# 1.人脸识别技术的简单介绍

人脸识别技术和扩展已应用于生活的各个方面，本文将介绍如何使用流形学习子空间方法来实现人脸识别技术，

## 1.1人脸识别的发展情况

### 1.1.1人脸识别的技术及应用简介

人脸识别是基于人脸特征信息的身份识别。人脸识别技术始于20世纪60年代，到现在已经有了近60年的发展史，在这60年中人脸识别技术得到突飞猛进的发展，相继出现了许多经典算法、经典思想和经典人脸库，如今人脸识别技术已经达到了比肉眼还要准确的地步，我相信在未来人脸识别技术一定能我们创造更加便利的生活条件。

### 1.1.2人脸识别技术的优势

人脸识别是生物识别的一种，生物识别因为具有易检测、唯一性和终身不变的特点，所以十分适合互联网时代用户对安全的需求，如今生物识别技术的识别速度更快，准确率更高，因此也更具有研究的价值。

目前主流的生物识别技术有人脸识别、指纹识别、虹膜识别、语音识别、静脉识别等。相对于其他的生物识别技术，人脸识别技术的成本更低、稳定性更好、准确率更高，因此也更具有发展优势，所以在多数应用场景中都会首选人脸识别技术。

### 1.1.3人脸识别技术的市场前景

市场前景是判断一项技术是否有研究价值的重要指标。

人脸识别市场前景非常好，发展势头也非常迅速，2017年全球人脸识别的市场规模超过40.5亿美元，预计2022年达到77.6亿美元，复合年增长率高达13.9%。

### 1.1.4人脸识别系统的研究内容

## 1.2流形学习

流形学习的本质在于根据有限的离散样本学习和发现嵌入在高维空间中的低维光滑流形,从而揭示隐藏在高维数据中的内在低维结构,以实现非线性降维或者可视化。

### 1.2.1流形学习的研究背景

随着互联网时代的发展，在各个研究领域，每时每刻都在快速生成大量的数据，而在这些数据背后的规律却难以发现，人们虽然获取了海量的信息却发现自己正处于“数据丰富，知识匮乏”的尴尬境地，因此如何从海量信息中提取自己所需的知识是当今各个领域共同面临的巨大挑战。

在许多实际应用中，尤其是在人脸识别中，往往要对成千上万张图片进行处理，而每张图片又有极高的维度，这种高维的特质往往隐藏了数据间关系的本质，对于传统的数据分析方法往往会造成“维数灾难”，这种情况就需要通过降维来把高维空间的数据间的关系映射到低纬度空间，这样就可以更加方便快速的处理数据。因此降维就成了这一任务吸引了许多科研人员的注意，也成为了如人脸识别、机器学习和数据挖掘等领域的热门研究问题。而流形学习就是解决这类问题的方案之一。

自从流形学习方法被提出到现在，研究它的工作就一直在进行着，特别是近年来随着数据挖掘和人脸识别等技术的高速发展，“维度灾难”的问题就成了相关研究领域的重大障碍。本文将利用流行学习方法解决这类问题。

### 1.2.2流形学习的发展情况

1984年斯坦福大学统计系的Hastie在一份技术报告中首次提出主曲线和主曲面的概念。1995年Bregler和Omohundro首次提出“流形学习”的概念，主要应用于图像插值和语音识别中。2000年Seung和Lee在《科学》杂志上发表《认知的流行模式》一文，提出了视觉感知的流行结构假说。就在同年同期的《科学》上还刊登了另外两篇著名的文章，它们提出了两个经典算法LLE和Isomap。随着后期学者不断地深入研究，又有许多经典算法相继被提出，比如Belkin和Niyogi提出拉普拉斯特征映射算法（LE），使高维空间相近的点映射到低微空间时也相近；Donoho和Grimes提出海森特征映射（HLLE），是对LLE算法的扩展；He和Hiyogi提出局部保持投影算法（LPP），是对LE算法的线性扩展；Zhang和Zha提出局部切空间校准算法（LTSA），基于“局部拟合，全局整合”的思想；Lin和Zha提出黎曼流行（RML），利用局部黎曼正交坐标系将高维空间的数据映射到低微本质空间中去。这些经典算法被提出后，又有许多学者为了弥补这些算法的缺陷而相继提出了很多经典的改进算法。

