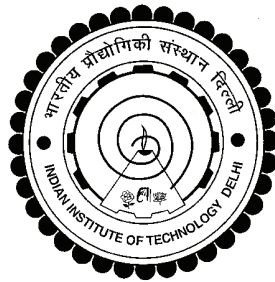


SENSING MATRIX OPTIMIZATION AND FEATURE EXTRACTION FOR IMAGE PROCESSING

Meenakshi



**DEPARTMENT OF ELECTRICAL ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY DELHI**

OCTOBER 2022

© Indian Institute of Technology Delhi (IITD), New Delhi, 2022

SENSING MATRIX OPTIMIZATION AND FEATURE EXTRACTION FOR IMAGE PROCESSING

by

Meenakshi

DEPARTMENT OF ELECTRICAL ENGINEERING

Submitted

*in fulfilment of the requirements of the degree of Doctor of Philosophy
to the*



INDIAN INSTITUTE OF TECHNOLOGY DELHI

OCTOBER 2022

CERTIFICATE

This is to certify that the thesis titled **Sensing matrix optimization and feature extraction for image processing**, submitted by **Meenakshi**, to the Indian Institute of Technology Delhi, for the award of the degree of **Doctor of Philosophy**, is a bonafide record of the research work done by her under my supervision.

The contents of this thesis, in full or in parts, have not been submitted to any other Institute or University for the award of any degree or diploma.

Prof. Seshan Srirangarajan

Thesis Supervisor

Department of Electrical Engineering

Indian Institute of Technology, Delhi

Hauz Khas New Delhi 110016

Place: New Delhi

ACKNOWLEDGEMENTS

This thesis has been made possible owing to the generous help of many people and I am delighted to acknowledge them here. I would like to thank Prof. Seshan Srirangarajan for accepting me into his group. I can never thank him enough for his guidance, encouragement and patience that I received during my graduate program.

I would further like to express my deepest appreciation to my family and friends for their endless support and for understanding me in the low moments. I have immense gratitude for my parents, I would not be where I am today without their help and support. I am deeply indebted to them for their unconditional love and support through this long arduous journey. To my wonderful husband, Vinay, thank you for your support, appreciation and for your unshaken belief in me.

Finally, I would like to send my heartfelt thanks to Prof. Shankar Prakriya, Prof. Monika Agarwal, and Prof. Brejesh Lall for accepting to be my student research committee members and for their valuable feedback during my end semester presentations.

ABSTRACT

At the heart of many pattern recognition and computer vision applications is the question of how to represent and recover data. Representing data in a manner which highlights its key properties can significantly affect the performance of pattern recognition and computer vision applications. Many approaches have been proposed for representing and recovering the high-dimensional data which tend to reside in low-dimensional intrinsic subspaces. Recently, there has been much research in optimally selecting informative samples from high-dimensional data via sensing matrix using compressed sensing. Compressed sensing allows us to exploit the sparse structure of the underlying phenomena for capturing incoherent measurements using a sensing matrix and then recovering the high-dimensional data. However, the sparse signal structure is not the only key aspect of compressed sensing; a robust incoherent frame structure to be used for the measurements is equally important.

In this thesis, we address the problem of reconstructing high-dimensional data (such as images) in the presence of noise using measurements via incoherent sensing matrix for data compression. We propose an incoherent and robust sensing matrix design based on an equiangular tight frame to ensure reduced pairwise correlation among the frame vectors and tightness of the frame. Experiments are performed on synthetic data as well as real images. The performance evaluation is carried out in terms of mutual coherence, signal reconstruction accuracy, and peak signal-to-noise ratio. In addition, it is observed that the high dimensional data samples widely used in computer vision and pattern recognition tasks may contain irrelevant and redundant features. This can result in increased computational complexity, memory requirements, as well as the possibility of overfitting which can greatly weaken the inherent structural relationship between features degrading the overall performance. Thus, finding compact low-dimensional representation and features with good generalization and discriminative abilities is very important.

