ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL

EXPLOITING OPTIMAL SUPPORTS IN ENHANCED MULTIVARIANCE PRODUCTS REPRESENTATION FOR LOSSY COMPRESSION OF HYPERSPECTRAL IMAGES

M.Sc. THESIS

Muhammed Enis ŞEN

Department of Computational Science and Engineering

Computational Science and Engineering Programme

JANUARY 2024

ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL

EXPLOITING OPTIMAL SUPPORTS IN ENHANCED MULTIVARIANCE PRODUCTS REPRESENTATION FOR LOSSY COMPRESSION OF HYPERSPECTRAL IMAGES

M.Sc. THESIS

Muhammed Enis ŞEN
(702211008)

Department of Computational Science and Engineering

Computational Science and Engineering Programme

Thesis Advisor: Assist. Prof. Dr. Süha TUNA

JANUARY 2024

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

HİPERSPEKTRAL GÖRÜNTÜLERİN ÇOKDEĞİŞKENLİLİĞİ YÜKSELTİLMİŞ ÇARPIMLAR GÖSTERİLİMİ DESTEK VEKTÖRLERİNİN OPTİMİZE EDİLEREK KAYIPLI SIKIŞTIRILMASI

YÜKSEK LİSANS TEZİ

Muhammed Enis ŞEN
(702211008)

Hesaplamalı Bilim ve Mühendislik Anabilim Dalı

Hesaplamalı Bilim ve Mühendislik Programı

Tez Danışmanı: Dr. Öğr. Üyesi Süha TUNA

OCAK 2024

Muhammed Enis ŞEN, an M.Sc. student of ITU Graduate School student ID 702211008 successfully defended the thesis entitled "EXPLOITING OPTIMAL SUPPORTS IN ENHANCED MULTIVARIANCE PRODUCTS REPRESENTATION FOR LOSSY COMPRESSION OF HYPERSPECTRAL IMAGES", which he prepared after fulfilling the requirements specified in the associated legislations, before the jury whose signatures are below.

Thesis Advisor :	**Assist. Prof. Dr. Süha TUNA** Istanbul Technical University	.....
Jury Members :	**Assoc. Prof. Dr. Burcu TUNGA** Istanbul Technical University	.....
	Assist. Prof. Dr. Ercan GÜRVİT Marmara University	.....

Date of Submission : 5 January 2024
Date of Defense : 31 January 2024

To my family,

FOREWORD

Completing this master's thesis has been an enlightening and inspiring journey. For his assistance and guidance through this experience and also for introducing a unique perspective to every student he came across, I owe my advisor Assist. Prof. Dr. Süha TUNA a debt of gratitude. I am truly fortunate and grateful to have had the chance to study under your guidance and mentorship. I can't thank my family enough for their continuous support in every step of this journey, encouraging me to fulfill my potential. Your presence has been a source of motivation and strength that lifted me up to this study and beyond. To all of my friends, I express my deepest thanks. Every new perspective that you introduced has added valuable layers to our lives and also forged lasting connections. I feel fortunate to have crossed paths with such amazing people.

Computing resources used in this work were provided by the National Center for High Performance Computing of Türkiye (UHeM) under grant number 1016472023.

January 2024 Muhammed Enis ŞEN

TABLE OF CONTENTS

ABBREVIATIONS

HSI	:	Hyperspectral Imaging
HDMR	:	High Dimensional Model Representation
EMPR	:	Enhanced Multivariance Products Representation
ADS	:	Averaged Directional Support
ADMM	:	Alternating Direction Method of Multipliers
CANDECOMP	:	Canonical Decomposition
PARAFAC	:	Parallel Factorization
CP	:	CANDECOMP/PARAFAC
ALS	:	Alternating Least Squares
TD	:	Tucker Decomposition
HSI	:	Hyperspectral Imaging
GD	:	Gradient Descent
PCA	:	Principal Component Analysis
ICA	:	Independent Component Analysis
MNF	:	Minimum Noise Fraction
NMF	:	Non-negative Matrix Factorization
OMP	:	Orthogonal Matching Pursuit
FISTA	:	Fast Iterative Shrinkage-Thresholding Algorithm
AVIRIS	:	Airborne Visible / Infrared Imaging Spectrometer
ROSIS	:	Reflective Optics System Imaging Spectrometer
MSE	:	Mean Squared Error
SNR	:	Signal-to-Noise Ratio
PSNR	:	Peak Signal-to-Noise Ratio
SSIM	:	Structural Similarity Index
CosSim	:	Cosine Similarity

SYMBOLS

$\mathbf{t}$	:	Vector of size n : $(t_1, t_2, \ldots, t_n)$
$\mathbf{t}^k$	:	k'th iteration of vector $\mathbf{t}$
$[\mathbf{t}]_j$	:	j'th element of vector $\mathbf{t}$
$\mathcal{H}$	:	3-dimensional data tensor
$\widehat{\mathcal{H}}$	:	3-dimensional approximation tensor
$\mathbf{H}_{(i)}$	:	i'th unfolding of tensor $\mathcal{H}$
h_0	:	0-dimensional (scalar) EMPR component
$\mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3$	:	1-dimensional (vector) EMPR components
$\mathbf{h}_{1,2}, \mathbf{h}_{1,3}, \mathbf{h}_{2,3}$	:	2-dimensional (matrix) EMPR components
$\mathbf{h}_{1,2,3}$	:	3-dimensional (tensor) EMPR component
$\mathbf{F}$	:	2-dimensional data tensor
$\widehat{\mathbf{F}}$	:	2-dimensional approximation tensor
$\mathbf{s}_i$	:	EMPR's support vectors
$\mathbf{z}_i$	:	Dual variables used in ADMM
$\mathbf{y}_i$	:	Lagrange multipliers
$\mathbf{B}_i$	:	i'th factor matrix of TD
λ	:	Regularization parameter / Weight parameter
ρ	:	Penalty parameter
α	:	Step size of the Gradient Descent Algorithm
$\circ$	:	Tensor outer product
$\overline{\times}_n$	:	n-mode tensor inner product
$\mathcal{L}$	:	Augmented Lagrangian function
$\nabla_{\mathbf{s}}$	:	Differentiation operation with respect to $\mathbf{s}$
$\mathcal{S}$	:	Soft Thresholding operator
f	:	Convex loss function
g	:	Convex regularization function

LIST OF TABLES

Page

LIST OF FIGURES

Page

EXPLOITING OPTIMAL SUPPORTS IN ENHANCED MULTIVARIANCE PRODUCTS REPRESENTATION FOR LOSSY COMPRESSION OF HYPERSPECTRAL IMAGES

SUMMARY

Data serves as an irreplacable foundation of modern society as it is the core element of numerous fields such as technological innovations, scientific advancements, and economic decisions. It enables insights into domains of knowledge and experience, assistance with decision-making tasks, and predictions for future outcomes. It has progressed since the very beginning of time from being knowledge and information kept in physical formats like carvings on cave walls and events conserved as inscriptions, to evolving into such mathematical structures that can be obtained by any interaction made through current technological devices like interactions on social media and observations acquired through the use of advanced tools owing to technological advancements. Data transforming into more detailed and complex structures poses efficiency challenges that entails computational methods which can handle the processing and handling of such structures.

For the data handling needs, many methods have been proposed and have been in use thus far. Each has its advantages as well as some drawbacks in the form of problems in either certain limitations or computational complexity issues. Some alternative workarounds have been suggested for these kinds of issues such as embracing an iterative approach rather than employing direct solutions and techniques have been customized to fit specific workflows. Moreover, some innovative approaches operate on representations of data that have undergone compression and transformation, rendering them into more easily processable structures. An important aspect of these practices is the preservation of the data's characteristic features. Compression methods execute this procedure in unique ways like exploiting the eigenvalues and eigenvectors or utilizing singular values. These techniques not only streamline the processing of data but also contribute to the efficiency and accuracy of analyses by retaining characteristic features throughout the compression process. In the field of data processing, an understanding of these diverse methodologies proves convenience in selecting the most effective solutions for the application under consideration.

Hyperspectral imaging is an area that requires such computational techniques to process the collected data due to its high dimensional workflow. It outputs 3-dimensional mathematical structures where the first two dimensions correspond to the spatial attributes of the captured area while the third dimension captures the spectral information with respect to the obtaining device's capacity of retrieving bands. As a result, the fibers in the data's third dimension relate to spectral signatures that empower the identification of objects and materials. The

ability to analyze these spectral data opens doors to multiple useful applications in numerous areas like remote sensing, agriculture, medical imaging, archaeology, and urban planning.

Recent studies in computational sciences for high-dimensional structures have adopted new methods that improve the overall processing performance and make more in-depth analyses possible. Considering the relational design in its third dimension, the High Dimensional Model Representation (HDMR) is a technique that hyperspectral imaging can benefit deeply thanks to its decorrelation properties. The aim of HDMR is to represent multivariate functions in terms of lower dimensional ones. But thanks to the way it was defined, this technique is also applicable on tensors, hence, it can be used to decompose a given tensor in terms of less dimensional entities where each element refers to the attitude of a certain combination of dimensions. This ability of HDMR addresses the decorrelation of each dimension of the given data. The decorrelation procedure enables reducing the noise and removing artifacts while preserving the high-frequency components. Hence, it can be said that HDMR is a suitable compression technique for high-dimensional data with strong relations on individual axes such as hyperspectral images. HDMR employs a set of weights and support vectors to represent data, consequently, necessitating calculation steps. These entities are either assigned certain values or arranged using techniques like Averaged Directional Supports (ADS) but the process of calculating the optimal entities can also be optimized by employing iterative methods such as the Alternating Direction Method of Multipliers (ADMM) where the entailments of HDMR could be used as constraints of ADMM. A sub-method of HDMR which is called the Enhanced Multivariance Products Representation (EMPR) specializes in optimizing the representation by focusing on the support vectors. The weights are assumed to be constant values or scalars and the support vectors are managed by the previously mentioned calculation techniques. As these methods employ the main data for the calculation of the support vectors, they introduce a more robust method EMPR compared to HDMR. Iterative approaches like ADMM can assist in properties of these support vectors such as enforcing sparsity for better representions and improving denoising capabilities.

This thesis work explores the hyperspectral imaging area and proposes a new perspective on the decomposition methods by bringing a tensor-based iterative approach to EMPR through the use of ADMM. The study compares the proposed method's performance and efficiency with some other well-known tensor decomposition techniques, namely CANDECOMP/PARAFAC Alternating Least Squares (CP-ALS) and Tucker Decomposition (TD), while also comparing the results to EMPR's regular application by ADS. Multiple tests are performed on hyperspectral datasets which are 3-dimensional and as a result, the proposed technique is arranged to be applicable on any 3-dimensional tensor especially data that can benefit the decorrelation properties of EMPR. As a result of EMPR, the relations in each dimension and the combinations of these dimensions are acquired through the support vectors.

Results from multiple metrics prove that the proposed method performs similarly to the mentioned tensor decomposition methods for specified ranks and the decorrelated dimensions are successfully represented by the 1-dimensional EMPR components. Tests also employ the 2-dimensional components to reveal the effect on final representations with comparisons to CP-ALS and TD aiming for multiple rank options. The key point of this proposed technique lies in EMPR's superior decorrelation ability. Not only does it demonstrate the capability of reconstructing high-dimensional data with similar accuracy but it also highlights its potential to reduce noise and artefacts in the process. These results are particularly promising for any lossy compression task including Cartesian geometry utilizing tensor decomposition techniques where accurate and efficient data processing is paramount. Furthermore, this performance advantage paves the way for advancements in lossy compression techniques, enabling researchers and practitioners to gain more precise insights from data.

HİPERSPEKTRAL GÖRÜNTÜLERİN ÇOKDEĞİŞKENLİLİĞİ YÜKSELTİLMİŞ ÇARPIMLAR GÖSTERİLİMİ DESTEK VEKTÖRLERİNİN OPTİMİZE EDİLEREK KAYIPLI SIKIŞTIRILMASI

ÖZET

Veri, teknolojik inovasyonlar, bilimsel gelişmeler ve ekonomik kararlar gibi pek çok alanın temel unsuru olması nedeniyle modern toplumda yeri doldurulamaz bir temeli oluşturmaktadır. Bilgiye ve deneyime ilişkin içgörüler sağlarken, bir kararın verilmesi gerektiği durumlarda geçmiş olaylara bağlı olarak öneriler sunabilir ve gelecekteki sonuçlara yönelik tahminler elde edilmesinde yardımcı olur. Zamanın başlangıcından bu yana, mağara duvarlarındaki oymalar ve yazıtlarda kaydedilen geçmiş olaylar gibi fiziksel formatlarda saklanırken, günümüz teknolojik cihazlarıyla yapılan herhangi bir etkileşimle elde edilebilecek algoritmik yapılara evrilene kadar ilerlemiştir ve değişmeye de devam etmektedir. Verilerin daha detaylı ve karmaşık yapılara dönüşmesi verimlilik noktasında sorunlar ortaya çıkarmaktadır ve bu tür yapıların işlenmesini ve sonraki aşamalarda kullanılabilecek hale getirilebilecek hesaplamalı yöntemlerin önemini ortaya koymaktadır.

Verinin verimli bir şekilde işlenebilmesi için bugüne kadar pek çok yöntem önerilmiş ve halihazırda birçok başarılı uygulaması kullanılmaktadır. Bu yöntemlerin sağladığı avantajlarının yanı sıra, belirli kısıtlamalar ya da hesaplama karmaşıklığı (ing: computational complexity) gibi konularda bazı dezavantajları da vardır. Bu tür sorunlar için doğrudan çözümler kullanmak yerine yinelemeli (ing: iterative) bir yaklaşımı benimsemek gibi bazı alternatif çözümler önerilmiş ve teknikler belirli iş akışlarına uyacak şekilde özelleştirilmiştir. Ayrıca bazı yenilikçi yaklaşımlar, sıkıştırılmış ve dönüştürülmüş verilerin temsilleri üzerinde çalışarak onları daha kolay işlenebilir yapılara dönüştürmektedir. Bu yöntemlerin dikkat edilmesi gereken önemli bir yanı da verinin karakteristik özelliklerinin korunmasıdır. Sıkıştırma yöntemleri, veriyi sıkıştırırken bahsedilen karakteristik özelliklerini korumak için özdeğerlerden ve özvektörlerden yararlanmak veya tekil değerleri kullanmak gibi yollarla başvurmaktadır. Bu teknikler yalnızca verilerin işlenmesini kolaylaştırmakla kalmaz, aynı zamanda sıkıştırma süreci boyunca karakteristik özellikleri koruyarak analizlerin verimliliğine ve ana veriye olan yakınlığın korunmasına katkıda bulunur. Veri işleme alanında, bu çeşitli metodolojilerin anlaşılması, hedeflenen uygulamalar için en etkili çözümlerin seçilmesinde kolaylık sağlar.

