, Flickrd4271, and Bungee jumper of 308×206. These images are shown in Figure 3.2. In this setting, each test image is considered as a 3rd order tensor with $d = 3$. In this experiment, the performance of the aforementioned algorithms is studied in four different cases: image completion with randomly missing values, text removal, scratched removal, and object removal.

In the case of image completion with randomly missing values, the inpainting input image is obtained by randomly setting $\rho\% = m/(n_1 \times n_2 \times n_3)$ of the overall number of image pixels to zero. This experiment is conducted for several values of ρ. For each value, the inpainting experiment is repeated for 10 trials and the average quality metrics are shown in Figures 3.3–3.5 for comparison, in accordance with the common practice in the literature. Then the average PSNRs, SSIMs, and RSEs of the all five images for each algorithm are reported in Table 3.1.

It is seen in Figures 3.3–3.5 and the Table 3.1 that the proposed CTT algorithm generally out-performs FaLRTC, HaLRTC, WTucker, STDC and FBCP algorithms in terms of all three metrics. This is particularly valid for missing ratios less than 70% and

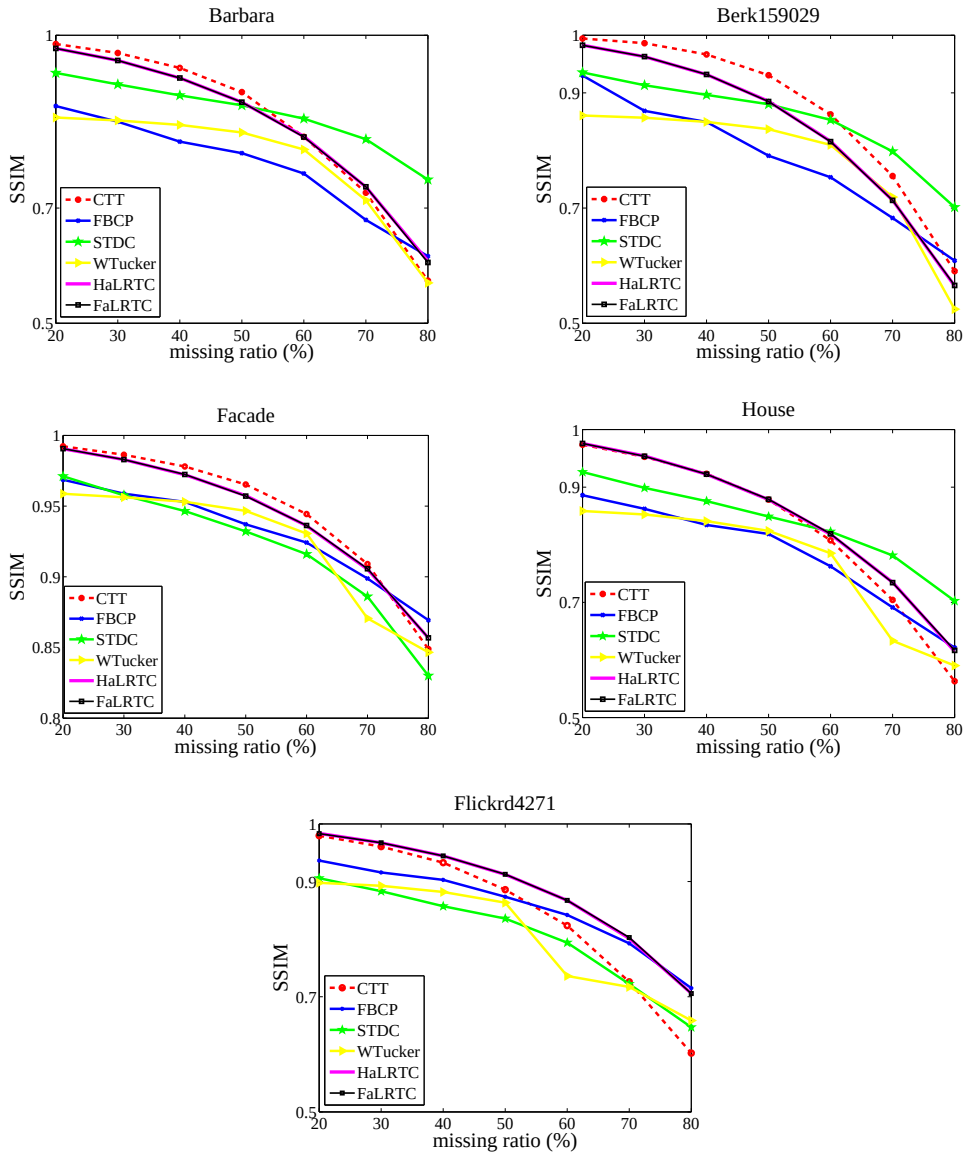


Figure 3.4: SSIM comparison of different algorithms on five color images at various missing ratios.

for structure-rich images. For missing ratios greater than 70%, STDC and FBCP out-perform CTT; nevertheless, CTT still out-performs HaLRTC, FaLRTC, and WTucker in terms of PSNR and RSE.

As a visual comparison, Figures 3.6 and 3.7 show the inpainting reconstructions of

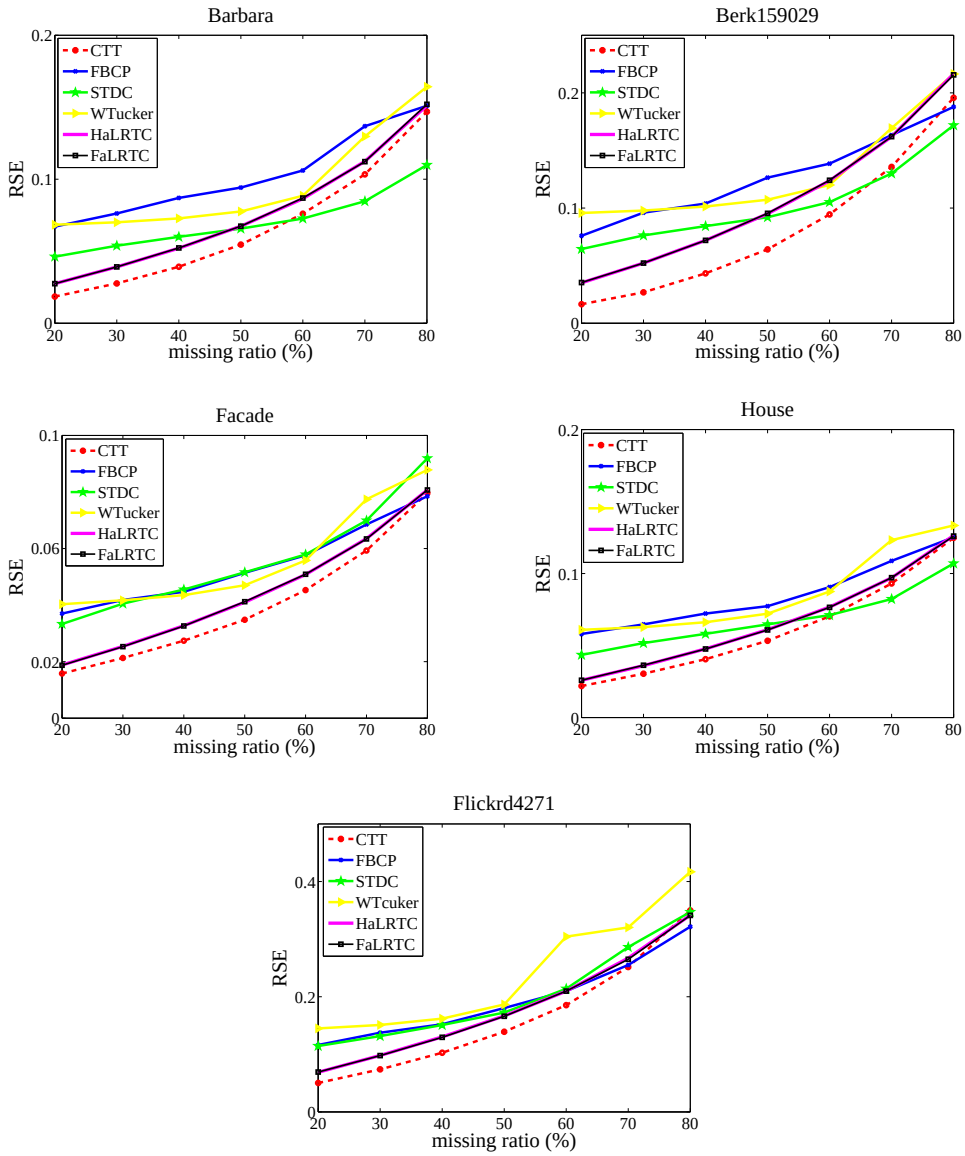


Figure 3.5: RSE comparison of different algorithms on five color images at various missing ratios.

Barbara and Facade images obtained with the aforementioned algorithms with $\rho = 70$. It is clear that CTT yields the best reconstruction compared to the FaLRTC, HaLRTC, and WTucker algorithms and comparable to STDC and FBCP algorithms. It is noted that the reconstruction obtained with the CTT algorithm contains more details and exhibits less visual artifacts compared to the other reconstructions, thanks to the ability of

Table 3.1: The average RSEs, PSNRs (dB), and SSIMs of all five color-images.

$\rho(\%)$	Measure	FaLRTC [65]	HaLRTC [65]	WTucker [66]	STDC [57]	FBCP [63]	CTT
20	RSE	0.0353	0.0352	0.0820	0.0604	0.0708	**0.0246**
	PSNR	36.75	36.76	29.33	32.02	31.05	**39.89**
	SSIM	0.9819	0.9819	0.8866	0.9346	0.9196	**0.9849**
30	RSE	0.0501	0.0501	0.0847	0.0708	0.0832	**0.0360**
	PSNR	33.74	33.74	29.06	30.58	29.96	**36.59**
	SSIM	0.9645	0.9646	0.8820	0.9135	0.8911	**0.9710**
40	RSE	0.0668	0.0669	0.0892	0.0798	0.0920	**0.0506**
	PSNR	31.28	31.27	28.65	29.58	29.25	**33.65**
	SSIM	0.9394	0.9393	0.8740	0.8943	0.8709	**0.9487**
50	RSE	0.0862	0.0862	0.0981	0.0893	0.1059	**0.0692**
	PSNR	29.09	29.09	27.91	28.63	28.13	**30.97**
	SSIM	0.9034	0.9035	0.8604	0.8750	0.8430	**0.9123**
60	RSE	0.1096	0.1095	0.1312	0.1041	0.1205	**0.0943**
	PSNR	27.01	27.03	26.01	27.48	27.19	**28.29**
	SSIM	0.8522	0.8523	0.8126	0.8481	0.8084	**0.8523**
70	RSE	0.1399	0.1404	0.1640	0.1306	0.1465	**0.1286**
	PSNR	24.91	24.89	23.49	**25.76**	25.74	25.63
	SSIM	0.7786	0.7783	0.7306	**0.8014**	0.7489	0.7642
80	RSE	0.1831	0.1833	0.2037	**0.1655**	0.1728	0.1793
	PSNR	22.57	22.57	21.84	23.56	**24.19**	22.80
	SSIM	0.6700	0.6700	0.6380	**0.7260**	0.6861	0.6357

the CTT to preserve structural coherence of the tensor. This visual comparison result is in line with the result of Figures 3.3–3.5. To provide a better visual inspection, we show the magnified subregion of the inpainted images in Figures 3.5 and 3.6. Inspection of these images show that the proposed CTT algorithm indeed offers improved reconstructions.

In demonstrating the potential of the proposed algorithm, the experiment is also conducted for text, scratched, and object removal cases. We prepared a text and scratched images of Barbara and Facade (shown in Figure 3.8). Reconstruction of Barbara image obtained with the aforementioned algorithms are shown in Figure 3.9. Although the low rank property can characterize random missing entries, the structural missing entry problem degrades the performance of these approaches. The restoration results

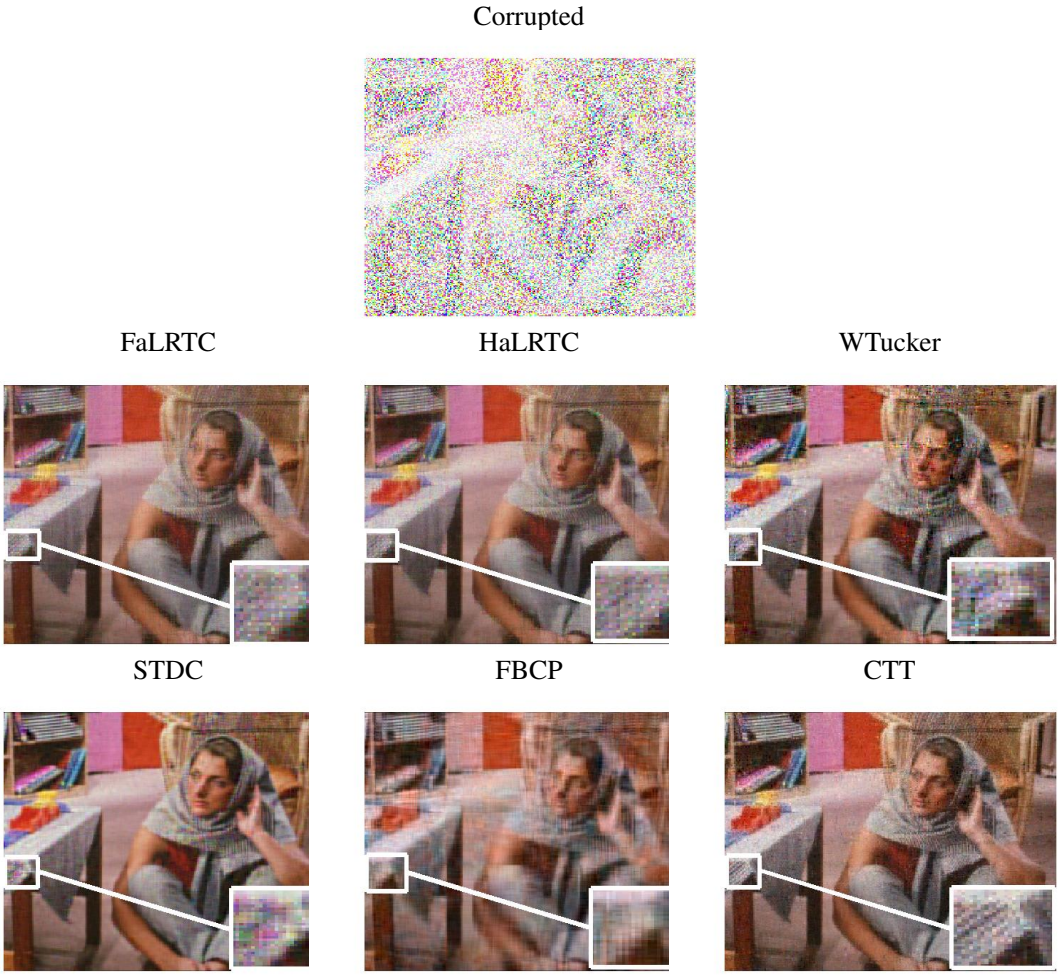


Figure 3.6: Visual results and the magnified subregions for Barbara image.

obtained by these algorithms have obvious artifacts. In contrast, these results confirm the effectiveness of the CTT completion algorithm. In the case of object removal (shown in Figure 3.10), the CTT algorithm recovers the image structure with acceptable visual effects and much better performances. In the CTT algorithm the repaired large region seems more consistent with the surrounding areas, however, the restoration results obtained by the state-of-the-art algorithms have visual fractures, and undesired color appears. The performance of the PSNR, SSIM, and RSE of the whole image is shown in Table 3.2, where the case of the object removal is not shown since the ground truth is not available. Accordingly, it can be concluded that the CTT algorithm

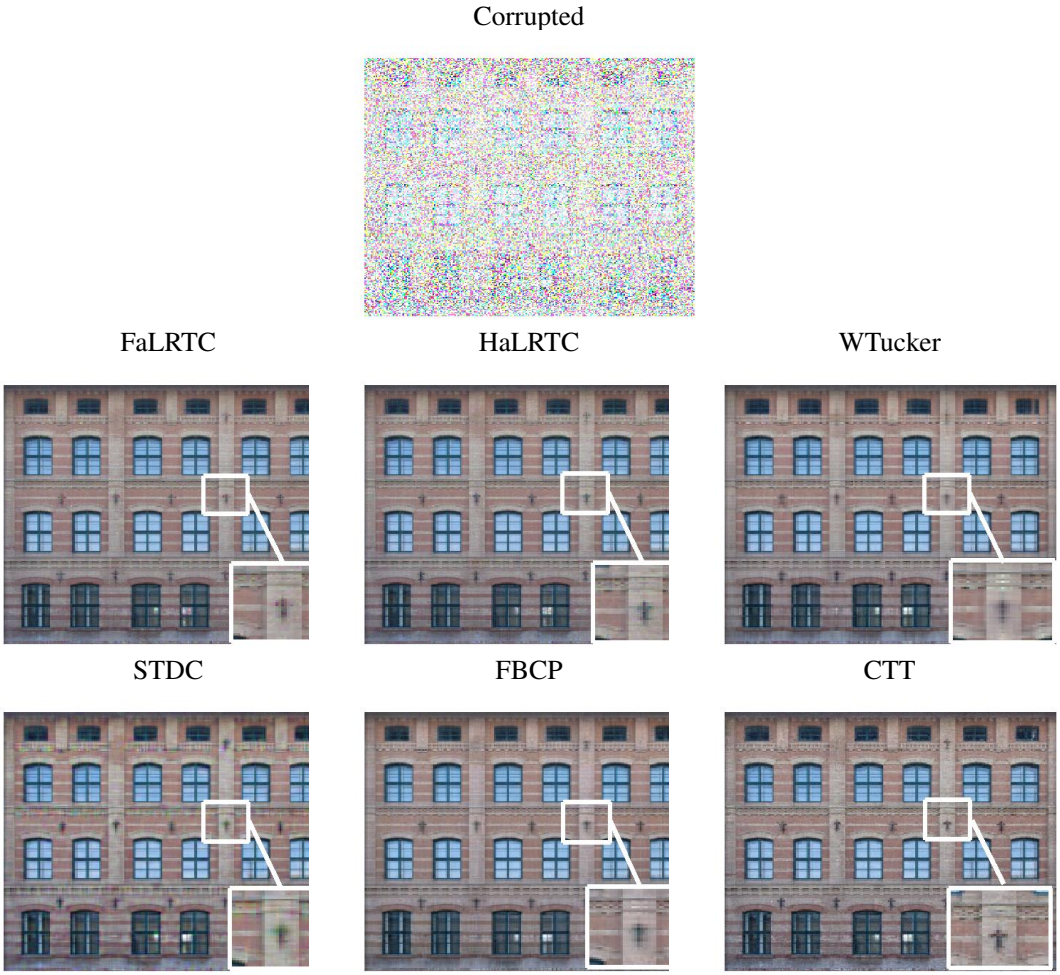


Figure 3.7: Visual results and the magnified subregions for Facade image.

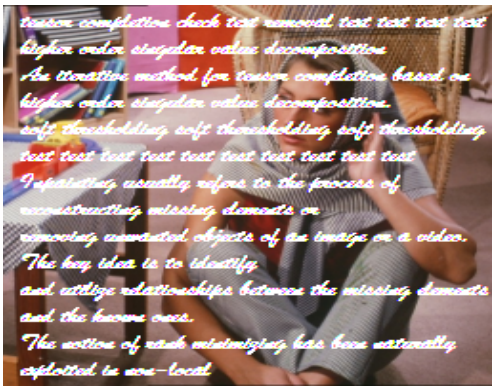
is comparable to the state-of-the-art algorithms in various image inpainting problems.

3.3.2 Video Sequence Inpainting

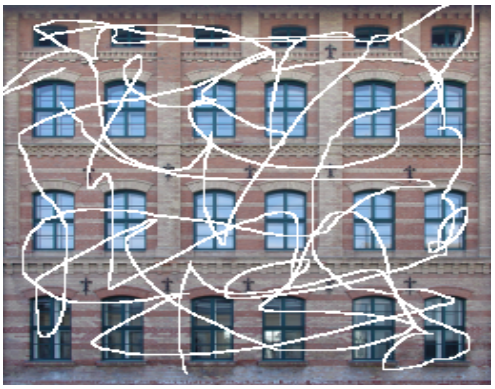
In this experiment, the performance of CTT in video inpainting is evaluated over the Foreman, Coastguard, and Salesman gray-scale benchmark video sequences. These video sequences are composed of 20, 40 and 50 frames, respectively. Thus, they give 3D tensors of the sizes $288 \times 352 \times 20$, $144 \times 176 \times 40$ and $288 \times 352 \times 50$, respectively. As in the case of color images, $\rho\%$ pixels are randomly selected from all the pixels in all frames and are set to be zero. These pixels are then recovered with the CTT algorithms.

Barbara-scratch

Barbara-text



Facade-scratch



Facade-text

Figure 3.8: Corrupted images: Barbara-scratch, Barbara-text (row 1), and Facade-scratch, Facade-text (row 2).

Quantitative performance comparison results are reported in Table 3.3. Examining the results, we see that CTT is generally superior to the other algorithms. In contrast to the case of color images, the improvement of CTT over the other algorithms is more significant with increased missing ratios. This result suggests the advantage of the CTT algorithm is dealing with tensors of large sizes, where directional structures tend to be more overwhelming.

Figure 3.11 shows the original and the corrupted of Foreman, Coastguard, and Salesman sample frames with $\rho = 70$. Figure 3.12 represents inpainted and the corre-

Figure 3.9: Visual results of algorithms applied to Babara-scratch.

sponding residual images of Foreman sample frame after the reconstruction with the aforementioned algorithms. It is evident that CTT yields significantly less artifacts compared to the other algorithms.

3.3.3 MR Image Inpainting

In this experiment, CTT is tested and compared to the aforementioned inpainting algorithms for the case of MR images. The images Spectral, Brain and PNEUMATIX of the sizes $205 \times 246 \times 96$, $196 \times 256 \times 90$, and $192 \times 256 \times 25$, respectively, are used.

Recovery results are reported in Table 3.4. It is noted that the proposed CTT algorithm

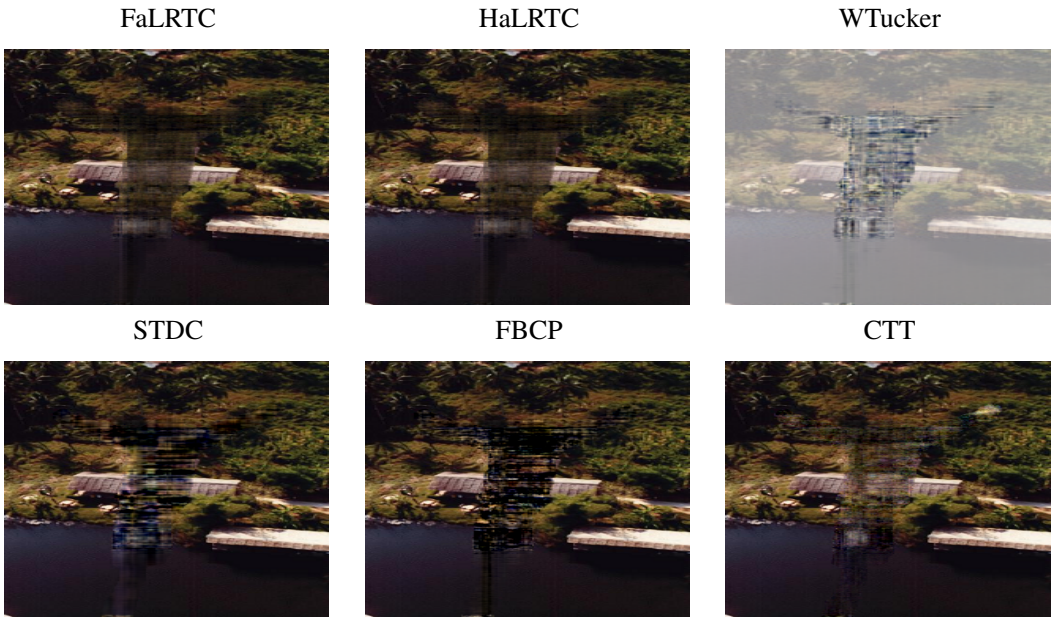


Figure 3.10: Visual results of algorithms applied to Bungee jumper.