此外，很多学者发现不同的流形学习算法之间存在着一定的联系，又提出了一些框架将多种流形学习算法纳入其中，比如经典的核主成分分析框架（KPCA）将MDS、LLE、LE、Isomap和谱聚类进行统一，GA框架在将LE、LLE、LPP和Isomap等流形学习方法纳入其中的同时又将PCA和LDA等传统线性降维方法统一进去。

### 1.2.3流形学习的定义

### 1.2.4本文用到的流形学习算法

本文主要用的NPE算法和IsoProjection算法并详细介绍算法实现原理，本文会在后面介绍正交的NPE和IsoProjection算法，并对比这些算法的优缺点。

# 2本文要用到的其他算法的详细介绍

NPE和IsoProjection算法的实现需要用到LGE算法，IsoProjection算法因为涉及最短路径选择所以用到迪杰斯特拉算法，而IsoProjection算法又是基于ISOMAP算法的改进，ISOMAP算法又涉及到MDS算法。通过LGE算法构造OLGE算法要用到规范正交基。

## 2.1规范正交基

定义：设是一个向量空间。

（1）若是向量空间V的一个基，且是两两正交的向量组，则称是向量空间V的正交基。

（2）若是向量空间V的一个基，两两正交，且都是单位向量，则称是向量空间V的一个规范正交基（或标准正交基）。

求法：设是向量空间V的一个基，要求V的一个规范正交基，也就是要找一组两两正交的单位向量，使与等价。这一过程称为把基规范正交化。

可以先进行施密特正交化在进行单位化。本节接下来将详细介绍施密特正交化和单位化的计算方法。

### 2.1.1施密特(Schmidt)正交化

令：

；

;

……

，

则易验证两两正交，且与等价。且满足：对任何，向量组与等价。

### 2.1.2单位化

令：

则是V的一个规范正交基。

施密特正交化过程中可将中的任一线性无关的向量组化为与之等价的正交向量组；在经过单位化，得到与等价的规范正交向量组.

## 2.2迪杰斯特拉(Dijkstra)算法

迪杰斯特拉(Dijkstra)提出了一个按路径长度递增的次序产生最短路径的算法。该算法是典型最短路径算法，用于计算一个节点到其他节点的最短路径。

它的主要特点是以起始点为中心向外层层扩展 (广度优先搜索思想)，直到扩展到终点为止。

### 2.2.1迪杰斯特拉(Dijkstra)算法原理

1. 通过 Dijkstra 计算图 G 中的最短路径时，需要指定起点 s（即从顶点 s 开始计算）。
2. 此外，引进两个集合 S 和 U。S 的作用是记录已求出最短路径的顶点 (以及相应的最短路径长度)，而 U 则是记录还未求出最短路径的顶点 (以及该顶点到起点 s 的距离)。
3. 初始时，S 中只有起点 s；U 中是除 s 之外的顶点，并且 U 中顶点的路径是“ 起点 s 到该顶点的路径”。然后，从 U 中找出路径最短的顶点，并将其加入到 S 中；接着，更新 U 中的顶点和顶点对应的路径。 然后，再从 U 中找出路径最短的顶点，并将其加入到 S 中；接着，更新 U 中的顶点和顶点对应的路径。重复该操作，直到遍历完所有顶点。

### 2.2.2迪杰斯特拉(Dijkstra)算法步骤

1. 初始时，S 只包含起点 s；U 包含除 s 外的其他顶点，且 U 中顶点的距离为“起点 s 到该顶点的距离”。例如，U 中顶点 v 的距离为 的长度，然后 s 和 v 不相邻，则 v 的距离为∞。
2. 从 U 中选出“ 距离最短的顶点 k”，并将顶点 k 加入到 S 中；同时，从 U 中移除顶点 k。
3. 更新 U 中各个顶点到起点 s 的距离。之所以更新 U 中顶点的距离，是由于上一步中确定了 k 是求出最短路径的顶点，从而可以利用 k 来更新其它顶点的距离；例如，的距离可能大于 的距离。
4. 重复步骤 2 和3，直到遍历完所有顶点。

