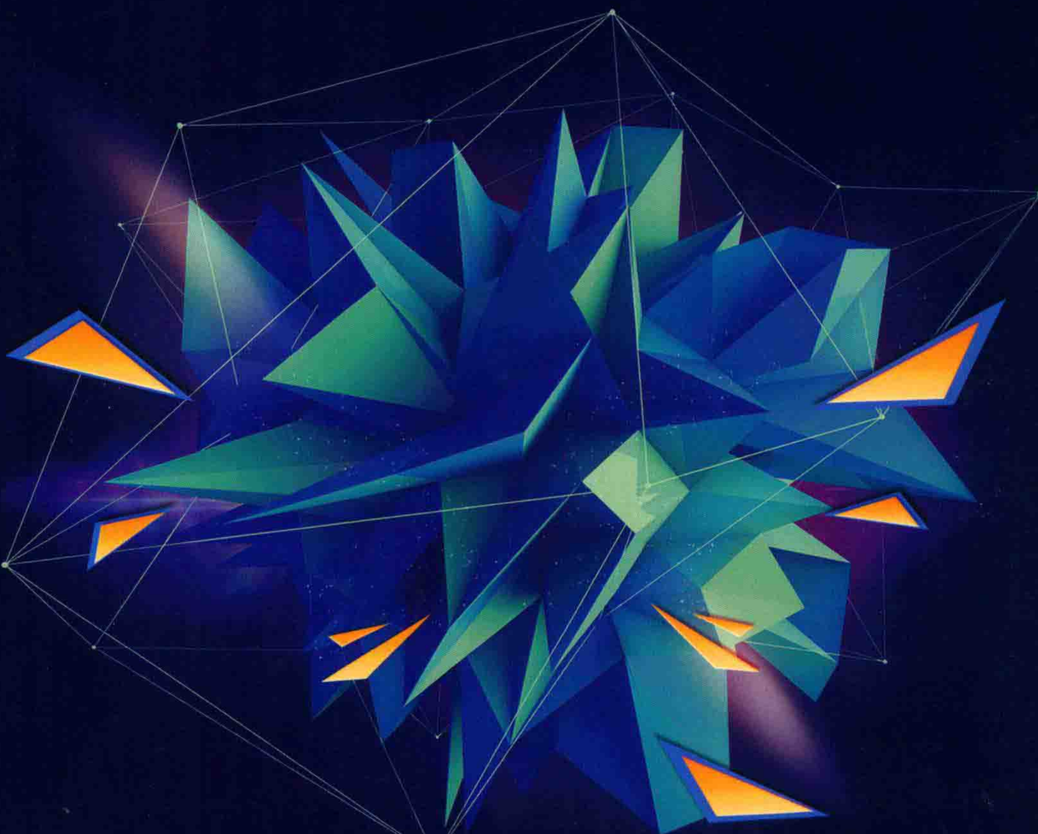
张量分析在计算机视觉中的应用研究

ZHANGLIANG FENXI ZAI JISUANJI SHIJUE
ZHONG DE YINGYONG YANJIU

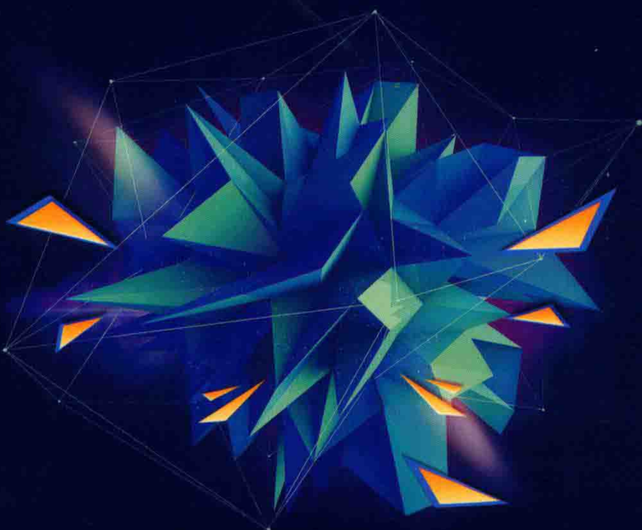
刘昶 著



四川大学出版社

张量分析在计算机视觉中的应用研究

ZHANGLIANG FENXI
ZAI JISUANJI SHIJUE ZHONG DE
YINGYONG YANJIU



ISBN 978-7-5690-3785-2

9 787569 037852 >

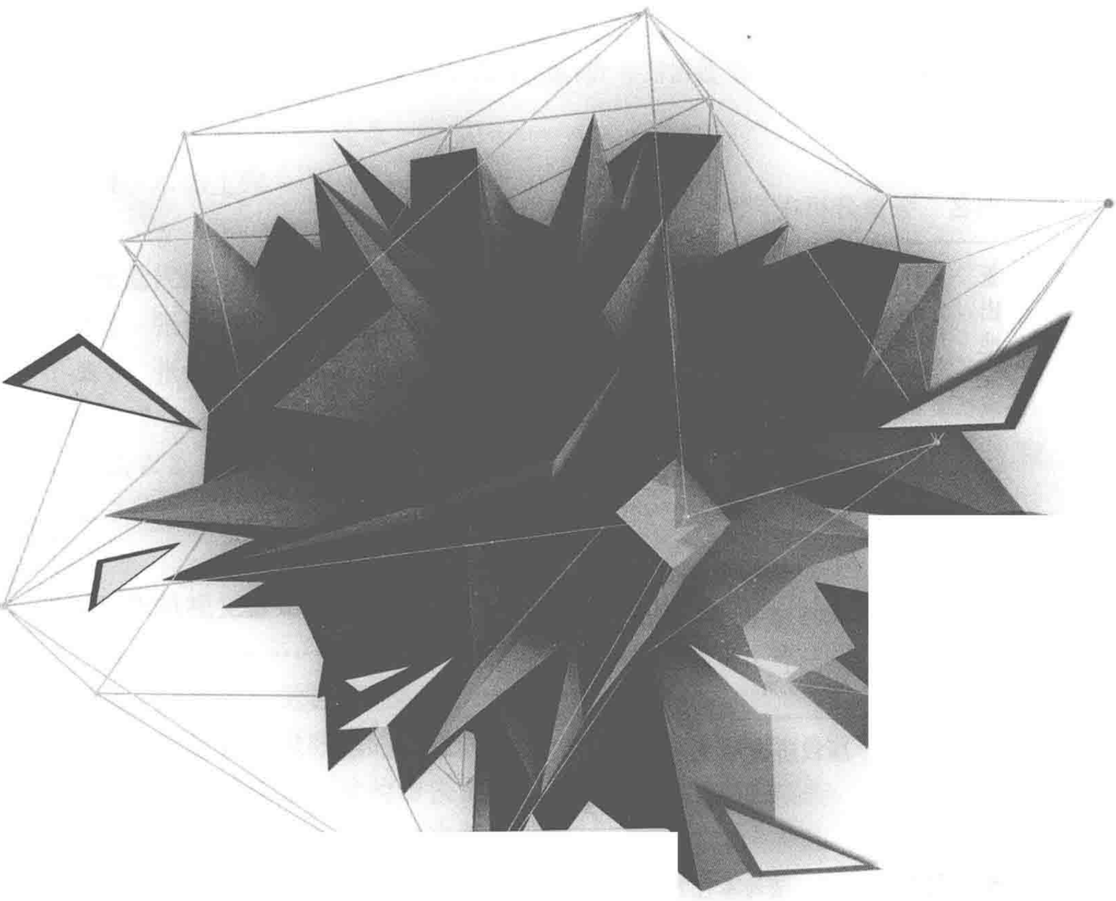
定价：49.00元

然科学基金青年基金项目(61502059)
科技厅应用基础面上项目(2018JY0272)
士后科学基金项目（2016M592655）
资助

张量分析在计算机视觉中的应用研究

ZHANGLIANG FENXI ZAI JISUANJI SHIJUE ZHONG DE YINGYONG YANJIU

刘昶　著



四川大学出版社

项目策划：梁　平
责任编辑：梁　平
责任校对：傅　奕
封面设计：璞信文化
责任印制：王　炜

图书在版编目（CIP）数据

张量分析在计算机视觉中的应用研究 / 刘昶著．—
成都 ： 四川大学出版社， 2020.9
ISBN 978-7-5690-3785-2

Ⅰ．①张…　Ⅱ．①刘…　Ⅲ．①张量分析－应用－计算机视觉　Ⅳ．① TP302.7

中国版本图书馆 CIP 数据核字（2020）第 113958 号

书名　张量分析在计算机视觉中的应用研究

著　　者	刘　昶
出　　版	四川大学出版社
地　　址	成都市一环路南一段 24 号（610065）
发　　行	四川大学出版社
书　　号	ISBN 978-7-5690-3785-2
印前制作	四川胜翔数码印务设计有限公司
印　　刷	四川盛图彩色印刷有限公司
成品尺寸	170mm×240mm
印　　张	9
字　　数	171 千字
版　　次	2021 年 1 月第 1 版
印　　次	2021 年 1 月第 1 次印刷
定　　价	49.00 元

