视觉作为人类的主要的感知机能之一，对人类感知世界的重要性不言而喻。计算机视觉的任务就是为计算机赋予接近甚至超过人类视觉的感知能力。图像作为计算机视觉任务的主要输入，与其它数据形式（如文本，语音等）相比蕴含了更多的信息。

另一方面尽管图像本身蕴含了丰富的信息但是如何运用这些信息，以及图像本身的一些问题（如视角变化大、光照变化剧烈、分辨率低等）也给视觉任务带来不小的挑战。与此同时，越来越多现实生活中的数据以集合的形式出现：视频监控数据、用户上传视频、主题相册、物体的多视角数据以及动作描述视频等在近年来都呈现出爆发式的增长；图像集合分类问题也在这样的背景下应运而生，针对集合数据呈现出的量大但质未必优的特点，图像集合分类问题的核心任务之一便是利用数据量大的特点以克服质低的问题。经过10多年的发展，根据图像集合的表示方式的不同，图像集合分类相关方法逐渐形成了以下的一些类别：1、子空间以及流形建模的方法；2、仿射包建模的方法；3、统计建模的方法；4、深度学习的方法；5、其它（稀疏编码，协同表示等）。

在众多方法中，统计建模的方法以其优越表现逐渐成为图像集合分类问题的主要方法之一，本文将以黎曼流形为工具对统计建模图像集合问题进行研究分析。本文的主要工作包含：（1）研究了矩阵函数与流形上的优化理论与方法，在对流形、矩阵函数等概念进行介绍的基础上，针对矩阵流形上的优化问题进行讨论与探究，并结合学位论文课题中提炼出的相关实例对矩阵流形优化进行介绍，一方面帮助读者理解并复现本文所提出的方法和结论，另一方面为解决类似流形优化问题提供借鉴。（2）提出了黎曼流形上的偏最小二乘回归方法，通过在流形上单一切空间构建子流形的方式将欧氏空间中的偏最小二乘回归（Partial Least Square Regression, PLS）扩展到黎曼流形空间；考虑到黎曼流形与欧氏空间的几何结构差别以及图像集合数据稀疏的问题，进一步设计了流形上多切空间构建子流形的方法，采用逐步回归的策略整合多个切空间中的结果；本文以非奇异协方差矩阵即对称正定矩阵（Symmetric Positive Definite, SPD）黎曼流形为实例，在集合数据分类问题上进行了实验验证，取得了与当前最优方法可比甚至更好的结果。（3）提出了低秩对称半正定矩阵（Low-Rank symmetric Positive Semi-Definite, PSD）建模图像集合的方法，解决采用样本协方差矩阵建模图像集合时由于数据稀疏带来的矩阵奇异（不满秩）、由于噪声带来的矩阵估计不准、以及对称正定矩阵表示时空开销大等问题，并采用图嵌入（Graph Embedding）的方法将判别信息内嵌到的低秩对称半正定矩阵表示中，最后在核判别分析（Kernel Discriminant Analysis, KDA）的框架下研究了该表示下的判别学习问题，最终实验验证了低秩对称半正定矩阵表示的有效性。

Vision functionality serves as one of the main abilities for human to perception the real world, and its importance goes without saying. The mission of CV (Computer Vision) is to endow computers with close to or even stronger ability than human to perceive the real world.

As the main input for CV tasks, images contain much more information than text, audio and so on, but how to make full use of the information becomes a problem. The variations of images bring great challenges to CV tasks. At the same time, data come more frequently in the form of image set, such as surveillance video, multi-view image sets and so on. Under these background, image set classification comes into being. Image sets usually contain a large amount of images in poor quality. So one major task in image set classification is to overcome the disadvantage of low quality and leverage the advantage of large quantity.

With more than ten years of development, a lot of methods have been proposed for this task. According to how to model an image set they can be divided into following categories: 1. Subspace/Manifold based methods, 2. Affine hull based methods, 3. Statistics model based methods, 4. Deep Learning based methods. 5. Others, like Dictionary/Sparse coding based method, Collaborative representation methods, etc.

Among the categories listed above, Statistics model based methods have attracted a lot attention with its excellent performance. This thesis takes Riemannian manifold as basic tool and tries to explore statistics model based methods. The main contributions include: 1) Studied matrix function and manifold optimization theory and methods. By introducing the basic concept of manifold and matrix function, along with the real-world problems extracted from the following research topics, optimization algorithms on the manifold have been studied in this thesis (Chapter 2). On the one hand it will help readers understand and implement methods proposed in this thesis, and on the other hand it can also provide basic instructions to solve other similar problems. 2) Proposed Partial Least Square Regression methods on Riemannian manifold with sub-manifold constructed from one tangent space (usually taking the tangent space of samples’ Karcher mean). Then in order to overcome the structure difference between Euclidean space and Riemannian manifold as well as the drawback of sparse sampling, multi-tangent space Partial Least Square Regression method has been designed. On the Symmetric Positive Definite (SPD) matrices manifold, image set classification experiment were designed to evaluate the proposed method and it is observed that the proposed method is comparable or even outperforms the state-of-art methods on the commonly used databases. 3) Proposed Low-Rank PSD matrices based image set model to overcome the rank-deficient and high dimension problems of sample covariance models as well as lack of scale information (eigenvalue) drawback in the subspace models. With Graph Embedding framework we encoded label information into Low-Rank PSD representations of image sets and then designed the discriminant learning methods with Kernel Discriminant Analysis framework. Experiments on the commonly used databases has shown to support our proposition.