In this thesis we address these issues to develop feature extraction models for image classi-

fication. Firstly, we present a representation-based supervised feature extraction model exploring the underlying global and local structure of the high-dimensional data. Next, we propose a regression-based feature extraction model to address the problem of insufficient inter-class margin and loss of the manifold structure of the data. To ensure the relative significance of the extracted features, an $\ell_{2,1}$ -norm based regularization is employed. Performance of the proposed models is evaluated on various data sets for face, scene, and object recognition. The results achieved are compared with state-of-the-art methods to validate the effectiveness of the proposed models.

सार

कई पैटर्न मान्यता और कंप्यूटर विज्ञान अनुप्रयोगों के केंद्र में का प्रश्न है डेटा का प्रतिनिधित्व और पुनर्प्राप्ति कैसे करें। डेटा को इस तरह से प्रस्तुत करना जो इसकी कुंजी को उजागर करता है गुण पैटर्न पहचान और कंप्यूटर दृष्टि के प्रदर्शन को महत्वपूर्ण रूप से प्रभावित कर सकते हैं अनुप्रयोग। उच्च का प्रतिनिधित्व करने और पुनर्प्राप्त करने के लिए कई दृष्टिकोण प्रस्तावित किए गए हैं- आयामी डेटा जो निम्न-आयामी आंतरिक उप-स्थानों में रहते हैं। हाल ही में, वहाँ उच्च-आयामी डेटा से सूचनात्मक नमूनों का बेहतर चयन करने में बहुत शोध किया गया है संवेदन मैट्रिक्स के माध्यम से संपीड़ित संवेदन का उपयोग करना। संपीड़ित संवेदन हमें इसका शोषण करने की अनुमति देता है असंगत मापों को कैप्चर करने के लिए अंतर्निहित परिघटनाओं की विरल संरचना एक संवेदन मैट्रिक्स और फिर उच्च-आयामी डेटा को पुनर्प्राप्त करना। हालांकि, विरल संकेत संरचना संकुचित संवेदन का एकमात्र प्रमुख पहलू नहीं है; एक मजबूत असंगत फ्रेम संरचना माप के लिए उपयोग किया जाना भी उतना ही महत्वपूर्ण है। इस थीसिस में, हम उच्च-आयामी डेटा के पुनर्निर्माण की समस्या का समाधान करते हैं (जैसे छवियों) डेटा के लिए असंगत संवेदन मैट्रिक्स के माध्यम से माप का उपयोग करते हुए शोर की उपस्थिति में संपीड़न। हम एक समानता के आधार पर एक असंगत और मजबूत सेंसिंग मैट्रिक्स डिजाइन का प्रस्ताव करते हैं- फ्रेम वेक्टर और जकड़न के बीच कम जोड़ीदार सहसंबंध सुनिश्चित करने के लिए गूलर टाइट फ्रेम फ्रेम का। सिंथेटिक डेटा के साथ-साथ वास्तविक छवियों पर प्रयोग किए जाते हैं। प्रदर्शन करने वाला- आपसी तालमेल, सिग्नल पुनर्निर्माण सटीकता के संदर्भ में मंस मूल्यांकन किया जाता है, और पीक सिग्नल-टू-शोर अनुपात। इसके अलावा, यह देखा गया है कि उच्च आयामी डेटा सम-कंप्यूटर दृष्टि और पैटर्न पहचान कार्यों में व्यापक रूप से उपयोग किए जाने वाले कार्यों में अप्रासंगिक और हो सकते हैं अनावश्यक विशेषताएं। इसके परिणामस्वरूप कम्प्यूटेशनल जटिलता में वृद्धि हो सकती है, मेमोरी की आवश्यकता होती है- साथ ही ओवरफिटिंग की संभावना जो अंतर्निहित संरचना को बहुत कमजोर कर सकती है समग्र प्रदर्शन को कम करने वाली सुविधाओं के बीच संबंध। इस प्रकार, कॉम्पैक्ट कम ढूँढना- अच्छे सामान्यीकरण और भेदभावपूर्ण क्षमताओं के साथ आयामी प्रतिनिधित्व और विशेषताएं बहुत ज़रूरी है।