◆ 读者邮购本书，请与本社发行科联系。
电话：(028)85408408/(028)85401670/
(028)86408023　邮政编码：610065
◆ 本社图书如有印装质量问题，请寄回出版社调换。
◆ 网址：http://press.scu.edu.cn

四川大学出版社
微信公众号

目　录

1　概论

1.1　引言

随着计算机计算能力的日益增强和存储容量的增长，数据量呈几何级数的增长。尤其是在近年来，高维结构化数据呈现明显增长势头，如图像、视频等数据。大规模高维数据的获取和使用带来了很多新的问题：一方面数据的计算、存储、传输给计算机硬件和网络造成极大的压力，甚至可能造成“维数灾难”，为人类社会带来了负面影响和潜在危机；另一方面是不能被人的感知直接理解及发现数据集内在的规律。

目前数据处理可以分为两大类。一方面不断发展硬件技术，缓解当前数据增长对硬件造成的压力。设计更强大的计算结构，如分布式计算、多核处理器和超级计算机等。但不断增长的数据量和数据复杂度似乎超出了计算能力的增长。另一方面利用原始数据常常包含很多冗余信息甚至是干扰信息，从数据本身出发，借助数据分析方法，对高维数据集采用某种维数约简方法进行降维处理，在降低数据复杂度的同时，保持数据的有用信息。

数据分析方法的通用性，使得数据降维具有跨学科的特性。目前，各个领域的研究者都关注于数据降维，如图像处理、数据挖掘、机器学习、数据压缩、计算机视觉、人工智能、模式识别，甚至神经科学等。这些传统数据分析方法常常要求将数据进行向量化，但是对于本身就是多维结构的数据而言，这种向量化操作必然会破坏数据的空间结构，导致局部结构信息的丢失，并且向量化操作导致数据维数的急剧增加，极易引起“维数灾难”。数据降维的最终目的是探究高维数据中的内在规律，同时有效避免“维数灾难”问题。传统的

统计数据分析将具有特定现象的实例作为观测样本。随着社会信息化的高速发展，不仅能获得越来越多的观测样本，并且一个样本由越来越多的变量组成，从而导致一个独立观测可能有成千上万甚至过亿的维数。经典的数据分析方法并不能很好地处理这种维数急剧增长的观测向量，因此，高维数据分析将变得越来越重要。对于高维数据的处理，仅仅通过调整数据结构使之适应现有的模型已经不可能实现对数据的分析和处理，而是需要通过对数据进行分析，挖掘数据的本征结构，设计表征数据的内在模型。高维数据分析将成为未来发展的热点，并将得到广泛的应用。

作为一门新的数学工具，张量代数被引入视觉数据分析。张量代数也称为多线性代数，是矩阵代数的高维扩展。2002 年，Vasilescu 首次将张量代数用于人脸分析，自此，研究者开始重视张量代数。然而，张量代数在计算机视觉中的应用时间尚短，基于张量代数的数据分析算法还需要进一步的完善和改进。

1.2　国内外研究现状

在计算机视觉中，如何找到一种保留数据的隐含结构的降维算法是一件亟待解决的难题。传统数据分析方法主要包括线性分析方法和非线性分析方法。线性分析方法计算简单，效率较高，但是不能用于挖掘隐藏在数据中的非线性结构。为了克服这一缺点而采用了非线性分析方法，但是其计算代价较高。传统数据分析方法都是基于向量空间，破坏了高维结构化数据的空间信息，导致了数据维数较高。随着张量代数的引入，出现了一类新的数据分析方法，称为张量分析。张量分析方法在保留高维数据结构信息的同时，一定程度上避免了“维数灾难”。

1.2.1　向量分析

1.2.1.1　线性分析方法

传统的线性降维方法常常假设输入空间可表示为特征子空间的线性组合，常用的线性降维方法包括主分量分析、线性判别分析、非负矩阵分解等。

1. 主分量分析

Jolliffe（2002）提出了主分量分析（Principal Component Analysis，PCA），运用统计分析的方法分析训练样本协方差矩阵，从数据能量的角度出发挖掘表征训练样本集的特征向量，丢弃最小特征向量，实现了数据降维。PCA 沿着数据能量的最大变化方向找到正交分量，并将高维向量压缩到低维空间，能实现重构，并且可以直接从数据中计算得到模型参数。