Table 3.2: The RSE, PSNR (dB), and SSIM of Barbara and Facade in the case of text and scratched removal.

Image		Measure	FaLRTC [65]	HaLRTC [65]	WTucker [66]	STDC [57]	FBCP [63]	CTT
Facade	scratched	RSE	0.0239	0.0239	0.1392	0.0311	0.0291	**0.0182**
		PSNR	38.17	38.17	22.87	35.89	36.46	**40.55**
		SSIM	0.9853	0.9853	0.7450	0.9767	0.9798	**0.9902**
	text	RSE	0.0298	0.0298	0.1917	0.0377	0.0312	**0.0226**
		PSNR	36.28	36.28	20.09	34.22	35.86	**38.67**
		SSIM	0.9768	0.9768	0.6787	0.9654	0.9749	**0.9845**
Barbara	scratched	RSE	0.0474	0.0474	0.1807	0.0551	0.0588	**0.0274**
		PSNR	32.72	32.72	21.09	31.40	31.77	**37.46**
		SSIM	0.9447	0.9447	0.5621	0.9286	0.9233	**0.9714**
	text	RSE	0.0569	0.0569	0.2102	0.0598	0.0622	**0.0313**
		PSNR	31.13	31.13	19.78	30.69	31.00	**36.32**
		SSIM	0.9228	0.9229	0.5246	0.9144	0.9056	**0.9578**

out-performs the other algorithms in terms of the RSE measure even with high missing ratios. As a visual comparison, Figure 3.13 shows the original and the corrupted of Spectral, Brain and PNEUMATIX sample frames with $\rho = 70$. Figure 3.14 represents inpainted and the corresponding residual images of Spectral sample slice after the reconstruction with the aforementioned algorithms.

Table 3.3: Video inpainting comparison in RSE.

Video	$\rho(\%)$	FaLRTC [65]	HaLRTC [65]	WTucker [66]	STDC [57]	FBCP [63]	CTT
Foreman	20	0.0158	0.0159	0.0581	0.0304	0.0592	**0.0142**
	50	0.0369	0.0369	0.0589	0.0424	0.0621	**0.0273**
	70	0.0594	0.0595	0.0605	0.0478	0.0688	**0.0385**
	80	0.0791	0.0791	0.0628	0.0590	0.0731	**0.0469**
Coastguard	20	0.0337	0.0337	0.0626	0.0442	0.0692	**0.0202**
	50	0.0656	0.0659	0.0660	0.0675	0.0726	**0.0399**
	70	0.0934	0.0936	0.0821	0.0849	0.0795	**0.0559**
	80	0.1131	0.1137	0.1236	0.1131	0.0883	**0.0670**
Salesman	20	0.0344	0.0344	0.1425	0.0675	0.1230	**0.0273**
	50	0.0840	0.0841	0.1438	0.0917	0.1256	**0.0535**
	70	0.1398	0.1395	0.1464	0.1059	0.1282	**0.0781**
	80	0.1855	0.1856	0.1511	0.1419	0.1353	**0.0968**

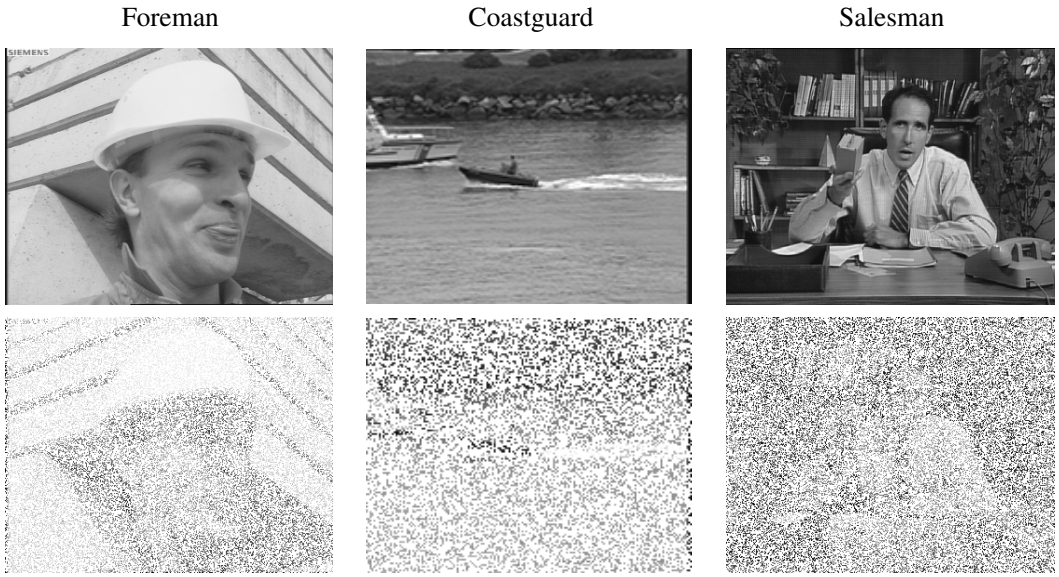


Figure 3.11: Original images (row 1), corrupted images (row 2).

3.3.4 Complexity Analysis

In terms of computational cost, when compared with the state-of-the-art algorithms, the proposed inpainting algorithm complexity is

$$\boldsymbol{O}(K((d\prod_{i=1}^{d} n_i) + (\sum_{i=1}^{d} r_i) + (\sum_{i=1}^{d} n_1 \dots n_i r_i \dots r_d))),$$

where the first term is the cost of HOSVD of full tensor [21], the second one comes from the computational cost of (2.13) and then the last one is for reconstructing the full tensor. The computational complexity of the algorithms in [65] (FaLRTC and

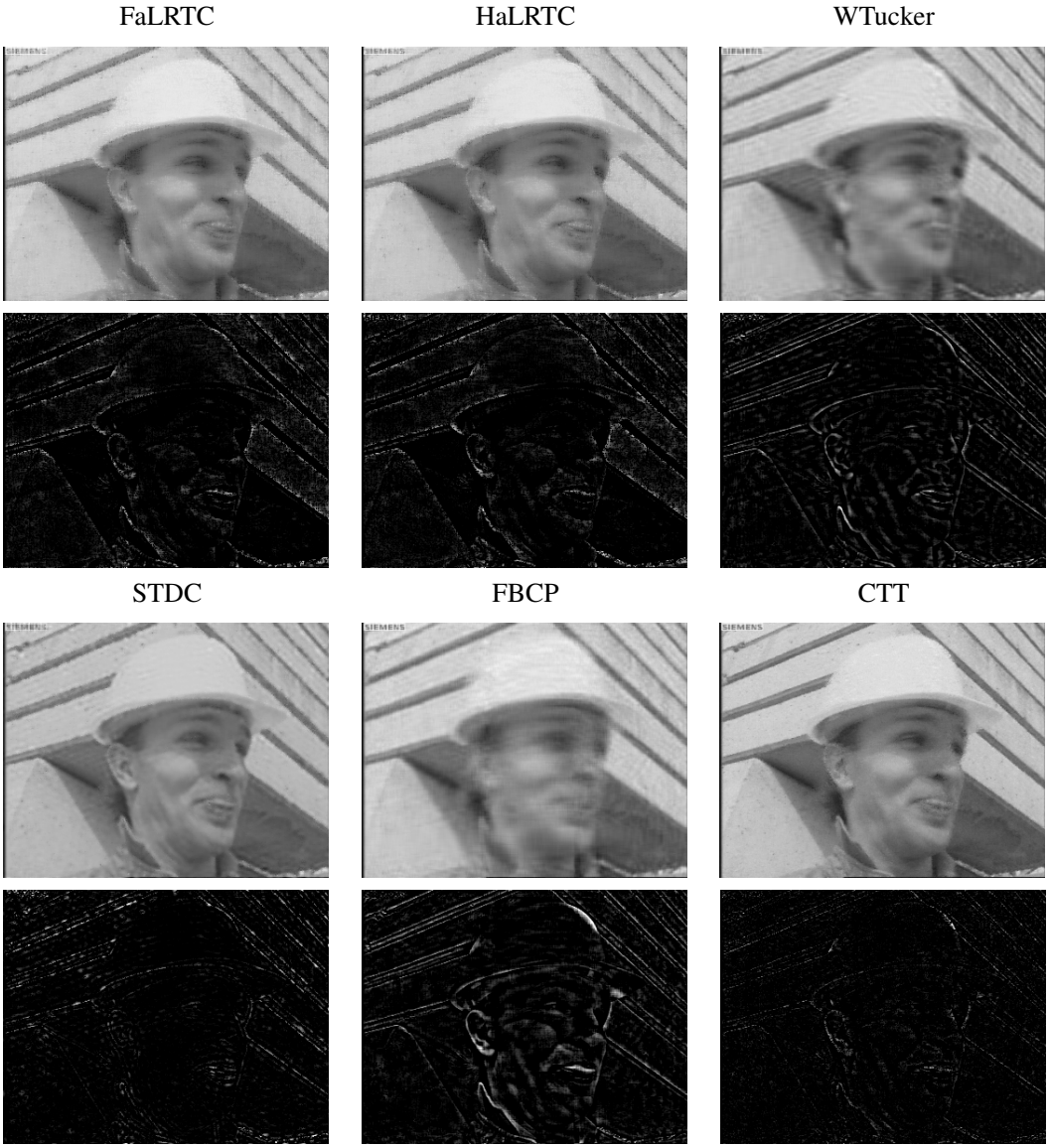


Figure 3.12: The inpainted and the corresponding residual images of Foreman after reconstruction by: FaLRTC, HaLRTC, WTucker, STDC, FBCP and CTT.

HaLRTC), [66] (WTucker), [57] (STDC), and [63] (FBCP) are, respectively

$$\boldsymbol{O}(K((2r_1 + r_2 + r_3)\prod_{i=1}^{d} n_i + \sum_{i=1}^{d} n_i \prod_{i=1}^{d} r_i)), \boldsymbol{O}(K(2(d+1)R\prod_{i=1}^{d} n_i)),$$

$$\boldsymbol{O}(K(dR^2m + R^3\sum_{i=1}^{d} n_i)), \text{ and } \boldsymbol{O}(K(dR^2m + R^3)).$$

It is observed that the computational cost of the proposed algorithm is comparable with the FaLRTC and HaLRTC algorithms and may become expensive comparing with the WTucker, STDC, and FBCP algorithms. This computational cost is required to

Table 3.4: MR image inpainting comparison in RSE.

MRI	$\rho(\%)$	FaLRTC [65]	HaLRTC [65]	WTucker [66]	STDC [57]	FBCP [63]	CTT
Spectral	20	0.0164	0.0164	0.0375	0.0258	0.0282	**0.0059**
	50	0.0332	0.0332	0.0382	0.0383	0.0288	**0.0093**
	70	0.0486	0.0487	0.0385	0.0450	0.0309	**0.0111**
	80	0.0601	0.0601	0.0432	0.0561	0.0342	**0.0126**
Brain	20	0.0743	0.0742	0.1912	0.1220	0.2391	**0.0635**
	50	0.1612	0.1621	0.1986	0.1670	0.2407	**0.1165**
	70	0.2457	0.2458	0.2192	0.2264	0.2439	**0.1579**
	80	0.3151	0.3138	0.2632	0.2875	0.2504	**0.1866**
PNEUMATIX	20	0.0412	0.0411	0.0976	0.0711	0.0777	**0.0290**
	50	0.1015	0.1003	0.1017	0.0941	0.0850	**0.0503**
	70	0.1765	0.1775	0.1156	0.1270	0.0952	**0.0681**
	80	0.2474	0.2491	0.2671	0.1869	0.1042	**0.0843**

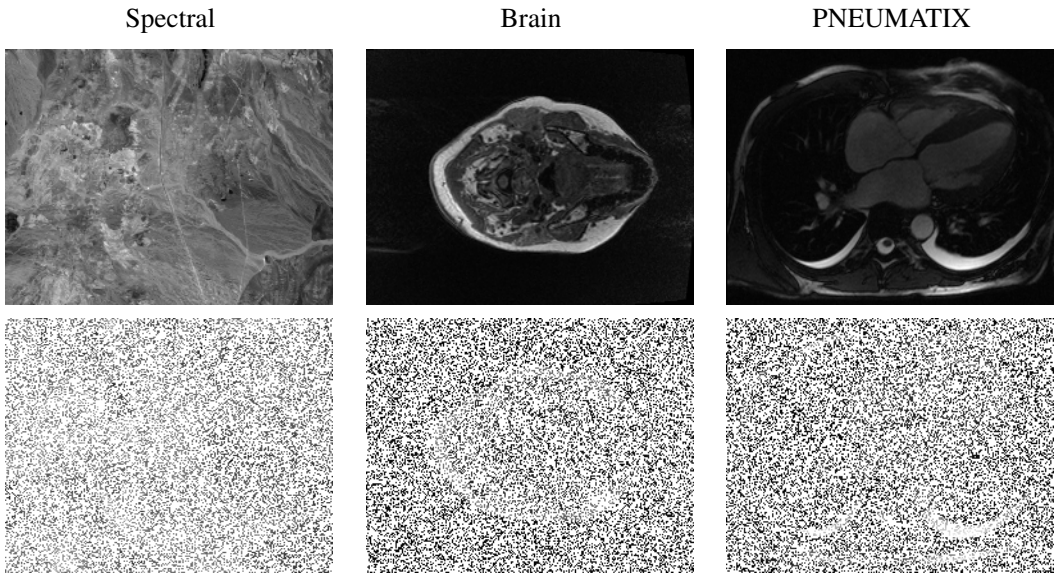


Figure 3.13: Original images (row 1), corrupted images (row 2).

generate the proposed quality improvement.

3.4 Conclusion

In this chapter, an iterative algorithm (referred to as CTT) is proposed for tensor inpainting. With the use of the higher-order singular value decomposition of tensors, while restoring the missing data, the algorithm is able to preserve structural coherence of the tensor in all directions. The CTT algorithm is compared with the state-of-the-art FaLRTC, HaLRTC, WTucker, STDC, and FBCP algorithms. Improved performances

of the CTT are demonstrated in our experiments on color images, video sequences, and MR images.

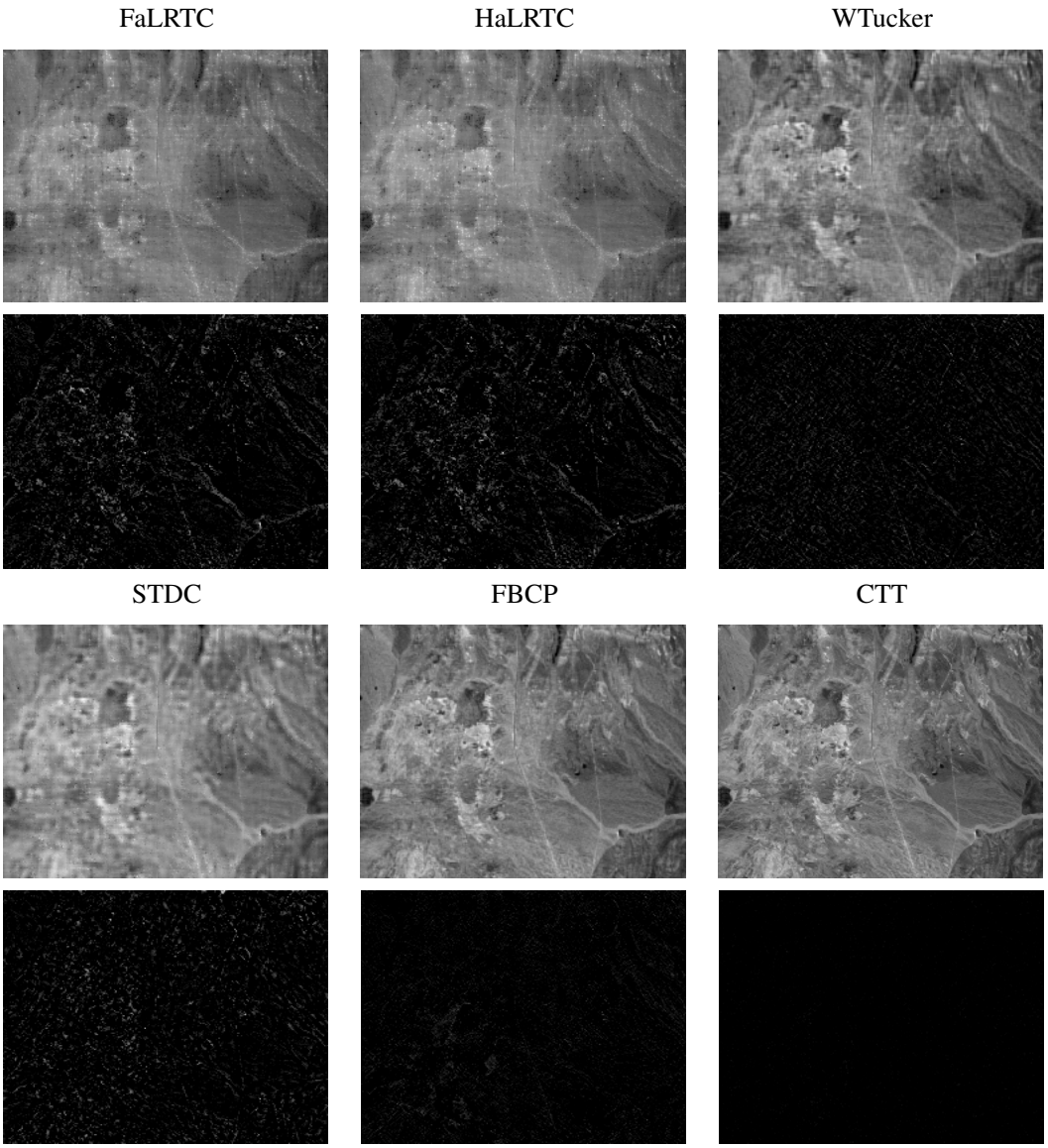


Figure 3.14: The inpainted and the corresponding residual images of Spectral after reconstruction by: FaLRTC, HaLRTC, WTucker, STDC, FBCP and CTT.

Chapter 4

NOISE REMOVAL OF MR IMAGES VIA ITERATIVE REGULARIZATION BASED ON HIGHER ORDER SINGULAR VALUE DECOMPOSITION

4.1 Introduction

Despite the success of magnetic resonance imaging techniques in many applications, acquisition noise is still a limiting factor for the quality and hence the usefulness of the techniques. In this chapter, we develop an iterative regularization algorithm for denoising magnetic resonance images based on higher-order singular value decomposition. The proposed algorithm first forms a single tensor from the noisy data. Next, higher-order singular value decomposition is applied on this tensor with respect to a set of learned orthogonal matrices over the corresponding tensor mode. Finally, soft thresholding is iteratively applied to the calculated coefficients, thereby suppressing the noise. The proposed algorithm is further enhanced with a post-process Wiener filtering. Experiments are conducted on synthetic and real magnetic resonance images to compare the performance of the proposed algorithm to state-of-the-art denoising approaches. The comparison is made quantitatively and qualitatively through visual comparisons. The results demonstrate the competitive performance of the proposed algorithm.

4.1.1 Related Work

MR imaging is one of the most widely-used imaging techniques in clinical diagnosis and scientific research. It can provide highly detailed images of organs and tissues in the human body. One key usage of MR images is to reveal pathological or other physiological alternations in living anatomy. The advantages of this imaging technique are well-acknowledged. Unfortunately, in addition to the long acquisition time, the noise accumulated through the acquisition phase remains a major limiting factor. Such noise can potentially mislead the diagnosis and limit the applicability of computerized analysis processes, such as classification and registration on MRIs. Therefore, MRI denoising is of crucial importance to the successful use of the MR technique. In the context of MRI denoising, and unlike the standard image denoising, a Rician noise model, instead of a Gaussian model, is usually used. This finding is observed because an MR image is typically obtained by applying an inverse Fourier transform on the acquired raw magnitude data. Because the real and the imaginary image noise can be typically modeled as Gaussians, the MR (magnitude) image, calculated as the square root of the sum of squares of these two components, can be better described by a Rician noise model [73, 74, 75, 76, 77, 78].

A trivial way to reduce the MR image noise is to obtain an image as an average of several acquisitions. However, this will result in lengthening the acquisition time and degrading the SNR. Therefore, denoising is often performed by post-processing methods. These denoising methods can be classified into three main categories. The first category is the filtering-based methods, and based on linear or non-linear filtering on

the image data to suppress the noise. Note however that filtering inherently blurs images and tends to be unfaithful to the high frequencies that represent details. Therefore, filtering is directed towards preserving edges and details [79, 80, 81, 82, 83, 84, 85]. The second category includes transform-domain denoising methods. The methods manipulate image coefficients calculated with respect to a specific transform basis. Manipulation comes in line with a certain reconstruction prior, or regularizer. Examples include using wavelet transforms [86, 87], curvelet transforms [88] and discrete cosine transforms [89]. The third category is based on statistical and probabilistic denoising approaches. Examples of such work conducted in [90, 91, 92] use linear minimum mean square error and maximum likelihood for the estimation of the Rician noise.

More recently, researchers have considered merging the non-local means (NLM) and the transform-domain approaches [93, 94, 95, 96]. The intent is to combine the advantages of non-local patch similarity and signal sparsity to better regularize the reconstruction process. This combination has been shown to render an outstanding performance. Along this line, Dabov *et al*. [93] proposed the block matching in three dimensions (BM3D) algorithm as an extension to the NLM filtering approach. The BM3D algorithm is based on grouping similar patches into 3D stacks. These stacks are assumed to admit sparse representation in the transform domain of certain basis functions. Hence, sparse regularization is applied on the transformed patches, and a patch is reconstructed by aggregating several reconstructions at each location. This sparsity constraint appears to be the reason behind the outstanding performance of this approach. In [95], Elahi *et al.* proposed a modification to the BM3D algorithm and ap-

plied it to magnetic resonance imaging denoising. Maggioni *et al*. proposed the block matching in four dimensions (BM4D) algorithm [96] as an extension of the BM3D filtering to volumetric data. The BM4D algorithm demonstrates the state-of-the-art performance in 3D MRI denoising [96].

From another viewpoint, the last decade has witnessed the success of tensor factorization in signal processing applications. This success is primarily due to the tensor's ability to exploit structural information of signals in an efficient and compact representation [2, 21]. This ability is especially valid for multidimensional signals, such as MR images. In the context of tensor analysis, tensor decomposition offers a natural means for applying the sparsity regularizer on the coefficients of tensors [97, 98]. Along this line, Rajwade *et al*. [97] proposed a patch-based image denoising algorithm using HOSVD. This algorithm forms 3D stacks of similar image patches on which the HOSVD is performed. Similar to [93], this method offers a simple and elegant means of exploiting sparsity among similar patches for denoising images. Note that while the algorithm of [93] uses fixed basis functions, the algorithm of [97] learns the HOSVD bases over image patches. Learning the basis functions allows for better signal representation and promises improved denoising performance. The denoising outcomes can possibly be further improved at the post-processing stage. For example, it was shown in [99] that improved denoising results can be achieved by applying a Wiener filter or a recursive regularization to the MR image.

4.2 HOSVD-Based Methods and the Proposed Algorithm

Let $\boldsymbol{X}$ with size $n_1 \times n_2 \times n_3$ denote the observed noisy MRI data. If $\boldsymbol{T}$ represents the underlying noise-free MRI data, then, one can model the relationship between $\boldsymbol{X}$ and $\boldsymbol{T}$ as follows (as in equation (1.1)):

$$\boldsymbol{X} = \boldsymbol{T} + \eta \tag{4.1}$$

where η is the noise accumulated in the acquisition process. The aim of a denoising process is to obtain the closest approximation to $\boldsymbol{T}$ based on the information provided by $\boldsymbol{Y}$ and any other information about the noise. The most commonly used noise model is the additive independent white Gaussian noise (AWGN) $\eta(0, \sigma)$, with zero-mean and standard deviation σ [see Appendix E].

As mentioned in the introduction of this chapter, an MR image is obtained by using the magnitude of a single complex raw data. Because the real and imaginary images are modelled as Gaussian noises, the distribution of noise in the MR (magnitude) image, calculated as the square root of the sum of squares of these two component images, can be modeled with a Rician distribution. Specifically, let the individual value of the complex MRI data be $t = p + iq$, then $x = (p + n_{re}) + i(q + n_{im})$, where the noises n_{re}, $n_{im} \sim \eta(0, \sigma)$ thus, the magnitude of the noisy raw data is $|x| = \sqrt{(p + n_{re})^2 + (q + n_{im})^2}$. As a result, the distribution of the $|x|$ is modeled with the Rician distribution [see Appendix F].