## 2.3 ISOMAP算法

Isomap 通过 “改造一种原本适用于欧氏空间的算法”，达到了“将流形映射到一个欧氏空间” 的目的。

Isomap 所改造的这个方法是 MDS，它的目的就是使得降维之后的点两两之间的距离尽量不变。只是 MDS 是针对欧氏空间设计的，对于距离的计算也是使用欧氏距离来完成的。如果数据分布在一个流形上的话，欧氏距离就不适用了。

Isomap 把 MDS 中原始空间中距离的计算从欧氏距离转换为了流形上的测地距离。当然，如果流形的结构事先不知道的话，这个距离是没法算的，于是 Isomap 通过将数据点连接起来构成一个邻接 Graph 来离散地近似原来的流形，而测地距离也相应地通过 Graph 上的最短路径来近似了。



所以低维嵌入流形上的测地线距离（红色）不能用高维空间的直线距离计算，但能用近邻距离来近似。

### 2.3.1 ISOMAP算法步骤

ISOMAP算法分为三步：

1. 构造邻域图。
2. 计算测地距离。
3. 数据嵌入。使用MDS算法。

### 2.3.2 ISOMAP算法流程

输入：样本集；近邻参数k；低维空间维数d。

过程：

确定的k近邻；

与k近邻之间的距离设置为欧式距离，与其他点的距离设置为无穷大；

调用最短路径算法（迪杰斯特拉算法）计算任意两样本点之间的距离；

将作为MDS算法的输入；

MDS算法的输出；

输出：样本集D在低维空间的投影。

### 2.3.3 ISOMAP算法的优点

ISOMAP算法最主要的优点就是使用 “测地距离”，而不是使用原始的欧几里得距离，这样可以更好的控制数据信息的流失，能够在低维空间中更加全面的将高维空间的数据表现出来。

### 2.3.4 ISOMAP算法的缺点

虽然相对于MDS算法而言，Isomap算法可以更好的保持和表示数据内部的几何结构，但Isomap算法还存在着一些不足之处，其中最严重的一点是算法本身的拓扑不稳定性，即Isomap算法在图中可能会建立错误的连接，这将严重的影响算法的执行。其次，Isomap算法针对的是凸形的流形，对于非凸形的流形将不能处理。再者，对于流形中含有“孔洞”的情况，一般的Isomap算法也不能很好的处理。

## 2.4 MDS算法

MDS 与 PCA 一样，是一种有效的降维方式，其可获得样本间相似性的空间表达。MDS 的原理可以简述为，利用样本的成对相似性，构建一个低维空间，使每对样本在高维空间的距离与在构建的低维空间中的样本相似性尽可能保持一致。

MDS 算法的核心思想是：降维前后，各自样本间的距离是不变的。由此可得到如下关系：原空间中的两样本之间的距离等于降维后这两样本之间的距离 ，即。所以：

，得到：

，，其中B为降维后样本的内积矩阵，。

对B做特征值分解 ，这里V是特征向量矩阵、A是由特征值构成的对角阵。此时，把A中的特征值排序后，把其中每个非0特征值拿出来构成对角矩阵，其对应的特征向量也需按特征值的大小改变排列顺序，组成新的特征向量矩阵。最终，通过这个公式完成降维操作：。

另外，在现实应用中为了有效降维，往往只需要降维后两样本间的距离应尽可能和原空间中两样本的距离相近就好了，不需要强行一致，因此上面特征值构成的对角阵和特征向量矩阵有了一些变化。