PCA 挖掘的特征向量是通过对所有训练样本进行分析计算，然而在实际应用中，如时间序列预测和模式识别，不可能预先得到所有的训练样本。为了解决新增训练数据特征向量的挖掘，必须丢弃先前获得的模型参数，将原始样本连同新样本重新训练，具有较高的时间和空间复杂度，这在一定程度上约束了 PCA 的实用性。为了解决该问题，提出了增量 PCA 算法，这些增量算法基本上分为两类。一类是从原始训练图像和新加入样本中重构重要的主分量（PCs）。当增加新的训练样本时，子空间的维数相应增加。由于权重向量和重构图像只是坐标系不同，所以可以利用原始图像的低维权重向量获得更新后的特征空间。由于特征空间的维数很小，因此该方法的计算效率较高。丢弃最不显著的 PC 分量，从而保持子空间的维数不变，但是这种方法常常存在拟合误差。为了降低拟合误差，Hall（2000）提出逐个加入新样本点。但是这种方式导致计算效率降低。Ozawa（2006）通过融合和分裂特征空间模型改进该方法，允许每次加入块数据，提出了基于数据块的增量主分量分析。已有文献证明，当数据矩阵满足 low-rank-plus-shift 结构时，即数据 $\boldsymbol{A}$ 满足$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$＝低秩矩阵＋$\alpha$×单位矩阵，采用奇异值分解（Singular Value Decomposition，SVD）更新算法和采用批处理方法都能获得相同的低秩拟合结果。但是，实际上，由于插入新样本时，数据矩阵会改变，从而改变矩阵的秩，因此很难证明数据矩阵具有 low-rank-plus-shift 结构。为此，基于奇异值分解更新算法，Zhao（2006）推导出增量算法和批处理模式之间的误差边界。在原始输入维数较高和样本点数量较大的情况下，分析协方差矩阵需要较大的存储空间，且计算代价较高。为此，Weng（2003）提出另一类算法——协方差无关算法，该算法不需要分析协方差矩阵，但会遭遇收敛问题。

PCA 对离群点敏感。然而，在计算机视觉中，由于噪声、光照、配准误差或遮蔽等因素，都有可能造成离群点的出现，因此研究者提出了各种改进方法。但是这些改进方法假设训练数据是理想的，并且预知视觉模型。在实际应用中，很难满足这些假设条件。因此，研究者对鲁棒学习进行了研究。然而，由于初始训练阶段建立的模型并不可靠，因此很难实现鲁棒学习。同时，由于

无法预知样本的离群点，这些算法都采用迭代算法，处理时间较长，导致计算效率降低。针对这些问题，Y. Li（2004）提出在增量 PCA 算法的每步中检测离群点，无需迭代计算，计算效率提高。

当数据维数较高或训练样本数较大时，协方差矩阵的维数很高，也导致了较高的计算代价。传统 PCA 采用奇异值分解（SVD）计算主分量，该算法将对称矩阵进行对角化，计算量较大。Ahn（2003）提出了计算效率比 EVD 略高的 EM 算法，由于在每次迭代中 E 步骤和 M 步骤都要求矩阵的逆，该算法不能收敛到全局最大，只能得到局部最大。Sharma（2007）提出了一种快速 PCA 算法，不需要任何初始设定，也不需要对角化任何对称矩阵和求解矩阵的逆，而是使用定点法计算主分量，并且实现在极少迭代次数内收敛。

PCA 尽可能地保留了数据的信息量，适用于数据压缩、低阶统计量的解等相关领域。

2. 线性判别分析

尽管 PCA 尽可能地保留了数据信息，但 PCA 计算得到的主分量没有包含类别信息。为了克服这个问题，Fisher（1936）提出了一种结合类别信息的类间分离的线性投影方法，解决了两类分类问题，该方法称为 Fisher 线性判别（Fisher Linear Discriminant，FLD）。其目标在于寻找一组最佳判别矢量使得各类的类内离散度最小，同时使类间离散度最大。Rao（1948）将 FLD 推广至多类问题分析，即线性判别分析（Linear Discriminant Analysis，LDA）。由于 LDA 具有较强的区分能力，因此在机器学习、数据挖掘、信息检索中具有广泛的应用。

当训练样本有限时，如人脸识别、图像检索，样本维数常常大于或接近样本个数，则类内离散度矩阵奇异或是病态的，传统的 LDA 方法很难直接计算或计算结果不稳定，这就是所谓的“小样本（Small Size Sample，SSS）问题”，简称 3S 问题。为了解决小样本问题，常常采用子空间方法，即将原始样本投影到更低维空间，从而保证类内离散度矩阵是满秩的。首先采用 PCA 选取主分量，使得类内离散度矩阵满秩。然后采用标准 LDA 方法。选取的主分量越多，保留的信息量越多，因此数据重建的效果越好，但 PCA 中较小特征向量对应高频信息（可能为噪声），如果将这些特征向量也作为主分量用于构造 PCA 子空间，在判别分析时由于存在噪声则可能导致过拟合现象。标准的 FLD 过程需要同时对类内和类间离散度矩阵进行对角化。首先对类内离散度矩阵进行白化变换，目的是进行归一化处理，然后对变换后数据的类间离散度矩阵进行 PCA 变换。由于在白化变换中，类内离散度矩阵过小的特征值会使