4.2.1 Image Denoising via HOSVD

Inspired by BM3D, several patch-based HOSVD methods for denoising have been proposed in recent years. The patch-based HOSVD method is designed for images

that are corrupted by Gaussian noise which is independent of the signal. Patch-based HOSVD methods, refer to the process that separates several similar patches of small size from the noisy image into a 3D stack, in which the group patches have similar local structures. In HOSVD-based methods, the 3D stack is transformed by the HOSVD transform to compute the HOSVD coefficients. Then, these coefficients are shrunk by using hard thresholding and the inverse of the transform is performed to produce the final filtered image. The hard thresholding of the coefficients is represented as follows:

$$\left[s_{i_1,i_2,i_3}\right]_\tau = \begin{cases} s_{i_1,i_2,i_3}, & \text{if } |s_{i_1,i_2,i_3}| > \tau \\ 0, & \text{otherwise} \end{cases} \tag{4.2}$$

The HOSVD method should not be directly applied to MR images that are generally corrupted by the Rician noise. Zhang *et al.* in [99], extended the HOSVD method to MR image denoising by using the advantages of the optimal forward and inverse variance stabilizing transformations. Furthermore, the denoising outcome can be further enhanced by a post-process, such as Wiener filtering [97]. Similarly, for better performance, the HOSVD method in [99] is augmented by a recursive regularization.

4.2.2 Proposed Algorithm

The proposed iterative higher order singular value decomposition algorithm (IHSVD) uses similarity across MRI slices as a regularizer for the reconstruction in addition to the sparsity of the representation forming the basic regularizer [100]. This sparsity is imposed by applying soft thresholding on the representation coefficients, which were calculated with respect to a transform basis. For this purpose, the proposed algorithm first forms a tensor that is composed of all the MRI slices, taking advantage of mutual information from different slices. In addition, it does not require patch selection as the

tensor is naturally formed by the slices. Next, orthogonal basis matrices are extracted over the tensor modes. We can subsequently apply an iterative soft thresholding to the HOSVD coefficients calculated with respect to these basis matrices. In a denoising context, iterative regularization considers the denoising outcome after each iteration as the input noisy image of the next iteration [31, 32, 101]. Thus, starting with $\boldsymbol{Y}^0$, iterative algorithm for denoising is of the form:

$$\begin{aligned} \boldsymbol{X}^k &= \boldsymbol{D}_\tau(\boldsymbol{Y}^{k-1}) \\ \boldsymbol{Y}^k &= \boldsymbol{Y}^{k-1} - \delta\,(\boldsymbol{X} - \boldsymbol{X}^k) \end{aligned} \tag{4.3}$$

This iterative procedure is continued until a stopping criterion is met. In this work, the stopping criterion is chosen to be $\|\boldsymbol{X} - \boldsymbol{X}^k\|_F \leq \sigma$. This stopping criterion is the same one used in the iterated total variation–based model [31]. In the iteration procedure, the proposed algorithm transforms the noisy data into the HOSVD domain to compute the HOSVD coefficients, manipulates the coefficients by soft thresholding, and inverts the HOSVD transform to get the final filtered data. This thresholding minimizes the contribution of noise in the HOSVD transform.

The aforementioned procedure represents the basic idea of the proposed algorithm IHSVD. Similar to other methods, our IHSVD is augmented with an additional stage to improve performance. In this stage, we apply a nonlocal Wiener filtering post-processing step to smooth the resulting images. The Wiener Filtering has two inputs: the noisy observation $\boldsymbol{X}$ and the denoised data $\hat{\boldsymbol{X}}$. Let the coefficients of $\boldsymbol{X}$ and $\hat{\boldsymbol{X}}$ on the HOSVD transform be defined as s and $\hat{s}$, respectively. The Wiener filtering replaces the thresholding output and give the final filtered version of the core tensor in

the following form

$$\hat{s} = \frac{\hat{s}^2}{\hat{s}^2 + \sigma^2} s \tag{4.4}$$

We term this augmented algorithm as IHSVD-W. In order to test the performance of the proposed algorithm, we consider both Gaussian and Rician distributed noise. The former one is outlined in Algorithm 3, and the flowchart of the respective algorithm is depicted in Figure 4.1.

Algorithm 3 The Proposed IHSVD Algorithm for Gaussian Noise

1: **INPUT:** Observed data $\boldsymbol{X}$, step size δ, shrinkage threshold τ, tolerance ϵ, standard deviation σ, and iteration number K.
2: **Initialize $\boldsymbol{Y}^0 = \boldsymbol{X}$.**
3: **for k= 0, 2, ..., K do**
Compute the HOSVD of $\boldsymbol{Y}^k$ as $\boldsymbol{Y}^k = \boldsymbol{S} \times_1 U_1 \times_2 U_2 \times_3 U_3$.
Calculate $\boldsymbol{D}_\tau(\boldsymbol{S})$ by applying soft thresholding on $\boldsymbol{S}$.
Construct $\boldsymbol{X}^k = \boldsymbol{D}_\tau(\boldsymbol{S}) \times_1 U_1 \times_2 U_2 \times_3 U_3$.
If $\|\boldsymbol{X} - \boldsymbol{X}^k\|_F \leq \sigma$ then break.
Update $\boldsymbol{Y}^k = \boldsymbol{Y}^{k-1} + \delta \ (\boldsymbol{X} - \boldsymbol{X}^k)$
Obtain an improved denoised image $\boldsymbol{X}^k$ as
$\hat{\boldsymbol{X}} \leftarrow \boldsymbol{X}^k$
4: **OUTPUT:** The denoised image $\hat{\boldsymbol{X}}$.

As Rician noise is the assumed noise model in MRI, a denoising algorithm requires a noise variance stabilization (VSt) step [102]. This finding is observed because the variance of Rician noise is dependent on the signal itself. In accordance with the common practices in MRI denoising, we add a VSt step to remove the dependence of the noise variance on the raw MR image intensities recorded before applying the denoising algorithm. One can express the overall outcome of the proposed algorithm as follows:

$$\hat{\boldsymbol{V}} = (\mathrm{VSt})^{-1}(\mathrm{IHSVD}(\mathrm{VSt}(\boldsymbol{X}, \sigma_e), \sigma_v), \sigma_e) \tag{4.5}$$

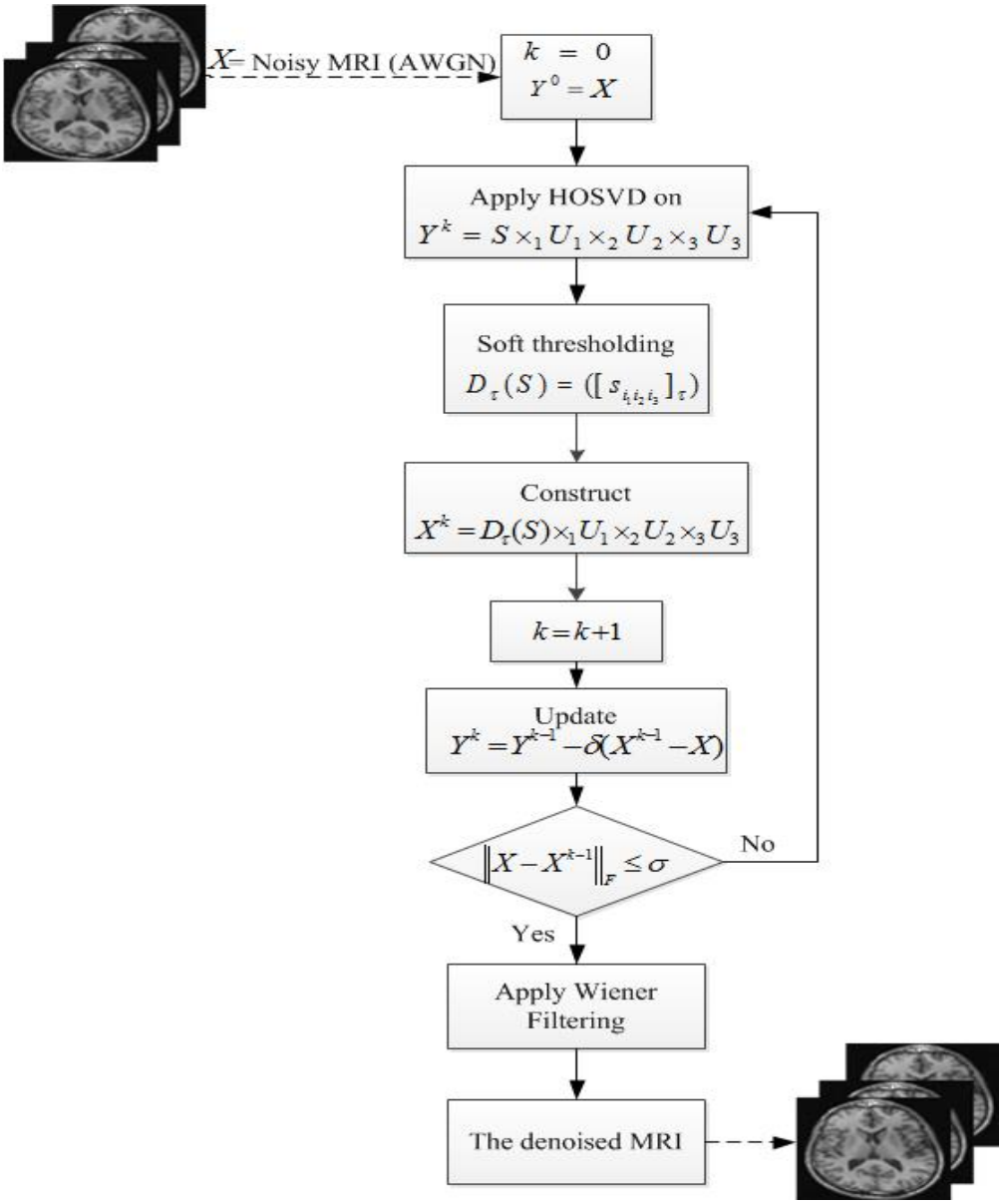


Figure 4.1: Flowchart of IHSVD algorithm.

where $(\mathrm{VSt})^{-1}$ denotes the inverse VSt transformation, $\boldsymbol{X}$ is the MR image corrupted by Rician noise, σ_e is the estimated standard deviation of the Rician noise, and σ_v is the stabilized standard deviation induced by the VSt. The proposed denoising algorithm for Rician distributed noise is outlined in Algorithm 4, and the relevant flowchart is shown in Figure 4.2.

Algorithm 4 The Proposed IHSVD Algorithm for Rician Noise

1: **INPUT:** Observed data $\boldsymbol{X}$, step size δ, shrinkage threshold τ, tolerance ϵ, standard deviation σ, estimated standard deviation σ_e, stabilized standard deviation σ_v, and iteration number K.
2: **Initialize** $\boldsymbol{Y}^0 = \mathrm{VSt}((\boldsymbol{X}, \sigma), \sigma_v)$.
3: **for k= 0, 2, ..., K do**
Compute the HOSVD of $\boldsymbol{Y}^k$ as $\boldsymbol{Y}^k = \boldsymbol{S} \times_1 U_1 \times_2 U_2 \times_3 U_3$.
Calculate $\boldsymbol{D}_\tau(\boldsymbol{S})$ by applying soft thresholding on $\boldsymbol{S}$.
Construct $\boldsymbol{X}^k = \boldsymbol{D}_\tau(\boldsymbol{S}) \times_1 U_1 \times_2 U_2 \times_3 U_3$.
If $\|\boldsymbol{X} - \boldsymbol{X}^k\|_F \leq \sigma$ then break.
Update $\boldsymbol{Y}^k = \boldsymbol{Y}^{k-1} + \delta \ (\boldsymbol{X} - \boldsymbol{X}^k)$
Apply an inverse VSt on the calculated $\boldsymbol{X}^k$ as
$\hat{\boldsymbol{X}} \leftarrow (\mathrm{VSt})^{-1}(\boldsymbol{X}^k)$
4: **OUTPUT:** The denoised image $\hat{\boldsymbol{X}}$.

4.3 Experimental Results

In this section, the performance of the proposed denoising method is compared to the HOSVDR [99], BM4D [96] and PES-TV [94] denoising algorithms. The HOSVDR algorithm is included in the comparison as it is based on the HOSVD transform, the BM4D and PES-TV algorithms are included as the state-of-the-art methods. Both versions (IHSVD and IHSVD-W) of the proposed algorithm are included in the comparison. Performance evaluation is carried out in terms of the PSNR, and the mean absolute difference (MAD) [see Appendix D] as error-based measures which are regularly employed in MRI evaluation [94], and the SSIM as a correlation-based measure. Note that a higher performance corresponds to having a higher PSNR, SSIM and a lower MAD.

In our experiments, we set the parameter τ to be

$$\tau = \sigma_k \times \sqrt{2 \times \log(n_1 \times n_2 \times n_3)} \tag{4.6}$$

where σ_k is the updated noise variance, which is calculated as $\sigma_k = \gamma \times \sqrt{\sigma - \|\boldsymbol{X} - \boldsymbol{X}^k\|_F}$

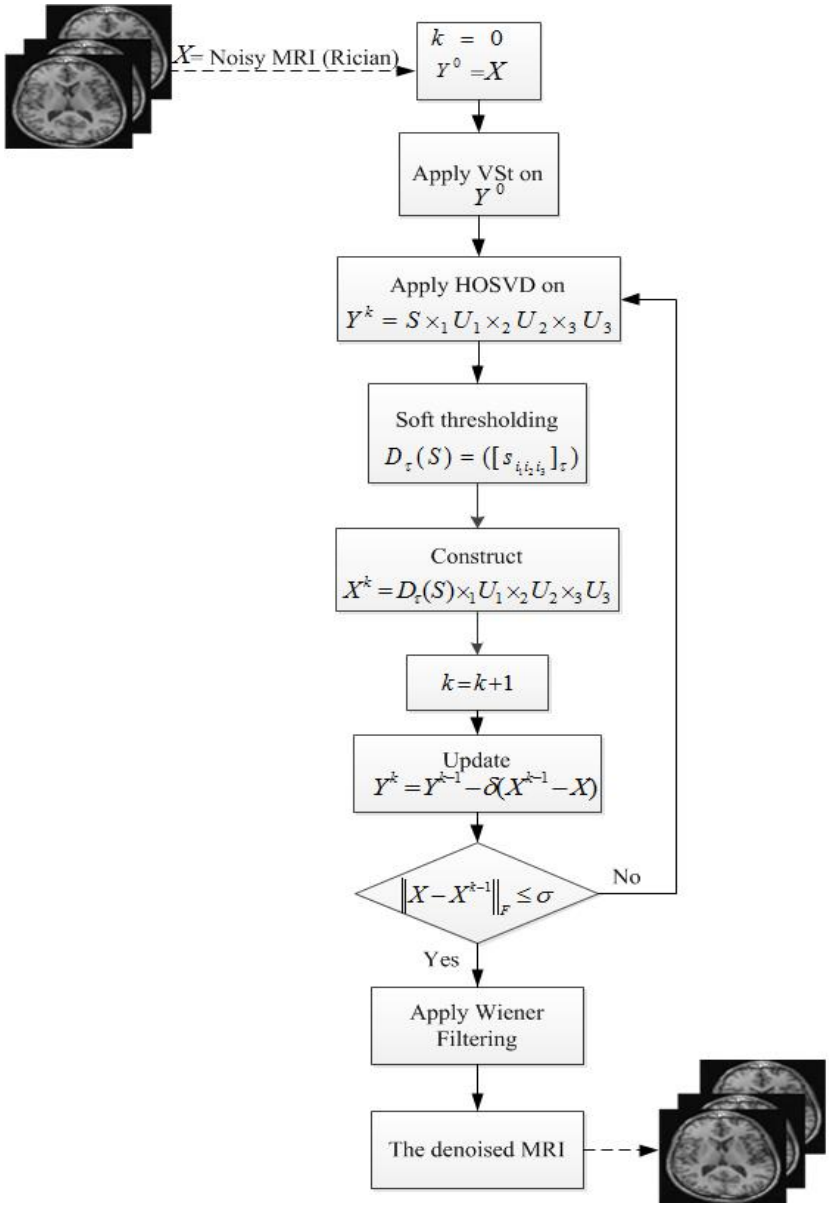


Figure 4.2: Flowchart of IHSVD-W algorithm.

[103] where a parameter $\gamma \in (0,\ 1)$ and σ being the noise standard deviation in the original data $\boldsymbol{X}$. And we set $\delta = 1/\sigma$.

In this section we present an assessment of the above-mentioned methods as applied to synthetic and real MR images. In accordance with the common practice in the literature, these experiments are made under both Gaussian and Rician distributed noises.

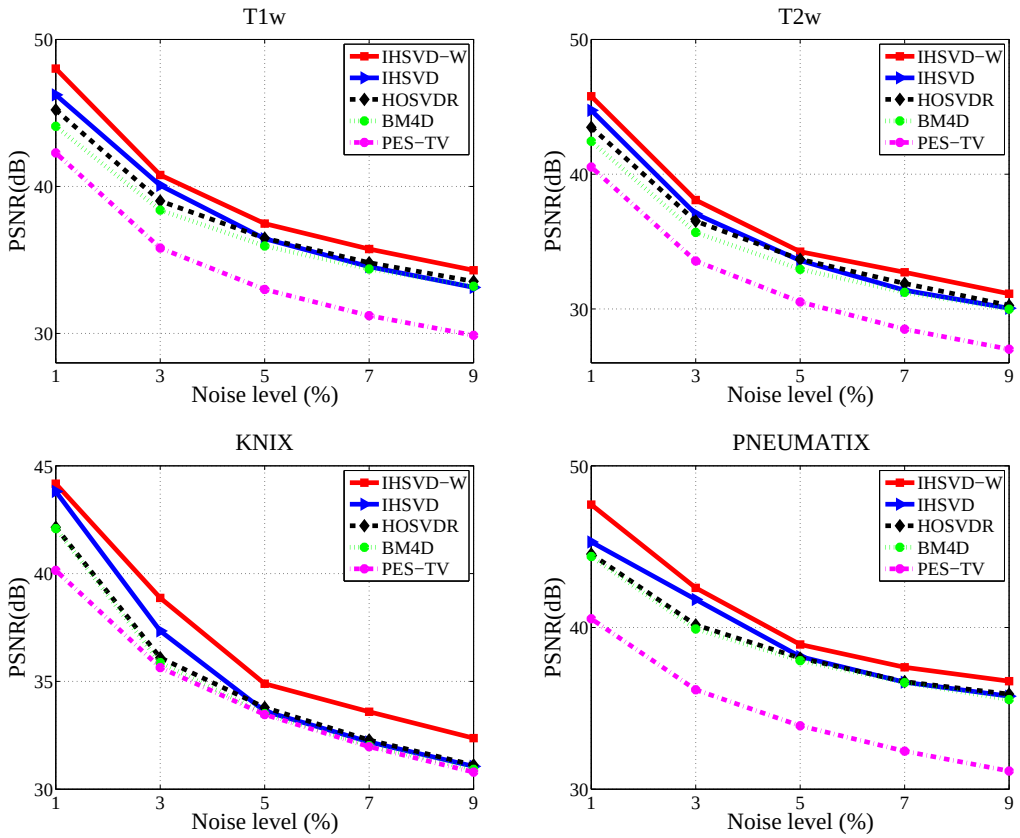


Figure 4.3: PSNR (dB) comparison on synthetic data under Gaussian noise of various levels [Numerical results are tabulated in Appendix G].

4.3.1 Synthetic MRI Datasets

In this experiment, the aforementioned test is applied on synthetic data along with four MR images: T1 weighted (T1w), and T2 weighted (T2w) from the Brainweb phantom resolution [104]. The parameters of these synthetic images are as follows: the size is $181 \times 217 \times 181$, the voxel resolution is 1mm^2 (8 bit quantization). The KNIX data set of size $256 \times 256 \times 26$ and PNEUMATIX data set of size $192 \times 256 \times 25$ from the OsiriX repository [105] are also used. We synthetically generate the noisy observations $\boldsymbol{X}$ using different values of standard deviation of the noise σ, ranging from 1% to 9%.

Figures 4.3–4.8 show the PSNR, SSIM, and MAD comparisons of the aforementioned

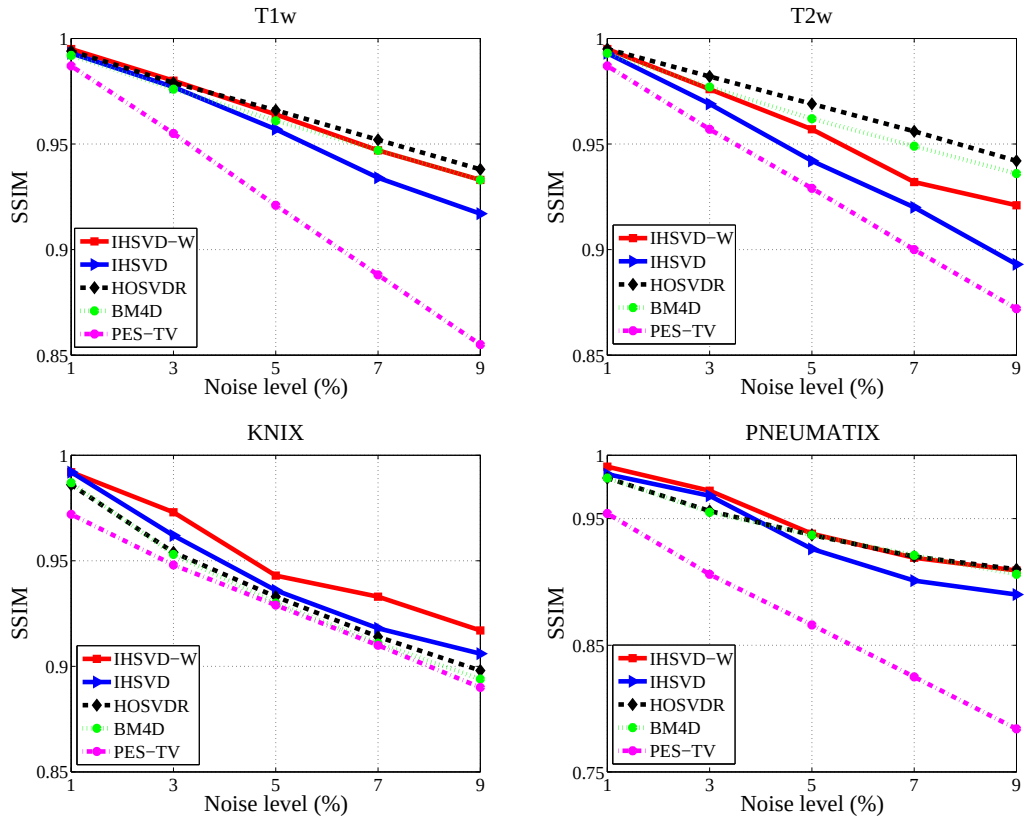


Figure 4.4: SSIM comparison on synthetic data under Gaussian noise of various levels [Numerical results are tabulated in Appendix G].

algorithms at various noise levels for both Gaussian and Rician distributed data, respectively. In PSNR and MAD, the IHSVD is competitive with the three denoising comparison methods, and the IHSVD-W considerably out-performs those three approaches. This is especially the case for noise levels of less than 5%. Note however, that the HOSVDR yields slightly higher SSIM than the proposed IHSVD-W when the noise level is greater than 5%. The proposed algorithms performed better in the PSNR and MAD quantitative metrics than in SSIM. In terms of computational cost, compared with state-of-the-art methods, the proposed algorithm spends less time to denote 3D MRI since it is a holistic algorithm dealing with a single cube. It is obvious that patch-based approaches will have to deal with many localized patches or cubes with

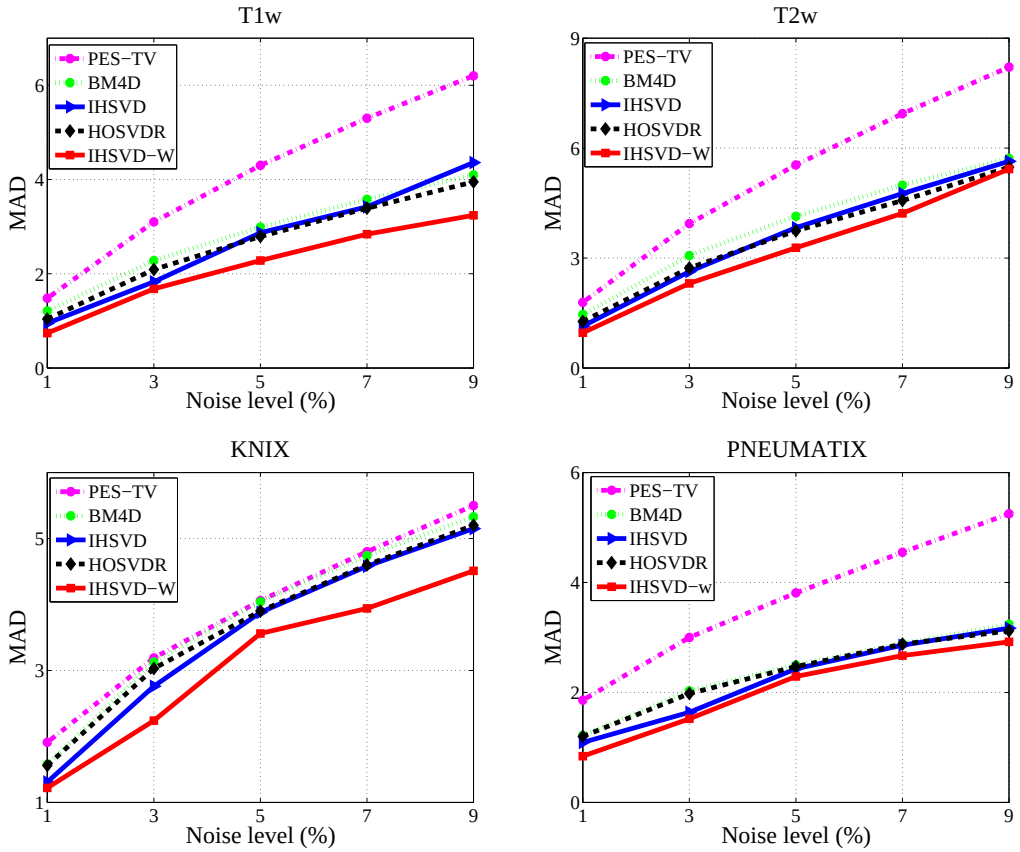


Figure 4.5: MAD comparison on synthetic data under Gaussian noise of various levels [Numerical results are tabulated in Appendix G].

higher computational burden. Overall, we can conclude that the proposed algorithms have competitive performance with the comparison denoising methods in the set of fabricated noisy images.