Hiperspektral görüntüleme, yüksek boyutlu iş akışı sebebiyle toplanan verilerin işlenmesi için bu tür hesaplama tekniklerinin kullanımını gerektiren bir alandır. İlk iki boyutun yakalanan alanın boyutuna karşılık geldiği, üçüncü boyutun ise alıcı sensörlerin kapasitesine bağlı olarak spektral bantlardaki yakaladığı bilgiyi, 3-boyutlu matematiksel yapılar cinsinden bir araya getirir. Sonuç olarak elde

edilen verinin üçüncü boyutundaki lifler olarak adlandırabileceğimiz numerik diziler, spektral imzalar olarak adlandırılır ve görüntünün incelenen bölgesindeki nesnelerin ve malzemelerin tanımlanmasını sağlar. Bu spektral verileri analiz etme yeteneği, uzaktan algılama, tarım, tıbbi görüntüleme, arkeoloji ve şehir planlama gibi birçok alanda çok sayıda yararlı uygulamaya kapı açmaktadır.

Yüksek boyutlu yapılara yönelik hesaplamalı bilimlerde yapılan son çalışmalar, genel işlem performansını artıran ve daha derinlemesine analizleri mümkün kılan yeni yöntemleri benimsemektedir. Hiperspektral görüntülerin üçüncü boyutunda görülen ilişkisel tasarım göz önüne alındığında Yüksek Boyutlu Biçe Gösterilimi (YBBG), ilintisizleştirim özellikleri sayesinde hiperspektral görüntüleme alanının derinlemesine faydalanabileceği bir tekniktir. YBBG'nin amacı, yüksek boyutlu fonksiyonların daha düşük boyutlu fonksiyonlar cinsinden idafe edilmesidir. Ancak tanımının yapıldığı şartların uyumlu olması sayesinde sadece fonksiyonlar üzerinde değil, tensörler üzerinde de uygulanabilmektedir ve bunun sonucunda bir tensörü, belirli boyutların kombinasyonu hakkındaki karakteristik bilgiye sahip olan daha az boyutlu öğeler cinsinden gösterebilmektir. YBBG'nin bu yeteneği, sahip olunan verilerin her boyutunun ilintisizleştirilmesini mümkün kılmaktadır. İlintisizleştirim prosedürü, yüksek frekanslı bileşenleri korurken verideki gürültünün azaltılmasını ve verinin toplanması sırasında ortaya çıkabilen hataların ortadan kaldırılmasını sağlar. Dolayısıyla YBBG'nin, hiperspektral görüntüler gibi bireysel eksenler üzerinde güçlü ilişkilere sahip yüksek boyutlu veriler için uygun bir sıkıştırma tekniği olduğu söylenebilir. Bir verinin YBBG'si elde edilirken, gösterim için bir dizi ağırlık ve destek vektörü kullanır, dolayısıyla bazı hesaplama adımları gerektirmektedir. Bu bileşenlerin belirlenmesi için çeşitli yöntemler bulunmaktadır. Belirli değerlerin atanmasıyla ya da Yönel Ortalamalı Destekler (YOD) kullanılarak belirlenebilirken, yinelemeli yöntemler ile de bu bileşenler için optimal değerleri hesaplanabilmektedir. ÇYÇG'nin başarılı sonuçlar elde etmesinde ve son yaklaştırım üzerinde destek vektörlerinin etkisi büyüktür. Sıkıştırılacak veri kullanılarak hesaplamaların gerçekleştirilmesi ve destek vektörlerinin bu şekilde hesaplanması sonucunda YBBG ile karşılaştırıldığında daha güçlü ve başarılı sonuçlar elde edilmektedir. Ancak destek vektörlerinin yaklaştırımın başarısında bu kadar etkili olması bu bileşenlere hesaplanırken kullanılan yöntemlerin ve optimal değerlere olan yakınlığın önemini ortaya koymaktadır. Bu noktada yinelemeli bir çözüm yöntemi olan Çarpanların Alternatif Yön Yöntemi'nin (ÇAYY) kullanımının avantajları vardır. YBBG'nin destek vektörlerini hesaplama sürecinde, direkt hesaplamalar gerçekleştirilmesi yerine, hesaplama karmaşıklığını azaltacak şekilde yinelemeli bir yolun izlenmesi ve YBBG'nin gerekliliklerinin ÇAYY'ın kısıtlamaları olarak kullanılarak da optimize edilmesi mümkündür. YBBG'nin alt yöntemlerinden biri olarak adlandırılabilecek olan Çokdegişkenliliği Yükseltilmiş Çarpımlar Gösterilimi (ÇYÇG), gösterimin destek vektörlerine odaklanarak daha başarılı bir final gösterim elde etmeyi hedefler. Bu noktada ağırlıklar sabit ya da skaler kabul edilirken destek vektörleri bahsedilen yöntemler yardımıyla hesaplanmaktadır. ÇAYY gibi yinelemeli yöntemler, ÇYÇG'nin destek vektörlerinin seyrekliği gibi hedeflenen belirli şartların sağlanmasında yardımcı olmaktadır.

Bu tez çalışması, hiperspektral görseller alanını araştırırken ÇAYY yardımıyla ÇYÇG'ye tensör tabanlı yinelemeli bir yaklaştırım getirerek tensör ayrıştırımı yöntemlerine yeni bir bakış açısı kazandırmayı hedeflemektedir. Bu çalışma içerisinde, yaygın olarak kullanılan tensör ayrıştırımı yöntemlerinden CANDECOMP/PARAFAC Alternatif En Küçük Kareler (ing: CANDECOMP/-PARAFAC Alternating Least Squares, CP-ALS) ve Tucker Ayrıştırımı (ing: Tucker Decomposition, TD) ile performansını ve verimliliğini karşılaştırırken aynı zamanda gğncel ÇYÇG hesaplarında kullanılan YOD yöntemi ile önerilen yöntem arasındaki farkın sonuçlarını incelemektadir. 3-boyutlu olan hiperspektral verisetleri üzerinde çeşitli metrikleri de dikkate alarak testler gerçekleştirilmiştir. Aynı zamanda önerilen yöntemin herhangi bir 3-boyutlu tensörde, özellikle de ÇYÇG'nin ilintisizleştirme özelliğinden faydalanabilecek verilerde uygulanabilirliği ortaya konmuştur. ÇYÇG'nin kullanımı sonucunda da boyutlar ve boyutlar arasındaki ilişkileri taşıyan destek vektörleri elde edilmiştir.

Çeşitli metrikler üzerinde yapılan testlerden elde edilen sonuçlar, önerilen yöntemin bahsedilen tensör ayrıştırma yöntemleriyle karşılaştırıldığında karşılaştırılan yöntemlerin belirtilen bazı ranklarda benzer performans sergilediğini ve boyutların ilintisizleştirilerek kendilerine özgü bilgilerinin de ÇYÇG'nin destek vektörlerinde başarıyla temsil edildiğini kanıtlar niteliktedir. Bu testlere ek olarak, ÇYÇG'nin 2-boyutlu öğelerinin de eklenmesiyle son yaklaştırımdaki etkisi gözlemlenmiştir ve bu sonuçlar diğer yöntemlerin farklı rank hedeflendiğinde elde ettiği sonuçlarla karşılaştırılmıştır. Önerilen bu tekniğin kilit noktası, ÇYÇG'nin üstün ilintisizleştirme yeteneğindedir. Yalnızca yüksek boyutlu verileri daha yüksek doğrulukla yeniden oluşturma yeteneğini göstermekle kalmayıp, aynı zamanda süreçteki gürültüyü ve hataları azaltma potansiyelini de ortaya koymaktadır. Bu sonuçlar, doğru ve verimli veri işlemenin çok önemli olduğu, tensör ayrıştırma yöntemleri yoluyla kayıplı sıkıştırmayı ve kartezyen geometrisini kullanan herhangi bir uygulama alanı için faydalı olmaktadır. Ayrıca bu kayıplı sıkıştırmalar içerisinde sağlayabileceği performans avantajı, bu alandaki ilerlemelerin önünü açarak araştırmacıların ve benzeri uygulamalardan faydalanan kişilerin, verilerden daha kesin içgörüler elde etmelerine olanak tanımaktadır.

1. INTRODUCTION

Regardless of the domain, data has been an essential part of our studies, our knowledge, our technology, our development, and even our daily lives. Any advancement that improves or concludes results, is an outcome of processed data and mechanisms that decide over defined rules related to its applications. Whether it is acquired knowledge or observed results, the size and detail of data is increasing over time, bringing certain necessities for handling it. Owing to advancements in technology, more in-depth observations into studies are obtainable but as a consequence, the size of collected data increases at tremendous speeds. Tasks such as storing, processing, and understanding data are becoming a challenge where nowadays any interaction is considered useful and inspectable. Thus, the need for processes that would support numerous operations through more efficient storage and better representations arises. Even though such advancements are being made in computational science, the aspect of efficiency is a concern point that needs to be taken into consideration.

This work focuses on providing an efficient computational method for the specified needs and explores the obtained results on applications performed on hyperspectral images which contain certain relationships between dimensions and their layers.

Hyperspectral Imaging (HSI) enables data to be captured from numerous adjacent and narrow spectral bands, reflecting the abovementioned need for development of methods in managing this data [2]. In contrast to conventional images that are only comprised of three bands, the broader scope of spectral bands in HSI, provides a more comprehensive spectral resolution. Each pixel that these images contain, stores a spectral signal formed by the intensity values over the recorded spectral bands. These spectral signals are used in tasks such as object identification and revelation of material properties, allowing them to be also

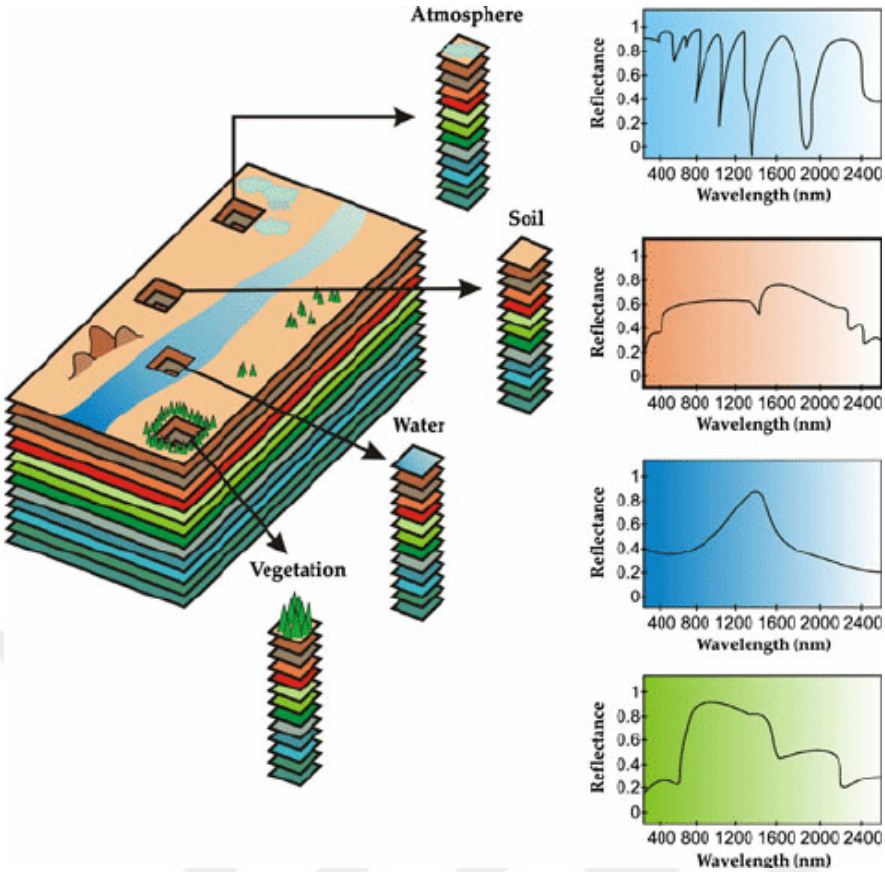

Figure 1.1 : Abstraction of Hyperspectral Images [1].

called spectral signatures. Hence, HSI has a vast range of applications spanning across various fields including remote sensing [3], geology [4], urban planning [5], agriculture [6], archaeology [7], and medical imaging [8].

An abstraction of a hyperspectral image is seen in Fig. 1.1.

1.1 Literature Review

In contemporary times, the structure of data has evolved to become high-dimensional and complex, presenting challenges and opportunities across various application areas. HSI is an area that has emerged where sensors capture a multitude of spectral bands, generating a 3-dimensional set of data points. Diverse and impactful practice areas such as crop health monitoring [9], identification of pollutants [10], and tissue analysis [11] benefit such detailed structures. However, these structures lead to numerous challenges in processing and storing needs [12], especially in restoration [13], compressive sensing [14],

anomaly detection [15], multispectral fusion [16], and spectral unmixing tasks [14]. Each of these tasks requires a unique perspective on how these challenges could be overcome. Various computational methods have been studied and due to their tensor-like structure, tensor decomposition methods have been found applicable to hyperspectral data. Past studies employ customized tensor decomposition techniques with respect to their specified needs for increasing processing, storing, and transmission efficiency.

Before touching on these tasks, a short introduction to tensor algebra would be useful. Tensor algebra is a branch of mathematics that generalizes linear algebra to become more capable, modifying the operations to satisfy the needs and constraints in higher dimensions. It handles the substructure of most operations in modern-day engineering and scientific research tasks, giving lead to improvements for more efficient processes and enabling before impossible procedures. Tensors, which could be referred to as an extension of scalar, vector, and matrix algebra, are proven to be indispensable in describing and manipulating multi-dimensional data especially owing to their ability to represent information in a structured and processable state. This advantage is what empowers most technological advancements and tasks. As stated in the previous paragraph, hyperspectral imaging is an area which highly benefits these tools and the mentioned challenges are now considered as forges for future solution methods and techniques.

We can now return to hyperspectral images and the tasks previously mentioned. Restoration is an essential study area in hyperspectral imaging as the recorded information are often compromised by various factors. Whether it is the variations in environmental factors or issues caused by equipment, hyperspectral images are prone to such errors during the process of acquisition and transformation. External elements tend to have such effects regarding the present state of the atmosphere. Occasionally, undesired artifacts might also occur as an outcome due to problems in sensor and transmitting equipment. These factors inevitably result in visual degradation including noise, blurring, and missing data. As a consequence, these artifacts affect the efficiency in later steps of operations,

necessitating methods that can reduce errors and improve the condition of the data.