得白化变换对错误的方差进行拟合，从而使得其对新数据的推广能力变差。针对该问题，Liu（2000）提出了增强的 FLD 模型（Enhanced FLD Model，EFM），充分考虑了过拟合问题，从而增强了分类能力。此外，当训练样本数较小时，LDA 类内离散度矩阵是奇异的，该问题也被称为小样本问题，即 3S 问题。Liu（1993）修改了传统的 Fisher 判别函数，提出使用总体离散度矩阵代替原始判别函数的类内离散度矩阵，由于特征值相互比较接近，可能使得投影向量的计算变得不稳定，此外该算法需要计算可逆矩阵，这将成为提高计算效率的瓶颈。

Chu（2010）认为类内离散度的零空间中包含了更多的鉴别信息，提出了零空间方法（Null-space Linear Discriminant Analysis，NLDA）。该算法先将样本集投影类内离散度矩阵的零空间，然后寻找满足类间离散度矩阵最大的判别向量。当样本数较大时，零空间的计算量较大，同时会丢失零空间外的判别信息。Huang（2002）证明了类内和类间离散度矩阵零空间的交集构成了总体离散度矩阵的零空间，因此移除总体离散度矩阵的零空间，不会丢失判别信息，同时提高了计算效率。H. Li（2004）将 Fisher 判别函数修改为差值形式，最大化类边界，具有更好的区分能力，能很好地解决 3S 问题。

虽然零空间方法能够解决小样本问题，但它不是在原始输入空间，而是在零空间中进行以最大化类间离散度，因此该方法是次优的。为了弥补 NLDA 没有考虑零空间外的任何信息的问题，Yu（2001）采用传统的 Fisher 准则，在类间离散度矩阵的列空间上寻找使得 Fisher 判别函数最小的判别向量，该方法称为 DLDA（Direct Linear Discriminant Analysis）算法。根据贝叶斯决策理论，Gao（2006）认为在保留全部鉴别矢量下，DLDA 并不是真正意义上的 LDA，而是一种特例。DLDA 想要避免移除类内协方差矩阵的零空间，但是类间协方差矩阵零空间的移除将导致类内协方差矩阵部分零空间被移除，从而导致信息的损失。X. Wang（2004）应用概率视觉模型（Probabilistic Visual Model），在类内离散度矩阵的零空间和主空间中寻找鉴别向量，在特征级上结合两个完备 LDA 子空间，提出了对偶空间 LDA（Dual-Space Linear Discriminant Analysis）。Yang（2005）提出完备 LDA 框架，在类内离散度矩阵的列空间和零空间中寻找判别向量。为了保留类内和类间离散度零空间的判别信息，Sharma（2008）提出了一种新的 DLDA 方法，称为梯度 LDA（Gradient Linear Discriminant Analysis，GLDA），它引入了梯度下降法，速度较快。

由于大多数 LDA 扩展方法得到的鉴别特征都是统计相关的，Jin（2001）

提出了 ULDA（Uncorrelated Linear Discriminant Analysis）来提取统计无关的特征。后来 Ye（2006）提出，当总体离散度矩阵非奇异时，传统 LDA 和 ULDA 是等价的，同时，他提出了 OLDA（Orthogonal Linear Discriminant Analysis），通过离散度矩阵的同时对角化，计算正交鉴别矢量集。该方法首先移除总体离散度矩阵的零空间，然后进行 ULDA 分析，最后进行正交化处理。早前就有很多基于 LDA 的正交变换，其中最著名的就是 FSLDA（Foley-Sammon Linear Discriminant Analysis）。Ye（2005）将 LDA 进行推广，提出了 GDA（Generalized Discriminant Analysis），其中 OLDA 和 ULDA 都是 GDA 的两种特殊情况。尽管 NLDA 和 OLDA 都会得到正交变换，但是，它们采取不同的方案。NLDA 在类内离散度矩阵的零空间计算正交变换，而 OLDA 通过计算离散度矩阵的同时对角化计算正交变换。Ye（2006）还对 NLDA 和 OLDA 进行了理论分析，并提出当总体离散度矩阵的秩等于类内和类间离散度矩阵秩之和时 NLDA 等价于 OLDA。