Figure 4.9 shows the original T1w, T2w, KNIX and PNEUMATIX MR images and corrupted images with Gaussian and Rician distributed noise of $\sigma = 9\%$. To provide a better visual inspection, we show the denoised images in Figure 4.10 and the corresponding residual images in Figure 4.11 in Gaussian distributed noise. The corresponding results for Rician distributed noise are in the Figures 4.12–4.13. One can notice that the reconstructions using the HOSVDR, BM4D and PES-TV algorithms

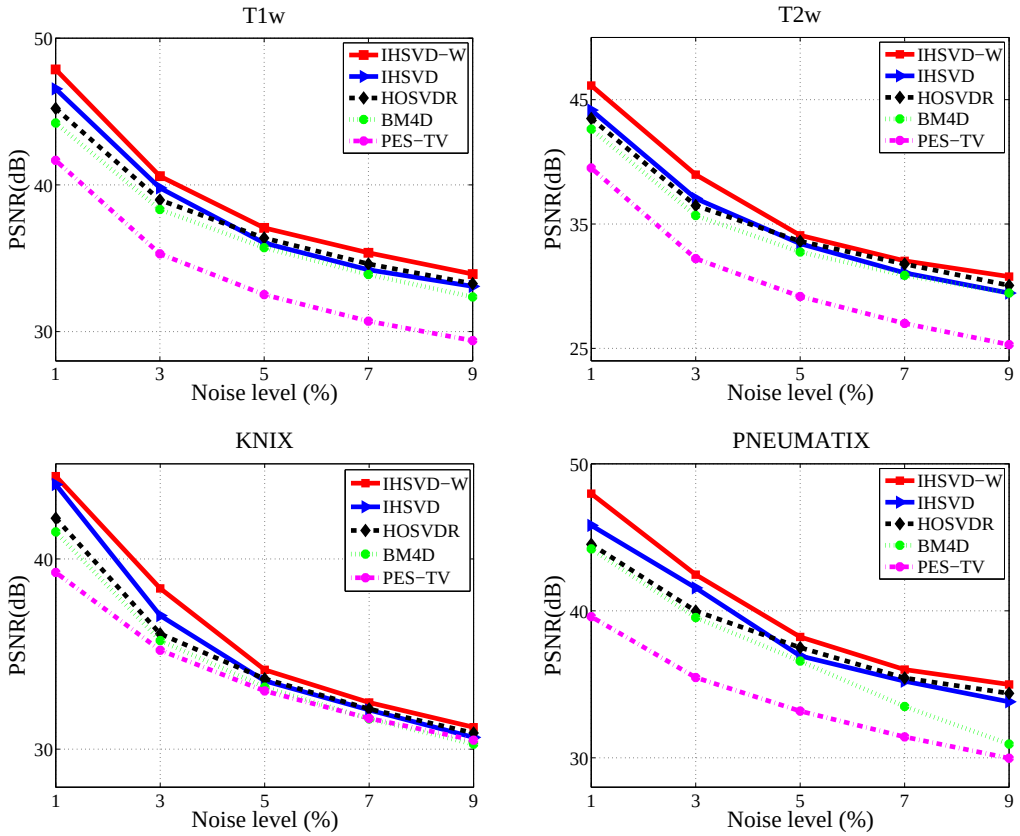


Figure 4.6: PSNR (dB) comparison on synthetic data under Rician noise of various levels [Numerical results are tabulated in Appendix H].

exhibit some artifacts. Inspection of these images show that the proposed IHSVD-W method indeed offers improved denoising. Clearly the proposed algorithm has the least amount of artifacts. This is particularly visible in the reconstruction of the high frequency-dependent details.

Therefore, while the IHSVD-W performs comparably with the HOSVDR in terms of SSIM, it provides better results compared to the other three algorithms in terms of PSNR, MAD, and visual comparisons. Accordingly, it can be concluded that the proposed algorithm is competitive with state-of-the-art methods.

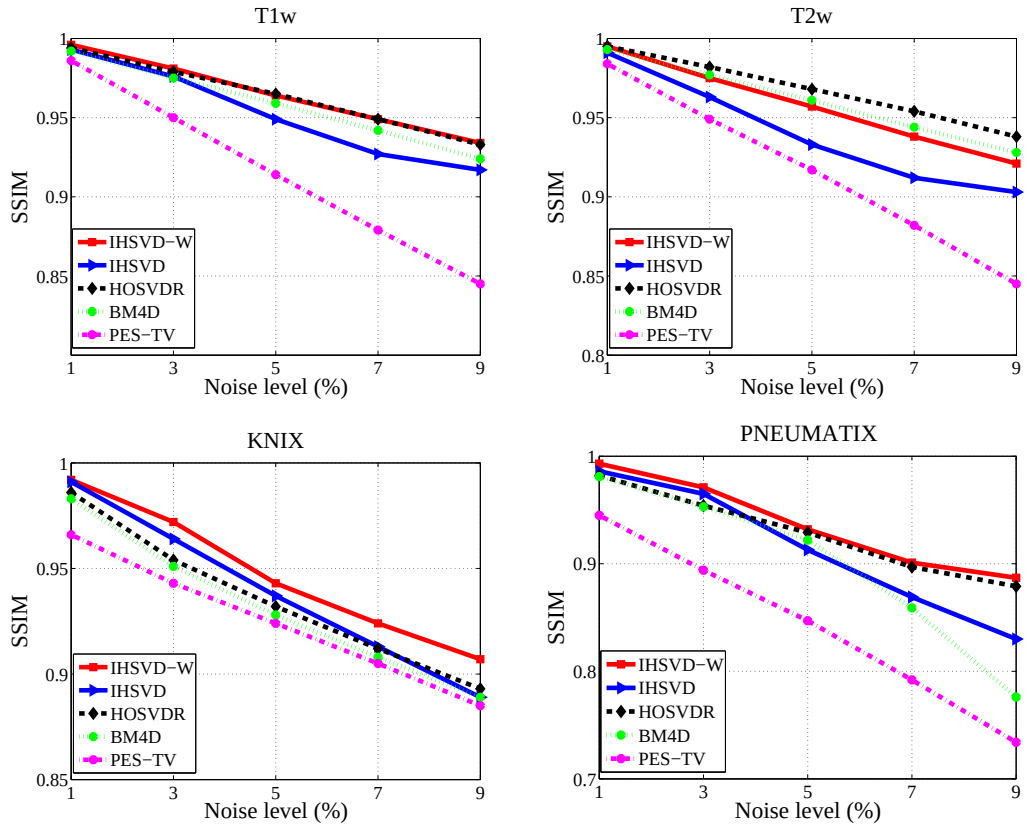


Figure 4.7: SSIM comparison on synthetic data under Rician noise of various levels [Numerical results are tabulated in Appendix H].

4.3.2 Real MRI Datasets

In this experiment, real MR images are considered. We use a real cross-sectional T1w MR brain images (OAS1-0112) from the publicly open access series of imaging studies (OASIS) database [106]. The images were acquired using a Siemens 1.5 T Vision scanner. The acquisition parameters of these images are as follows: the repetition time is 9.7ms, the echo time is 4ms, the flip angle is 4^o, the inversion time is 20ms, the duration time is 200ms, the tensor dimension is $256 \times 256 \times 128$ and the voxel resolution is $1 \times 1 \times 1.25\text{mm}^3$. The noise levels in these images are calculated to be approximately 3%. We used the method in [102] for this noise level estimation. Because the truly noiseless images are unknown, we compare the reconstructions only visually. Recon-

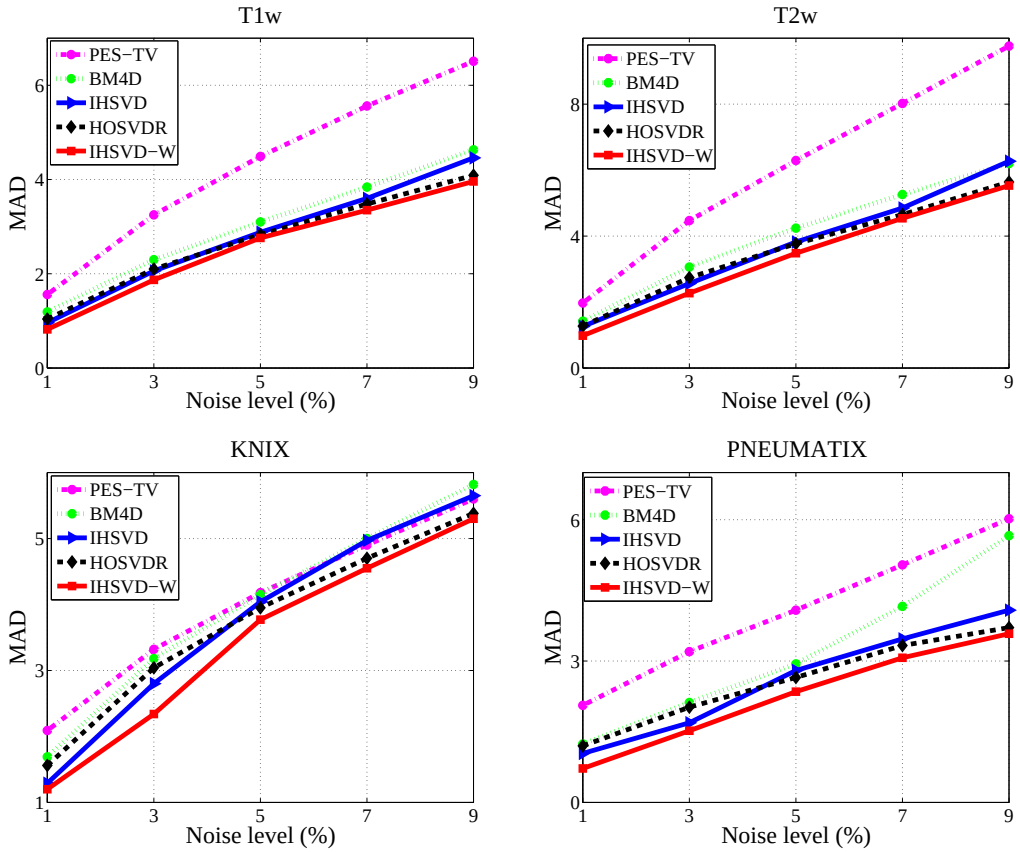


Figure 4.8: MAD comparison on synthetic data under Rician noise of various levels [Numerical results are tabulated in Appendix H].

structions of the four aforementioned algorithms are shown in Figure 4.14. From this figure, one can conclude that both versions of the proposed algorithms achieve better reconstructions compared to the others. This is especially visible when comparing the enlarged insets in Figure 4.14. It is evident that the proposed algorithm is better able to preserve detail features and boundaries, whereas the HOSVDR, BM4D and PES-TV algorithms blur the regions and hence reduce the contrast.

4.4 Conclusion

In this chapter, we presented a denoising algorithm for MR images. The proposed algorithm forms a single tensor from the observed noisy MR image slices. This tensor undergoes an HOSVD, where its sparse representation coefficients are calculated with respect to a set of orthogonal basis matrices. These matrices are extracted over the

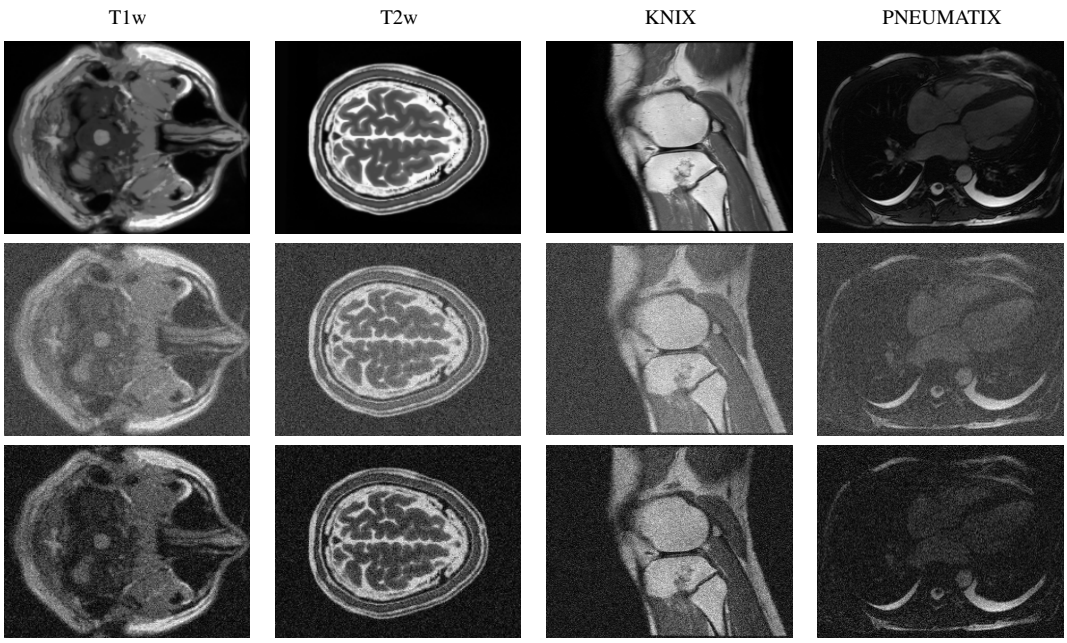


Figure 4.9: Original images (row 1), images corrupted by Gaussian noise (row 2) and Rician noise (row 3) at noise level of $\sigma = 9\%$.

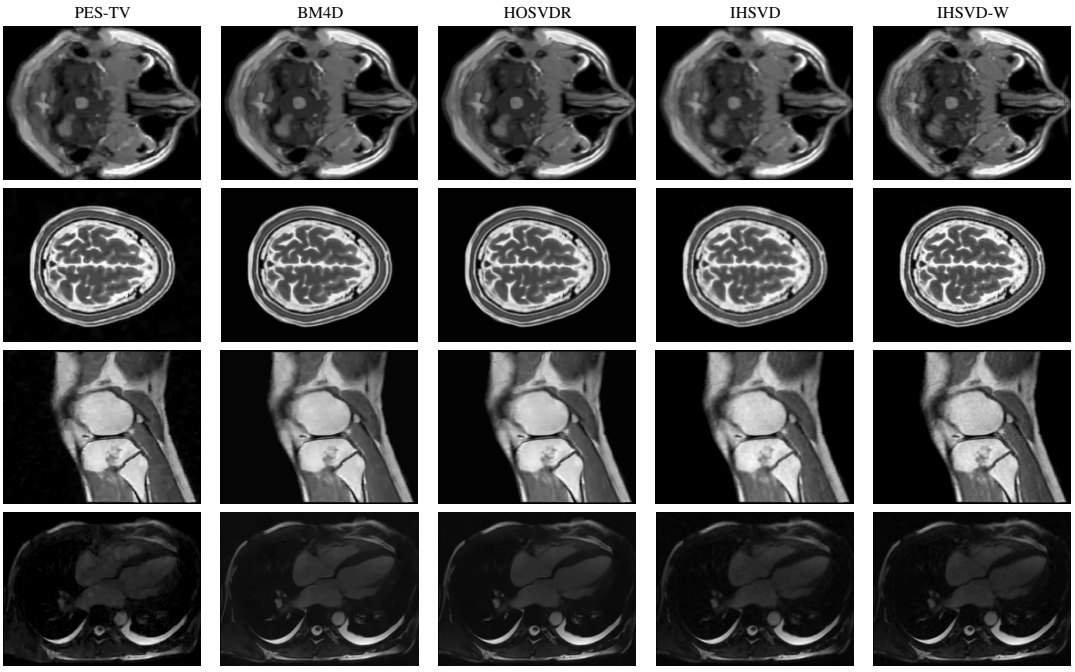


Figure 4.10: Denoising results of algorithms applied to T1w (row 1), T2w (row 2), KNIX (row 3) and PNEUMATIX (row 4) images corrupted by Gaussian noise at noise level of $\sigma = 9\%$.

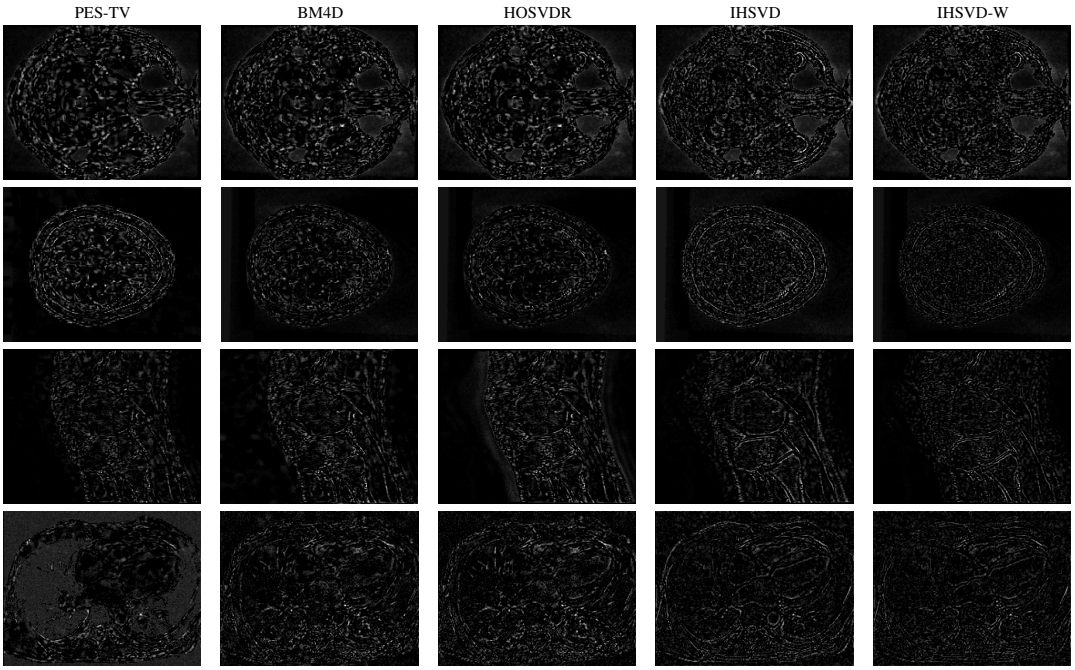


Figure 4.11: The residual images after denoising the Gaussian noise: T1w (row 1), T2w (row 2), KNIX (row 3) and PNEUMATIX (row 4).

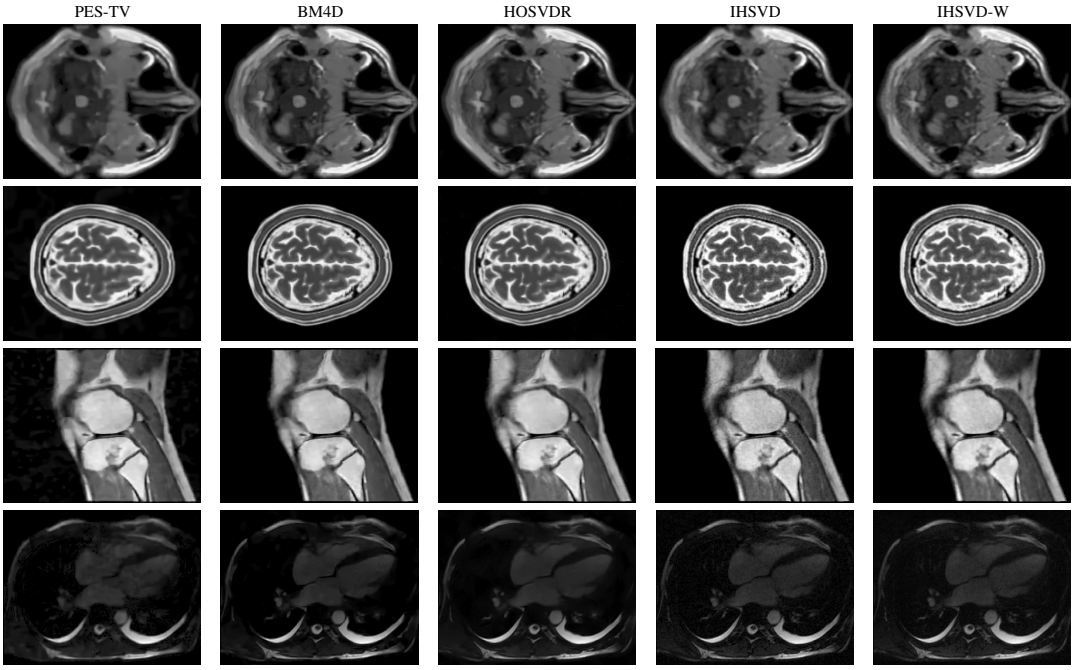


Figure 4.12: Denoising results of algorithms applied to T1w (row 1), T2w (row 2), KNIX (row 3) and PNEUMATIX (row 4) images corrupted by Rician noise at noise level of $\sigma = 9\%$.