Methods aiming to achieve an absolved result generally employ tensor decomposition techniques alongside image denoising, deblurring, inpainting, and destriping degradation operations, modifying these decomposition algorithms to work coherently. Thus, for a degraded hyperspectral image $\mathcal{T}$ defined as

$$\mathcal{T} = M(\mathcal{X}) + S + N \tag{1.1}$$

where $\mathcal{T}$, $\mathcal{X}$, S, and N stand for the observed hyperspectral image, the restored hyperspectral image, the sparse error, and the additive noise, respectively. The function M is defined as the degradation operation that is used for the specific problem. Hence, the restoration problem can be generalized for $\mathcal{T}$ defined as

$$\min_{\mathcal{X}} \frac{1}{2}\|\mathcal{T} - M(\mathcal{X}) - S\|_F^2 + \tau f(\mathcal{X}) + \lambda g(S) \tag{1.2}$$

The parameters τ and λ in (1.2) are regularization parameters and the functions f and g ensure the regularization on the recovered tensors $\mathcal{X}$ and S with respect to their desired properties. The Frobenius norm term aims to minimize the noise, eventually converging to a final clear tensor $\mathcal{X}$.

Another significant field of study focuses on retrieving collected information from sensing equipment over long distances. During the acquisition of hyperspectral data, the sampling rate is required to be larger than double its maximum frequency component to assure the clarity of collected information. This is a well-established principle in the literature of signal processing, which is also known as the Nyquist Theorem [17], underlining the conditions that help avoiding distortions. However, this necessity also increases the need for computing and storage extensively. The growing spectral resolution of HS images contributes to the heightened costs and diminished efficiency of transmitting data from airborne and spaceborne platforms to ground stations. Compressive Sensing aims to overcome these challenges by reducing the cost of signal storage and

transmission through the process of compressively sampling and reconstructing signals. This procedure is generally done by combining a downsampling operator, a permutation matrix, and a Walsh-Hadamard transform [18] and applying this unified operation on a hyperspectral recording $\mathcal{X}$ to obtain compressive measurements y which is later used for the reconstructions on ground stations.

$$y = f(\mathcal{X}) \tag{1.3}$$

where f denotes the unified operation applied on $\mathcal{X}$. Thus, the compressive sensing problem of hyperspectral image $\mathcal{X}$ can be defined as

$$\min_{\mathcal{X}} \|y - f(\mathcal{X})\|_F^2 + \lambda h(\mathcal{X}) \tag{1.4}$$

Similar to (1.2), the additional term $\lambda h(\mathcal{X})$ is added for regularization. Through a suitable minimization process, efficiently transmittable compressive measurements can be attained and the received information can be used in a clear reconstruction of the hyperspectral image.

These study areas utilize a number of tensor decomposition methods in solving their minimization problems with as high efficiency as possible [12]. Considering each task obligates unique constraints, computational methods are customized to fit specified needs. Owing to their flexibility, CANDECOMP/PARAFAC and Tucker Decomposition are among the widely used tensor decomposition techniques.

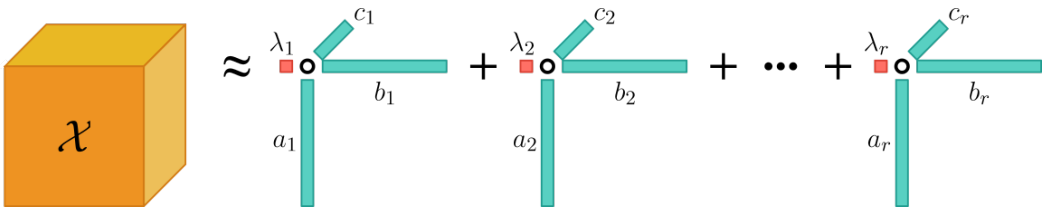

Figure 1.2 : 3-Dimensional Visualization of CANDECOMP/PARAFAC.

Canonical Decomposition/Parallel Factor Analysis (CP), also known as CANDECOMP/PARAFAC, is a factorization technique, widely used in signal and image processing tasks. It is applied to high-dimensional data and used for

representing the data in terms of rank 1 tensors of the same size. A generalization of CP decomposition application on a 3-dimensional tensor $\mathcal{X}$ could be made by

$$\mathcal{X} = \sum_{r=1}^{N} \lambda_r \, a_r \circ b_r \circ c_r \tag{1.5}$$

where N stands for the predefined final rank of approximation, a_r, b_r, and c_r are 1-dimensional components used for the calculation of rank 1 tensors and λ_r are the non-zero weight parameters. It is also important to mention that the $\circ$ operation stands for the tensor outer product. This procedure can also be visualized as in Figure 1.2.

This process of calculating the optimal components a_r, b_r, and c_r and parameters λ_r are achieved via various methods furthermore one of the common ways of computing these entities is through iterative approaches such as Alternating Least Squares (ALS) and Gradient Descent (GD). ALS employs an alternating calculation process where each vector component undergoes a minimization problem while the rest are assumed constants. After the vectorial components complete an iteration, the weight parameters are updated with the help of a specified optimization problem. When a pre-defined tolerance value between contiguous iterations is satisfied, the process terminates and a final approximation is established.

Tucker Decomposition (TD) is a widely used tensor decomposition method, especially for the abovementioned hyperspectral imaging tasks. It employs factor matrices and a core tensor to represent the main data tensor. The level of approximation can be managed by adjusting the targeted rank. Below can be found the equation

$$\mathcal{X} = \mathcal{A} \times_1 \mathbf{B}_1 \times_2 \mathbf{B}_2 \cdots \times_N \mathbf{B}_N \tag{1.6}$$

where $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \cdots \times I_N}$ is an N-dimensional tensor to be decomposed. $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \cdots \times I_N}$ holds as the core tensor and $\mathbf{B}_i$ are the factor matrices. This operation can also be visualized as in Figure 2.1 for a 3-dimensional data tensor.

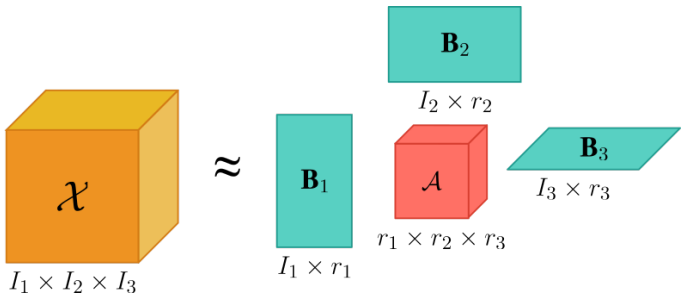

Figure 1.3 : Visualization of Tucker Decomposition for 3-D case.

Similar to CANDECOMP/PARAFAC, the optimal values in factor matrices and the core matrix are generally attained via iterative approaches, especially ALS. The factor matrices are updated for each mode individually with the core matrix being fixed. After each iteration, the core matrix is updated with another ALS problem. This process is repeated until the convergence criteria is met, hence, yielding a final approximation.

The concept of efficient storage and processing of high-dimensional data, heavily relies on computational techniques. Besides CP decomposition and Tucker Decomposition, several other methods such as Principal Component Analysis (PCA) [19]–[21], Independent Component Analysis (ICA) [22], Minimum Noise Fraction (MNF) [23], and Non-negative Matrix Factorization (NMF) [24] are among the commonly used computational techniques that have been developed for efficient storing purposes without sacrificing the characteristics of the data which are important aspects for further tasks.

In this regard, PCA aims to represent high-dimensional data in terms of decorrelated 1-dimensional orthogonal components by employing transformations through the use of eigenvalues and eigenvectors. Through these operations, it projects data points onto a different hyperplane which is also more beneficial for feature extraction tasks. ICA is a technique developed for separating complex mixed signals into statistically independent components, aiming to uncover the structure of the data under certain assumptions. MNF is a transformation tool that is designed to decorrelate bands through prioritization of features. It is exploited in remote sensing and image-processing tasks, enabling analysis of complex multi-band data. Lastly, NMF is a dimensionality reduction and

feature extraction method applied in image and text data analysis tasks. It reveals meaningful patterns by exploiting the data's basis vectors and coefficients, factoring the data matrix into two lower-dimensional matrices while enforcing non-negativity constraints.

Although these methods provide representations with low error rates, their computational complexity is generally high due to their underlying matrix inversion-based mathematical operations. Studies have developed alternative techniques to overcome such issues by utilizing iterative solution processes.

Some other techniques build on algorithms referred to as dictionary learning [25] and sparse coding [26,27]. In dictionary learning-based algorithms, dictionary elements are updated according to the results of present calculations and the desired outcome until a representation at the desired level is achieved or is performed for over a predetermined number of iterations. On the other hand, sparse coding refers to representing data in terms of fewer dictionary elements. This also implies minimizing the amount of non-zero elements in components, in other words, employs calculations of the $\mathbf{L}_0$-norm. Minimization of the $\mathbf{L}_0$-norm is a problem of NP-hard, and thus the process is relaxed by minimizing the problem with respect to $\mathbf{L}_1$-norm instead of $\mathbf{L}_0$-norm.

Methods based on sparse coding are divided into $\mathbf{L}_0$-norm minimization methods (greedy pursuit) [28], $\mathbf{L}_p$-norm regularization-based algorithms [29], and iterative shrinkage algorithms [30]. These subcategories also have situations that could pose problems within themselves, and various solution methods have been developed.

The Orthogonal Matching Pursuit (OMP) [31] method was introduced to prevent the selection of the same dictionary element on different iterations of greedy pursuit algorithms, while Alternating Direction Method of Multipliers (ADMM) [**?**] has been developed to allow an iterative process for minimization problems involving the $\mathbf{L}_1$-norm. Although iterative shrinkage algorithms are useful for dense large matrices, their convergence rates are known to be slow. As a solution

for this problem the Fast Iterative Shrinkage-Thresholding Algorithm (FISTA) [32] has been proposed to increase the convergence rates.

In addition to all these methods, recent studies have been adopting High Dimensional Model Representation (HDMR) [33]–[36] which is a method used for dimensionality reduction and feature extraction tasks. It enables multidimensional datasets to be analyzed and examined by expressing them in terms of lower-dimensional components. Similar to PCA and its variations, these lower-dimensional components are arranged to be orthogonal to each other. This ensures that these methods capture the maximum variance in the data, furthermore, it makes it possible for each to component capture a unique and non-redundant aspect of the information in the data. Having components representing unique information corresponding to behaviours in different dimensions, the need for dimensionality reduction in certain tasks is satisfied and is achieved without losing much relevant information. This is the result of orthogonality which ensures that these directions are not mixed or correlated, simplifying the interpretation of each component's contribution to the overall structure of data. Orthogonality also simplifies the computational side of numerous tasks and improves the numerical stability. This can be important when dealing with data that may have noise or small numerical errors such as sensor-related errors or weather-caused variations in hyperspectral images. It can also be mentioned that orthogonality simplifies the reconstruction steps to be straightforward and involve simple linear combinations. This makes it computationally efficient.

HDMR aims to resolve a tensor in terms of lower or equal dimensional entities which are weight components, support vectors, and HDMR components. These components are restricted to fit certain constraints, hence for the ease of calculations, numerous ways for calculating approximations with high accuracies have been used. The weight vectors are expected to be arranged to be of length 1 and for the support vectors, some of these methods assign them as vectors of ones $\mathbf{1}_n$ where n corresponds to the size of the support vector, coherent with the main

tensor sizes while other methods either employ Averaged Directional Supports (ADS) [37] or iterative techniques.

After finalizing these entities, an HDMR expansion of a 3-dimensional tensor $\mathcal{H} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ could be expressed as,

$$\begin{aligned}\mathcal{H} = & h_0 \left(\mathbf{w}_1 \circ \mathbf{w}_2 \circ \mathbf{w}_3\right) * \left(\mathbf{s}_1 \circ \mathbf{s}_2 \circ \mathbf{s}_3\right) + \mathbf{h}_1 \circ \left(\mathbf{w}_2 \circ \mathbf{w}_3\right) * \left(\mathbf{s}_2 \circ \mathbf{s}_3\right) \\ & + \mathbf{h}_2 \circ \left(\mathbf{w}_1 \circ \mathbf{w}_3\right) * \left(\mathbf{s}_1 \circ \mathbf{s}_3\right) + \left(\mathbf{w}_1 \circ \mathbf{w}_2\right) * \left(\mathbf{s}_1 \circ \mathbf{s}_2\right) \circ \mathbf{h}_3 \\ & + \mathbf{h}_{1,2} \circ \mathbf{w}_3 * \mathbf{s}_3 + \mathbf{h}_{1,3} \circ \mathbf{w}_2 * \mathbf{s}_2 + \mathbf{w}_1 * \mathbf{s}_1 \circ \mathbf{h}_{2,3} + \mathbf{h}_{1,2,3}\end{aligned} \tag{1.7}$$

where the $*$ operation is the Hadamartd product, the $\mathbf{s}_i$ components are the support vectors and the h_0, $\mathbf{h}_i$, $\mathbf{h}_{i,j}$, $\mathbf{h}_{i,j,k}$ components are the HDMR components. The HDMR components are calculated with the help of the weight vectors $\mathbf{w}_1$, $\mathbf{w}_2$, and $\mathbf{w}_3$ which are also required to satisfy the constraint $\|\mathbf{w}_i\|_1 = 1$.

As previously mentioned, these support vectors can also be arranged by the ADS which employs a set of equations for the calculations of these components. An example calculation for the support vector $\mathbf{s}_2$ can ben done via Eqn. (1.8).

$$\left[\mathbf{s}_2\right]_j = \sum_{i=1}^{n_1} \sum_{k=1}^{n_3} \left[\mathbf{w}_1\right]_i \left[\mathbf{w}_3\right]_k \mathbf{H}_{ijk} \tag{1.8}$$

where the subscripts i, j, and k denote the element indices of the specified tensors.