避免 LDA 类内离散度矩阵的奇异问题也可以利用线性代数方法，如 RDA（Regularized Discriminant Analysis）、LDA/QR、LDA/GSVD、PLDA（Pseudo-inverse Linear Discriminant Analysis）。其中，正则化方法使用对角矩阵的尺度缩放保证离散度矩阵非奇异。Guo（2006）对类内离散度矩阵进行正则化处理。该方法是一种纯正的 LDA 方法，没有采用任何降维，但是该方法的计算代价较高并且很难确定最优尺度缩放因子，同时由于 Fisher 判别函数没有严格最大，使得该方法是次优的。Bensmail（1996）提出了正则化的高斯判别分析（Regularized Gaussian Discriminant Analysis，RGDA），它提供了明确的分类法则，效果较好。LDA/GSVD 是 PLDA 的一种特例，如果将伪逆用到类间离散度矩阵，则 Ye（2005）提出的 LDA/QR 是 PLDA 的一种特例。但是 PLDA 不能保证 Fisher 判别函数的最优性。

LDA 对满足同方差高斯模型样本的统计分析才是最优的。在现实生活中，常常由于缺乏充分的样本或存在离群值，使得样本估计常常同真实矩阵分布不一致。对于不满足同方差的数据，研究者提出了异方差方法。Marais（2011）对每类协方差单独建模，选择判别子空间，开展了异方差判别分析（Heteroscedastic Discriminant Analysis，HDA）。Duin（2004）引入 Chernoff 准则函数，提出了 HFLDA（Heteroscedastic Fisher Linear Discriminant Analysis），该方法考虑了类协方差信息，能够处理异方差问题。在处理异方差问题上 HFLDA 比 HDA 更有效。然而，这些方法通常最大化平均类间距离，容易导致"类分离"（Class Separation）问题。为了解决该问题，Su

(2015) 结合异方差和最大最小距离分析 (Max-Min Distance Analysis, MMDA) 思想，提出了异方差最大最小距离分析 (Heteroscedastic Max-Min Distance Analysis, HMMDA)。对于不满足高斯分布的数据或不能预知分布的数据，研究者提出了非参数方法进行分析。Li (2009) 定义了非参数类间离散度矩阵，并提出了非参数判别分析 (Nonparametric Discriminant Analysis, NDA)。但是，NDA 仍然存在小样本问题。于是，Z. Li (2005) 提出了 PNSA (Principal Nonparametric Subspace Analysis)，在类内离散度矩阵的主空间提取非参数判别特征，从而提高了识别率。针对 PNSA 丢弃了类内离散度矩阵的零空间，文章又提出了另一种零空间非参数子空间分析 (Null-space Nonparametric Subspace Analysis, NNSA)。由于 PNSA 利用了类内离散度的主空间，而 NNSA 利用了其零空间，因此 PNSA 和 NNSA 是互补的。后来，研究者提出了 LDA 的加权方法，如 aPAC (approximate Pairwise Accuracy Criterion) 或 PDA (Penalized Discriminant Analysis)，在矩阵定义时引入权重减轻不稳定样本的作用。PDA 不仅可以克服 3S 问题，而且能平滑判别向量，使得其具有更好的解释。此外，Zhu (2006) 还提出了 SDA (Subclass Discriminant Analysis) 将不同类型分布划分成多个子类，定义了一种新的子类间离散度矩阵，并提出了两种准则函数求解子空间。

LDA 的 HOG 模型假设还导致了其最大特征数为 $C-1$，其中 C 是类别数，因此，当 C 较小时，不能得到足够多的特征，并且，当投影子空间的维数严格小于 $C-1$ 时，子空间投影可能合并那些在原始特征空间中互相靠近的类。为此，Xiang (2006) 提出 RFLD (Recursive Fisher Linear Discriminant)，在 RFLD 中，最大特征数跟类别数无关。此外 NDA 和一些零空间方法的最大特征数也不局限于 $C-1$。

对于 C 类分类问题，当 $C>2$ 时，Fisher 判别准则是次优的。显然，对于 C 类问题分解为多个两类问题的和，这种将多类问题分解为多个两类问题的方法虽然使得 LDA 变换很好地保留了那些本来已经很好分开的类的距离，但造成邻近的类有较大的重叠。一种解决方法是在准则中引入加权体系，使得最优准则在变换空间中更能表示类可分性。但是由于计算复杂度和小样本问题，仍然不能直接用于高维数据。而且这些加权 LDA 不能保证得到统计无关的判别特征。Liang (2007) 充分利用了类内离散度矩阵零空间外的有用信息，提出了一种统计无关的加权方法。Sharma (2008) 在降维前引入旋转变换，将原始空间的数据绕着其类质心进行旋转，使得原始空间中类别的重叠部分在特征空间中尽可能小，提出了旋转 LDA (Rotational Linear Discriminant Analysis, RLDA)。