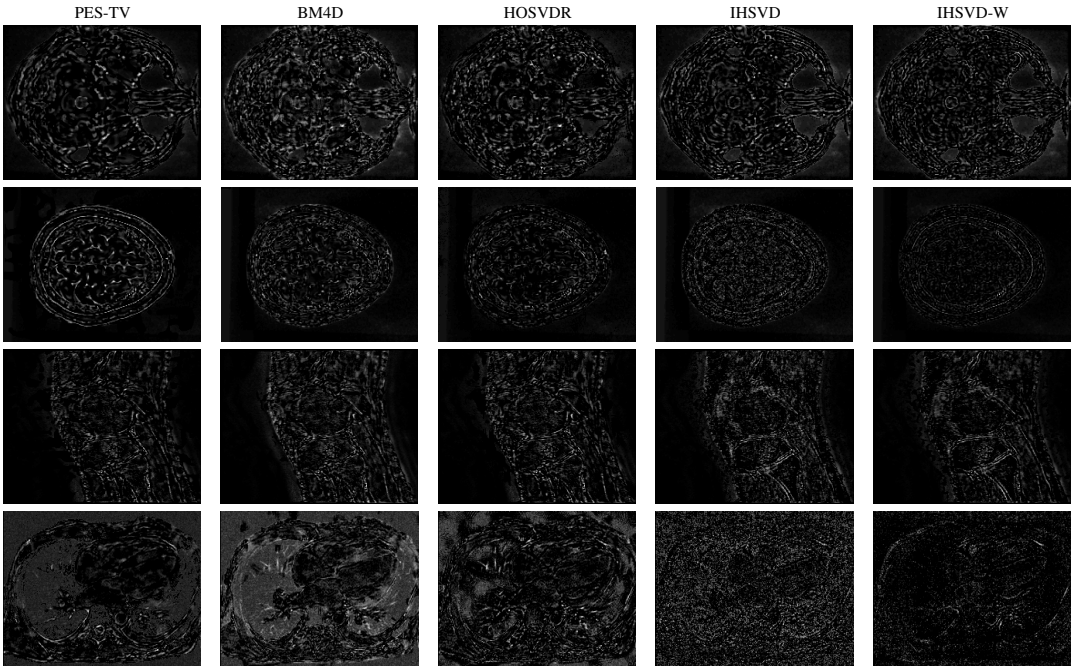


Figure 4.13: The residual images after denoising the Rician noise: T1w (row 1), T2w (row 2), KNIX (row 3) and PNEUMATIX (row 4).

tensor data to render improved representational quality. In addition, the composition of the observed data into one tensor allows better exploitation of image self-similarity across the slices of the MR image. Denoising is achieved by iteratively applying soft thresholding on the calculated sparse representation coefficients. With the use of the HOSVD of tensors, while denoising the data iteratively, the algorithm is able to preserve structure coherence of the tensor in all directions and enhance the computational cost. The IHSVD algorithm is further enhanced with Wiener filtering. The extended version IHSVD-W is compared with three recent state-of-the-art algorithms for MR image denoising. The comparison is made quantitative in terms of PSNR, SSIM, and MAD as objective quality metrics. The experiments demonstrate the competitive performance of the proposed algorithms in terms of these metrics. Visual comparisons are in line with the quantitative conclusions.

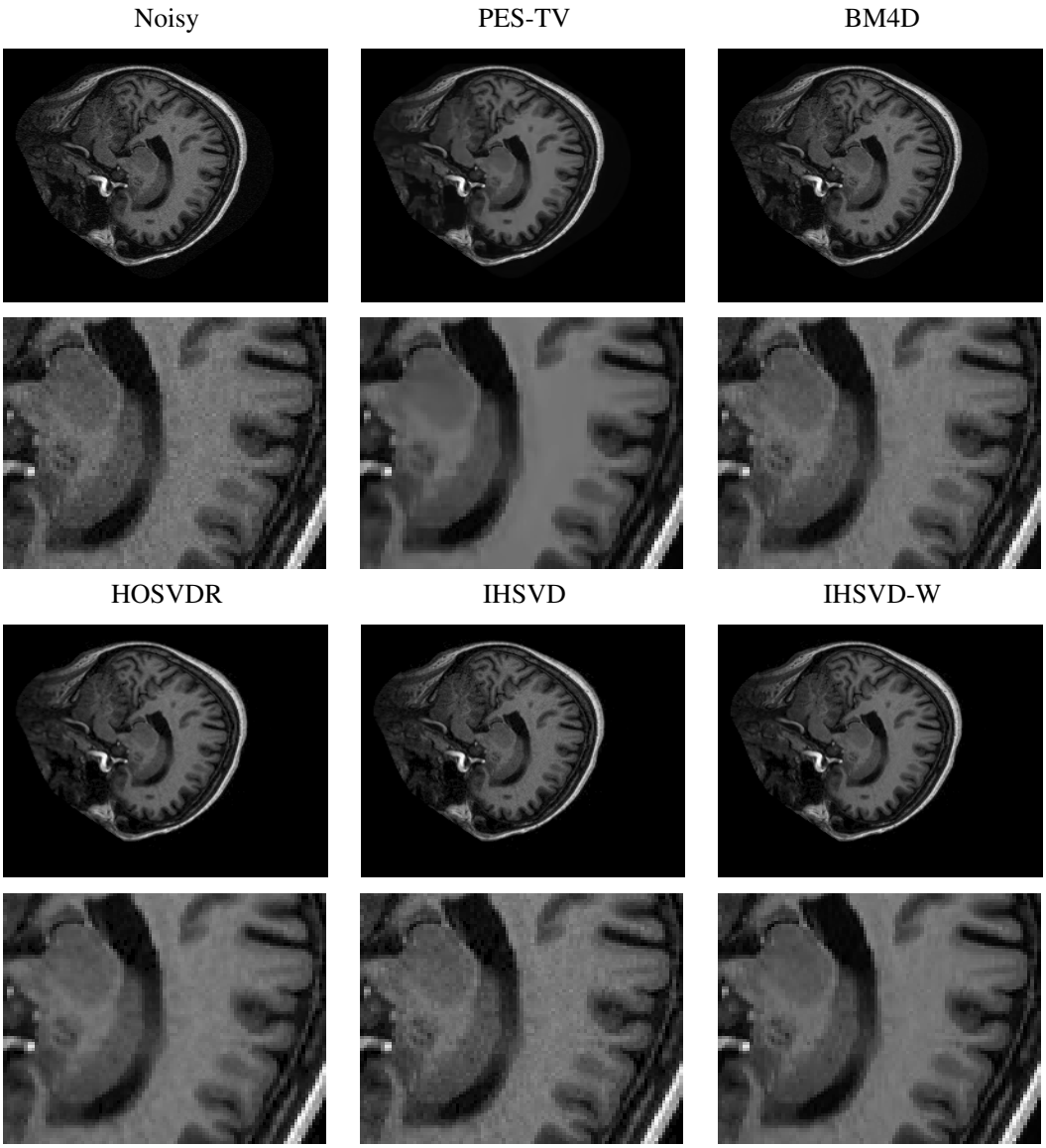


Figure 4.14: Denoising results and the corresponding enlarged insets of OAS1-0112.

Chapter 5

A PATCH-BASED ITERATIVE HYPERSPECTRAL IMAGE DENOISING METHOD VIA HIGHER ORDER SINGULAR VALUE DECOMPOSITION

5.1 Introduction

Denoising is the essential task in hyperspectral image (HSI) preprocessing that can improve the performance of the subsequent applications. In the previous chapter, we developed an MR image denoising method based on iterative regularization under the HOSVD framework, which exhibits good performance in denoising MR images. In this chapter, on the one hand we consider the denoising of the HSIs by employing the proposed iterative HOSVD to improve the performance of the subsequent applications. But on the other hand, we go one step further and propose an iterative denoising algorithm which utilizes the advantages of the patch-based HOSVD sparse model and the iterative regularization technique. The experiments with both synthetic noisy and real HSI reveal that the proposed iterative algorithm improves the HSI quality in terms of both quality metrics and visual inspection. The subsequent classification results further validate the effectiveness of the proposed HSI noise reduction algorithm.

5.1.1 Related Work

HSI is one of the most widely-used imaging techniques in the field of remote sensing to observe the earth surface at various contiguous spectral bands using ground, airborne or spaceborne hyperspectral sensors. With the wealth of spectral information, HSI provides a large amount of information in wide range of applications such as mineral detection and exploration, military surveillance, and other related fields [107, 108, 109, 110].

Unfortunately, due to equipment limitation, such as sensor sensitivity, photon effects, and thermal electronics the HSI is often corrupted by noises in the acquisition process. Noise in HSI can degrade the visual quality and limit the applicability of computerized analysis processes such as unmixing, classification, and object detection [111, 112, 113, 114, 115, 116]. Therefore, to improve the image quality prior to data analysis, HSI denoising is of crucial importance to the successful use of the HSI preprocessing technique.

The noises in HSI can be grouped into two categories namely: fixed pattern noise and random noise. Fixed pattern noise such as striping noise can be reduced by a suitable denoising method. However, random noise due to its stochastic nature cannot be removed entirely. One widely used model for random noise model in HSI is the signal independent AWGN along both spectral and spatial dimensions which is reasonable when the thermal noise is dominant. However, with improvement in the new generation of hyperspectral sensors, in some cases, the dominant noise source is a mixture of

signal independent, signal dependant noise, and fixed pattern noise [111, 112, 117].

Generally, the signal independent AWGN is the situation found in HSI, many methods have been proposed based on this model for the denoising of HSI. The traditional HSI denoising methods adopt the classical 1D and 2D denoising methods to reduce the noise in HSI pixel by pixel, or band by band [117, 118, 119, 120, 121, 122, 123]. However, these kinds of methods do not deal with the spatial and spectral information simultaneously and generally cannot reach good performance in real application. In order to treat the HSI as a whole entity, to remove the noise efficiently, in recent years tensor algebra is brought to jointly analyze the HSI [124].

In HSI processing, images are modelled as a 3D data, i.e., two spatial dimensions and one spectral dimension, which can be considered as a tensor on multidimensional space. Based on the tensor algebra, some recent methods utilize multilinear algebra to analyze the whole HSI directly [124, 125, 126, 127, 128, 129, 130, 131]. Along this research line, based on the Tucker decomposition, Renard *et al*. [125] presented an effective low rank tensor approximation to obtain the low rank approximation of the input HSI. In [126] multidimensional Wiener filtering is used for HSI denoising based on Tucker decomposition. Most recently, Peng *et al*. [130] presented an effective tensor based dictionary learning method based on Tucker model with group block sparsity constraint over the core tensor. Here, each set of grouped similar cubic patches are estimated by low rank tensor approximation. Apart from Tucker decomposition, Liu *et al*. [127] designed the PARAFAC model by utilizing the parallel factor analysis to

denote the HSI for the first time. Moreover, Guo *et al*. [128] proposed a method for HSI noise reduction base on the high order rank-1 tensor decomposition by separating signal dominant component and noise component.

5.2 Proposed Method

The HSI is composed by many bands of 2D images, which can be naturally regarded as a 3D tensor. The aim of the HSI denoising is to reconstruct the clean HSI $\boldsymbol{\mathcal{T}}$ from its noisy measurement $\boldsymbol{\mathcal{X}}$ (as in equation (1.1)). Suppose measurement $\boldsymbol{\mathcal{X}}$ with size $n_1 \times n_2 \times n_3$ is acquired in the presence of noise $\eta(0, \sigma)$. In this model, the noise is assumed to be AWGN, which is the main perturbing noise in HSI.

In the following, the two steps of the iterative HOSVD method are described: First, the global iterative thresholding step which considers the whole HSI as a single cube, is explained. Then, the nonlocal iterative HOSVD which exploits the sparsity inside the grouped similar cubic patches is discussed.

5.2.1 Step One: The Global Iterative Higher Order Singular Value Decomposition

In this section, we employ the proposed iterative regularization method based on the HOSVD transform for HSI denoising. For this purpose, the proposed method first forms a single cube which is composed of all the HSI bands that is able to consider both the spatial and spectral information simultaneously. Similar to the previous chapter for MR image denoising, the denoising is achieved by iterating a thresholding process on the HOSVD sparse representation coefficients which are calculated with respect to orthogonal matrices. These orthogonal basis matrices are learned over the tensor data to

render improved representational quality. In this setting, the sparsity of the core tensor in the HOSVD domain is exploited to reconstruct the denoised HSI. Thus, information from all the spatial and spectral bands will be contributed to the reconstruction.

A trivial assumption for the AWGN in subspace analysis is that the useful data components in HSI are highly correlated between the spectral bands, and the noise is less correlated due to its random distribution. Coefficients of HOSVD can be employed to express the correlation between rank-1 tensors and the data components, therefore to distinguish the dominant component and the noise, we propose to extract the dominant component from noisy measurements by keeping the largest absolute coefficients in the core tensor and setting the rest of them to zero rather than finding the tensor rank of noisy measurement.

The pseudo-code of the global HSI denoising method is like as Algorithm 3 in previous chapter for MRI denoising, except the $\boldsymbol{Y}^0$ is initialized by the noisy HSI.

For better denoising performance, similar to the previous chapter, the method is augmented with a second stage. For clarity, in this chapter this augmented method is referred to as GIHSVD-W. The flowchart of the respective proposed method for HSI denoising is shown in Figure 5.1.

5.2.2 Step Two: Nonlocal Patch-Based Iterative Higher Order Singular value Decomposition

The advantages of the proposed GIHSVD-W method are that it took the correlation between spatial and spectral information into consideration, and tries to eliminate the

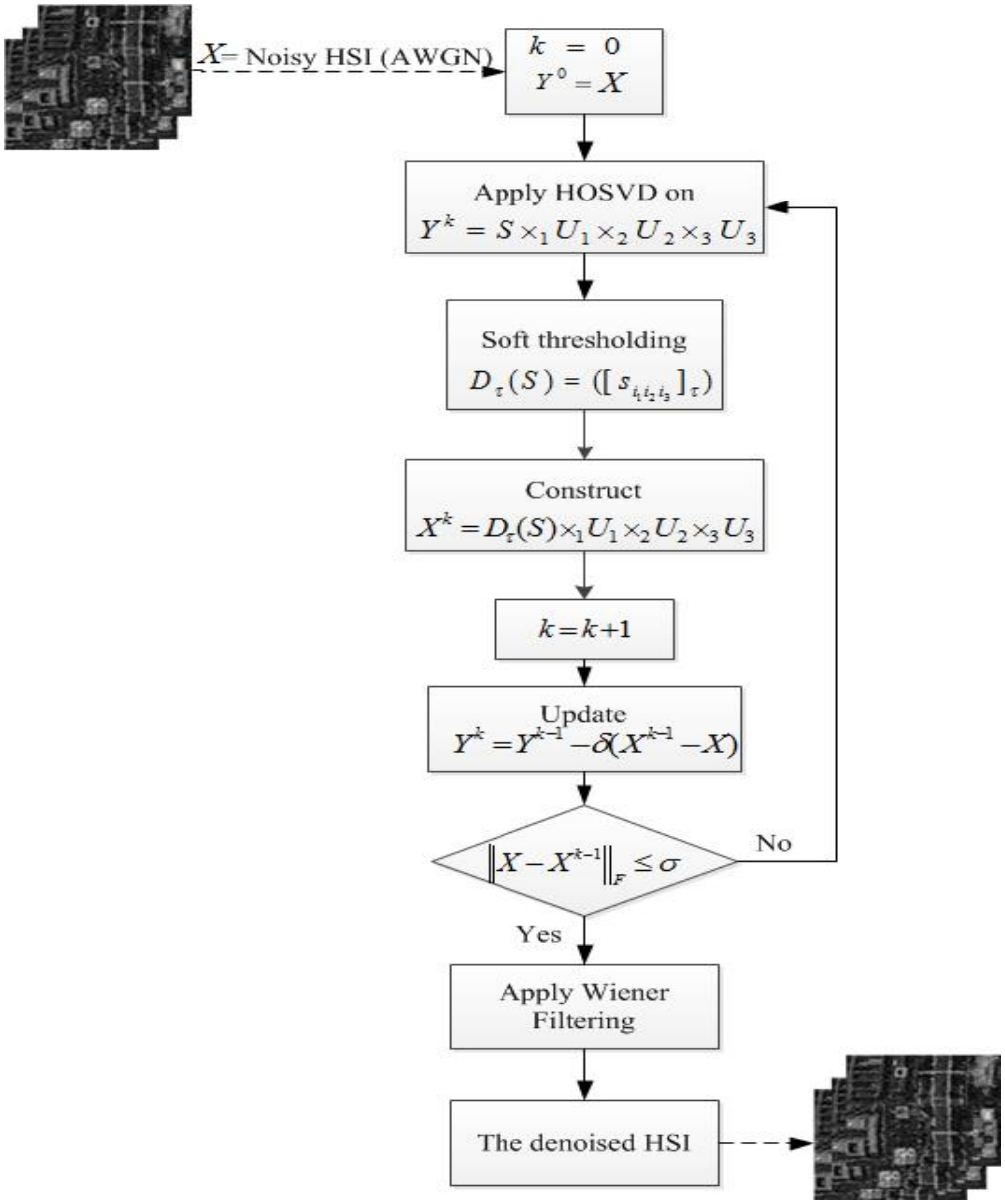


Figure 5.1: Flowchart of GIHSVD-W algorithm.

spectral redundancy of HSI. However, it has not used the nonlocal similarity property of HSI. Therefore, the new method is proposed to estimate the noise free HSI by combining the nonlocal HOSVD sparse model and the iterative regularization framework. By grouping similar cubic patches together to form 4D stack, it can achieve a more enhanced sparse representation in HOSVD domain, which ensures its outstanding de-

noising performance.

The proposed nonlocal iterative HOSVD regularization method scheme based on the discussions in the previous section has two steps: the first step gets an initial estimation of the clean HSI $\boldsymbol{T}$, and the second step further enhances the denoising result of the first step.

The general idea of iterative regularization is to add a part of filtered noise to the denoised data, since in each iteration some detail information in HSI is also removed with the denoising processing; i.e.,

$$\boldsymbol{Y}^k = \boldsymbol{Y}^{k-1} + \delta \left(\boldsymbol{X} - \boldsymbol{X}^k \right) \tag{5.1}$$

where $\boldsymbol{Y}^k$ is the noisy data in the kth iteration. Similar to GIHSVD-W method, $\boldsymbol{Y}^0$ is initialized by the noisy HSI $\boldsymbol{X}$. The proposed patch-based HOSVD method, refer to the process that separates t similar cubic patches (including the reference patch) of size $p \times p \times p$ from the noisy image into a 4D stack, in which the group patches have similar local structures [93, 97].

Next, each stack is projected onto the HOSVD domain. The HOSVD sparse coefficients are calculated with respect to orthogonal matrices, different from the other algorithms [93, 94, 95, 96], the HOSVD bases (factor matrices) is learned from data which are more adapted to the data content, whereas the other algorithms use fixed bases. The HOSVD shrinkage exploits the sparsity of the HOSVD coefficients, to estimate the desired denoised stack. The denoised HSI $\boldsymbol{X}^k$ data at kth iteration is achieved

by aggregating all filtered patches.

As in the previous chapter, this iterative procedure is continued until a stopping criterion is met. The nonlocal HOSVD iterative shrinkage is more general than the nonlocal HOSVD shrinkage. It provides a principled, iterative restoration method for a better chance of getting better results. We term this augmented method as NLIHSVD-W. The complete procedure of the proposed nonlocal HOSVD iterative regularization method is summarized in Algorithm 5 and depicted as the flowchart in Figure 5.2.

The mentioned denoising method can be formulated as in equation (3.2),

$$\hat{\boldsymbol{X}}^t = \arg\min \frac{1}{2}\|\boldsymbol{X}^t - \boldsymbol{T}^t\|_F^2 + \tau\psi(\boldsymbol{X}^t) \tag{5.2}$$

where $\hat{\boldsymbol{X}}^t$ is the denoised cubic patches, $\boldsymbol{T}^t$ is the original cubic patches corresponding to $\boldsymbol{X}^t$, $\psi(\boldsymbol{X}^t)$ is a sparse regularization term in the HOSVD transform domain, and τ is the regularization parameter used to balance the relative contribution between the two terms. Then, equation (5.2) can be formulated for each group as:

$$\hat{\boldsymbol{X}}^t = \arg\min \frac{1}{2}\|\boldsymbol{X}^t - \boldsymbol{T}^t\|_F^2 + \tau\|\boldsymbol{X}^t\|_1 \tag{5.3}$$

5.3 Experimental Results

In this section, the performance of the proposed denoising methods are evaluated on both synthetic and real datasets, and compared with several state-of-the-art denoising methods, including the HOSVDR [99], BM4D [96] and PARAFAC [127]. Performance evaluation is carried out in terms of two quantitative measures, the first measure is the PSNR, and the second measure is the SSIM. Moreover, the parameters (i.e., τ,

Algorithm 5 The Proposed NLIHSVD-W Algorithm for HSI

1: **INPUT:** Observed noisy HSI $\boldsymbol{X}$, threshold τ, standard deviation σ, step size δ, $\gamma \in (0, 1)$, and iteration number K.
2: **OUTPUT:** The denoised HSI $\hat{\boldsymbol{X}}$.
3: **Initialize $\boldsymbol{Y}^0 = \boldsymbol{X}$.**
4: **for k= 0, 2, ..., K do**
Update $\boldsymbol{Y}^k = \boldsymbol{Y}^{k-1} + \delta\ (\boldsymbol{X} - \boldsymbol{X}^k)$
Search for the similar cubic patches from $\boldsymbol{Y}^k$ to cluster them and create 4D tensor $\boldsymbol{Y}^t$.
Update the σ_t as $\sigma_t = \gamma \times \sqrt{\sigma - \|\boldsymbol{X} - \boldsymbol{X}^k\|_F}$
Calculate the HOSVD of each 4D stack as $\boldsymbol{Y}^t = \boldsymbol{S} \times_1 U_1 \times_2 U_2 \times_3 U_3 \times_4 U_4$.
Update the threshold τ_t as equation $\tau_t = \sigma_t \times \sqrt{2 \times \log(n_1 \times n_2 \times n_3 \times n_4)}$
Apply soft thresholding on $\boldsymbol{S}$.
Construct $\boldsymbol{X}^t = \boldsymbol{D}_\tau(\boldsymbol{S}) \times_1 U_1 \times_2 U_2 \times_3 U_3 \times_4 U_4$.
Aggregate all the reconstructed cubic patches to restore clean HSI $\boldsymbol{X}^k$
If $\|\boldsymbol{X} - \boldsymbol{X}^k\|_F \leq \sigma$
break.
5: **end for**
Obtain the final denoised HSI after applying Wiener filtering on $\boldsymbol{X}^k$
$\hat{\boldsymbol{X}} \leftarrow \boldsymbol{X}^k$.

δ, and γ, K) are set as in chapter 4.

5.3.1 Synthetic HSI Dataset

In this experiment, the performance of the aforementioned methods is evaluated on two synthetic HSI datasets: Washington DC Mall [132] and RemoteImage [133] with size of $200 \times 200 \times 160$ and $200 \times 200 \times 89$, respectively. The experiment is made under Gaussian distributed noise, and the noisy observations $\boldsymbol{X}$ is generated using different values of standard deviation of the noise σ, ranging from 1% to 11%.

Tables 5.1 and 5.2 show the PSNR, and SSIM comparisons of the aforementioned methods at various noise levels. Figure 5.3 illustrates the PSNR and SSIM of each band at a noise level of $\sigma = 11\%$.