इस थीसिस में हम छवि वर्गीकरण के लिए फीचर निष्कर्षण मॉडल विकसित करने के लिए इन मुद्दों को संबोधित करते हैं। सबसे पहले, हम उच्च-आयामी डेटा की अंतर्निहित वैश्विक और स्थानीय संरचना की खोज करते हुए एक प्रतिनिधित्व-आधारित पर्यवेक्षित सुविधा निष्कर्षण मॉडल प्रस्तुत करते हैं। अगला, हम अपर्याप्त अंतर-वर्ग मार्जिन और डेटा की कई गुना संरचना के नुकसान की समस्या को हल करने के लिए एक प्रतिगमन-आधारित सुविधा निष्कर्षण मॉडल का प्रस्ताव करते हैं। निकाली गई सुविधाओं के सापेक्ष महत्व को सुनिश्चित करने के लिए, एक $\ell_{2,1}$ -मानक आधारित नियमितीकरण नियोजित है। चेहरे, दृश्य और वस्तु की पहचान के लिए विभिन्न डेटा सेटों पर प्रस्तावित मॉडलों के प्रदर्शन का मूल्यांकन किया जाता है। प्राप्त परिणामों की तुलना प्रस्तावित मॉडलों की प्रभावशीलता को प्रमाणित करने के लिए अत्याधुनिक विधियों से की जाती है।

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	i
ABSTRACT	ii
LIST OF TABLES	viii
LIST OF FIGURES	xi
ABBREVIATIONS	xii
NOTATION	xiv
1 Introduction	1
1.1 Overview of Compressive Sensing	1
1.1.1 Closer Look at Frame Theory	4
1.2 Overview of Feature Extraction	5
1.3 Main Hypothesis and Contributions	6
1.4 Summary of Contributions	7
1.5 Thesis Outline	7
2 Literature Review	9
2.1 Sensing Matrix Optimization	9
2.2 Representation-based Feature Extraction	11
2.3 Regression-based Feature Extraction	14
2.4 Research Gaps	17
2.5 Motivation and Objectives	18
2.6 Data Sets	19
3 Incoherent Sensing Matrix Design	23
3.1 Introduction	23

3.2	Problem Statement	25
3.2.1	Incoherent Frame Design	26
3.3	Experimental Results	28
3.4	Chapter Summary	31
4	Incoherent and Robust Sensing Matrix Design	32
4.1	Introduction	32
4.2	Proposed ETF Design Method	33
4.2.1	Structural Constraint	33
4.2.2	Spectral Constraint	34
4.3	Robust Sensing Matrix Design	36
4.3.1	Solving the Optimization Problem	38
4.3.2	Computational Complexity	40
4.4	Experimental Results	41
4.4.1	Performance Evaluation of IR-PMD Model on Synthetic Data . . .	41
4.4.2	Performance Evaluation on Real Data	50
4.5	Chapter Summary	54
5	Locality-Aware Subspace Learning	55
5.1	Introduction	55
5.2	Problem Formulation	56
5.2.1	Objectives	57
5.2.2	LADSL Model Formulation	59
5.2.3	Solving the Optimization Problem	60
5.3	Experimental Results	66
5.3.1	Classification Rate	68
5.3.2	Precision, Recall, and F1-score	72
5.3.3	Classification Performance using SVM Classifier	75
5.3.4	Pixel Corruption and Block Occlusion	76
5.3.5	Experimental Analysis	78
5.3.6	Remarks and Discussion	83
5.4	Chapter Summary	84

6	Relaxed Least Squares Regression	85
6.1	Introduction	85
6.2	Proposed LR-DRR Algorithm	87
6.2.1	Solving the Optimization Problem	88
6.3	Proposed TRLSR algorithm	90
6.3.1	Twin Relaxed Least Squares Regression	92
6.3.2	Optimization	94
6.4	Experimental Results of LR-DRR	97
6.5	Experimental Results of TRLSR	102
6.5.1	Classification Rate	104
6.5.2	Effectiveness of Twin Matrix	106
6.5.3	Significance of Structure Consistency	107
6.5.4	Image Recognition on Large Data Sets using Deep Features	108
6.6	TRLSR Model Analysis	110
6.6.1	Comparison with Deep Learning	110
6.6.2	Parameter Selection	113
6.6.3	Computational Complexity	115
6.6.4	Convergence Study	115
6.6.5	Remarks	116
6.7	Chapter Summary	117
7	Conclusion and Future Directions	118
7.1	Future Directions	120
	REFERENCES	122
	LIST OF PUBLICATIONS	133
	CURRICULUM VITAE	134

