EXPLORING GEOMETRICAL STRUCTURES IN IMAGE AND VIDEO DATA FOR CLASSIFICATION: FROM SUBSPACES TO MATRIX MANIFOLDS

A Thesis

submitted by

KRISHAN SHARMA

in partial fulfillment for the degree of

DOCTOR OF PHILOSOPHY

School of Computing and Electrical Engineering

Indian Institute of Technology Mandi

India - 175005

September, 2020

DECLARATION

I hereby certify that the work which is being presented in the thesis titled **"Exploring Geometrical Structures in Image and Video Data for Classification: from Subspaces to Matrix Manifolds"** in the requirement for the award of the degree of **DOCTOR OF PHILOSOPHY**, and submitted in the School of Computing and Electrical Engineering, Indian Institute of Technology Mandi, is an authentic record of my own work carried with the guidance of Dr. Renu Rameshan, Indian Institute of Technology Mandi, India. Due acknowledgment has been made wherever the work described is based on the findings of the other investigators or is done in collaboration with other researchers. The work presented in this thesis has not been submitted by me for the award of any other degree of this or any other institute.

Krishan

Mandi, 175 005 **Krishan Sharma**

THESIS CERTIFICATE

This is to certify that the thesis titled "**Exploring Geometrical Structures in Image and Video Data for Classification: from Subspaces to Matrix Manifolds**", submitted by **Krishan Sharma**, to the **Indian Institute of Technology Mandi**, for the award of the degree of **Doctor of Philosophy**, is a bonafide record of the research work done by him under my guidance. 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.

Mandi, 175 005

Dr. Renu Rameshan
(Advisor)

Acknowledgments

During the period of research we get inspired by many people and feel grateful towards them. This page is dedicated to all such unforgettable contributions in my research and in my life.

To start with, I thank profoundly to my thesis supervisor Dr. Renu Rameshan. I cannot thank her enough, certainly not in a few lines here. I believe that the deep respect for her in my heart is no secret to her. I owe her everything that I have learned in the last five years, be it developing an aptitude of thinking like a research scholar or critically reading and analyzing. I had shortcomings; still she showed her confidence in my abilities. I learned how to keep working with patience and perseverance in face of failures. Most importantly I learned that there are no shortcuts on the way that leads to success, one must work hard. And in this journey I also evolved as a better human being. Now only I understood what the true definition of a Guru is.

Further, I would like to thank my Doctoral Committee members Dr. Dileep A.D., Dr. Arnav Bhavasar, and Dr. Prasanth P. Jose for their constant support and valuable remarks during the DC meetings. I would also like to thank the SCEE chair professor for being a part of the DC Committee. The office staff of the School of Computing and Electrical Engineering, IIT Mandi has been a great help. My sincere thanks to them as well.

I am thankful to MANAS Lab, SCEE, IIT Mandi for providing me workspace and all necessary research related facilities. For financial support, I am thankful to the Ministry of Human Resource and Development (MHRD), Govt. of India for providing me a research fellowship. I thank Finance dept. of IIT Mandi for timely release of fellowships. I also would like to acknowledge the open source softwares LATEX and Ubuntu which made this thesis possible in a quick time.

I owe sincere thanks to my colleagues Jyoti Nigam, Rahul Mishra, Kuldeep, Mohana and Dr. Amit Unde for valuable discussions. I would specially like to thank Ms. Jyoti Nigam for hosting wonderful lunches and dinners during my stay on the campus. Now its time to mention some wonderful friends who have a big contribution in my thesis.

During my journey as a research scholar I happened to meet few wonderful friends. They stood with me in my difficult times and lent me their unconditional support in various capacities. Shivam, thank you for the moral support you gave me during my hostel days. You were the only and therefore by default the best room mate I have ever lived with. Thank you for helping me in improving my language skills and motivating me through your high level of patience.

Shikha, with your excellent choice of colors for diagrams you have made the illustrations in my thesis all the more attractive. Also, I enjoyed insightful sessions with you, when we together discussed technical ideas and tried to find solutions of challenging problems. Thank you very much for your inputs as a friend and as a research scholar.

Munender, Rohit, Ajay, Karan, Shivam Gujral, Dauood Saleem and Sujeet, it was always relaxing talking to you, especially after hours of intense work. Your experiences kept me motivated and the leisure time activities with you kept me in good mood. Thank you for precious company.

The list goes on and it is difficult to mention each and every name here, for whom I feel a sense of gratitude. My apologies to those whose names could not be mentioned here.

Last but not the least I wish to thank my parents and my dear wife Kajal. Their love, understanding and support has been the strongest source of energy which kept me going, despite challenges. They never complained about my absence or anything else, which gave me a great deal of mental peace to focus on my work. I have been away from them for last couple of years. Now, after completion of my work, I am happy that I can live and spend quality time with them.

- *Krishan*

Abstract

One of the important tasks in computer vision and machine learning is data classification, which needs data to be represented as feature vectors that are compact, informative as well as discriminative. This thesis makes an attempt to generate such representations using data geometry. The geometry being referred to here is either subspace or manifold geometry.

Two types of data are being considered; those which can be modeled as points in a union of subspaces and those which lie on a matrix manifold. For the former, kernel based and autoencoder based transformation learning approaches are described. The methods are based on the fact that real data models usually deviate by a large extent from an ideal union of subspace structure; thereby resulting in non-linearities. These non-linearities are handled by the two above-mentioned approaches.

In manifold geometry, distance measures existing over the manifolds are useful for measuring similarity between two points. Same class data points tend to be nearby points in manifolds while the points from different classes would be far away. This thesis uses distance functions over matrix manifolds to come up with distance based positive definite kernels defined on a manifold space.

For all the kernel based approaches, a discriminative representation corresponding to each data point is obtained by diagonalizing the kernel-gram matrix while for the autoencoder based approach, representation is directly obtained as the encoder output. These representations are the outcome of the assumed underlying geometrical models. In order to test the performance of the models designed for image classification task, images are considered as points in a union of subspaces. Similarly, image sets and videos are modeled as points on matrix manifolds for their respective classification tasks. While affine Grassmann manifold (AGM) geometry is considered for image sets, two geometrical models are proposed for videos - product Grassmann manifold (PGM) and product manifold of symmetric positive semi-definite (SPSD) matrices. The obtained results show the significance of these geometries in data classification.

Contents