In view of these tables and figure, as the level of noise increases, it can be observed

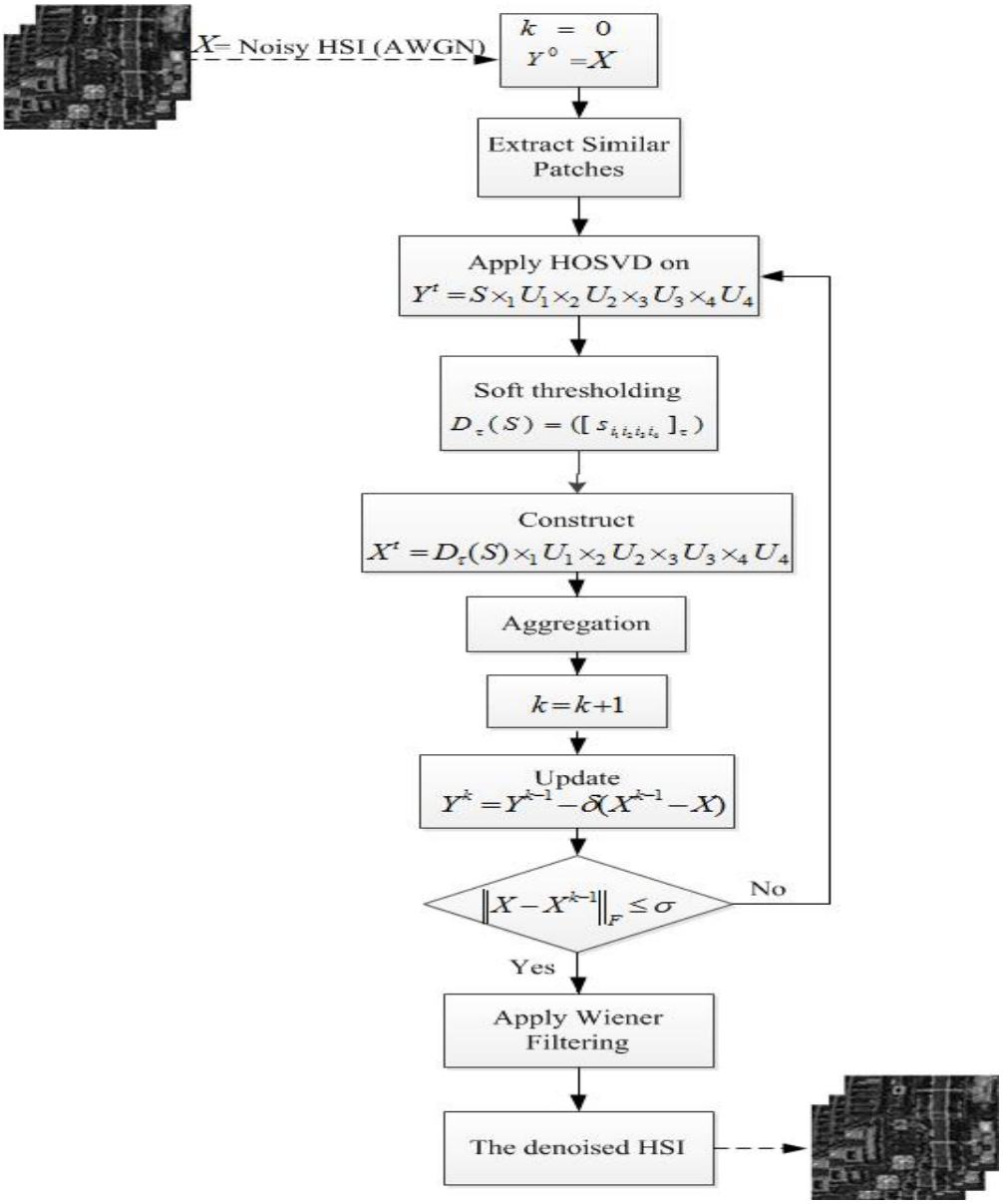


Figure 5.2: Flowchart of NLIHSVD-W algorithm.

Table 5.1: PSNR (dB) and SSIM comparisons of different methods on Washington DC Mall.

Algorithm	1		3		5		7		9		11	
PARAFAC [127]	34.63	0.947	34.28	0.943	33.86	0.937	33.37	0.929	32.84	0.919	32.26	0.907
BM4D [96]	46.51	0.996	39.48	0.979	36.23	0.958	34.12	0.936	32.56	0.913	31.34	0.890
HOSVDR [99]	46.74	0.995	40.15	0.980	36.71	0.958	34.47	0.934	33.37	0.919	32.12	0.898
GIHSVD-W	47.86	0.997	42.03	0.988	38.73	0.976	36.53	0.961	34.80	0.944	33.41	0.925
NLIHSVD-W	**48.53**	**0.997**	**42.63**	**0.988**	**39.31**	**0.976**	**37.00**	**0.962**	**35.21**	**0.947**	**33.72**	**0.930**

Table 5.2: PSNR (dB) and SSIM comparisons of different methods on RemoteImage.

Algorithm	1		3		5		7		9		11	
PARAFAC [127]	31.03	0.870	30.83	0.865	30.74	0.864	30.56	0.859	30.23	0.849	29.96	0.841
BM4D [96]	44.05	0.992	37.26	0.962	34.13	0.921	32.06	0.873	30.58	0.825	29.48	0.781
HOSVDR [99]	44.05	0.991	37.34	0.962	34.12	0.920	32.03	0.872	30.98	0.842	29.84	0.799
GIHSVD-W	44.45	0.992	38.72	0.971	36.62	0.957	34.50	0.932	32.91	0.904	31.65	0.876
NLIHSVD-W	**45.26**	**0.993**	**39.60**	**0.977**	**36.86**	**0.957**	**34.95**	**0.936**	**33.52**	**0.913**	**32.23**	**0.885**

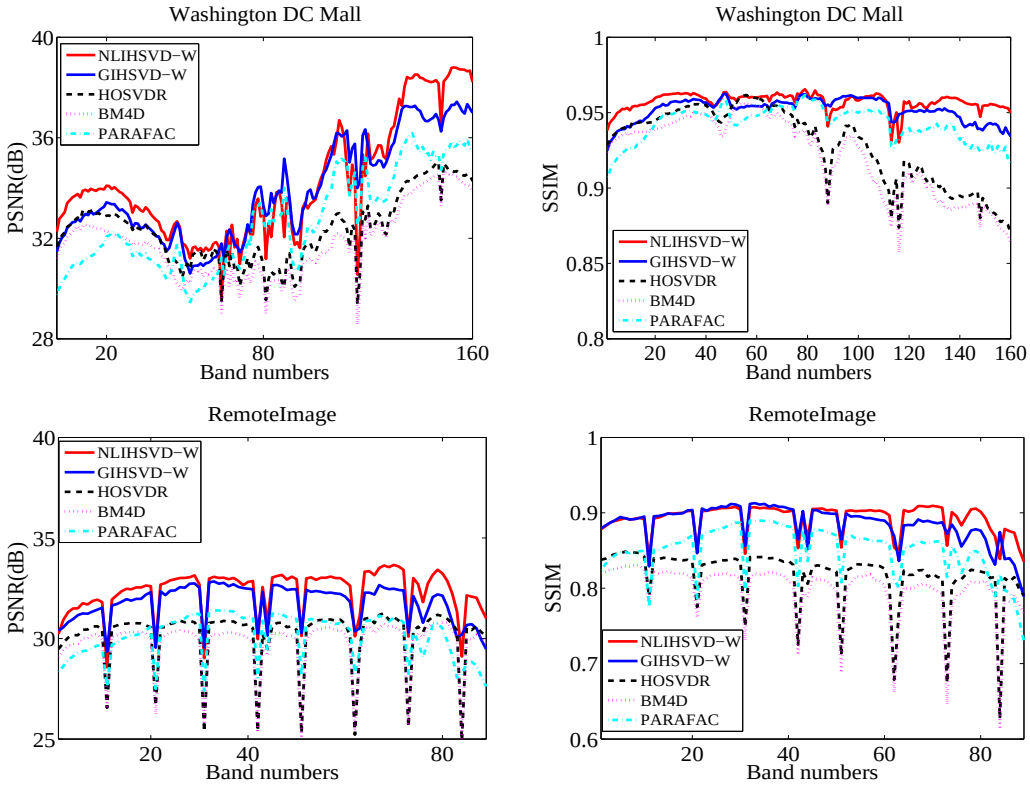


Figure 5.3: PSNR (dB) and SSIM comparisons of different methods on each band of Washington DC Mall (row 1) and RemoteImage (row 2) at noise level of $\sigma = 11\%$.

that the proposed methods gives a promising performance in terms of PSNR and SSIM when compared to the HOSVDR, PARAFAC, and BM4D methods. Note however that, the performance of the NLIHSVD-W shows significant improvement over the GIHSVD-W and the other three denoising methods.

To examine the denoising results visually, Figure 5.4 shows the original and noisy of

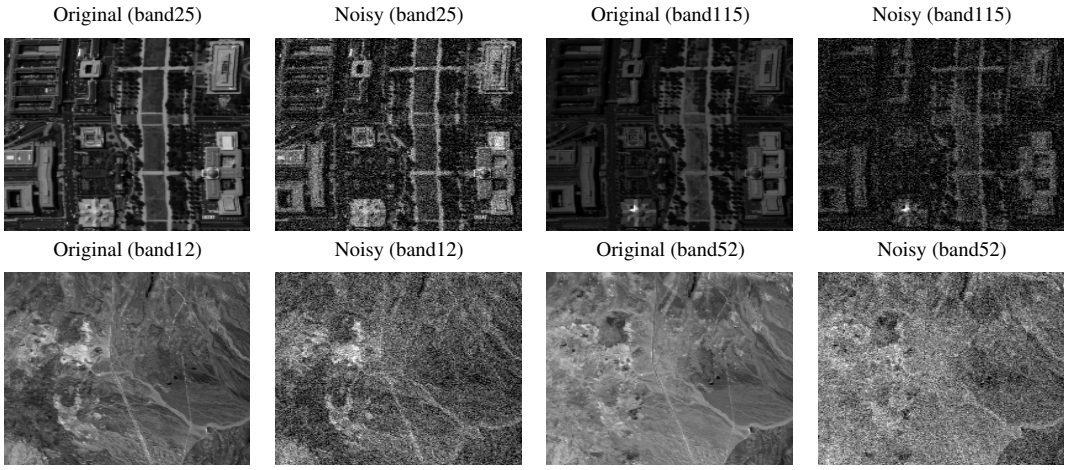


Figure 5.4: Original and noisy images of Washington DC Mall (band 25 and band 115) (row 1); Original and noisy images of RemoteImage (band 12 and band 52) (row 2).

Washington DC Mall, and RemoteImage images. Figure 5.5 illustrates their reconstructions obtained by using the five aforementioned methods for Gaussian distributed noise at a noise level of $\sigma = 11\%$. In terms of visual quality, the proposed HSI denoising methods are comparable and even superior to the state-of-the-art denoising methods. However, one can easily notice that the reconstructions using the HOSVDR, BM4D and PARAFAC methods exhibit some blurring artifacts. In contrast, the reconstructions using the proposed methods does not suffer from that blurring. The detailed features in the HSIs are well preserved by the GIHSVD-W and NLIHSVD-W methods compared with other three denoising methods. As can be seen from Figures 5.5, the results of the NLIHSVD-W are more visually pleasant than those of other compared methods. To provide a better visual inspection, we opt to compare residual images as shown in Figure 5.6. The residual images illustrate that the proposed NLIHSVD-W method offers improved denoising. Overall, we can conclude that the GIHSVD-W and NLIHSVD-W methods preserve well in details.

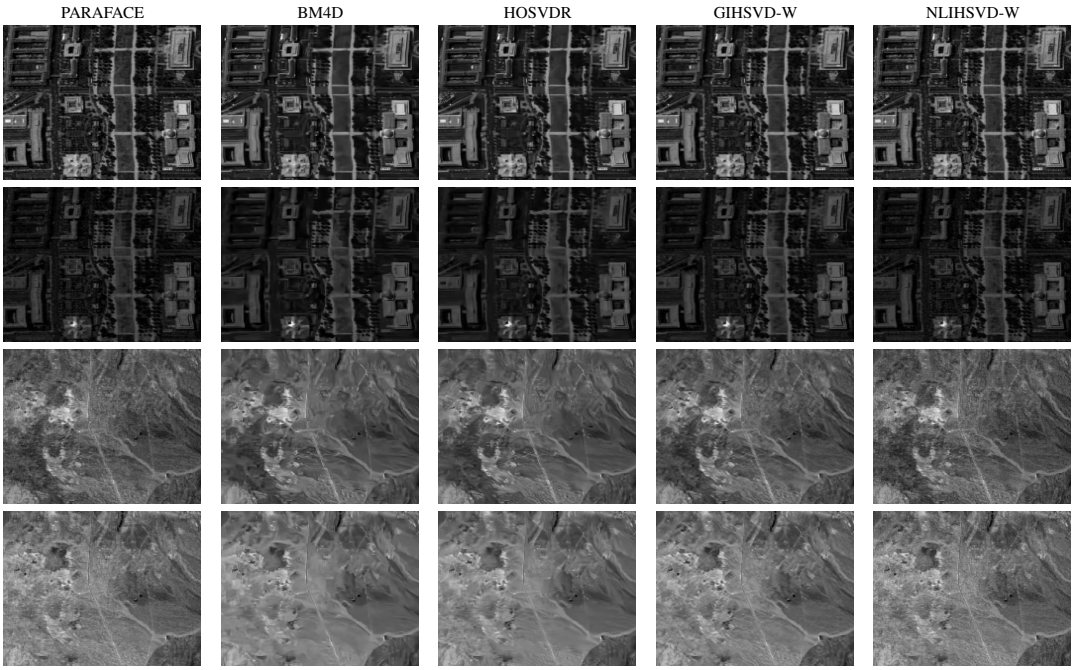


Figure 5.5: Denoising results of methods applied to Washington DC Mall, band 25 (row 1), band 115 (row 3), and the RemoteImage, band 12 (row 3), band 52 (row 4).

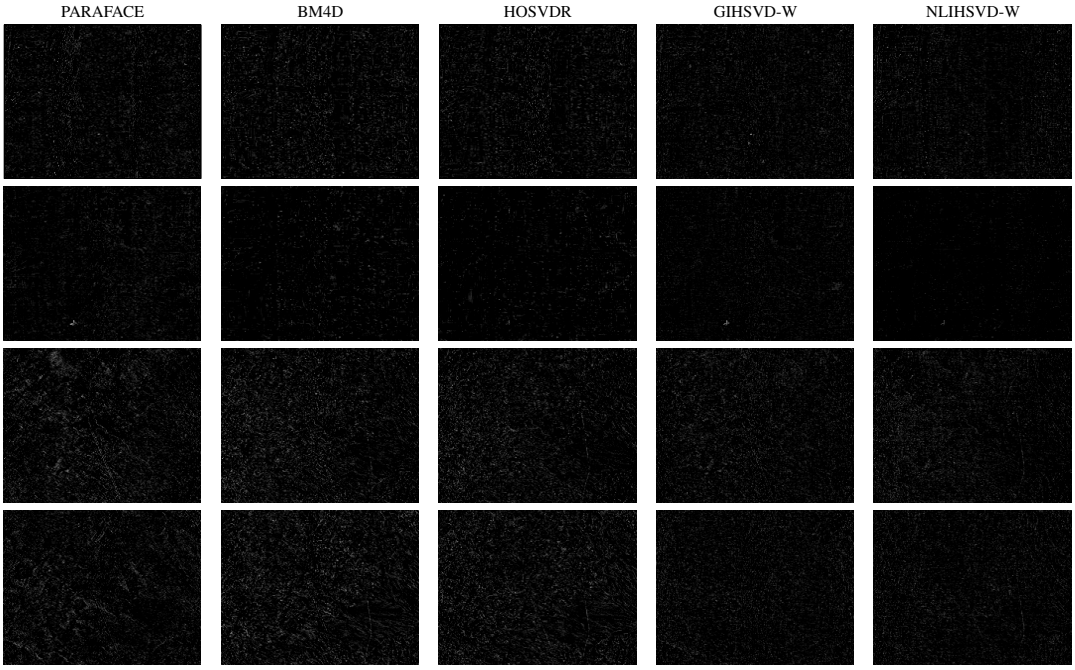


Figure 5.6: The residual images after denoising of Washington DC Mall, band 25 (row 1), band 115 (row 2), and the RemoteImage, band 12 (row 3), band 52 (row 4).

5.3.2 Real HSI Dataset

In this experiment, the consistency of the proposed method is verified on real HSI dataset. The effectiveness of the proposed methods on real HSI dataset is verified on the Indian Pines dataset which was collected by the AVIRIS sensor over the Indian Pines region in Northwestern Indiana in 1992. The acquisition parameters of these images are as follows: the image size is 145×145, there are 220 bands in the image, covering the wavelength range of $0.4 - 2.5\mu\mathrm{m}^3$. The noise level in some bands is high and there is nearly no useful information in these bands, therefore, the number of bands is reduced to 200 by abandoning the bands covering the region of water absorption: $104 - 108$, $150 - 163$, and 220.

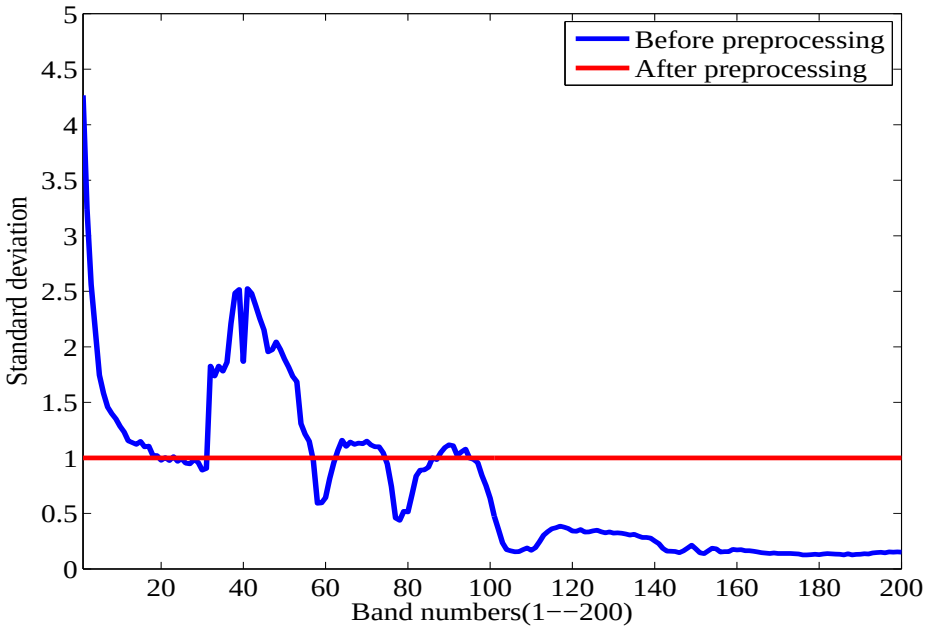


Figure 5.7: The standard deviation of each band of Indian Pines before and after preprocessing step.

In real HSI, the noise level in each spectral band is different. However, the proposed method depends on the assumption that the noise is uniformly distributed in each band.

In order to apply the proposed denoising method, a preprocessing stage is necessary to make the noise level uniformly. Here, the preprocessing stage involves the following steps: 1) The standard deviation of each band is estimated in wavelet domain as [134]

$$\sigma_i = \frac{\text{median}(|W_i^{(L+1)}|)}{0.6745} \tag{5.4}$$

where $W_i^{(L+1)}$ is the wavelet coefficients for the $(L+1)$th subband (HH) in the ith band. Note that we use Daubechies wavelet with 10 coefficients in 3 level decompositions. 2) The uniform σ is obtained by taking the mean value of the all estimated standard deviations. 3) The standard deviation of each band is changed to σ by the following formula,

$$X(:,:,i) = \frac{\sigma}{\sigma_i} X(:,:,i) \tag{5.5}$$

where $X(:,:,i)$ is the intensity in the ith band. 4) At the end, the process is inverted after applying the proposed denoising method on the data,

$$\tilde{X}(:,:,i) = \frac{\sigma_i}{\sigma} \hat{X}(:,:,i) \tag{5.6}$$

where $\hat{X}(:,:,i)$ is the final recovered result in the ith band, and $\tilde{X}(:,:,i)$ is the intensity in each denoised band. The standard deviation of noise in each band before and after this preprocessing is shown in Figure 5.7. Note that, the truly noiseless Indian Pines dataset are unknown. Therefore, classification is implemented to evaluate the performance of the aforementioned denoising methods. Here, the support vector machine algorithm (SVM) [135] is used as the classifier, and cross validation (CV) is employed to set the parameters. The SVM has been widely used in HSI classification [136, 137, 138]. The main idea is to project the data to a higher dimensional space and use a hyperplane to get a better separation. This is done by using a kernel

Table 5.3: Number of total, training, and test samples used in the classification of the Indian Pines dataset.

Name of class	Total	Training	Testing
Alfalfa	46	15	31
Corn-notill	1428	100	1328
Corn-mintill	830	100	730
Corn	237	50	187
Grass-pasture	483	50	433
Grass-trees	730	100	630
Grass-pasture-mowed	28	15	13
Hay-windrowed	478	50	428
Oats	20	15	5
Soybean-notill	972	100	872
Soybean-mintill	2455	150	2305
Soybean-cleantill	593	50	543
Wheat	205	50	155
Woods	1265	100	1165
Buildings-Grass-Trees-Drives	386	50	336
Stone-Steel-Towers	93	50	43

Table 5.4: Classification accuracy (%) for Indian Pines datasets with different denoising methods.

Name of class	Noisy	PARAFAC	BM4D	HOSVDR	GIHSVD-W	NLIHSVD-W
Alfalfa	74.19	100	93.55	77.42	96.77	87.10
Corn-notill	73.57	80.35	84.04	84.79	83.13	83.06
Corn-mintill	70.14	79.45	86.99	90.14	88.63	91.92
Corn	77.01	90.91	87.17	93.58	95.72	99.47
Grass-pasture	89.38	85.91	94.69	90.99	96.30	96.30
Grass-trees	92.06	99.21	98.25	97.62	97.78	97.94
Grass-pasture-mowed	100	92.31	100	92.31	100	100
Hay-windrowed	97.66	97.43	99.07	99.53	100	99.53
Oats	100	100	100	100	100	100
Soybean-notill	68.35	89.91	81.77	84.06	91.86	92.20
Soybean-mintill	65.03	81.69	85.86	81.74	88.72	90.89
Soybean-cleantill	78.64	72.56	90.42	93.00	93	86
Wheat	95.48	90.32	98.71	98.71	99.35	100
Woods	88.84	94.16	96.82	97.51	96.65	99.14
Buildings-Grass-Trees-Drives	74.11	83.63	82.74	80.65	97.62	95.54
Stone-Steel-Towers	97.67	95.35	97.67	97.67	100	100
Overall Accuracy	76.65	85.82	89.07	88.65	**91.51**	**92.41**

function like the quadratic function. To avoid unnecessary deviation, the number of the training and testing samples are chosen randomly (shown in Table 5.3), and the classification experiments are repeated 30 times. The efficiency of aforementioned

methods is evaluated in term of the overall accuracy. Table 5.4 illustrates the overall accuracy using the proposed methods and the other three methods. It is evident that the proposed GIHSVD-W and NLIHSVD-W methods can be better able to improve the classification performance of HSI than other three algorithms methods. In summary, the proposed method is better for detail preservation and is more effective than the other three methods in classification case. Reconstructions of the five aforementioned methods are shown in Figure 5.8. Thus, from the view of visual effect, it can be seen that the performance of the proposed methods outperforms the other three methods.

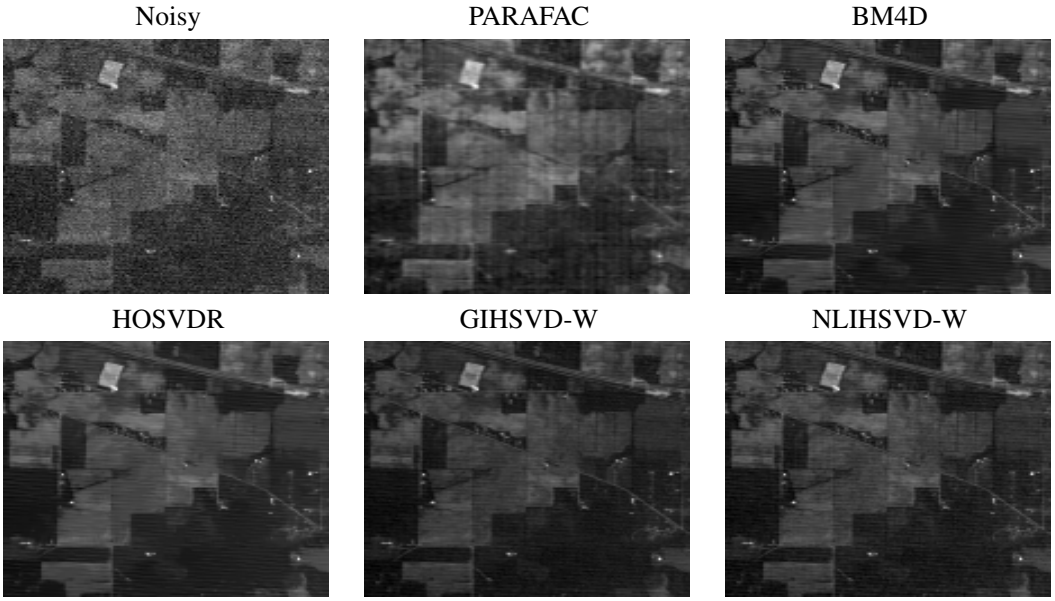


Figure 5.8: Denoising results of Indian Pines (band 2).

5.4 Conclusion

In this chapter, two iterative HSI denoising methods based on HOSVD transform were proposed. One method is simply an extension of the global iterative regularization based on HOSVD to HSIs, named as GIHSVD-W. However, due to the curse of dimensionality, proposing a global denoising method for the processing the whole 3D HSI data is often found to be too hard. Therefore, in the second method for further im-

provement, we presented a simple and efficient method for HSI denoising, which takes the advantage of the patch-based HOSVD sparse model and the iterative regularization framework by integrating the proposed global iterative method into the nonlocal similar cubic patches group. For better denoising performance, the proposed methods are further enhanced with a post process Wiener filtering.