Originally defined as a method for multivariate functions, there are several sub-HDMR methods where each method focuses on certain aspects of the procedure. Some of these methods are named as Plain HDMR, Factorized HDMR [38], Logarithmic HDMR [39], Hybrid HDMR [39]–[41], and Generalized HDMR [42]. The HDMR method that has just been explained is the general definition of Plain HDMR. It outputs well-received results under functions that have an additive nature. Functions that have a multiplicative nature produce better results by using Factorized, Logarithmic, and Hybrid HDMR as they can be expressed as a product of less-dimensional functions. These methods can be used interchangeably in situations where a condition is not met and the other version handles such situations better. For some multivariate functions Factorized HDMR may yield results that don't meet the expected level of representation, for example for higher ranks of approximations the recent additions might not comply

with the lower-dimensional functions, thus, entailed the studies for Logarithmic HDMR which was introduced as an alternative for such problems. On the other hand, problems that don't have either an additive nature or a multiplicative nature, necessitate a different approach and lead up to the introduction of Hybrid HDMR. With this technique, Plain HDMR and a multiplicative nature based method are combined with coefficients, employing both kinds of methods. Lastly, Generalized HDMR is proposed as an alternative for these options as it shows a different approach on the weight components. Instead of conducting the products of 1-dimensional weight components, it employs a non-multiplicative weight function that can be generalized to achieve a more qualitative operation. Studies conducted on this topic have been employing the Fluctuationlessness Approximation Theorem [43] for the purpose of optimizing the weight parameter that manages the contribution level of Plain HDMR and multiplicative nature based HDMR expansions.

As an expansion extended from HDMR, Enhanced Multivariance Products Representation (EMPR) [37,44]–[48] is a recently adopted method for similar tasks.

It aims to represent the main tensor in terms of lower-dimensional components and optimize the approximation by adjusting the support vectors. It is also important to note that the support vectors are calculated by utilizing certain techniques that employ the main data tensor. This condition assists EMPR in becoming a more robust and successful version of HDMR. Hence, these support vectors play a significant role in the success of the method. Therefore, optimizing the support vectors becomes crucial for the success of the representation obtained by the method. There are several ways of calculating or initiating these components, either by assigning pre-determined values or by arranging them to fit the basis of ADS. Based on these facts, both studies and experiences considering tests performed on HDMR and its variants show that the success of the method increases when the support vectors are optimized which also indicates the benefits of using iterative methods.

Similar to HDMR, the EMPR expansion of a 3-dimensional tensor $\mathcal{H} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ can be expressed as,

$$\begin{aligned}\mathcal{H} = h_0\, \mathbf{t}_1 \circ \mathbf{t}_2 \circ \mathbf{t}_3 + \mathbf{h}_1 \circ \mathbf{t}_2 \circ \mathbf{t}_3 + \mathbf{t}_1 \circ \mathbf{h}_2 \circ \mathbf{t}_3 + \mathbf{t}_1 \circ \mathbf{t}_2 \circ \mathbf{h}_3 \\ + \mathbf{h}_{1,2} \circ \mathbf{t}_3 + \mathbf{h}_{1,3} \circ \mathbf{t}_2 + \mathbf{t}_1 \circ \mathbf{h}_{2,3} + \mathbf{h}_{1,2,3}\end{aligned} \tag{1.9}$$

where $\mathbf{t}_i$ are the support vectors, the h_0, $\mathbf{h}_i$, $\mathbf{h}_{i,j}$, $\mathbf{h}_{i,j,k}$ components are the EMPR components.

EMPR will be explained with a more in-depth understanding in the next chapter alongside other useful methods that were used in this thesis study.

1.2 Purpose of the Thesis

This thesis work explores the lossy compression area and compares the widely known tensor decomposition methods with a recently introduced method utilizing the aforementioned techniques. This work also aims to propose a competent method that represents data at a similar or better level compared to widely used tensor decomposition techniques, CANDECOMP/PARAFAC and Tucker Decomposition. It is aimed to utilize optimization methods like the Alternating Direction Method of Multipliers (ADMM) that can assist in an iterative manner with efficient solutions resulting in low error rates instead of performing direct solutions. This is the result of ADMM aiming to optimize solutions while ensuring the robustness of results that fit into predefined constraints.

The literature review section lays emphasis on the importance of optimizing the support vectors. Hence, the method this work proposes involves the use of ADMM in optimizing the lower dimensional support vectors in EMPR to achieve an efficient approximation of hyperspectral tensors. Supporting EMPR's decorrelation and denoising capability with the superior converging properties of ADMM enables determining a well-established representation of the original tensor. Therefore, this new method is capable of elevating the representation quality and decorrelation capacity of EMPR in further tasks such as compression and classification.

2. MATHEMATICAL BACKGROUND

This chapter consists of the main mathematical operations and computational methods that were used in the implementation of the proposed method. The chapter will start with a short introduction into tensor algebra and the equivalents of inner product from linear algebra, namely the tensor inner product. Then the Enhanced Multivariance Products Representation will be explained in detail as it stands as the core method of the proposed method with some modifications which will be seen in Chapter 3. The Alternating Direction Method of Multipliers will assist the optimization process as an iterative numerical approach with flexibilities that serve for the enforcement of certain constraints. As the last method, Gradient Descent is introduced and explained.

2.1 Tensor Operations

Tensors are algebraic objects, often referred to by their elements through defined coordinate systems. These objects can be scalars, vectors, matrices, or high-dimensional matrices. They form the mathematical structure in various operations where data is the source to be processed. Tensors support similar algebraic operations to their linear algebraic counterparts through some adjustments and this work builds upon the use of the tensor n-mode inner Product [49] which is used extensively in tensor decomposition tasks. This operation provides the ability to multiply tensors of higher dimensions under certain notations. The n-mode inner tensor product brings the multiple steps of unfolding, Kronecker product, Khatri-Rao product, Hadamard product, and folding together into a single operation [49].

An n-mode inner tensor product between an n-dimensional tensor $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_r}$ and a vector $\mathbf{b} \in \mathbb{R}^{n_k}$ is denoted as $\mathcal{A} \overline{\times}_k \mathbf{b}$ and defined explicitly

as follows

$$(\mathcal{A} \,\overline{\times}_k\, \mathbf{b})_{i_1 \dots i_{k-1} i_{k+1} \dots i_r} = \sum_{i_k=1}^{n_k} a_{i_1 \dots i_r} b_{i_k} \tag{2.1}$$

where the resulting tensor is $(n-1)$-dimensional and of size $n_1 \times \cdots \times n_{k-1} \times n_{k+1} \times \cdots \times n_r$. This operation enables the multiplication on a single dimension and can be repeated successfully when one wants to execute the same task on other dimensions.

Thus, for a tensor $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_r}$ and a matrix $\mathbf{B} \in \mathbb{R}^{n_k \times n_\ell}$, we denote the multiplication operation $\overline{\times}$ on dimensions k and ℓ as

$$(\mathcal{A} \,\overline{\times}_{k,\ell}\, \mathbf{B})_{i_1 \dots i_{k-1} i_{k+1} \dots i_{\ell-1} i_{\ell+1} \dots i_r} = \sum_{i_k=1}^{n_k} \sum_{i_\ell=1}^{n_\ell} a_{i_1 \dots i_r} b_{i_k i_\ell} \tag{2.2}$$

where the resulting tensor is of size $n_1 \times \cdots \times n_{k-1} \times n_{k+1} \times \cdots \times n_{\ell-1} \times n_{\ell+1} \times \cdots \times n_r$.

2.2 Enhanced Multivariance Products Representation

Enhanced Multivariance Products Representation (EMPR) is a statistics-based method which represents multivariate functions in terms of lower-dimensional ones. As a consequence of its definition, this technique is also applicable to tensors as well as their properties fit the constraints of the representation [37].

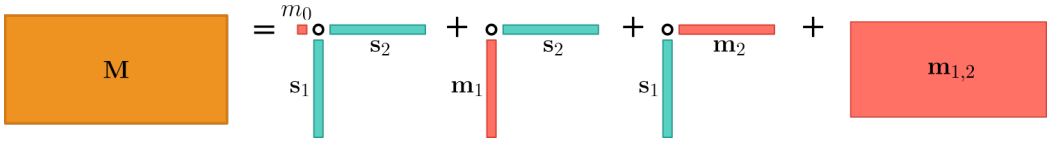

Figure 2.1 : Visualization of EMPR on a matrix $\mathbf{M}$.

By employing EMPR, a 2-dimensional tensor (matrix) $\mathbf{M}$ of size $n_1 \times n_2$ could be rewritten in terms of vector outer products as follows

$$\mathbf{M} = m_0\, \mathbf{s}_1 \mathbf{s}_2^T + \mathbf{m}_1 \mathbf{s}_2^T + \mathbf{s}_1 \mathbf{m}_2^T + \mathbf{m}_{1,2} \tag{2.3}$$

where m_0, $\mathbf{m}_1$, $\mathbf{m}_2$ are the scalar and 1-dimensional EMPR components, respectively. Other entities in the right-hand side of the representation in Eqn. (2.3), namely $\mathbf{s}_1$ and $\mathbf{s}_2$ are the accompanying support vectors. The last term, that is $\mathbf{m}_{1,2}$, is a matrix and stands as the residual term for the original matrix $\mathbf{M}$. These terms can be calculated uniquely and explicitly by engaging support vectors

$\mathbf{s}_1$ and $\mathbf{s}_2$ appropriately. Thus, the scalar EMPR component m_0 is computed as follows

$$m_0 = \frac{1}{n_1 n_2} \mathbf{s}_1^T \mathbf{M}\, \mathbf{s}_2 \tag{2.4}$$

By following the theory [37], the first 1-dimensional EMPR component, $\mathbf{m}_1$ is determined as

$$\mathbf{m}_1 = \frac{1}{n_2} \mathbf{M}\, \mathbf{s}_2 - m_0 \mathbf{s}_1 \tag{2.5}$$

while the second 1-dimensional EMPR component, $\mathbf{m}_2$, is calculated as

$$\mathbf{m}_2 = \frac{1}{n_1} \mathbf{M}^T \mathbf{s}_1 - m_0 \mathbf{s}_2, \tag{2.6}$$

respectively. The abovementioned residual term is obtained by subtracting all calculated EMPR terms from the raw matrix as follows

$$\mathbf{m}_{1,2} = \mathbf{M} - m_0 \mathbf{s}_1 \mathbf{s}_2^T - \mathbf{m}_1 \mathbf{s}_2^T - \mathbf{s}_1 \mathbf{m}_2^T. \tag{2.7}$$

Therefore, all EMPR components are determined uniquely by exploiting an appropriate support vector team, $\{\mathbf{s}_1, \mathbf{s}_2\}$.

As one can easily verify, the right-hand side of Eqn. (2.3) consists of one-rank terms except for the residual. By neglecting the residual term, it is obvious that the original matrix $\mathbf{M}$ can be approximated with a rank two matrix structure. Accordingly, it is evident that EMPR can be assessed as a low-rank approximation technique for matrices having a finite number of expansion terms.

2.3 Alternating Direction Method of Multipliers

Alternating Direction Method of Multipliers (ADMM) is an algorithm used in optimization tasks, an effective solution method for problems consisting of separable functions and linear constraints. It is built on the foundations of Augmented Lagrangians and combined with the method of multipliers, ADMM brings together dual ascent's decomposability and method of multipliers' convergence properties. Problems that ADMM aims to solve are of the form :

$$\begin{aligned} \min \quad & f(\mathbf{x}) + g(\mathbf{z}) \\ \text{s.t.} \quad & A\mathbf{x} + B\mathbf{z} = \mathbf{c} \end{aligned} \tag{2.8}$$

where $f(\mathbf{x})$ and $g(\mathbf{z})$ are convex functions, $\mathbf{x}$ and $\mathbf{z}$ are the tensors being aimed to optimize and A, B, and $\mathbf{c}$ are tensors managed accordingly with respect to the constraints of the problem. The solution is an iterative process, involving an auxiliary variable $\mathbf{y}$ to transform the problem into an Augmented Lagrangian problem and updating $\mathbf{x}$, $\mathbf{z}$, and $\mathbf{y}$. Below, can be seen the steps for solving such problems.

(1) Form the Augmented Lagrangian function.

$$\mathcal{L}_\rho(\mathbf{x}, \mathbf{z}, \mathbf{y}) = f(\mathbf{x}) + g(\mathbf{z}) + \mathbf{y}^T(A\mathbf{x} + B\mathbf{z} - \mathbf{c}) + \frac{\rho}{2}\|A\mathbf{x} + B\mathbf{z} - \mathbf{c}\|_2^2 \tag{2.9}$$

(2) Solve the problem for each variable's latest iteration respectively.

$$\mathbf{x}^{k+1} := \underset{x}{\operatorname{argmin}}(\mathcal{L}_\rho(\mathbf{x}, \mathbf{z}^k, \mathbf{y}^k)) \tag{2.10}$$

$$\mathbf{z}^{k+1} := \underset{z}{\operatorname{argmin}}(\mathcal{L}_\rho(\mathbf{x}^{k+1}, \mathbf{z}, \mathbf{y}^{k+1})) \tag{2.11}$$

$$\mathbf{y}^{k+1} := \mathbf{y}^k + \rho(A\mathbf{x}^{k+1} + B\mathbf{z}^{k+1} - \mathbf{c}) \tag{2.12}$$

where the superscripts k and $k+1$ are iteration counters.

(3) Depending on the convergence rates, either execute (2.10), (2.11) and (2.12) respectively until the constraint satisfies

$$A\mathbf{x}^{k+1} + B\mathbf{z}^{k+1} - \mathbf{c} \leq tol \tag{2.13}$$

where *tol* is a predefined tolerance value or execute (2.10), (2.11) and (2.12) respectively for a certain number of iterations.

2.4 Gradient Descent

The Gradient Descent (GD) is an optimization algorithm widely used for minimization of functions and attaining local minimums or maximums. It is applied in numerous areas including machine learning, scientific computation, and optimization tasks. It performs by iteratively taking steps in the direction of the function's negative gradient and updating the parameters accordingly.

Let $f(\mathbf{x})$ denote the function that is intended to be minimized with respect to variable $\mathbf{x}$. Hence, the $\mathbf{x}$ that manages to minimize the function with respect

to its latest parameters is calculated by performing the following update process for a specified number of iterations or until a certain level of tolerance between consequent iterations is met. The corresponding iteration formula for the GD is given as follows,

$$\mathbf{x}^{k+1} = \mathbf{x}^k - \alpha \nabla_{\mathbf{x}} f(\mathbf{x}^k); \qquad k = 0, 1, 2, \ldots \tag{2.14}$$

where the superscripts k and $k+1$ are the iteration counters and α is the predetermined learning rate.

3. IMPLEMENTATION

This thesis study focuses on implementing an efficient lossy compression technique on hyperspectral images which are 3-dimensional data tensors. The mentioned techniques and operations from Section 2 are adjusted to be compatible with 3-dimensional tensors and the necessary explanations are provided in the context.