LIST OF TABLES

3.1	Performance evaluation of various CS systems ($M = 25$, $N = 80$, $L = 100$ AND $\text{SNR} = 20$ dB)	31
4.1	Coherence measures and MSE for various CS systems ($M = 20$, $N = 60$, $L = 120$).	47
4.2	Coherence measures and MSE for various CS systems ($M = 20$, $N = 60$, $L = 120$).	48
4.3	MSE for various CS systems ($M = 25$, $N = 60$, $L = 80$).	48
5.1	Mean classification rate (%) of LADSL method on PubFig83 data set	68
5.2	Mean classification rate (%) of LADSL method for ORL data set	69
5.3	Mean classification rate (%) of LADSL method for Ex Yale B data set	69
5.4	Mean classification rate (%) using deep features for Caltech 101 data set . .	70
5.5	Mean classification rate (%) of LADSL method for AR data set	71
5.6	Mean classification rate (%) of LADSL method for COIL-100 data set	71
5.7	Mean classification rate (%) of LADSL method for CMU PIE data set	72
5.8	Precision, recall, and F1-score for ORL data set	73
5.9	Precision, recall, and F1-score for AR data set	73
5.10	Precision, Recall, and F1-score for Ex Yale B data set	74
5.11	Precision, Recall, and F1-score for CMU PIE data set	75
5.12	Mean classification rate (%) for Ex Yale B data set using SVM classifier . .	75
5.13	Mean classification rate (%) for ORL data set using SVM classifier	76
5.14	Mean classification rate (%) for AR data set using SVM classifier	76
5.15	Ablation study of the LADSL model on different data sets	78
6.1	Mean classification rate (%) of LR-DRR method for Ex Yale B data set. . .	97
6.2	Mean classification rate (%) of LR-DRR method for Caltech-101 data set .	98
6.3	Comparison of mean classification rates (in %) of LR-DRR method.	98
6.4	Mean classification rate (%) of LR-DRR method for COIL-20 dataset	99
6.5	Mean classification rate (%) of TRLSR method for Ex Yale B data set	103

6.6	Mean classification rate (%) of TRLSR method for CMU PIE data set . . .	103
6.7	Mean classification rate (%) using TRLSR method for 15-Scene data set . .	104
6.8	Comparison of mean classification rates (in %) of TRLSR method for PubFig83 data set	104
6.9	Mean classification rate (%) of TRLSR method for AR data set.	105
6.10	Mean classification rate (%) of TRLSR method for Caltech-101 data set. . .	105
6.11	Mean classification rate (%) of TRLSR method for large scale data sets. . .	109
6.12	Mean classification rate (%) using deep features for Caltech-256 data set. .	109
6.13	Mean classification rate (%) using deep features for Caltech-101 data set. .	110
6.14	Mean classification rate (%) of deep learning methods on Ex Yale B data set.	113

LIST OF FIGURES

1.1	Illustration of thesis organization.	8
2.1	Sample images from the data sets used in this work for face recognition. . .	20
2.2	Some of the standard images used for image reconstruction.	20
2.3	Sample images from the COIL-20, PubFig83, Caltech-101, and 15-Scene data sets used in this work for image classification.	21
3.1	Histogram of the absolute off-diagonal values of the optimized Gram matrix ($N = 80$, $L = 100$, and $M = 25$).	29
3.2	Evolution of coherence parameters for $N = 80$, $L = 100$, $M = 25$, and $t = \mu_W = 0.4$, for 100 trials. Note that for cumulative coherence, m is taken as sparsity s (refer (1.4) in Chapter 1).	29
3.3	Recovery performance error e_r along varying sparsity s and varying number of measurements M for $N = 80$, $L = 100$, $M = 25$ with optimized sensing matrix	30
4.1	Flowchart of the proposed IR-PMD model.	40
4.2	Convergence behavior of IFD and IR-PMD algorithms ($M = 20$, $L = 120$, $N = 60$, and SNR = 10 dB).	43
4.3	Performance of IR-PMD model for synthetic data ($N = 60$, $M = 20$, $L = 120$, and $s = 4$)	44
4.4	Performance of IR-PMD model for Caltech-101 data set.	44
4.5	Evolution of averaged mutual coherence over 1000 trials ($M = 20$, $L = 120$, $s = 4$, and SNR = 10 dB).	45
4.6	Averaged mutual coherence as a function of M for (a) frame $\mathbf{D}$ and (b) projection matrix Φ ($L = 120$, SNR = 10 dB).	45
4.7	Evolution of averaged mutual coherence with the number of iterations ($L = 120$, SNR = 10 dB).	46
4.8	Histogram of the absolute off-diagonal elements of the optimized Gram matrix ($N = 60$, $L = 120$, and $M = 20$).	47
4.9	Reconstruction error (NMSE, e_r) as a function of sparsity s (SNR = 10 dB).	48
4.10	Recovery performance of IR-PMD model in terms of NMSE (e_r) as a function of the number of measurements M	49