The proposed denoising methods for HSI are compared with three recent state-of-the-art methods. The comparison is made in terms of PSNR and SSIM as objective quality metrics. Experimental results indicate the competitive performance of the proposed algorithms in terms of these metrics. Visual comparisons are in line with the quantitative conclusions. The real dataset is used to evaluate the denoising and classification results of the aforementioned methods. The results of the denoised real dataset are classified by SVM, the classification accuracy shows improvements compared to the results with the other three state-of-the-art methods in the experiments.

Chapter 6

CONCLUSIONS AND FUTURE WORK

6.1 Conclusion

The aim of this thesis was towards restoration of multidimensional data, including tensor inpainting, MRI denoising and, HSI denoising. In this thesis, we proposed an iterative regularization algorithm by employing the HOSVD transform and exploiting the sparsity of the multidimensional signals. At each iteration, the HOSVD is performed; the core tensor obtained is then modified via soft thresholding. While preserving the structural information and acknowledging contribution of large entries in the core tensor, we further obtain approximate solution of the restoration problem by synthesizing the HOSVD using modified core tensor. This corresponds to the minimization of the ℓ_1 of the core tensor coefficients. Accordingly, the proposed restoration algorithm would lead to a tensor with fewer rank-1 terms and yields an approximate solution to sparse problem. The conclusions and future works made through this study will be summarized as follows.

6.1.1 Tensor Inpainting

Chapter 3 developed an iterative sparse HOSVD algorithm (referred to as CTT) to the problem of tensor inpainting which is quite different in nature. The performance of the CTT algorithm is evaluated in three applications, inpainting of multi-channel images, video sequences, and magnetic resonance images (MRI), since they have dif-

ferent characteristics. The visual and quantitative results are compared with some state-of-the-art tensor inpainting algorithms and show that the CTT algorithm can be effective in inpainting applications (especially for large size tensors), thus signifying the algorithm proposed.

In the proposed CTT approach, instead of unfolding or separating the slices of the tensor, the tensor is projected onto the HOSVD transform. In this approach the iterative soft thresholding is performed to regularize the result of the inpainting, leading to a sparse representation of the core tensor. Accordingly, the advantages of the proposed CTT algorithm are that it takes the advantages of global property of the observed data properly and is able to preserve structural coherence of tensor in all directions.

6.1.2 MRI Denoising

Chapter 4 presented the application of iterative sparse HOSVD to denoise a magnetic resonance (MR) image. Here, the proposed denoising algorithm (named as IHSVD) combines the noisy image slices into a single tensor, thereby exploiting non-local image similarity across slices; and transforms into the HOSVD domain. The evaluated HOSVD coefficients are subjected to a soft shrinkage operator to minimize the noise contribution in MRI. With the use of the HOSVD of tensors, while denoising the data iteratively, the algorithm is able to preserve structure coherence of the tensor in all directions. The proposed denoising algorithm is further enhanced with a post-process Wiener filtering (referred to as IHSVD-W).

In order to test the performance of the proposed algorithm, both Gaussian and Rician distributed noise were considered. The performance of the proposed algorithm is eval-

uated on synthetic and real datasets. The experiments demonstrate the competitive performance of the proposed algorithm with the state-of-the-art denoising approaches. And, it can be concluded that the proposed algorithm preserves well in details. In terms of computational cost, compared with the state-of-the-art algorithms, the proposed algorithm spends less time to denoise 3D MRI since it is a holistic algorithm dealing with a single cube.

6.1.3 HSI Denoising

Chapter 5 considered the denoising of the HSIs by employing the global iterative HOSVD (named as GIHSVD-W). The advantages of the GIHSVD-W algorithm are that it took the correlation between spatial and spectral information into consideration, and tries to eliminate the spectral redundancy. However, due to the curse of dimensionality, in this work a new iterative denoising method was proposed which utilizes the advantages of the patch-based HOSVD sparse model and the iterative regularization technique (named as NLIHSVD-W).

The experiments with both synthetic noisy data and real HSI reveal that the proposed iterative algorithm improves the HSI quality in terms of both quality metrics and visual inspection. The real dataset is used to evaluate the denoising and classification results of the proposed HSI denoising algorithm. The results of the denoised real dataset are classified by SVM, the subsequent classification results further validate the effectiveness of the proposed HSI noise reduction algorithm. In terms of the computational cost, the proposed NLIHSVD-W requires more computational cost than the other denoising methods, since both patch-based HOSVD sparse model and iterative regularization

techniques are time consuming techniques, whereas the GIHSVD-W spends less time to denoise 3D HSI since it is holistic method dealing with a single cube.

The superior denoising performance of the NLIHSVD-W method can be attributed to some aspects. First, the iterative nonlocal HOSVD is more general than the nonlocal HOSVD shrinkage. Next, it represents a stack of similar cubic patches by using learned orthogonal matrices. These matrices are extracted over the tensor data which are more adaptable to different data types and can achieve a more efficient and sparse representation than using fixed bases. Lastly, the NLIHSVD-W provides iterative regularization technique to improve the denoising performance. Similar to GIHSVD-W it benefits from the manipulation of better orthogonal matrices from the combined imaged in each iteration, compared to with orthogonal bases from the original noisy data in the first iteration.

6.2 Future work

The following is an account for possible future extensions that can be made on the work conducted in this thesis.

The proposed restoration algorithm can be applied to other medical images such as positron emission tomography and Fundus images which can be considered a good future direction. One limitation of the proposed algorithm is that several filtering parameters are determined heuristically. Automated determination of the parameters will be studied in our future work.

The proposed restoration algorithm is based on the assumption of signal independent noise. Therefore, it is useful for the AWGN. Meanwhile, real MRI and HSI may contain different types of noise in addition to the AWGN, and the proposed method can be improved to handle the mixture noise.

The patch-based tensor approaches can be future direction in tensor inpainting [70, 71, 97, 98]. Adaptation of soft thresholding and iterative regularization procedure into patch-based tensor approaches may help improving reconstruction quality [34, 139]. Moreover, in this direction we would like to show that tensor inpainting is provably accurate even when the few observed entries are corrupted with a small amount of noise.

The quality of the proposed restoration algorithm can be improved by using the concept of the tensor block-sparsity, and the hierarchical Tucker (HT) decompositions. Combination of these special decompositions with patch-based tensor approaches and iterative regularization may also help improving the reconstruction performance.

REFERENCES

[1] H. Lu, K.N. Plataniotis, and A. Venetsanopoulos. “Multilinear Subspace Learning: Dimensionality Reduction of Multidimensional Data.” *CRC Press*, (2013).

[2] T.G. Kolda, and B.W. Bader, “Tensor decomposition and application.” *SIAM Review*, vol. 51, no. 3, pp. 455–500, (2009).

[3] L.R. Tucker. “Implications of factor analysis of three-way matrices for measurement of change.” *Problems in Measuring Change*, pp. 122–137, (1963).

[4] L.R. Tucker. “The extension of factor analysis to three-dimensional matrices.” *Contributions to Mathematical Psychology*, pp. 110–119, (1964).

[5] J.F. Cardoso, “Eigen-structure of the fourth-order cumulant tensor with application to the blind source separation problem.” *Acoustics, Speech, and Signal Processing, 1990. ICASSP-90., 1990 International Conference on IEEE*, pp. 2655–2658, (1990).

[6] P. Comon, “Independent component analysis, a new concept?” *Signal Processing* vol. 36, no. 3, pp. 287–314, (1994).

[7] R. Bro. “Review on multiway analysis in chemistry–2000–2005.” *Critical Reviews in Analytical Chemistry*, vol. 38, no. 3–4, pp. 279–293, (2006).

[8] A. Smilde, R. Bro, and P. Geladi. “Multi-way Analysis: Applications in the Chemical Sciences.” *John Wiley and Sons*, (2005).

[9] L.D. Lathauwer. “Signal processing based on multilinear algebra.” *Katholieke Universiteit Leuven*, (1997).

[10] A. Cichocki, D. Mandic, L.D. Lathauwer, G. Zhou, Q. Zhao, C. Caiafa, and H.A. Phan. “Tensor decompositions for signal processing applications: From two-way to multiway component analysis.” *IEEE Signal Processing Magazine*, vol. 32, no. 2, pp. 145–163, (2015).

[11] M.A.O. Vasilescu and D. Terzopoulos. “Multilinear subspace analysis for image ensembles” *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR 03)*, (2003).

[12] D. Muti and S. Bourennane, “Multidimensional filtering based on a tensor approach.” *Signal Processing*, vol. 85, no. 12, pp. 2338–2353, (2005).

[13] A.G. Mahyari, D. Zoltowski, E. Bernat, and S. Aviyente. “A tensor decomposition based approach for detecting dynamic network states from EEG” *IEEE Transactions on Biomedical Engineering*, (2016).

[14] T.G. Kolda, B. W. Bader, and J. P. Kenny. “Higher-order web link analysis using multilinear algebra” *Proceedings of the 5th IEEE International Conference on Data Mining (ICDM 05)*, (2005).

[15] A. Shashua and T. Hazan. “Non-negative tensor factorization with applications to statistics and computer vision.” *Proceedings of the 22nd International Conference on Machine Learning*, pp. 792–799, (2005).

[16] A.M. Bruckstein, D.L. Donoho, and M. Elad. “From sparse solutions of systems of equations to sparse modeling of signals and images.” *SIAM Review*, vol. 51, no. 1, pp. 34–81, (2009).

[17] P. Comon and C. Jutten. “Handbook of blind source separation: independent component analysis and applications” *Academic Press*, (2010).

[18] A. Cichocki, R. Zdunek, A.H. Phan, and S.I. Amari. “Nonnegative Matrix and Tensor Factorizations: Applications to Exploratory Multi-way Data Analysis and Blind Source Separation.” *John Wiley and Sons*, pp. 792–799, (2009).

[19] D. Muti and S. Bourennane. “Survey on tensor signal algebraic filtering” *Signal Processing.*, vol. 87, no. 2, pp. 237–249, (2007).

[20] A. Cichocki, D. Mandic, L. D. Lathauwer, G. Zhou, Q. Zhao, C. Caiafa, and H. A. Phan. “Tensor decompositions for signal processing applications: From

two-way to multiway component analysis." *IEEE Signal Processing Magazine*, vol. 32, no. 2, pp. 145–163, (2015).

[21] L.D. Lathauwer, B.D. Moor, and J. Vandewalle. "Multilinear singular value decomposition." *SIAM Journal on Matrix Analysis and Applications*, vol. 21, no. 4, pp. 1253–1278, (2000).

[22] L.R. Tucker. "Some mathematical notes on three-mode factor analysis." *Psychometrika*, vol. 31, no. 3, pp. 279–311, (1966).

[23] H. Lu, K.N. Plataniotis, and A.N. Venetsanopoulos. "A survey of multilinear subspace learning for tensor data." *Pattern Recognition*, vol. 44, no. 7, pp. 1540–1551, (2011).

[24] A.H. Phan, and A. Cichocki. "Tensor decompositions for feature extraction and classification of high dimensional datasets." *Nonlinear Theory and its Applications, IEICE*, vol. 1, no. 1, pp. 37–68, (2010).

[25] M.A.O. Vasilescu, and D. Terzopoulos. "Multilinear analysis of image ensembles: Tensorfaces." *European Conference on Computer Vision*, pp. 447–460, (2002).

[26] T. Jiang, and N.D. Sidiropoulos. "Kruskal's permutation lemma and the identification of CANDECOMP/PARAFAC and bilinear models with constant modulus

constraints." *IEEE Transactions on Signal Processing*, vol. 52, no. 9, pp. 2625–2636, (2004).

[27] A. Ozdemir, M.A. Iwen, and S. Aviyente. "Multiscale analysis for higher-order tensors." *Numerical Analysis*, preprint, (2017).

[28] I. Daubechies, M. Defrise, and C.D. Mol. "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint." *Communications on Pure and Applied Mathematics*, vol. 57, no. 11, pp. 1413–1457, (2004).

[29] H. Pan, and T. Blu. "An iterative linear expansion of thresholds for ℓ_1-Based Image Restoration." *IEEE Transactions on Image Processing*,vol. 22, no. 9, pp. 3715–3728, (2013).

[30] A. Bruckstein, D. Donoho, and M. Elad. "From sparse solutions of systems of equations to sparse modeling of signals and images." *SIAM Review*, vol. 51, no. 1, pp. 34–81, (2009).

[31] S. Osher, M. Burger, D. Goldfarb, J. Xu, and W. Yin. "An iterative regularization method for total variation-based image restoration." *Multiscale Modeling and Simulation*, vol. 4, no. 2, pp. 460–489, (2005).

[32] J. Xu, and S. Osher. “Iterative regularization and nonlinear inverse scale space applied to wavelet based denoising.” *IEEE Transaction on Image Processing*, vol. 16, no. 2, pp. 534–544, (2007).

[33] G. Papamakarios. “Robust Low-Rank Modelling on Matrices and Tensors.” *Imperial College London, Department of Computing, MSc Theses*, (2014).

[34] C.F. Caiafa, and A. Cichocki. “Multidimensional compressed sensing and their applications” *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery.*, vol. 3, no. 6, pp. 355–380, (2013).

[35] C.F. Caiafa, and A. Cichocki. “Computing sparse representations of multidimensional signals using kronecker bases.” *Neural Computation*, vol. 25, no. 1, pp. 186–220, (2013).

[36] E.J. Candes and B. Recht. “Exact matrix completion via convex optimization.” *Foundations of Computational Mathematics*, vol. 9, no. 6, (2009).

[37] J.F. Cai, E.J. Candes, and Z. Shen. “A singular value thresholding algorithm for matrix completion.” *SIAM Journal on Optimization*, vol. 20, no. 4, pp. 1956–1982, (2010).

[38] I. Domanov and L.D. Lathauwer. “Multiple invariants and generalized rank of a P-way matrix or tensor.” *Studies in Applied Mathematics*, vol. 7, no. 1–4, pp. 39–79, (1928).

[39] F.L. Hitchcock. “The expression of a tensor or a polyadic as a sum of products.” *Studies in Applied Mathematics*, vol. 6, no. 1–4, pp. 164–189, (1927).

[40] F.L. Hitchcock. “The expression of a tensor or a polyadic as a sum of products. Analysis of individual differences in multidimensional scaling via an N-way generalization of Eckart-Young decomposition.” *Psychometrika*, vol. 35, no. 3, pp. 283–319, (1970).

[41] R.A. Harshman. “Foundations of the PARAFAC procedure: models and conditions for an explanatory multimodal factor analysis” *UCLA Working Papers Phonetics.*, vol. 16, (1970).

[42] J.B Kruskal. “Three-way arrays: rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics.” *Linear Algebra and its Applications*, vol. 18, no. 2, pp. 95–138, (1977).

[43] I. Domanov and L.D. Lathauwer. “On the uniqueness of the canonical polyadic decomposition-part i: Basic results and uniqueness of one factor matrix.” *SIAM Journal on Matrix Analysis and Applications*, vol. 34, no. 3, pp. 855–875, (2013).

[44] I. Domanov and L.D. Lathauwer. “On the uniqueness of the canonical polyadic decomposition-part ii: Overall uniqueness.” *SIAM Journal on Matrix Analysis and Applications*, vol. 34, no. 3, pp. 876–903, (2013).

[45] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester, “Image inpainting.” *Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques*, pp. 417–424, (2000).

[46] C. Guillemot and O.L. Meur. “Image inpainting: overview and recent advances.” *IEEE Signal Processing Magazine*, vol. 31, no. 1, pp. 127–144, (2014).

[47] M. Bertalmio, A. Bertozzi, and G. Sapiro. “Navier-stokes, fluid dynamics, and image and video inpainting.” *Proceedings of the 2001 IEEE Computer Society Conference*, vol. 1, pp. 355–362, (2001).

[48] A. Criminisi, P. Perez, and K. Toyama. “Region filling and object removal by exemplar-based inpainting.” *IEEE Transactions on Image Processing*, vol. 13, no. 9, pp. 1200–1212, (2004).

[49] M. Elad, J.L. Starck, P. Querre, and D.L. Donoho. “Simultaneous cartoon and texture image inpainting using morphological analysis (MCA).” *Applied and Computational Harmonic Analysis*, vol. 19, no. 3, pp. 340–358, (2005).

[50] M.J. Fadili, J.L. Starck, and F. Murtagh. “Inpainting and zooming using sparse representation.” *The Computer Journal*, vol. 52, no. 1, pp. 64–79, (2007).

[51] J.F Cai, H. Ji, F. Shang, and Z. Shen. “Inpainting for compressed images.” *Applied and Computational Harmonic Analysis*, vol. 29, no. 3, pp. 368–381, (2010).

[52] M.V. Alfonso, J.M. Bioucas-Dias, and M.A.T. Figueiredo. “Fast image recovery using variable splitting and constrained optimization.” *IEEE Transactions on Image Processing*, vol. 19, no. 9, pp. 2345–2356, (2010).

[53] S. Ma, D. Goldfard, and L. Chen. “Fixed point and Bregman iterative methods for matrix rank minimization.” *Mathematical Programming*, vol. 128, no. 1, pp. 321–353, (2011).

[54] S.F. Yeganli, and R. Yu, “Image inpainting via singular value thresholding.” *Signal Processing and Communications Applications Conference (SIU)*, Kyrenia, Cyprus, pp. 1–4, (2013).

[55] M. Signoretto, R. Van De Plas, B. De Moor, and J.A.K. Suykens. “Tensor versus matrix completion: A comparison with application to spectral data.” *IEEE Signal Processing Letter*, vol. 18, no. 7, pp. 403–406, (2011).

[56] J. Liu, P. Musialski, P. Wonka, and J. Ye. “Tensor completion for estimating missing values in visual data.” *IEEE 12th International Conference on Computer Vision (ICCV)*, pp. 2114–2121, (2009).

[57] Y.L. Chen, C.T. Hsu, and H.Y.M. Liao. “Simultaneous tensor decomposition and completion using factor priors.” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 3, pp. 577–591, (2014).

[58] S. Gandy, B. Recht, and I. Yamada. “Tensor completion and low-n-rank tensor recovery via convex optimization.” *Inverse Problems*, vol. 27, no. 2, (2011).

[59] Y. Li, J. Yan, Y. Zhou, and J. Yang. “Optimum subspace learning and error correction for tensors.” *Computer Vision-ECCV*, pp. 790–803, (2010).

[60] J. Liu, P. Musialski, P. Wonka, and J. Ye. “Tensor completion for estimating missing values in visual data.” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 1, pp. 208–220, (2013).

[61] Y. Liu, F. Shang, W. Fan, J. Cheng, and H. Cheng. “Generalized higher-order orthogonal iteration for tensor decomposition and completion.” *In Advances in Neural Information Processing Systems*, pp. 1763–1771, (2014).

[62] H. Tan, B. Cheng, W. Wang, Y.J. Zhang, and B. Ran. “Tensor completion via a multi-linear low-n-rank factorization model.” *Neurocomputing*, vol. 133, pp. 161–169, (2014).

[63] Q. Zhao, L. Zhang, and A. Cichocki. “Bayesian CP factorization of incomplete tensors with automatic rank determination.” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, no. 9, pp. 1751–1763, (2015).

[64] J. Zhou, S. Liu, G. Qiu, F. Zhang, and J. Sun. “Tensor missing value recovery with tucker thresholding method.” *In Intelligent Networking and Collaborative Systems (INCoS)*, pp. 716–720, (2013).

[65] H. Rauhut, R. Schneider, and Z. Stojanac. “Low-rank tensor recovery via iterative hard thresholding.” *Linear Algebra and its Applications*, (2013).

[66] M. Filipovic and A. Jukic. “Tucker factorization with missing data with application to low-n-rank tensor completion.” *Multidimensional Systems and Signal Processing*, vol. 26, no. 3, pp. 677–692, (2015).

[67] G. Favier, and A.L.F. de Almeida. “Overview of constrained PARAFAC models.” *EURASIP Journal on Advances in Signal Processing*, vol. 1, no. 142, (2014).

[68] A. Cichocki, D. Mandic, L.D. Lathauwer, G. Zhou, Q. Zhao, C. Caiafa, and H.A. Phan. "Tensor decompositions for signal processing applications from two-way to multiway component analysis." *IEEE Signal Processing Magazine*, vol. 32, no. 2, pp. 145–163, (2015).

[69] M. Ghorai and B. Chanda. "An image inpainting algorithm using higher order singular value decomposition" *22th International Conference on Pattern Recognition.*, pp. 2867–2872, (2014).

[70] M. Ghorai, S. Mandal, and B. Chanda, B., "A Two-step image inpainting algorithm using tensor SVD." *Computer Vision-ACCV 2014 Workshops*, vol. 2, pp. 63–77, (2014).

[71] S.F. Yeganli, R. Yu, and H. Demirel. "An iterative method for tensor inpainting based on higher order singular value decomposition." *Circuits, Systems, and Signal Processing*, (2018).

[72] B. Jiang, F. Yang, and S. Zhang. "Tensor and its Tucker core: the invariance relationships" *Numerical Linear Algebra with Applications*, vol. 24, no. 3, pp. 1253–1278, (2017).

[73] A.O. Rodriguez. "Principles of magnetic resonance imaging." *Revista Mexicana de Fisica*, vol. 50, no. 3, pp. 272–286, (2004).

[74] G.A. Wright. “Magnetic resonance imaging.” *IEEE Signal Processing Magazine*, vol. 14, no. 1, pp. 56–66, (1997).

[75] H. Zhu, Y. Li, J.G. Ibrahim, X. Shi, H. An, Y. Chen, W. Gao, W. Lin, D.B. Rowe, and B.S. Peterson. “Regression models for identifying noise sources in magnetic resonance images” *Journal of the American Statistical Association.*, vol. 104, no. 486, pp. 623–637, (2009).

[76] H. Gudbjartsson, and S. Patz. “The Rician distribution of noisy MRI data.” *Magnetic Resonance in Medicine*, vol. 34, no. 6, pp. 910–914, (1995).

[77] J. Rajan, D. Poot, J. Juntu, and J. Sijbers. “Noise measurement from magnitude MRI using local estimates of variance and skewness.” *Physics in Medicine and Biology*, vol. 55, no. 16, pp. 441–449, (2010).

[78] J. Mohan, V. Krishnaveni, and Y. Guo. “A survey on the magnetic resonance image denoising methods.” *Biomedical Signal Processing and Control*, vol. 9, pp. 56–69, (2014).

[79] K. Krissian, and S.A. Fernandez. “Noise-driven anisotropic diffusion filtering of MRI.” *IEEE Transactions on Image Processing*, vol. 18, no. 10, pp.2265–2274, (2009).

[80] M. Heydari, M.R. Karami, and A. Babakhani. “A new adaptive coupled diffusion PDE for MRI Rician noise.” *Signal, Image and Video Processing*, vol. 10, no. 7, pp. 1211–1218, (2016).

[81] P. Coupe, P. Yger, S. Prima, P. Hellier, C. Kervrann, and C. Barillot. “An optimized blockwise non local means denoising filter for 3D magnetic resonance image.” *IEEE Transactions on Medical Imaging*, vol. 27, no. 4, pp. 425–441, (2008).

[82] J.V. Manjon, P. Coupe, L. Marti-Bonmati, D.L. Collions, and M.Robels. “Adaptive nonlocal means denoising of MR images with spatially varying noise levels.” *Journal of Magnetic Resonance Imaging*, vol. 31, no. 1, pp. 192–203, (2010).

[83] A. Buades, B. Coll, and J.M. Morel. “A review of image denoising algorithms, with a new one.” *Multiscale Modeling and Simulation*, vol. 4, no. 2, pp. 490–530 (2005).

[84] R. Riji, J. Rajan, J. Sijbers, and M.S. Nair. “Iterative bilateral filter for Rician noise reduction in MR images.” *Signal, Image and Video Processing*, vol. 9, no. 7, pp. 1543–1548, (2015).

[85] P.V. Sudeep, P. Palanisamy, C. Kesavadas, and J. Rajan. “Nonlocal linear minimum mean square error methods for denoising MRI.” *Biomedical Signal Processing and Control*, vol. 20, pp. 125–134, (2015).