Consider a 3 dimensional tensor $\mathcal{H} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, where $\mathcal{H}$ is comprised of $n_1 \times n_2$ and n_3 spectral bands. Hence the EMPR application on $\mathcal{H}$ is denoted as,

$$\begin{aligned} \mathcal{H} = h_0\, \mathbf{t}_1 \circ \mathbf{t}_2 \circ \mathbf{t}_3 + \mathbf{h}_1 \circ \mathbf{t}_2 \circ \mathbf{t}_3 + \mathbf{t}_1 \circ \mathbf{h}_2 \circ \mathbf{t}_3 + \mathbf{t}_1 \circ \mathbf{t}_2 \circ \mathbf{h}_3 \\ + \mathbf{h}_{1,2} \circ \mathbf{t}_3 + \mathbf{h}_{1,3} \circ \mathbf{t}_2 + \mathbf{t}_1 \circ \mathbf{h}_{2,3} + \mathbf{h}_{1,2,3} \end{aligned} \tag{3.1}$$

where $\mathbf{t}_1$, $\mathbf{t}_2$, and $\mathbf{t}_3$ are the corresponding EMPR support vectors, the h_0, $\mathbf{h}_i$, $\mathbf{h}_{i,j}$, $\mathbf{h}_{1,2,3}$ components are the EMPR components and $\circ$ denotes the outer product.

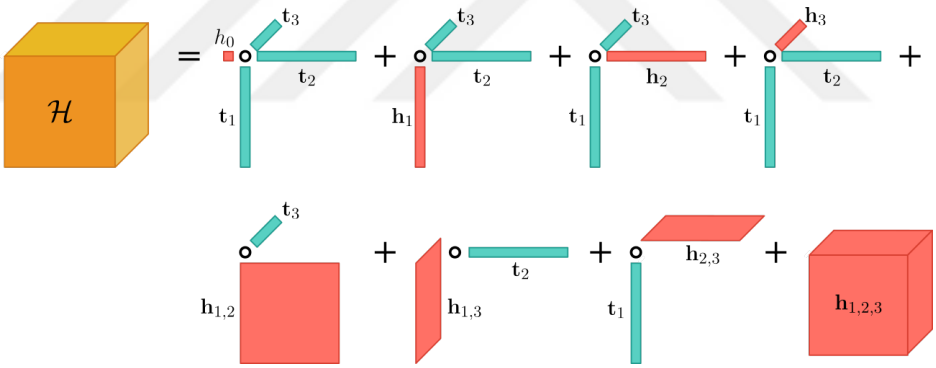

Figure 3.1 : Graphical demonstration of EMPR on a 3-dimensional tensor.

The terms with 2 or higher dimensional entities are truncated from the equality to obtain a first-order EMPR approximation $\widehat{\mathcal{H}}$ of $\mathcal{H}$ and it is denoted as,

$$\mathcal{H} \approx \widehat{\mathcal{H}} = h_0\, \mathbf{t}_1 \circ \mathbf{t}_2 \circ \mathbf{t}_3 + \mathbf{h}_1 \circ \mathbf{t}_2 \circ \mathbf{t}_3 + \mathbf{t}_1 \circ \mathbf{h}_2 \circ \mathbf{t}_3 + \mathbf{t}_1 \circ \mathbf{t}_2 \circ \mathbf{h}_3 \tag{3.2}$$

The scalar EMPR component, h_0, is calculated through the use of the n-mode product rule in (2.1) as follows,

$$h_0 = \frac{1}{n_1 n_2 n_3} \mathcal{H} \,\overline{\times}_1 \mathbf{t}_1 \,\overline{\times}_2 \mathbf{t}_2 \,\overline{\times}_3 \mathbf{t}_3. \tag{3.3}$$

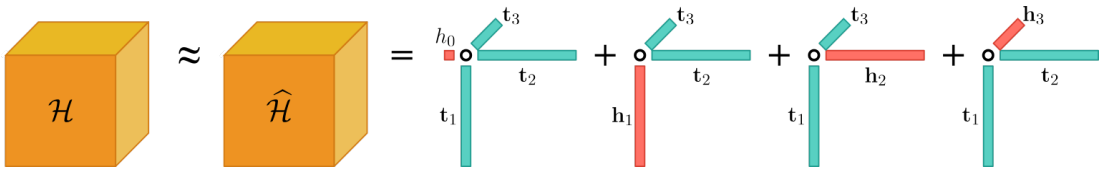

Figure 3.2 : The EMPR approximation $\widehat{\mathcal{H}}$ of $\mathcal{H}$.

Additionally, the corresponding 1-dimensional EMPR components, namely, $\mathbf{h}_1$, $\mathbf{h}_2$, and $\mathbf{h}_3$ are calculated as follows,

$$\begin{aligned}
\mathbf{h}_1 &= \frac{1}{n_2 n_3}\mathcal{H}\,\overline{\times}_2\mathbf{t}_2\overline{\times}_3\mathbf{t}_3 - h_0\mathbf{t}_1,\\
\mathbf{h}_2 &= \frac{1}{n_1 n_3}\mathcal{H}\,\overline{\times}_1\mathbf{t}_1\overline{\times}_3\mathbf{t}_3 - h_0\mathbf{t}_2,\\
\mathbf{h}_3 &= \frac{1}{n_1 n_2}\mathcal{H}\,\overline{\times}_1\mathbf{t}_1\overline{\times}_2\mathbf{t}_2 - h_0\mathbf{t}_3.
\end{aligned} \tag{3.4}$$

The approach that this work pursues, performs EMPR on 3-dimensional tensors. Although, it is possible to proceed with a 3-dimensional EMPR application on $\mathcal{H}$, moving forward with 2-dimensional EMPR applications on $\mathcal{H}$'s matrix

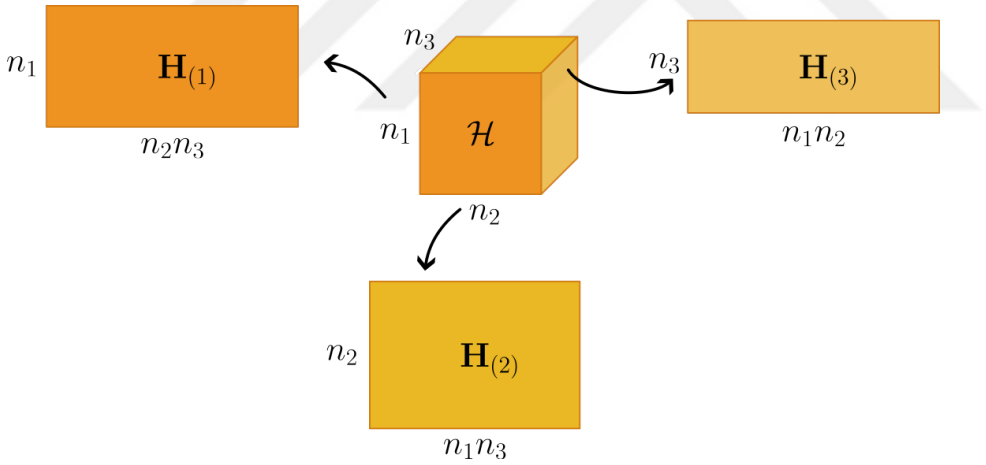

Figure 3.3 : Unfoldings of tensor $\mathcal{H}$.

unfoldings [50] provides better insights into the characteristics and features of each dimension. Thus, the remaining operations are defined and employed by means of a 2-dimensional EMPR procedure.

3.1 The Proposed Method

Let $\mathbf{F}$ denote the unfolding of $\mathcal{H}$ with respect to its i-th dimension, which is

$$\mathbf{F} = \mathbf{H}_{(i)}; \qquad i = 1, 2, 3 \tag{3.5}$$

Then, the EMPR approximation $\widehat{\mathbf{F}}$ of $\mathbf{F}$ is expressed as follows

$$\mathbf{F} \approx \widehat{\mathbf{F}} = f_0\, \mathbf{s}_1\mathbf{s}_2^T + \mathbf{f}_1\mathbf{s}_2^T + \mathbf{s}_1\mathbf{f}_2^T \tag{3.6}$$

where $\widehat{\mathbf{F}} = \mathbf{F}(\mathbf{s}_1, \mathbf{s}_2)$. It should be noted that $\mathbf{s}_1$ stores information on $\mathcal{H}$'s i-th dimension.

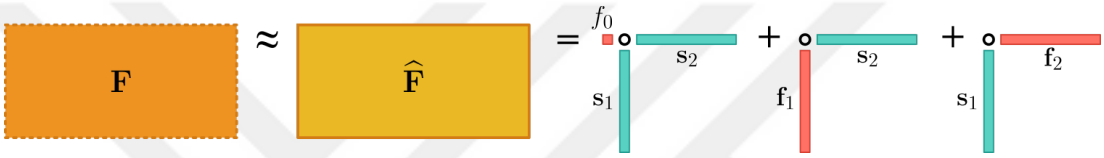

Figure 3.4 : The EMPR approximation $\widehat{\mathbf{F}}$ of $\mathbf{F}$.

Optimal support vectors $\mathbf{s}_1$ and $\mathbf{s}_2$ can be found through a gradient descent based procedure but rather than proceeding with such an iterative process, embedding ADMM for managing the iteration process improves the convergence rate and enables better control over the variables' properties. As stated in Section 2.3, ADMM requires convex functions and constraints to perform.

Hence, to address the distortion between $\mathbf{F}$ and its approximation $\widehat{\mathbf{F}}$, the corresponding convex loss function f and the convex function g that ensures the sparsity of vectors $\mathbf{z}_1$ and $\mathbf{z}_2$ are defined as follows

$$f(\mathbf{s}_1, \mathbf{s}_2) = \frac{1}{2}\|\mathbf{F} - \widehat{\mathbf{F}}\|_F^2 \tag{3.7}$$

$$g(\mathbf{z}_1, \mathbf{z}_2) = \lambda_1\|\mathbf{z}_1\|_1 + \lambda_2\|\mathbf{z}_2\|_1 \tag{3.8}$$

As the aim of the method is to optimize the weight vectors $\mathbf{s}_1$ and $\mathbf{s}_2$, the constraints might be arranged as

$$A\mathbf{s}_i + B\mathbf{z}_i = \mathbf{c}; \qquad i = 1, 2 \tag{3.9}$$

where $A = I$, $B = -I$, $c = \mathbf{0}$ with an expected convergence of $\mathbf{s}_i$ and $\mathbf{z}_i$ to common optimal vectors.

Therefore, the ADMM problem would be of the form

$$\begin{aligned} \min_{\mathbf{s}_1,\mathbf{s}_2,\mathbf{z}_1,\mathbf{z}_2} \quad & f(\mathbf{s}_1,\mathbf{s}_2) + g(\mathbf{z}_1,\mathbf{z}_2) + \frac{\rho_1}{2}\|\mathbf{s}_1 - \mathbf{z}_1\|_2^2 + \frac{\rho_2}{2}\|\mathbf{s}_2 - \mathbf{z}_2\|_2^2 \\ \text{s.t.} \quad & \mathbf{s}_1 - \mathbf{z}_1 = \mathbf{0},\ \mathbf{s}_2 - \mathbf{z}_2 = \mathbf{0} \end{aligned} \tag{3.10}$$

By transforming this ADMM problem into its Augmented Lagrangian form, the following loss function $\mathcal{L}$ is provided explicitly;

$$\begin{aligned} \mathcal{L}_{\rho_1,\rho_2}(\mathbf{s}_1,\mathbf{s}_2,\mathbf{z}_1,\mathbf{z}_2,\mathbf{y}_1,\mathbf{y}_2) = \frac{1}{2}\|\mathbf{F} - \widehat{\mathbf{F}}\|_F^2 + \lambda_1\|\mathbf{z}_1\|_1 + \lambda_2\|\mathbf{z}_2\|_1 \\ + \mathbf{y}_1^T(\mathbf{s}_1 - \mathbf{z}_1) + \mathbf{y}_2^T(\mathbf{s}_2 - \mathbf{z}_2) + \frac{\rho_1}{2}\|\mathbf{s}_1 - \mathbf{z}_1\|_2^2 + \frac{\rho_2}{2}\|\mathbf{s}_2 - \mathbf{z}_2\|_2^2 \end{aligned} \tag{3.11}$$

Optimizing the support vectors requires calculating the respective partial derivatives of $\widehat{\mathbf{F}}$ which can be seen below,

$$\frac{\partial\widehat{\mathbf{F}}}{\partial\mathbf{s}_1} = \mathbf{I}_{n_1} \circ \mathbf{f}_2 + \mathbf{s}_1 \circ \left(\frac{\mathbf{F}^T}{n_1} - \mathbf{s}_2\left(\frac{\mathbf{F}\mathbf{s}_2}{n_1 n_2}\right)^T\right)^T \tag{3.12}$$

$$\frac{\partial\widehat{\mathbf{F}}}{\partial\mathbf{s}_2} = \mathbf{f}_1 \circ \mathbf{I}_{n_2} + \left(\frac{\mathbf{F}}{n_2} - \mathbf{s}_1\left(\frac{\mathbf{F}^T\mathbf{s}_1}{n_1 n_2}\right)\right) \circ \mathbf{s}_2 \tag{3.13}$$

where $\mathbf{I}_{n_1}$ and $\mathbf{I}_{n_2}$ stand as the identity matrices of size $n_1 \times n_1$ and $n_2 \times n_2$ while $\partial\widehat{\mathbf{F}}/\partial\mathbf{s}_1 \in \mathbb{R}^{\,n_1\times n_2\times n_1}$ and $\partial\widehat{\mathbf{F}}/\partial\mathbf{s}_2 \in \mathbb{R}^{\,n_1\times n_2\times n_2}$, respectively.

Consequently, the corresponding gradients which will be employed throughout the iteration steps of $\mathbf{s}_i$ are calculated as

$$\nabla_{\mathbf{s}_i}\mathcal{L}(\mathbf{s}_1,\mathbf{s}_2) = -\frac{\partial\widehat{\mathbf{F}}}{\partial\mathbf{s}_i} \overline{\times}_{1,2} (\mathbf{F} - \widehat{\mathbf{F}}) + \mathbf{y}_i^k + \rho_i\left(\mathbf{s}_i^k - \mathbf{z}_i^k\right); \qquad i = 1,2 \tag{3.14}$$

Due to their explicit structures, dissociating $\mathbf{s}_i$'s in Eqn. (3.14) is not possible using elementary operations. Thus, the Gradient Descent Algorithm from Section 2.4 can be employed for the minimization task which will leads to determining the optimal support vector for each $\mathbf{s}_i$. Then, the iterative optimization procedure is expressed mathematically as follows

$$\mathbf{s}_i^{k+1} = \mathbf{s}_i^k - \alpha\nabla_{\mathbf{s}_i}\mathcal{L}(\mathbf{s}_1,\mathbf{s}_2); \qquad \begin{matrix} i = 1,2 \\ k = 0,1,2,\ldots \end{matrix} \tag{3.15}$$

where α is an appropriate learning rate.