4.11	Recovery performance error NMSE (e_r) as a function of the number of dictionary atoms L	50
4.12	MSE and PSNR as a function of SNR for optimized projections ($N = 64$, $L = 100$, and $M = 20$).	51
4.13	MSE and PSNR as a function of SNR for optimized projections ($N = 64$, $L = 100$, and $M = 20$) using Caltech 101 data set.	51
4.14	Reconstruction performance on the standard <i>Cameraman</i> , <i>Pirate</i> , and <i>Elaine</i> images with $M = 80$, $s = 16$, and $N = 256$	52
4.15	Comparison of the proposed IR-PMD method with the $\ell_{2,1}$ regularization and Frobenius norm regularization (CS_{WF}).	52
5.1	Sample corrupted images from the Ex Yale B data set.	77
5.2	Mean classification rate (%) for Ex Yale B data set with varying levels of data corruption.	77
5.3	Mean classification rate of LADSL as a function of the parameter values (α , β , λ , and τ) for the CMU PIE, PubFig83, and ORL data sets.	79
5.4	Mean classification rate of LADSL as a function of parameter values.	81
5.5	Mean classification rate of LADSL as a function of the projected subspace dimension (d).	82
5.6	Objective function value and classification rate of LADSL as a function of the number of iterations.	83
5.7	Residual of the LADSL method as a function of the number of iterations on different data sets.	83
6.1	Flowchart of the proposed TRLSR model.	94
6.2	t-SNE visualization of the features extracted on the Extended Yale B data set, where 8 random samples from each subject are used for validation.	100
6.3	Visualization of the matrix $\mathbf{V}$ in the LR-DRR problem formulation.	100
6.4	Classification rate of the LR-DRR method as a function of the parameters α , β and τ on the Extended Yale B data set ($t_r = 10$).	100
6.5	Classification rate of the proposed LR-DRR method as a function of the parameters α , β , γ , and τ on different data sets.	101
6.6	Evolution of objective function value and classification rate with number of iterations for the LR-DRR method.	102
6.7	Comparison of the proposed method with and without the twin matrix on different data sets.	106
6.8	Visualization of the matrix $\mathbf{V}$ for different data sets.	107
6.9	Mean classification rate using deep features for the ImageNet data set.	111

6.10	Mean classification rate (%) of the proposed TRLSR method as a function of the weight parameter γ and τ on different data sets	112
6.11	Mean classification rate (%) of the proposed TRLSR method as a function of the weight parameter γ on different data sets.	113
6.12	Mean classification rate (%) of the proposed TRLSR method as a function of the parameters α and τ on different data sets.	114
6.13	Evolution of the objective function value and classification rate with the number of iterations for the proposed TRLSR method.	114
6.14	t-SNE visualization of the original samples and transformed features obtained using TRLSR method for Ex Yale B ($t_r = 10$, validation = 8 samples/person), CMU PIE ($t_r = 15$, validation = 10 samples/person), and PubFig83 ($t_r = 15$, validation = 8 samples/person) data sets.	116
7.1	Summary of the research contributions and publications from this thesis. . .	119