各类之间的距离或相似关系是数据集分类的重要信息。在统计方法中，常常采用欧式距离、马氏距离（Mahalanobis Distance）等，反映两类数据在特征空间中的分隔程度。由于多类问题存在互相靠近的类别时常常容易混淆，因此需要在特征提取阶段加以处理。那些远离其他类的类别称为离群类（Outlier Class）。由于类间离散度矩阵常常跟分类精度没有直接关联，因此不能保证任意两类在输出空间的可分性，此外类间离散度矩阵平等对待每类，因此离群类可能会对类间协方差矩阵的估计产生负面影响。Lotlikar（2000）提出了 F-LDA（Fractional-step LDA），该方法认为，在输出空间中较为靠近的类别可能导致较高的错分率，因此，对于这些类别，需要在类间离散度中加大权重。为了将数据集从 D 维空间变换到 d 维空间，首先找到最优的 $D-1$ 子空间，更新各类的关系，并重新计算权重，然后找到 $D-2$ 维子空间，通过 $D-d$ 次迭代，最终找到 d 维空间。该算法需要迭代实现，每次使类间离散度矩阵减少一维，从而得到 d 维低维表示，当 $d=C-1$，该结果是最优的。但是，当样本数和维数的比率较小时，F-LDA 可能失效。为了解决这个问题，Lu（2003）提出了 DFLDA（Direct Fractional LDA），该算法先将数据投影到类间离散度矩阵的子空间，从而降低了复杂度。Tang（2005）通过分析传统 LDA、aPAC 和 FLDA 的关系，提出 FLDA 是对 aPAC 算法的改进。需要注意的是，aPAC 和 FLDA 都需要将数据进行白化变换作为预处理。这种变换旋转和尺度化坐标轴，将类内离散度矩阵归一化为单位矩阵，因此类间离散度矩阵包含了所有的判别信息。

为了尽可能克服传统 LDA 存在的问题，研究者将 LDA 与其他方法结合。Su（2011）提出了最优判别投影寻踪算法（Optimal Discriminatory Projection Pursuit，ODPP）。不同于传统的 LDA 方法，该方法首先使用最近类间边界样本构建候选投影集（Candidate Projection Set，CPS），然后根据类可分性标准，使用 AdaBoost 选取最优判别投影。投影的边界性质和 AdaBoost 学习过程保证了该算法具有较好的性能。该算法试图解决传统 LDA 存在的多个问题，如 3S 问题、高斯分布假设问题、贝叶斯最优问题、特征数受限问题等，但是该算法缺乏理论分析。Hu（2009）提出了 DMM（Discriminant Multidimensional Mapping）算法，将 MDS 和 LDA 结合，克服 3S 问题。类似 LPP，Hu（2008）首先构建邻域图反映人脸流形，并引入保邻的类间和类内离散度矩阵，使得局部结构能够在低维空间中保留。因此，该算法继承了 LDA 较强的判别能力，同时保留了数据样本的本征几何结构。类似地，

Sugiyama（2007）结合 LDA 和 LPP，提出了局部 Fisher 判别分析（Local Fisher Discriminant Analysis，LFDA），最大化类间可分性，同时保留类内局部结构，有效地解决了多模态数据（Multimodal Data）问题，并能够得到任意维的特征空间。

为了解决新增样本问题，研究人员在 LDA 中引入了增量学习。Ye（2005）在 ILDA（Incremental Linear Discriminant Analysis）中，每次只引入一个新样本。当样本类别数 C 很大时，该方法涉及 $C \times C$ 维离散度矩阵的特征值分解问题，从而增加了计算代价。Pang（2005）提出更新类内和类间离散度矩阵的方法。Hiraoka（2000）提出了基于梯度的增量学习算法。但是该方法的缺点是每次只包含一个新样本，并且要求用户设置学习效率。针对离散度矩阵的增量学习，Yan（2004）提出了计算类间和类内离散度矩阵的差异，从而修改判别函数，但是这可能导致两个离散度矩阵的正则化问题。基于特征空间的分裂和融合模型，Kim（2011）提出了一种新的增量算法。该算法的优点在于能够高效处理多类大数据集，并且该增量算法的结果和批量 LDA 算法一致。在该算法中，最大化类间和总体离散度矩阵的比率，首先更新两个离散度矩阵的主分量，然后根据计算的主分量计算判别分量。由于在每步中应用了充分生成集（Sufficient Spanning Sets），高效地解决了特征计算问题，并且只需要存储两个离散度矩阵的主分量，从而保证了该算法在时间和空间上的高效性。针对 LDA 增量学习中难以处理类内离散度矩阵逆的问题，Zhao（2008）提出了 GSVD-ILDA，动态增加人脸图像，增量学习自适应子空间，而不是重新计算 LDA/GSVD，从而节约了计算成本。