$\mathcal{L}_{\rho_i}$ differentiates with respect to $\mathbf{z}_i$ as follows.

$$\nabla_{\mathbf{z}_i}\mathcal{L}_{\rho_i} = \lambda_1 \frac{\partial}{\partial \mathbf{z}_i}(\|\mathbf{z}_i\|_1) - \mathbf{y}_i^k - \rho_i(\mathbf{s}_i^k - \mathbf{z}_i^k) \tag{3.16}$$

By (3.16), the dissociation of $\mathbf{z}_i$ can be rewritten as

$$\mathbf{z}_i^{k+1} = S_{\lambda_i/\rho_i}(\mathbf{s}_i^{k+1} + \frac{\mathbf{y}_i^k}{\rho_i}) \tag{3.17}$$

where S_κ is the soft-thresholding operator $S_\kappa \colon \mathbb{R}^m \longrightarrow \mathbb{R}^m$ defined as follows.

$$S_\kappa(\alpha) = \begin{cases} \alpha - \kappa & , \quad \alpha > \kappa \\ 0 & , \quad |\alpha| \leq \kappa \\ \alpha + \kappa & , \quad \alpha < -\kappa \end{cases} \tag{3.18}$$

Conducting soft-thresholding on the derivative of $\mathbf{L}_1$ norm, involves the use of subgradient operations. Below can be seen the subgradient soft-thresholding operation with respect to vector $\mathbf{z}_i$.

$$[\mathbf{z}_i]_j = \begin{cases} [\mathbf{s}_i]_j + \frac{[\mathbf{y}_i]_j - \lambda_i}{\rho_i} & , \quad [\mathbf{s}_i]_j > \frac{\lambda_i - [\mathbf{y}_i]_j}{\rho_i} \\ [\mathbf{s}_i]_j + \frac{[\mathbf{y}_i]_j + \lambda_i}{\rho_i} & , \quad [\mathbf{s}_i]_j < \frac{-\lambda_i - [\mathbf{y}_i]_j}{\rho_i} \\ 0 & , \quad otherwise \end{cases} \tag{3.19}$$

where $[\mathbf{z}_i]_j$ corresponds to j'th element of $\mathbf{z}_i$.

To keep track of the difference in subsequent iterations, $\mathbf{y}_i$ is updated as follows,

$$\mathbf{y}_i^{k+1} = \mathbf{y}_i^k + \left(\mathbf{s}_i^{k+1} - \mathbf{z}_i^{k+1}\right) \tag{3.20}$$

The update formulas for the support vectors $\mathbf{s}_i$ in Eqn. (3.15), the vectors $\mathbf{z}_i$ in Eqn. (3.19), and vectors $\mathbf{y}_i$ in Eqn. (3.20) are performed for each unfoldings of $\mathcal{H}$, namely $\mathbf{H}_{(1)}$, $\mathbf{H}_{(2)}$ and $\mathbf{H}_{(3)}$. For each unfolding, two support vectors are obtained with the help of the optimization process. The $\mathbf{s}_1$ vectors for each unfolding is employed as $\mathbf{t}_1$, $\mathbf{t}_2$ and $\mathbf{t}_3$, respectively. Therefore, the construction of a suitable approximation for the 3-dimensional hyperspectral tensor $\mathcal{H}$ is achieved by following the structure in Eqn. (3.2).

The proposed method can be replicated by following the pseudocode in Algorithm 1.

Algorithm 1 Calculate $\widehat{\mathcal{H}}$

Initialize α, λ_1, λ_2, ρ_1, ρ_2, MAX_ITER, and GRAD_ITER
for $i \leftarrow 1, 2, 3$ **do**
    $\mathbf{F} \Leftarrow \mathbf{H}_{(i)}$
    Initialize $\mathbf{s}_j, \mathbf{z}_j, \mathbf{y}_j$ for $j = 1, 2$
    for $j = 1$ to MAX_ITER **do**
        Calculate $\partial\widehat{\mathbf{F}}/\partial\mathbf{s}_1$ via (3.12)
        for $k = 1$ to GRAD_ITER **do**
            Calculate $\widehat{\mathbf{F}}$ via (3.6)
            Calculate $\nabla_{\mathbf{s}_1}\mathcal{L}$ via (3.14)
            $\mathbf{s}_1 \Leftarrow \mathbf{s}_1 - \alpha\nabla_{\mathbf{s}_1}\mathcal{L}$
        end for
        Update $\mathbf{z}_1$ via (3.17) and (3.19)
        Update $\mathbf{y}_1$ via (3.20)
        Calculate $\partial\widehat{\mathbf{F}}/\partial\mathbf{s}_2$ via (3.13)
        for $k = 1$ to GRAD_ITERS **do**
            Calculate $\widehat{\mathbf{F}}$ via (3.6)
            Calculate $\nabla_{\mathbf{s}_2}\mathcal{L}$ via (3.14)
            $\mathbf{s}_2 \Leftarrow \mathbf{s}_2 - \alpha\nabla_{\mathbf{s}_2}\mathcal{L}$
        end for
        Update $\mathbf{z}_2$ via (3.17) and (3.19)
        Update $\mathbf{y}_2$ via (3.20)
    end for
    $\mathbf{t}_i \Leftarrow \mathbf{s}_1$
end for
Calculate h_0, $\mathbf{h}_1$, $\mathbf{h}_2$, and $\mathbf{h}_3$ via (3.3) and (3.4)
Calculate $\widehat{\mathcal{H}}$ via (3.2)

With the help of Algorithm 1, an efficient lossy compression technique for a given hyperspectral tensor $\mathcal{H}$ is achieved.

3.2 Possible Improvements

As a possible improvement for the proposed algorithm, some terms from Eqn. (3.1), more specifically the terms with 2-dimensional sub-components, are added to the final approximation from Eqn. (3.2) to attain a better approximation $\widehat{\mathcal{H}}_{improved}$. The calculation of this improved result is done by the following equation,

$$\begin{aligned}\widehat{\mathcal{H}}_{improved} \approx h_0\,\mathbf{t}_1 \circ \mathbf{t}_2 \circ \mathbf{t}_3 + \mathbf{h}_1 \circ \mathbf{t}_2 \circ \mathbf{t}_3 + \mathbf{t}_1 \circ \mathbf{h}_2 \circ \mathbf{t}_3 + \mathbf{t}_1 \circ \mathbf{t}_2 \circ \mathbf{h}_3 \\ + \mathbf{h}_{1,2} \circ \mathbf{t}_3 + \mathbf{h}_{1,3} \circ \mathbf{t}_2 + \mathbf{t}_1 \circ \mathbf{h}_{2,3}\end{aligned} \tag{3.21}$$

Considering the extended equation which contains 2-dimensional EMPR components, it should be noted that the work done in Section 3.1 holds and only requires the addition of terms $\mathbf{h}_{1,2} \circ \mathbf{t}_3$, $\mathbf{h}_{1,3} \circ \mathbf{t}_2$, and $\mathbf{t}_1 \circ \mathbf{h}_{2,3}$. The 2-dimensional EMPR entities $\mathbf{h}_{1,2}$, $\mathbf{h}_{1,3}$, and $\mathbf{h}_{2,3}$ are calculated by,

$$\begin{aligned}
\mathbf{h}_{1,2} &= \frac{1}{n_3}\mathcal{H} \,\overline{\times}_3 \mathbf{t}_3 - h_0\, \mathbf{t}_1 \circ \mathbf{t}_2 - \mathbf{h}_1 \circ \mathbf{t}_2 - \mathbf{t}_1 \circ \mathbf{h}_2, \\
\mathbf{h}_{1,3} &= \frac{1}{n_2}\mathcal{H} \,\overline{\times}_2 \mathbf{t}_2 - h_0\, \mathbf{t}_1 \circ \mathbf{t}_3 - \mathbf{h}_1 \circ \mathbf{t}_3 - \mathbf{t}_1 \circ \mathbf{h}_3, \\
\mathbf{h}_{2,3} &= \frac{1}{n_1}\mathcal{H} \,\overline{\times}_1 \mathbf{t}_1 - h_0\, \mathbf{t}_2 \circ \mathbf{t}_3 - \mathbf{h}_2 \circ \mathbf{t}_3 - \mathbf{t}_2 \circ \mathbf{h}_3.
\end{aligned} \tag{3.22}$$

where the support vectors $\mathbf{t}_1$, $\mathbf{t}_2$, and $\mathbf{t}_3$ and the 1-dimensional EMPR components $\mathbf{h}_1$, $\mathbf{h}_2$, and $\mathbf{h}_3$ are carried over from the proposed algorithm.

Algorithm 2 Calculate $\widehat{\mathcal{H}}_{improved}$

Perform Algorithm 1
Calculate $\mathbf{h}_{1,2}$, $\mathbf{h}_{1,3}$, and $\mathbf{h}_{2,3}$ via (3.22)
Calculate $\widehat{\mathcal{H}}_{improved}$ via (3.21)

This improved method can be executed by performing the calculations from Algorithm 1 and calculating the additional terms as stated in Algorithm 2.

4. RESULTS

4.1 Datasets

To highlight the robustness and applicability of the proposed method, a selection of hyperspectral datasets were used for the implementations. Indian Pines and Salinas datasets which were collected by AVIRIS sensors and Pavia University dataset which is acquired by ROSIS sensors were used in testing the proposed method and comparing it to other well-known techniques [51]. The sizes of these datasets were adjusted to accommodate memory constraints, thus the specifications are provided in Table 4.1.

Table 4.1 : Dataset specifications

Dataset	Sensor	Original Size	Crop Size
Indian Pines	AVIRIS	145×145×224	75×75×200
Salinas	AVIRIS	512×217×224	70×70×204
Pavia Uni	ROSIS	610×340×103	101×101×96

These datasets are colorized and the false color images of them can be seen in Fig. 4.1, 4.2, and 4.3.

4.2 Initializations

The algorithm requires initial starting values for support vectors $\mathbf{s}_1$ and $\mathbf{s}_2$, and for vectors $\mathbf{z}_1$, $\mathbf{z}_2$, $\mathbf{y}_1$, and $\mathbf{y}_2$, as well. $\mathbf{s}_1$ and $\mathbf{s}_2$ are initialized with positive random numbers and the supports are normalized under the constant weight factors which are $1/n_1$ and $1/n_2$ for each unfolding matrix $\mathbf{H}_{(i)}$, respectively. Vectors $\mathbf{z}_1$, $\mathbf{z}_2$, $\mathbf{y}_1$, and $\mathbf{y}_2$ are initialized as zero vectors with appropriate sizes. Parameters α, λ_1, λ_2, ρ_1, and ρ_2 are selected through a grid search process that resulted in the lowest error rates. It should also be mentioned that normalizing the main tensor $\mathcal{H}$ with respect to its maximum element helps with the calculation process and

Figure 4.1 : Three-band false color image of Indian Pines Dataset.

Figure 4.2 : Three-band false color image of Salinas Dataset.

avoids undesired calculation errors. The final approximation $\widehat{\mathcal{H}}$ are multiplied with the maximum element to finalize the process.

Figure 4.3 : Three-band false color image of Pavia University Dataset.

4.3 Metrics

The proposed approach is compared by the test results with two widely used tensor decomposition techniques to address its efficiency. These techniques are CANDECOMP/PARAFAC Alternating Least Squares (CP-ALS) and Tucker Decomposition (TD), respectively.

For achieving a fair comparison, Mean Squared Error (MSE), Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index Measure (SSIM), and Cosine Similarity are utilized to measure the performance of each technique. These metrics can be computed as follows,

- MSE value between $\mathcal{H}$ and $\widehat{\mathcal{H}}$ by using the Frobenius norm is defined as

$$MSE(\mathcal{H}, \widehat{\mathcal{H}}) = \frac{1}{2}\|\mathcal{H} - \widehat{\mathcal{H}}\|_F^2 \tag{4.1}$$

where $\mathcal{H}$ stands for the original data and $\widehat{\mathcal{H}}$ is the final approximation.

- The peak signal-to-noise ratio utilizes the mean squared error and the maximum pixel value by calculating the following formula.

$$PSNR(\mathcal{H}, \widehat{\mathcal{H}}) = 10 \cdot \log_{10} \left(\frac{MAX^2}{MSE(\mathcal{H}, \widehat{\mathcal{H}})} \right) \tag{4.2}$$

The MAX used in the equation corresponds to the highest pixel value that the tensor can contain. In this case, MAX is equal to 255 for grayscale images as it is the highest value each pixel can contain. For a more general explanation, the MAX value is calculated by considering the tensor's bits per sample value. Let tensor $\mathcal{H}$ have a bits per sample value of B. Thus, the maximum possible pixel value can be calculated via

$$MAX(\mathcal{H}) = s^B - 1 \tag{4.3}$$

- Structural Similarity Index Measure is a metric generally used in determining the perceived quality of images or videos. It differs from MSE and $PSNR$ by considering the interdependencies between close pixels instead of solely relying on error calculations. This metric employs the pixel sample means as well as variances and covariances of the compared tensors.

$$SSIM(\mathcal{H}, \widehat{\mathcal{H}}) = \frac{\left(2\mu_{\mathcal{H}}\mu_{\widehat{\mathcal{H}}} + c_1\right)\left(2\sigma_{\mathcal{H}\widehat{\mathcal{H}}} + c_2\right)}{\left(\mu_{\mathcal{H}}^2 + \mu_{\widehat{\mathcal{H}}}^2 + c_1\right)\left(\sigma_{\mathcal{H}}^2 + \sigma_{\widehat{\mathcal{H}}}^2 + c_2\right)} \tag{4.4}$$

μ denotes the pixel sample mean while $\sigma_{\mathcal{H}}$, $\sigma_{\widehat{\mathcal{H}}}$ and $\sigma_{\mathcal{H}\widehat{\mathcal{H}}}$ denote the variances and covariance, respectively. c_1 and c_2 are used for stabilizing the division with weak denominators.

- Cosine Similarity calculates the similarity of compared tensors by employing the cosine theorem. The resulting cosine value determines how close these tensors are where a final value of 1 corresponds to identical tensors, and 0 corresponds to dissimilar tensors.

$$CosSim(\mathbf{x}, \widehat{\mathbf{x}}) = \frac{\mathbf{x} \cdot \widehat{\mathbf{x}}}{\|\mathbf{x}\|\|\widehat{\mathbf{x}}\|} \tag{4.5}$$

where $\mathbf{x}$ denotes individual fibers from the original data tensor and $\widehat{\mathbf{x}}$ denotes the individual fibers from the final approximation tensor. This calculation process is executed on each fiber couple and the average of these metric calculations is taken for a finalized comparison value.