ABBREVIATIONS

CS	Compressed Sensing
SEE	Sparse Encoding Error
OMP	Orthogonal Matching Pursuit
MP	Matching Pursuit
WB	Welch Bound
RIP	Restricted Isometry Property
RIC	Restricted Isometry Constant
ETF	Equiangular Tight Frame
UNTF	Unit Norm Tight Frame
PCA	Principal Component Analysis
LDA	Linear Discriminant Analysis
NPE	Neighbour Preserving Embedding
LPP	Locality Preserving Projection
SSS	Small Sample Size
OLDA	Orthogonal Linear Discriminant Analysis
MPDA	Manifold Partition Discriminant Analysis
SRC	Sparse Representation-based Classification
CRC	Collaborative Representation-based Classification
LLC	Locality-constrained Linear Coding
LRR	Low-rank Representation
LatLRR	Latent Low-rank Representation
LRE	Low-rank Embedding
RsLDA	Robust Sparse Linear Discriminant Analysis
SALPL	supervised approximate low-rank projection learning
LLRSE	Latent Low-rank and Sparse Embedding
CDPL	Constrained Discriminative Projection Learning
LR	Linear Regression

LSR	Least square regression
DLSR	Discriminative Least Squares Regression
ReLSR	Retargeted Least Squares Regression
GReLSR	Groupwise Retargeted Least Squares Regression
RLSL	Robust Latent Subspace Learning
HOG	Histogram of Oriented Gradients
LBP	Local Binary pattern
ISMD	Incoherent Sensing Matrix Design
UNF	Unit Norm Frame
IFD	Incoherent Frame Design
IR-PMD	Incoherent and Robust Projection Matrix Design
SVD	Singular Value Decomposition
SNR	Signal-to-noise Ratio
NMSE	Normalized Mean Square Error
MSE	Mean Square Error
PSNR	Peak Signal-to-noise Ratio
LADSL	Locality-aware Discriminative Subspace Learning
ADMM	Alternating Direction Method of Multipliers
K-NN	K-Nearest Neighbour
SVM	Support Vector Machine
LRSI	Low-rank Model with Structure Incoherence
LRRR	Low-rank Ridge Regression
SLRR	Sparse low-rank Regression
LRSI	Low-rank Linear Regression
LR-DRR	Low rank Based Double Relaxed Regression
TRLRSR	Twin Relaxed Least Squares Regression
RLSR	Relaxed Least Squares Regression

NOTATION

x	scalar
$\mathbf{x}$	vector
$\mathbf{A}$	Matrix
$\mathbf{A}^T$	Matrix Transpose
$\mathbf{A}^{-1}$	Matrix Inverse
$\text{tr}(\mathbf{A})$	trace of matrix $\mathbf{A}$
$\mathbf{A}^*$	Conjugate transpose of $\mathbf{A}$
$\mathbf{a}_i$ ($\mathbf{a}^i$)	i^{th} column (row) of matrix $\mathbf{A}$
$\mathbf{A}_{i,j}$	Element at the intersection of i^{th} row and j^{th} column of matrix $\mathbf{A}$
$\mathbf{A} \odot \mathbf{B}$	Hadamard product of two matrices $\mathbf{A}$ and $\mathbf{B}$
$\ \mathbf{A}\ _F$	$\sqrt{\sum_j \ \mathbf{a}_j\ _2^2}$ (Frobenius norm)
$\ \mathbf{A}\ _{2,1}$	$\sum_i \sqrt{\sum_j \mathbf{A}_{i,j} ^2}$ ($\ell_{2,1}$ -norm)
$[N]$	set $\{1, 2, \dots, N\}$
$\text{card}(\mathcal{P})$	Cardinality of the set $\mathcal{P}$
$\mathbf{A}_{\mathcal{P}}$	Sub-matrix of $\mathbf{A}$ containing the columns indexed by $\mathcal{P}$.
$\text{supp}(\mathbf{x})$	Support of $\mathbf{x} \in \mathbb{R}^N$, i.e. $\text{supp}(\mathbf{x}) = \{j \in [N] : x_j \neq 0\}$ the index set of its non-zero elements.
$\text{diag}[\mathbf{x}]$	Diagonal matrix with the vector $\mathbf{x}$ as its diagonal elements