3. 非负矩阵分析

将高维数据投影到低维子空间是信号处理和模式识别中一个根本问题。在很多实际应用中，获得的数据常常是非负的，如图像、频谱等，常常需要对数据进行非负投影。但是，大多数传统方法如 PCA、LDA 常常通过 SVD 分解求解问题，因此不能保证解的非负性，并且 SVD 分解的因子不能对原始数据进行解释，因此，一种新的矩阵分解方法被提出来，这就是 Lee（1999）提出的非负矩阵分解（Nonnegative Matrix Factorization，NMF）。非负矩阵分解能够保留原始数据中的非负特性，这种非负分解可以解释为对整体的感知是由对各个组成部分的感知构成的，符合人类直观的认识——整体是由部分组成的，便于理解和解释，而且人类大脑神经元的点火率是非负的，这也从心理学和神经学的角度解释了非负约束的可行性。

NMF 与 PCA 类似，都可以将图像表示为一系列基图像的线性组合，与

PCA 不同的是，NMF 不允许基图像 $\boldsymbol{W}$ 或权重矩阵 $\boldsymbol{H}$ 中存在负值。一幅图像可以表示为加权的基图像的加性组合，其中基图像对应图像的各个部分，因此 NMF 的基是局部化特征，更好地对应图像各个部分。这种非负约束也在一定程度上导致了表示的稀疏性，节约了存储空间。为了确定 NMF 是否得到各个部分的正确分解，Donoho（2004）对此进行了深入的研究，并定义了大量的前提条件，但是这些前提条件极为苛刻，并不能适用所有的图像。

根据不同的应用需要，不同研究工作对 NMF 的目标函数增加了各种约束。首先为了提高稀疏度，Li（2001）增加了局部约束，得到了局部非负矩阵分解（Local Nonnegative Matrix Factorization，LNMF）。其原因在于，虽然 NMF 的基图像是基于部分的，但是基于部分的表示不具有局部性。LNMF 对目标函数增加了更多的约束，从而增加了空间局部性和基图像的稀疏性。研究发现，较之 NMF 和 PCA，LNMF 得到更好的分类性能。Buciu（2004）发现，对于人脸表情识别，LNMF 比 NMF 具有更高的识别率。Liu（2003）提出了稀疏非负矩阵分解（Non-negative Matrix Factorization with Sparseness Constraints，NMFSC）。针对 LNMF 和 NMFSC 都不能显式控制稀疏度的问题，Hoyer（2004）提出了显式调节稀疏参数的算法。Pascual-Montano（2006）增加了非光滑约束，得到了非光滑非负矩阵分解（nonsmooth Nonnegative Matrix Factorization，nsNMF）。Essid（2013）在非负分解过程中加入光滑约束。针对 NMF 的解不唯一的问题，Tolié（2018）将正交约束引入非负矩阵分解，提出了正交非负矩阵分解。为了反映数据空间的几何结构，Cai（2008）提出了流形上的非负分解。Ablin（2018）结合流形结构和非负约束，提出了正交流形上的非负矩阵分解算法。

心理学和生理学已经证实了大脑对事物的理解过程是将其分解为各个部分再分别理解的过程，但是很多非负数据分解的算法是无监督的，并且都是从数据重构的角度执行的，因此仍然无法很好地解释大脑学习和分类的功能。因此，为了进一步加强非负矩阵分解的判别能力，Cui（2015）引入判别信息，更好地改进了分类强度。Jia（2004）在 NMF 目标函数的基础上，加入 Fisher 准则约束。Kotsia（2007）采用投影梯度方法求解判别非负矩阵分析。Buciu（2004）在 LNMF 的基础上，增加了 Fisher 准则，得到判别非负矩阵分解（Discriminant Nonnegative Matrix Factorization，DNMF）。

通常，非负矩阵分解的基图像 $\boldsymbol{W}$ 和权重矩阵 $\boldsymbol{H}$ 的初始值是随机非负值，而初始值对于算法的收敛有至关重要的影响。研究发现，很多 NMF 算法对于 $\boldsymbol{W}$ 和 $\boldsymbol{H}$ 的初始值较为敏感，因此研究者对 NMF 及其扩展方法的初始条件进