4.4 Performance Comparisons

The metric calculations are executed after each method establishes its optimal components. The final tests are performed on the approximation which is calculated over 250 iterations. This iteration count is determined after observing the convergence behavior in metric results, specifically concerning the MSE.

The results of the proposed method are compared with CANDECOMP/PARAFAC-ALS and Tucker Decomposition's results to gain a clear understanding of how well it performs at different rank levels. The proposed method is also compared to EMPR with its support vectors being calculated via Averaged Directional Supports. This comparison will acquaint us about the potential of using iterative solutions for the optimization of the support vectors. The calculated results also contain test results of the proposed method as well as its ADS applied version counterpart to investigate the effect of the added terms.

Tables 4.2, 4.3, and 4.4 present the results of the metric calculations. The algorithm is utilized in each dataset whose specifications are presented in Table 4.1.

Table 4.2 presents the test results of the mentioned techniques on the cropped part of the Indian Pines dataset. The test results show a similar level of performance between the tensor decomposition techniques applied with lower ranks 2 and 5 and the main method of the proposed method, implying a different output from the expected results. A closer inspection through a debugging process reveals that the soft-thresholding process during the optimization of the support vectors does not perform as expected. soft-thresholding is used as the result of differentiating $\mathbf{L}_1$-norms and it's parameters are arranged with the intended

Table 4.2 : Results on Indian Pines Dataset

Method	Rank	SNR	PSNR	SSIM	Cos.Sim.	MSE
CP-ALS	2	18.771432	45.173589	0.643620	0.995492	0.340579
	5	21.514204	47.916361	0.728089	0.997912	0.248358
	10	23.946030	50.348188	0.788861	0.998576	0.187710
	15	25.470259	51.872417	0.819490	0.998860	0.157498
TD	2	19.307530	45.709687	0.662698	0.996133	0.320194
	5	22.061614	48.463771	0.747603	0.998019	0.233189
	10	25.004040	51.406198	0.811073	0.998840	0.166183
	15	26.757828	53.159985	0.853358	0.999170	0.135799
EMPR-ADS	3	0.189933	26.592090	0.001168	0.994787	2.892628
	3+	2.378978	28.781136	0.146231	0.995333	2.248228
EMPR-ADMM	3	18.317528	44.719686	0.629272	0.994785	0.358850
	3+	20.822221	47.224379	0.756039	0.995793	0.268954

use of sparsification of the support vectors. The debugging process reveals that the support vectors, despite the expected increase in the number of zero-valued elements, aren't sparsified, hence, the proposed method can't achieve the expected level of results.

Table 4.3 : Results on Salinas Dataset

Method	Rank	SNR	PSNR	SSIM	Cos.Sim.	MSE
CP-ALS	2	15.617441	47.746480	0.756081	0.992210	0.268681
	5	20.481271	52.610310	0.818436	0.998252	0.153477
	10	23.519517	55.648556	0.848448	0.998986	0.108176
	15	26.143033	58.272072	0.885262	0.999351	0.079975
TD	2	16.135203	48.264242	0.764124	0.993602	0.253133
	5	21.076551	53.205590	0.819501	0.998370	0.143311
	10	24.900040	57.029079	0.861081	0.999195	0.092280
	15	27.875721	60.004760	0.906163	0.999609	0.065512
EMPR-ADS	3	0.014150	31.787518	0.000035	0.960477	1.711902
	3+	0.616961	32.390330	0.032719	0.975594	1.597123
EMPR-ADMM	3	14.443628	46.572667	0.726994	0.988784	0.307558
	3+	17.656066	49.785105	0.786195	0.992009	0.212474

Similar to the results in Table 4.2, the exhibited behaviour is carried over to both tests series on datasets Salinas and Pavia University. Due to not being able

to take advantage of the sparsification of the support vectors, the performance of the proposed method and its improved version falls in between the results of CP-ALS and TD's results considering ranks 2 and 5. On the other hand, the applications performed on the Pavia University dataset reveal an interesting behaviour. For the SSIM test results, the proposed method when applied with its improved version can surpass the tensor decomposition methods. Still, the ADS-applied EMPR method achieves a higher result compared to the proposed method, indicating the deficiency caused by the soft-thresholding algorithm.

The obtained test results for datasets Salinas and Pavia University can be seen in Tables 4.3 and 4.4.

Table 4.4 : Results on Pavia Uni Dataset

Method	Rank	SNR	PSNR	SSIM	Cos.Sim.	MSE
CP-ALS	2	7.685556	38.517698	0.122731	0.956340	0.785473
	5	9.390002	40.222144	0.204745	0.963309	0.645519
	10	12.200862	43.033004	0.315987	0.980330	0.467053
	15	13.800936	44.633078	0.387519	0.984494	0.388474
TD	2	7.685556	38.517698	0.122731	0.956341	0.785473
	5	9.976316	40.808458	0.208531	0.971923	0.603383
	10	12.893479	43.725621	0.338052	0.983634	0.431256
	15	14.571643	45.403785	0.417126	0.988179	0.355490
EMPR-ADS	3	0.031075	30.863217	0.000011	0.941355	1.896075
	3+	0.851196	31.683338	0.018618	0.948399	1.725240
EMPR-ADMM	3	6.697513	37.529655	0.097921	0.941273	0.880103
	3+	11.177258	42.009400	0.563073	0.954037	0.525468

After these observations, it is time to conduct the MSE experiments to compare the proposed method with the other tensor decomposition methods.

In light of these information, experiments were conducted on the mentioned datasets, comparing the performance of the proposed method to other decomposition methods.

In Fig. 4.4, the MSE values were visualized for the Indian Pines dataset by the proposed method, alongside the CP-ALS and TD techniques on varying ranks. It is observed that the proposed method performs similar to CP-ALS and TD

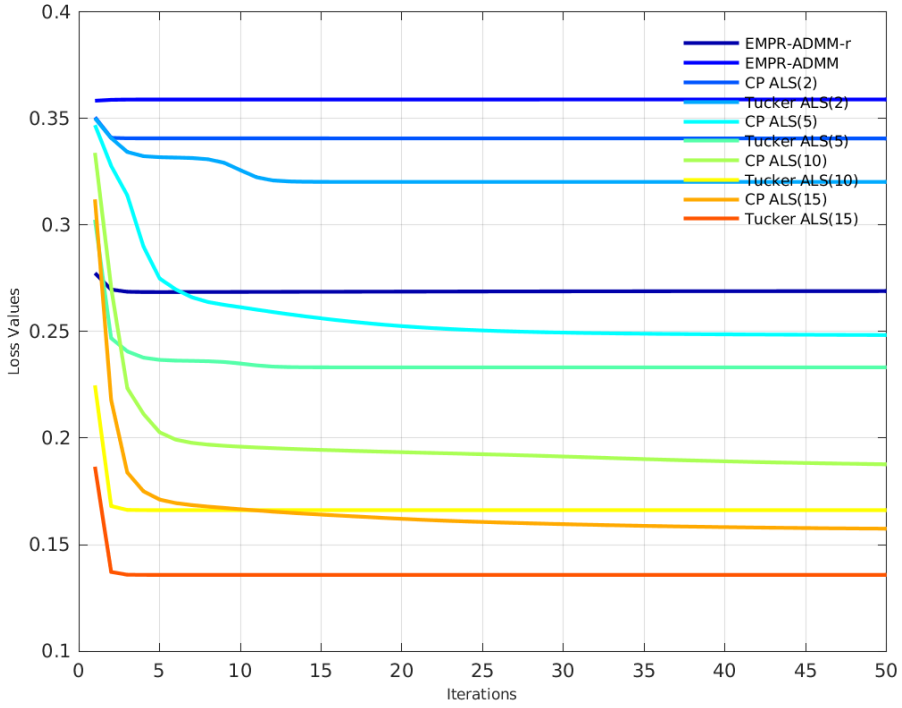

Figure 4.4 : MSE scores through iterations on dataset Indian Pines.

at lower ranks. Between CP-ALS and TD, it is seen that CP-ALS starts at a higher error level but converges after 5 to 10 iterations and it takes even less iterations for TD to converge to an optimal result. On the other hand, the proposed method achieves its optimal result generally in less then 5 iterations. Since GD is employed through the support vector optimization, an expected behaviour would be the score undergoing a converging process but due to the previously mentioned issue, such a result is not achieved.

Fig. 4.5 depicts the experiments on the cropped part of the Salinas dataset. It is seen that CP-ALS and TD exhibit similar performance to their test results on the Indian Pines dataset. TD converges faster with respect to CP-ALS and it takes CP-ALS more iterations to converge to an optimal solution. On the other hand, the studied method converges slower compared to its behaviour in its previous tests on Indian Pines.

Finally, results from the comparisons on tests performed on the cropped segment of Pavia University are shown in Fig. 4.6. CP-ALS is the slowest to converge, whereas TD converges in the first 10 iterations. The proposed method when

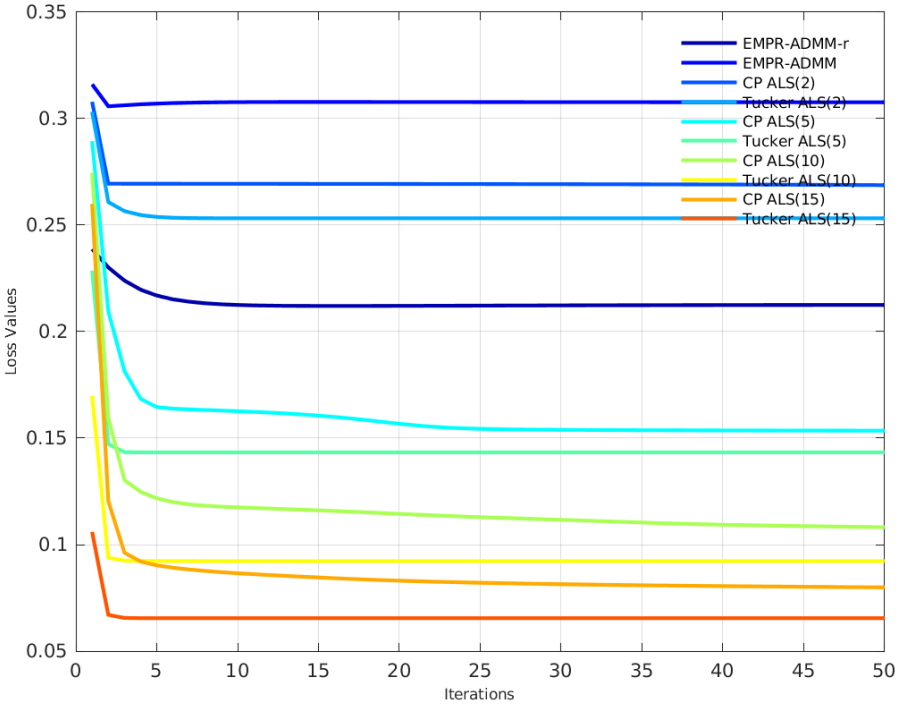

Figure 4.5 : MSE scores through iterations on dataset Salinas.

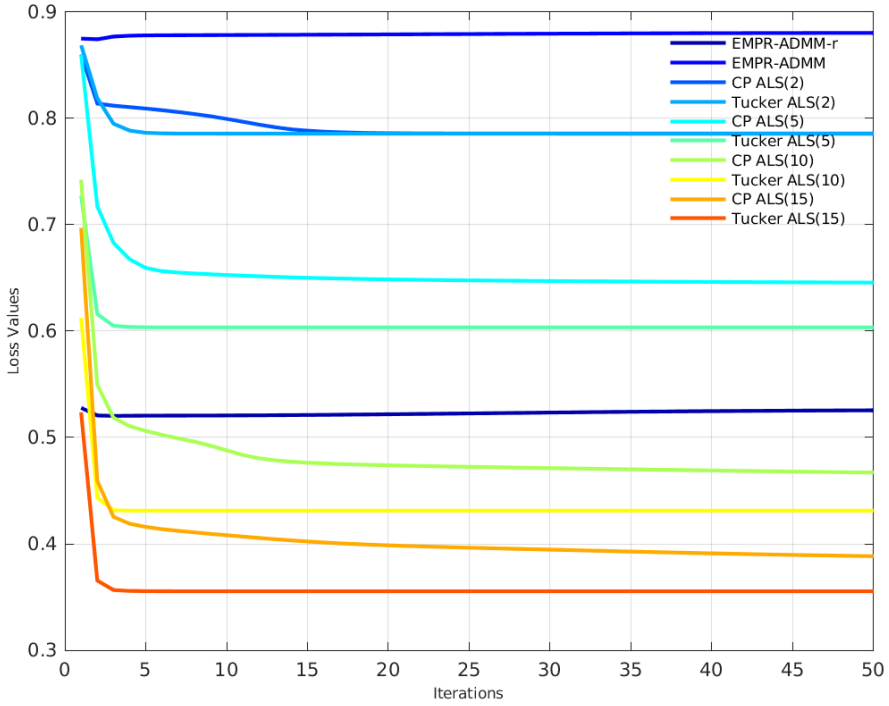

Figure 4.6 : MSE scores through iterations on dataset Pavia University.

tested with the improvements from Section 3.2 starts with a lower error rate,

similar to TD with given rank 15 but falls behind as the other tests converge to a lower error rate.

Through the figures and their deductions, it can be concluded that the proposed method performs at a similar level to both techniques at lower ranks. The studied method also achieves these results in a shorter iteration process compared to CP-ALS and TD.

5. CONCLUSIONS

This chapter is organized as the last section where the methods and their capabilities will be discussed, considering the theoretical and practical aspects of these techniques. In addition to all the current findings, future work that can improve the performance and efficiency will also be discussed.

As previously stated, this thesis work aims to present a new perspective in the area of compressing high dimensional data, more specifically 3-dimensional hyperspectral data for efficiency in further processing tasks and minimizing the data without losing its characteristic features. Hyperspectral data are 3-dimensional tensors that host high correlations throughout its spatial and spectral band information. The Enhanced Multivariance Product Representation's superior decorrelation and denoising abilities enable an effective compression while preserving the core features. This process is employed by optimizing the support vectors of the EMPR expansion through an iterative approach. Instead of directly assigning certain values through assumptions or calculations via the Averaged Directionally Supports (ADS), the Alternating Direction Method of Multipliers is employed for its superior converging properties. Combining these decorrelation, denoising, and converging properties into a new method contributes to the literature with a new capable method for increased efficiency.

Other tensor decomposition methods in the literature of compression and feature extraction, namely Canonical Decomposition/Parallel Factorization and Tucker Decomposition, are widely used and modified in accordance with the needs of the task at hand. These methods are compared with the proposed method, operating through the use of the Alternating Least Squares technique for component calculations.