[86] Y. Wang, X. Che, and S. Ma. “Nonlinear filtering based on 3D wavelet transform for MRI.” *EURASIP Journal on Advances in Signal Processing*, vol. 40, (2012).

[87] C.S. Anand and J.S. Sahambi. “Wavelet domain nonlinear filtering for MRI denoising.” *Magnetic Resonance Imaging*, vol. 28, no. 6, pp. 842–861, (2010).

[88] J. Ma, and G. Plonka. “Combined curvelet shrinkage and nonlinear anisotropic diffusion.” *IEEE Transactions on Image Processing*, vol. 16, no. 9, pp. 2198–2206, (2007).

[89] J.V. Manjon, P. Coupe, A. Buades, D.L. Collins, and M. Robles. “New methods for MRI denoising based on sparsness and self similarity.” *Medical Image Analysis*, vol. 16, no. 1, pp. 18–27, (2012).

[90] S.A. Fernandez, C.A. Lopez, and C.F. Westin. “Noise and signal estimation in magnitude MRI and Rician distributed images: a LMMSE approach.” *IEEE Transactions on Image Processing*, vol. 17, no. 8, pp. 1383–1398, (2008).

[91] J. Rajan, J. Veraart, J.V. Audekerke, M. Verhoye, and J. Sijbers. “Nonlocal maximum likelihood estimation method for denoising multiple coil magnetic reso-

nance images." *Magnetic Resonance Imaging*, vol. 30, no. 10, pp. 1512–1518, (2012).

[92] J. Rajan, B. Jeurissen, M. Verhoye, J.V. Audekerke, and J. Sijbers. "Maximum likelihood estimation-based denoising of magnetic resonance images using restricted local neighborhoods." *Physics in Medicine and Biology*, vol. 56, no. 16, pp. 5221–523, (2011).

[93] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian. "Image denoising by sparse 3-D transform domain collaborative filtering." *IEEE Transactions on Image Processing*, vol. 16, no. 8, pp. 2080–2095, (2007).

[94] M. Tofighi, K. Kose, and A.E. Cetin. "Denoising images corrupted by impulsive noise using projections onto the epigraph set of the total variation function (PES-TV)." *Signal Image and Video Processing*, vol. 9, no. 1, pp. 41–48, (2015).

[95] P. Elahi, S. Beheshti, and M. Hashemi. "BM3D MRI denoising equipped with noise invalidation technique" *IEEE International Conference in Acoustics, Speech and Signal Processing (ICASSP).*, pp. 6612–6616, (2014).

[96] M. Maggioni, V. Katkovnik, K. Egiazarian, and A. Foi. "Nonlocal transform domain filter for volumetric data denoising and reconstruction." *IEEE Transactions on Image Processing*, vol. 22, no. 1, pp. 119–133, (2013).

[97] A. Rajwade, A. Rangarajan, and A. Banerjee. “Image denoising using the higher order singular value decomposition.” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 4, pp. 849–862, (2013).

[98] A. Rajwade, A. Rangarajan, and A. Banerjee. “Using the higher order singular value decomposition for video denoising.” *In Energy Minimization Methods in Computer Vision and Pattern Recognition*, pp. 344–354, (2011).

[99] X. Zhang, Z. Xu, N. Jia, W. Yang, Q. Feng, W. Chen, and Y. Feng. “Denoising of 3D magnetic resonance images by using higher order singular value decomposition” *Medical Image Analysis.*, vol. 19, no. 1, pp. 75–86, (2015).

[100] S.F. Yeganli, H. Demirel, and R. Yu. “Noise removal from MR images via iterative regularization based on higher-order singular value decomposition.” *Signal Image and Video Processing*, vol. 11, no. 8, pp. 1477–1484, (2017).

[101] E.J. Candes and Y. Plan. “Matrix completion with noise.” *Proceedings of the IEEE*, vol. 98, no. 6, pp. 925–936, (2010).

[102] A. Foi. “Noise estimation and removal in MR imaging: the variance-stabilization approach.” *In Biomedical Imaging: from Nano to Macro, 2011 IEEE International Symposium on*, pp. 1809–1814, (2011).

[103] W. Dong, G. Shi, and X. Li. "Nonlocal image restoration with bilateral variance estimation: a low-rank approach." *IEEE Transactions on Image Processing*, vol. 22, no. 2, pp. 700-711, (2013).

[104] D.L. Collins, A.P. Zijdenbos, V. Kollokian, J.G. Sled, N.J. Kabani, C.J. Holmes, and A.C. Evans. "Design and construction of a realistic digital brain phantom." *IEEE Transactions on Medical Imaging*, vol. 17, pp. 463-468 (1998).

[105] OsiriX. DICOM sample image sets repository. http: //www.osirix-viewer.com, http.//pubimage.hcuge.ch:8080/.

[106] The Open Access Series of Imaging Studies. http://www.oasis-brains.org/.

[107] P.W. Yuen, and M. Richardson. "An introduction to hyperspectral imaging and its application for security, surveillance and target acquisition." *The Imaging Science Journal*, vol. 58, no. 5, pp. 241–253, (2010).

[108] C.A. Bishop, J.G. Liu, P.J. Mason. "Hyperspectral remote sensing for mineral exploration in Pulang, Yunnan Province, China." *International Journal of Remote Sensing*, vol. 32, no. 9, pp. 2409–2426, (2011).

[109] L. Alberotanza, F. Braga, R.M. Cavalli, S. Pignatti, and F. Santini. "Hyperspectral techniques for water quality monitoring: Application to the Sacca di

Goro Italy." *In Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, vol. 1, pp. 1–4, (2010).

[110] Z. Liu, H. Wang, and Q. Li. "Tongue tumor detection in medical hyperspectral images." *Sensors*, vol. 12, no. 1, pp. 162–174, (2012).

[111] J. Kerekes and J. Baum. "Full-spectrum spectral imaging system analytical model." *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 3, pp. 571–580, (2005).

[112] T. Skauli. "Sensor noise informed representation of hyperspectral data, with benefit for image storage and processing." *Optics Express*, vol. 19, no. 14, pp. 13031–13046, (2011).

[113] B. Aiazzi, L. Alprone, A. Barducci, S. Baronti, and I. Pippi. "Information theoretic assessment of sampled hyperspectral images." *IEEE Transactions on Geoscience and Remote Sensing*, vol. 39, no. 7, pp. 1447–1458, (2001).

[114] J.M. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Q. Du, P. Gader, and J. Chanussot. "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 5, no. 2, pp. 354–379, (2012).

[115] L. Ma, M. Crawford, X. Yang, and Y. Guo. "Local-manifold-learning-based graph construction for semisupervised hyperspectral image classification." *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, pp. 2832–2844, (2015).

[116] H. Li, C. Li, C. Zhang, Z. Liu, and C. Liu. "Hyperspectral image classification with spatial filtering and $\ell_{(2,1)}$ norm." *Sensors*, vol. 17, no. 2, pp. 1–19, (2017).

[117] L. Alparone, M. Selva, B. Aiazzi, S. Baronti, F. Butera, and L. Chiarantini. "Signal-dependent noise modelling and estimation of new-generation imaging spectrometers." *Hyperspectral Image and Signal Processing: Evolution in Remote Sensing*, pp. 1–4, (2009).

[118] J. Ma, J. Zhao, J. Tian, A. Yuille, and Z. Tu. "Robust point matching via vector field consensus." *IEEE Transaction Image Processing*, vol. 23, pp. 4399–4412, (2014).

[119] A.A. Green, M. Berman, P. Switzer, and M.D. Craig. "A transformation for ordering multispectral data in terms of image quality with implications for noise removal." *IEEE Transactions on Geoscience and Remote Sensing*, vol. 26, no. 1, pp. 65–74, (1988).

[120] H. Othman and S.E. Qian. "Noise reduction of hyperspectral imagery using hybrid spatialspectral derivative-domain wavelet shrinkage." *IEEE Transactions*

on Geoscience and Remote Sensing, vol. 44, no. 2, pp. 397–408, (2006).

[121] S.L. Chen, X.Y. Hu, and S.L. Peng. "Hyperspectral imagery denoising using a spatial spectral domain mixing prior." *Journal of Computer Science and Technology*, vol. 27, no. 4, pp. 851-861, (2012).

[122] B. Rasti, J.R. Sveinsson, and M.O. Ulfarsson. "Wavelet-based sparse reduced-rank regression for hyperspectral image restoration." *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 10, pp. 6688–6698, (2014).

[123] Y.Q. Zhao and J. Yang. "Hyperspectral image denoising vis sparse representation and low-rank constraint." *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 1, pp. 296–308, (2015).

[124] T. Lin and S. Bourennane. "Survey of hyperspectral image denoising methods based on tensor decomposition." *EURASIP Journal on Advances in Signal Processing*, vol. 2013, no. 1, pp. 186–196, (2013).

[125] N. Renard, S. Bourennane, and J.B. Talon. "Denoising and dimensionality reduction using multilinear tools for hyperspectral images." *IEEE Geoscience and Remote Sensing Letters*, vol. 5, no. 2, pp. 138–142, (2008).

[126] D. Letexier and S. Bourennane. “Noise removal from hyperspectral images by multidimensional filtering.” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46, no. 7, pp. 2061–2069, (2008).

[127] X. Liu, S. Bourennane, and C. Fossati. “Denoising of hyperspectral images using the PARAFAC model and statistical performance analysis.” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 10, pp. 3717–3724, (2012).

[128] X. Guo, X. Huang, L. Zhang, and L. Zhang. “Hyperspectral image noise reduction based on rank-1 tensor decomposition.” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 83, pp. 50–63, (2013).

[129] C. Li, Y. Ma, J. Huang, X. Mei, and J. Ma. “Hyperspectral image denoising using the robust low-rank tensor recovery.” *JOSA A*, vol. 32, no. 9, pp. 1604–1612, (2015).

[130] Y. Peng, D. Meng, Z. Xu, C. Gao, Y. Yang, and B. Zhang. “Decomposable nonlocal tensor dictionary learning for multispectral image denoising.” *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2949–2956, (2014).

[131] Y. Qian, Y. Shen, M. Ye, and Q. Wang. "3-D nonlocal means filter with noise estimation for hyperspectral imagery denoising." *IEEE International Geoscience Remote Sensing Symposium (IGARSS'12)*, pp. 1345–1348, (2012).

[132] https://engineering.purdue.edu/ biehl/multispec/hyperspectral.html.

[133] http://peterwonka.net/Publications/code/LRTC Package Ji.zip.

[134] S.G. Chang, B. Yu, and M. Vetterli. "Adaptive wavelet thresholding for image denoising and compression." *IEEE Transactions on Image Processing*, vol. 9, no. 9, pp. 1532–1546, (2000).

[135] C. Chang, and C. Lin. "Libsvm: A library for support vector machines." (2001), http://www.csie.ntu.edu.tw/ cjlin/libsvm.

[136] M. Fauvel, Y. Tarabalka, J. A. Benediktsson, J. Chanussot, and J. Tilton. "Advances in spectral-spatial classification of hyperspectral images." *Proceedings of the IEEE*, vol. 101, no. 3, pp. 652–675, (2013).

[137] M. Pedergnana, P.R. Marpu, M.D. Mura, J.A. Benediktsson, and L. Bruzzone. "A novel technique for optimal feature selection in attribute profiles based on genetic algorithms." *IEEE Transactions Geoscience and Remote Sensing*, vol. 51, no. 6, pp. 3514–3528, (2013).

[138] P.R. Marpu, M. Pedergnana, M.D. Mura, J.A. Benediktsson, and Lorenzo Bruzzone. “Automatic generation of standard deviation attribute profiles for spectral-spatial classification of remote sensing data.” *IEEE Geoscience and Remote Sensing Letters*, vol. 10, no. 2, pp. 293–297, (2013).

[139] J. Ballani, L. Grasedyck, and M. Kluge. “Black box approximation of tensors in hierarchical Tucker format.” *Linear Algebra and its Application*, vol. 438, no. 2, pp. 639–657, (2013).

[140] Z. Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli. “Image quality assessment: from error visibility to structural similarity.” *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, (2004).

APPENDIX

Appendix A: Peak Signal-to-Noise Ratio (PSNR)

The peak signal-to-noise ratio (PSNR) between a vectorized tensor $\boldsymbol{T}$ and its estimated $\boldsymbol{X}$ is calculated in dB as

$$\text{PSNR} = 10(\log_{10} 255^2/\text{MSE}) \tag{1}$$

where MSE is the mean squared error which is calculated as

$$\text{MSE}(\boldsymbol{X}, \boldsymbol{T}) = \frac{1}{N}\|\boldsymbol{X} - \boldsymbol{T}\|^2 \tag{2}$$

with N being length of the vectorized tensor.

Appendix B: Relative Squared Error (RSE)

The relative squared error (RSE) between a vectorized tensor $\boldsymbol{T}$ and its estimated $\boldsymbol{X}$ is calculated as

$$\text{RSE} = \frac{\|\boldsymbol{X} - \boldsymbol{T}\|_F}{\|\boldsymbol{T}\|_F} \tag{3}$$

Appendix C: Structural Similarity Index (SSIM)

The structural similarity index (SSIM) between a tensor $\boldsymbol{T}$ and its estimated $\boldsymbol{X}$, is calculated as:

$$\text{SSIM}(\boldsymbol{T}, \boldsymbol{X}) = \frac{(2\mu_{\boldsymbol{T}}\mu_{\boldsymbol{X}} + c_1)(2\sigma_{\boldsymbol{TX}} + c_2)}{(\mu_{\boldsymbol{T}}^2 + \mu_{\boldsymbol{X}}^2 + c_1)(\sigma_{\boldsymbol{T}}^2 + \sigma_{\boldsymbol{X}}^2 + c_2)}, \tag{4}$$

where $\mu_{\boldsymbol{T}}$, $\mu_{\boldsymbol{X}}$, $\sigma_{\boldsymbol{T}}$, $\sigma_{\boldsymbol{X}}$ denotes the mean value and the standard deviation. $\sigma_{\boldsymbol{TX}}$ defines the cross covariance between two data. $c_1 = 0.01$ and $c_2 = 0.03$ are two constants [140]. The SSIM is between 0 and 1, where 1 is when the two data are the same.

Appendix D: Mean Absolute Difference (MAD)

The mean absolute difference (MAD) between a vectorized tensor $\boldsymbol{T}$ and its estimated $\boldsymbol{X}$ is calculated as [92]

$$\text{MAD} = \frac{1}{N}\|\boldsymbol{X} - \boldsymbol{T}\|_1 \tag{5}$$

where N is length of the vectorized data.

Appendix E: Gaussian Noise Distribution

Gaussian distribution of a noisy random variable y with μ-mean and standard deviation σ is defined as the following equation

$$P(y) = \frac{1}{(\sigma\sqrt{2\pi})} e^{\frac{-(y-\mu)^2}{2\sigma^2}} \tag{6}$$

Appendix F: Rician Noise Distribution

Let denote an image intensity in the absence of noise by t and the noisy measured one by y. The Rician distribution of a noisy intensity y is defined as the following equation

$$P(y) = \frac{y}{\sigma^2} e^{\frac{y^2+t^2}{2\sigma^2}} I_0(\frac{yt}{\sigma^2}) \tag{7}$$

where I_0 denotes the modified Bessel function of zeroth order, and σ denotes the standard deviation of the Gaussian noise in the real and the imaginary images (which is assumed to be equal).

Appendix G: Numerical Results of PSNR, SSIM and MAD Comparisons of Different Methods on Synthetic Data under Gaussian noise of various levels [Figures 4.3-5].

Table 1: Results of T1w.

σ (%)	Measure	PES-TV [94]	BM4D [96]	HOSVDR [99]	IHSVD	IHSVD-W
1	PSNR	42.28	44.09	45.21	46.23	48.02
	SSIM	0.987	0.992	0.994	0.993	0.995
	MAD	1.48	1.21	0.95	1.04	0.74
3	PSNR	35.81	38.39	39.01	40.07	40.78
	SSIM	0.955	0.976	0.979	0.977	0.980
	MAD	3.10	2.28	1.83	2.09	1.68
5	PSNR	32.99	35.95	36.48	36.44	37.47
	SSIM	0.921	0.961	0.966	0.957	0.964
	MAD	4.30	2.99	2.87	2.79	2.28
7	PSNR	31.21	34.38	34.81	34.54	35.74
	SSIM	0.888	0.947	0.952	0.934	0.947
	MAD	5.30	3.58	3.42	3.39	2.28
9	PSNR	29.87	33.21	33.54	33.13	34.30
	SSIM	0.855	0.933	0.938	0.917	0.933
	MAD	6.20	4.10	4.36	3.95	3.24

Table 2: Results of T2w.

σ (%)	Measure	PES-TV [94]	BM4D [96]	HOSVDR [99]	IHSVD	IHSVD-W
1	PSNR	40.53	42.44	43.48	44.74	45.78
	SSIM	0.987	0.993	0.995	0.993	0.995
	MAD	1.79	1.46	1.14	1.27	0.96
3	PSNR	33.56	35.68	36.51	37.05	38.08
	SSIM	0.957	0.977	0.982	0.969	0.976
	MAD	3.94	3.07	2.63	2.73	2.31
5	PSNR	30.52	32.94	33.68	33.59	34.25
	SSIM	0.929	0.962	0.969	0.942	0.957
	MAD	5.54	4.14	3.83	3.74	3.28
7	PSNR	28.50	31.23	31.89	31.39	32.72
	SSIM	0.900	0.949	0.956	0.920	0.932
	MAD	6.94	4.99	4.76	4.57	4.22
9	PSNR	27.01	29.98	30.26	30.05	31.13
	SSIM	0.872	0.936	0.942	0.893	0.921
	MAD	8.21	5.72	5.64	5.48	5.43

Table 3: Results of KNIX.

σ (%)	Measure	PES-TV [94]	BM4D [96]	HOSVDR [99]	IHSVD	IHSVD-W
1	PSNR	40.15	42.09	42.15	43.81	44.17
	SSIM	0.972	0.987	0.986	0.992	0.992
	MAD	1.91	1.58	1.31	1.56	1.22
3	PSNR	35.64	35.87	36.10	37.34	38.87
	SSIM	0.948	0.953	0.954	0.962	0.973
	MAD	3.19	3.13	2.76	3.03	2.24
5	PSNR	33.46	33.51	33.79	33.62	34.89
	SSIM	0.929	0.930	0.933	0.936	0.943
	MAD	4.06	4.04	3.88	3.90	3.56
7	PSNR	31.97	32.02	32.28	32.19	33.59
	SSIM	0.910	0.911	0.914	0.918	0.933
	MAD	4.80	4.74	4.58	4.60	3.94
9	PSNR	30.79	30.92	31.12	31.06	32.37
	SSIM	0.890	0.894	0.898	0.906	0.917
	MAD	5.50	5.33	5.15	5.20	4.51

Table 4: Results of PNEUMATIX.

σ (%)	Measure	PES-TV [94]	BM4D [96]	HOSVDR [99]	IHSVD	IHSVD-W
1	PSNR	40.54	44.39	44.54	45.29	47.60
	SSIM	0.954	0.982	0.982	0.985	0.991
	MAD	1.86	1.22	1.09	1.20	0.84
3	PSNR	36.15	39.91	40.16	41.73	42.45
	SSIM	0.906	0.955	0.956	0.968	0.972
	MAD	3.00	2.03	1.64	1.98	1.52
5	PSNR	33.92	37.95	38.11	38.20	38.95
	SSIM	0.866	0.937	0.937	0.926	0.938
	MAD	3.81	2.50	2.43	2.47	2.29
7	PSNR	32.35	36.59	36.65	36.62	37.54
	SSIM	0.825	0.921	0.920	0.901	0.919
	MAD	4.55	2.89	2.86	2.88	2.67
9	PSNR	31.12	35.54	35.88	35.75	36.68
	SSIM	0.784	0.906	0.910	0.890	0.909
	MAD	5.35	3.24	3.17	3.12	2.92

Appendix H: Numerical Results of PSNR, SSIM and MAD Comparisons of Different Methods on Synthetic Data under Rician noise of various levels [Figures 4.6-8].

Table 5: Results of T1w.

σ (%)	Measure	PES-TV [94]	BM4D [96]	HOSVDR [99]	IHSVD	IHSVD-W
1	PSNR	41.67	44.22	45.21	46.54	47.87
	SSIM	0.986	0.992	0.994	0.993	0.996
	MAD	1.56	1.19	0.96	1.04	0.82
3	PSNR	35.30	38.32	38.97	39.79	40.59
	SSIM	0.950	0.975	0.979	0.976	0.981
	MAD	3.25	2.30	2.06	2.10	1.87
5	PSNR	32.52	35.72	36.38	36.04	37.07
	SSIM	0.914	0.959	0.965	0.949	0.964
	MAD	4.49	3.10	2.88	2.83	2.76
7	PSNR	30.71	33.89	34.61	34.21	35.37
	SSIM	0.879	0.942	0.949	0.927	0.949
	MAD	5.56	3.84	3.60	3.48	3.35
9	PSNR	29.38	32.36	33.27	33.08	33.92
	SSIM	0.845	0.924	0.933	0.917	0.934
	MAD	6.51	4.63	4.46	4.09	3.96

Table 6: Results of T2w.

σ (%)	Measure	PES-TV [94]	BM4D [96]	HOSVDR [99]	IHSVD	IHSVD-W
1	PSNR	39.50	42.63	43.48	44.16	46.13
	SSIM	0.984	0.993	0.995	0.991	0.995
	MAD	1.97	1.42	1.26	1.27	0.98
3	PSNR	32.22	35.68	36.48	37.05	38.97
	SSIM	0.949	0.977	0.982	0.963	0.975
	MAD	4.47	3.06	2.56	2.74	2.27
5	PSNR	29.17	32.75	33.61	33.42	34.09
	SSIM	0.917	0.961	0.968	0.933	0.957
	MAD	6.29	4.24	3.82	3.77	3.48
7	PSNR	27.01	30.87	31.78	31.07	32.04
	SSIM	0.882	0.944	0.954	0.912	0.938
	MAD	8.02	5.26	4.85	4.65	4.54
9	PSNR	25.30	29.45	30.06	29.45	30.77
	SSIM	0.845	0.928	0.938	0.903	0.921
	MAD	9.76	6.20	6.27	5.64	5.53

Table 7: Results of KNIX.

σ (%)	Measure	PES-TV [94]	BM4D [96]	HOSVDR [99]	IHSVD	IHSVD-W
1	PSNR	39.29	41.42	42.14	43.91	44.35
	SSIM	0.966	0.983	0.986	0.991	0.992
	MAD	2.09	1.69	1.29	1.56	1.20
3	PSNR	35.20	35.71	36.06	37.02	38.45
	SSIM	0.943	0.951	0.954	0.964	0.972
	MAD	3.32	3.18	2.80	3.04	2.34
5	PSNR	33.06	33.26	33.70	33.59	34.17
	SSIM	0.924	0.928	0.932	0.937	0.943
	MAD	4.18	4.15	4.04	3.95	3.77
7	PSNR	31.62	31.59	32.11	32.06	32.45
	SSIM	0.905	0.908	0.912	0.913	0.924
	MAD	4.90	5.00	4.97	4.70	4.55
9	PSNR	30.47	30.26	30.85	30.62	31.15
	SSIM	0.885	0.889	0.893	0.889	0.907
	MAD	5.60	5.82	5.65	5.38	5.30

Table 8: Results of PNEUMATIX.

σ (%)	Measure	PES-TV [94]	BM4D [96]	HOSVDR [99]	IHSVD	IHSVD-W
1	PSNR	39.60	44.21	44.52	45.81	47.97
	SSIM	0.945	0.981	0.982	0.986	0.993
	MAD	2.06	1.24	1.04	1.20	0.72
3	PSNR	35.46	39.54	39.98	41.56	42.46
	SSIM	0.894	0.953	0.954	0.965	0.971
	MAD	3.20	2.12	1.69	2.02	1.52
5	PSNR	33.18	36.58	37.50	36.93	38.22
	SSIM	0.847	0.922	0.929	0.913	0.932
	MAD	4.08	2.94	2.80	2.65	2.35
7	PSNR	31.42	33.49	35.43	35.21	36.00
	SSIM	0.792	0.859	0.897	0.869	0.901
	MAD	5.04	4.16	3.47	3.33	3.07
9	PSNR	29.97	30.93	34.38	33.81	34.99
	SSIM	0.734	0.776	0.879	0.830	0.887
	MAD	6.02	5.66	4.08	3.71	3.58