The comparison of the proposed method and the mentioned tensor decomposition is performed on the hyperspectral datasets Indian Pines, Salinas, and Pavia University. These datasets are used in performance comparisons for hyperspectral image processing tasks and the datasets are pre-processed for the tests. The results obtained from these tests indicate that the proposed method not only keeps pace with two important tensor decomposition techniques in metrics SNR, PSNR, SSIM, CosSim, and MSE, it also demonstrates its capability of compressing high dimensional data at a similar level to the well-known tensor decomposition methods at specified ranks by reconstructing high-dimensional data with comparable accuracy. This result of the proposed method highlights its potential to reduce noise and artefacts in the process by neglecting the residual term. The key point of the method lies in the combination of EMPR's superior decorrelation ability and ADMM's superior converging features.

These results are particularly promising for any lossy compression task including Cartesian geometry utilizing tensor decomposition techniques where accurate and efficient data processing is paramount. Furthermore, this performance advantage paves the way for advancements in lossy compression techniques, enabling researchers and practitioners to gain more precise insights from data.

REFERENCES

[1] **Plaza, A.**, **Plaza, J. and Vegas, H.** (2010). Improving the Performance of Hyperspectral Image and Signal Processing Algorithms Using Parallel, Distributed and Specialized Hardware-Based Systems, *Signal Processing Systems*, *61*, 293–315.

[2] **Audebert, N.**, **Le Saux, B. and Lefèvre, S.** (2019). Deep learning for classification of hyperspectral data: A comparative review, *IEEE geoscience and remote sensing magazine*, *7*(2), 159–173.

[3] **Lu, B.**, **Dao, P.**, **Liu, J.**, **He, Y. and Shang, J.** (2020). Recent Advances of Hyperspectral Imaging Technology and Applications in Agriculture, *Remote Sensing*, *12*(16), 2659, `http://dx.doi.org/10.3390/rs12162659`.

[4] **Gadea, O.C.A. and Khan, S.D.** (2023). Detection of Bastnäsite-Rich Veins in Rare Earth Element Ores Through Hyperspectral Imaging, *IEEE Geoscience and Remote Sensing Letters*, *20*, 1–4.

[5] **Jörges, C.**, **Vidal, H.S.**, **Hank, T. and Bach, H.** (2023). Detection of Solar Photovoltaic Power Plants Using Satellite and Airborne Hyperspectral Imaging, *Remote Sensing*, *15*(13), 3403.

[6] **Adão, T.**, **Hruška, J.**, **Pádua, L.**, **Bessa, J.**, **Peres, E.**, **Morais, R. and Sousa, J.** (2017). Hyperspectral Imaging: A Review on UAV-Based Sensors, Data Processing and Applications for Agriculture and Forestry, *Remote Sensing*, *9*(11), 1110, `http://dx.doi.org/10.3390/rs9111110`.

[7] **Malegori, C.**, **Sciutto, G.**, **Oliveri, P.**, **Prati, S.**, **Gatti, L.**, **Catelli, E.**, **Benazzi, S.**, **Cercatillo, S.**, **Paleček, D.**, **Mazzeo, R.** *et al.* (2023). Near-infrared hyperspectral imaging to map collagen content in prehistoric bones for radiocarbon dating, *Communications Chemistry*, *6*(1), 54.

[8] **Karim, S.**, **Qadir, A.**, **Farooq, U.**, **Shakir, M. and Laghari, A.A.** (2023). Hyperspectral imaging: a review and trends towards medical imaging, *Current Medical Imaging*, *19*(5), 417–427.

[9] **Thomas, S.**, **Kuska, M.T.**, **Bohnenkamp, D.**, **Brugger, A.**, **Alisaac, E.**, **Wahabzada, M.**, **Behmann, J. and Mahlein, A.K.** (2018). Benefits of hyperspectral imaging for plant disease detection and plant

protection: a technical perspective, *Journal of Plant Diseases and Protection*, *125*, 5–20.

[10] **Ai, W.**, **Liu, S.**, **Liao, H.**, **Du, J.**, **Cai, Y.**, **Liao, C.**, **Shi, H.**, **Lin, Y.**, **Junaid, M.**, **Yue, X.** *et al.* (2022). Application of hyperspectral imaging technology in the rapid identification of microplastics in farmland soil, *Science of The Total Environment*, *807*, 151030.

[11] **Halicek, M.**, **Fabelo, H.**, **Ortega, S.**, **Callico, G.M. and Fei, B.** (2019). In-vivo and ex-vivo tissue analysis through hyperspectral imaging techniques: revealing the invisible features of cancer, *Cancers*, *11*(6), 756.

[12] **Wang, M.**, **Hong, D.**, **Han, Z.**, **Li, J.**, **Yao, J.**, **Gao, L.**, **Zhang, B. and Chanussot, J.** (2023). Tensor decompositions for hyperspectral data processing in remote sensing: A comprehensive review, *IEEE Geoscience and Remote Sensing Magazine*.

[13] **Chang, Y.**, **Yan, L.**, **Fang, H.**, **Zhong, S. and Liao, W.** (2018). HSI-DeNet: Hyperspectral image restoration via convolutional neural network, *IEEE Transactions on Geoscience and Remote Sensing*, *57*(2), 667–682.

[14] **Li, C.**, **Sun, T.**, **Kelly, K.F. and Zhang, Y.** (2011). A compressive sensing and unmixing scheme for hyperspectral data processing, *IEEE Transactions on Image Processing*, *21*(3), 1200–1210.

[15] **Matteoli, S.**, **Diani, M. and Corsini, G.** (2010). A tutorial overview of anomaly detection in hyperspectral images, *IEEE Aerospace and Electronic Systems Magazine*, *25*(7), 5–28.

[16] **Yokoya, N.**, **Grohnfeldt, C. and Chanussot, J.** (2017). Hyperspectral and multispectral data fusion: A comparative review of the recent literature, *IEEE Geoscience and Remote Sensing Magazine*, *5*(2), 29–56.

[17] **Hege, E.K.**, **O'Connell, D.**, **Johnson, W.**, **Basty, S. and Dereniak, E.L.** (2004). Hyperspectral imaging for astronomy and space surveillance, *Imaging Spectrometry IX*, volume5159, SPIE, pp.380–391.

[18] **Menon, V.**, **Du, Q. and Fowler, J.E.** (2016). Hadamard-Walsh random projection for hyperspectral image classification, *2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, IEEE, pp.5141–5144.

[19] **Zabalza, J.**, **Ren, J.**, **Ren, J.**, **Liu, Z. and Marshall, S.** (2014). Structured covariance principal component analysis for real-time onsite feature extraction and dimensionality reduction in hyperspectral imaging, *Applied Optics*, *53*(20), 4440–4449.

[20] **Ren, J.**, **Zabalza, J.**, **Marshall, S. and Zheng, J.** (2014). Effective Feature Extraction and Data Reduction in Remote Sensing Using Hyperspectral Imaging, *Signal Processing Magazine, IEEE*, *31*, 149–154.

[21] **Zabalza, J.**, **Ren, J.**, **Yang, M.**, **Zhang, Y.**, **Wang, J.**, **Marshall, S. and Han, J.** (2014). Novel Folded-PCA for improved feature extraction and data reduction with hyperspectral imaging and SAR in remote sensing, *ISPRS Journal of Photogrammetry and Remote Sensing*, *93*, 112–122.

[22] **Wang, J. and Chang, C.I.** (2006). Independent component analysis-based dimensionality reduction with applications in hyperspectral image analysis, *Geoscience and Remote Sensing, IEEE Transactions on*, *44*, 1586 – 1600.

[23] **Amato, U.**, **Cavalli, R.**, **Palombo, A.**, **Pignatti, S. and Santini, F.** (2009). Experimental Approach to the Selection of the Components in the Minimum Noise Fraction, *Geoscience and Remote Sensing, IEEE Transactions on*, *47*, 153 – 160.

[24] **Huang, X. and Zhang, L.** (2008). An Adaptive Mean-Shift Analysis Approach for Object Extraction and Classification From Urban Hyperspectral Imagery, *IEEE T. Geoscience and Remote Sensing*, *46*, 4173–4185.

[25] **Soltani-Farani, A.**, **Rabiee, H.R. and Hosseini, S.A.** (2014). Spatial-aware dictionary learning for hyperspectral image classification, *IEEE Transactions on geoscience and remote sensing*, *53*(1), 527–541.

[26] **Charles, A.S.**, **Olshausen, B.A. and Rozell, C.J.** (2011). Learning sparse codes for hyperspectral imagery, *IEEE Journal of Selected Topics in Signal Processing*, *5*(5), 963–978.

[27] **Lin, X.**, **Liu, Y.**, **Wu, J. and Dai, Q.** (2014). Spatial-spectral encoded compressive hyperspectral imaging, *ACM Transactions on Graphics (TOG)*, *33*(6), 1–11.

[28] **Chen, Y.**, **Nasrabadi, N.M. and Tran, T.D.** (2011). Sparse representation for target detection in hyperspectral imagery, *IEEE Journal of Selected Topics in Signal Processing*, *5*(3), 629–640.

[29] **Xie, Y.**, **Qu, Y.**, **Tao, D.**, **Wu, W.**, **Yuan, Q. and Zhang, W.** (2016). Hyperspectral image restoration via iteratively regularized weighted schatten p-norm minimization, *IEEE Transactions on Geoscience and Remote Sensing*, *54*(8), 4642–4659.

[30] **Yuan, X.**, **Tsai, T.H.**, **Zhu, R.**, **Llull, P.**, **Brady, D. and Carin, L.** (2015). Compressive hyperspectral imaging with side information, *IEEE Journal of selected topics in Signal Processing*, *9*(6), 964–976.

[31] **Wang, J.**, **Kwon, S. and Shim, B.** (2012). Generalized orthogonal matching pursuit, *IEEE Transactions on signal processing*, *60*(12), 6202–6216.

[32] **Beck, A. and Teboulle, M.** (2009). A fast iterative shrinkage-thresholding algorithm for linear inverse problems, *SIAM journal on imaging sciences*, *2*(1), 183–202.

[33] **Li, G.**, **Rosenthal, C. and Rabitz, H.** (2001). High dimensional model representations, *The Journal of Physical Chemistry A*, *105*(33), 7765–7777.

[34] **Li, G. and Rabitz, H.** (2012). General formulation of HDMR component functions with independent and correlated variables, *Journal of mathematical chemistry*, *50*, 99–130.

[35] **Rabitz, H. and Aliş, Ö.F.** (1999). General foundations of high-dimensional model representations, *Journal of Mathematical Chemistry*, *25*(2-3), 197–233.

[36] **Sobol', I.M.** (2003). Theorems and examples on high dimensional model representation, *Reliability Engineering and System Safety*, *79*(2), 187–193.

[37] **Tuna, S.**, **Töreyin, B.**, **Demiralp, M.**, **Ren, J.**, **Zhao, H. and Marshall, S.** (2020). Iterative Enhanced Multivariance Products Representation for Effective Compression of Hyperspectral Images, *IEEE Transactions on Geoscience and Remote Sensing*.

[38] **Tunga, M.A. and Demiralp, M.** (2005). A factorized high dimensional model representation on the nodes of a finite hyperprismatic regular grid, *Applied Mathematics and Computation*, *164*(3), 865–883.

[39] **Tunga, B. and Demiralp, M.** (2009). A novel hybrid high-dimensional model representation (HDMR) based on the combination of plain and logarithmic high-dimensional model representations, *Advances in Numerical Methods*, 101–111.

[40] **Chowdhury, R. and Rao, B.** (2009). Hybrid high dimensional model representation for reliability analysis, *Computer Methods in Applied Mechanics and Engineering*, *198*(5-8), 753–765.

[41] **Tunga, M.A. and Demiralp, M.** (2006). Hybrid high dimensional model representation (HHDMR) on the partitioned data, *Journal of Computational and Applied Mathematics*, *185*(1), 107–132.

[42] **Tunga, M.A. and Demiralp, M.** (2012). Multivariate data modelling through Piecewise generalized HDMR method, *Journal of Mathematical Chemistry*, *50*, 1711–1726.

[43] **Demiralp, M.** (2009). Fluctuationlessness theorem to approximate univariate functions' matrix representations, *WSEAS Transactions on Mathematics*, *8*(5), 258–267.

[44] **Tunga, B. and Demiralp, M.** (2010). The influence of the support functions on the quality of enhanced multivariance product representation, *Journal of Mathematical Chemistry - J MATH CHEM*, *48*, 827–840.

[45] **Korkmaz Özay, E. and Demiralp, M.** (2014). Reductive enhanced multivariance product representation for multi-way arrays, *Journal of Mathematical Chemistry*, *52*.

[46] **Tuna, S. and Tunga, B.** (2013). A novel piecewise multivariate function approximation method via universal matrix representation, *Journal of Mathematical Chemistry*, *51*.

[47] **Tunga, A. and Demiralp, M.** (2013). A novel method for multivariate data modelling: Piecewise Generalized EMPR, *Journal of Mathematical Chemistry*, *51*.

[48] **Sukhanov, A.**, **Tuna, S. and Töreyin, B.** (2016). Lossy compression of hyperspectral images by using Enhanced Multivariance Products Representation (EMPR) method, pp.1925–1928.

[49] **Kolda, T. and Bader, B.** (2009). Tensor Decompositions and Applications, *SIAM Review*, *51*, 455–500.

[50] **Sidiropoulos, N.D.**, **De Lathauwer, L.**, **Fu, X.**, **Huang, K.**, **Papalexakis, E.E. and Faloutsos, C.** (2017). Tensor decomposition for signal processing and machine learning, *IEEE Transactions on signal processing*, *65*(13), 3551–3582.

[51] **Graña, M.**, **Veganzons, M. and Ayerdi, B.** (2023). *Hyperspectral Datasets*, `https://www.ehu.eus/ccwintco/index.php/Hyperspectral\_Remote\_Sensing\\\_Scenes`.

CURRICULUM VITAE

Name SURNAME: Muhammed Enis ŞEN

EDUCATION:

- **B.Sc.:** 2021, Bogazici University, Faculty of Arts and Sciences, Department of Mathematics

PROFESSIONAL EXPERIENCE AND REWARDS:

- December 2023 - Present, AI Engineer

PUBLICATIONS, PRESENTATIONS AND PATENTS ON THE THESIS:

- **Şen M.**, Tuna S. (2023). Exploiting Optimal Supports in Enhanced Multivariance Products Representation for Lossy Compression of Hyperspectral Images. *International Conference on Electrical and Electronics Engineering (ELECO)*, November 30 - December 2, 2023.