变化为：本来特征值取的是所有非0的特征值排序。现在变成排序好后，从大到小取特征值与其对应的特征向量。比如要降维到5维，就从大到小取5个特征值。

## 2.4线性图嵌入算法(LGE)

为基于图的子空间学习提供一般框架。该算法将由NPE和IsoProjection调用。

### 2.4.1 LGE算法推导

数据集，n表示样本数量，m表示样本维度。LGE算法采用无向有权图描述数据集的流形结构，顶点集对应数据，，表示边的权重，W矩阵是对称的。LGE算法在保持图的邻接关系的前提下，寻找X的低维表示。令表示从图到实线的映射，则LGE的目标函数如下：

变换后可得：

L是图的拉普拉斯矩阵，D是对角矩阵，对应W矩阵第列（行）所有元素之和，即。

令（消除嵌入时的量化影响），则：

若图到实线的映射为线性，则，则：

则最优向量对应于以下最小特征值所对应的特征向量。

矩阵通常是奇异的，所以LGE算法要先把X集投影到PCA空间，使变成可逆矩阵，然后求解特征值问题。经过PCA处理后上式可通过奇异值分解进行求解：

可解出，设为PCA空间特征向量，则LGE算法的解为：

矩阵是非对称的，所以LGE算法的解往往是非正交的。本节会在接下来介绍正交的LGE算法(OLGE)。

许多常用的线性子空间算法如LPP等，都可以通过定义其权重矩阵W将该算法统一于LGE框架下，通过进行求解。

TODO：参考论文：直接线性图嵌入算法及其人脸识别中的应用

### 2.4.2 LGE算法流程

## 2.5正交线性图嵌入算法(OLGE)

为基于图的子空间学习（正交基矢量）提供一般框架。该算法将由ONPE和OIsoProjection调用。

在LGE算法的基础添加正交迭代获得。

3等距映射算法（IsoProjection）

IsoProjection是ISOMAP的线性近似，

## 3.1 IsoProjection算法推导

设为输入数据集，其中，寻找投影矩阵W使得，其中。Y就是降维后的数据集。

定义一个距离矩阵D，表示到之间的测地距离，定义矩阵S令，令，I为单位矩阵，E为元素全为1的列向量，则可得内积矩阵：。则IsoProjection的目标函数如下：

使用拉格朗日算法将上式的解向量问题转化为求解下式的特征值和特征向量的问题。

所以要求是非奇异矩阵，因此IsoProjection的最佳投影向量就是上式最大特征值所对应的特征向量。

## 3.2 IsoProjection算法流程

IsoProjection算法分五步执行：

1. 先使用PCA算法对样本进行降维。因为是奇异矩阵，所以要进行去奇异处理。使用PCA算法的主要目的是将变为非奇异矩阵，然后在低维空间使用IsoProjection算法。
2. 使用K-近邻方法构造近邻图G。
3. 使用迪杰斯特拉算法计算最短路径矩阵D。
4. 计算最佳投影矩阵W。
5. 特征提取。令，则。

## 3.4 IsoProjection算法缺点

IsoProjection算法求出来的投影向量往往是非正交的，这种非正交的性质会在从高维空间向低维空间的投影过程中扰乱数据的流形结构，而且会使算法本身对降维后的子空间的维数十分敏感，很难估计样本的内蕴维数。

# 4领域保护嵌入算法（NPE）

## 4.1 NPE算法推导

令NPE算法的权重矩阵为S，w为一投影向量，使得满足。则NPE算法的目标函数如下：

对该目标函数做出如下推导：

上式中，令。则简化后的目标函数如下：

则NPE算法的最佳投影向量就是以下特征方程的最小特征值所对应的特征向量。

4.2 NPE算法流程

NPE算法的实现分为三步：

1. 构造近邻图。使用K-近邻法寻找与数据点欧氏距离最近的个近邻点。
2. 确定权值。用近邻对各个数据点进行重构。
3. 计算特征映射。在低维空间中，对各个进行数据点重构，即在保持重构权值不变的情况下，使重构误差最小，计算出降维矩阵。

## 4.3 NPE算法优点

## 4.4 NPE算法缺点

# 5正交的IsoProjection (OIsoProjection)和NPE(ONPE)

## 5.1 OIsoProjection

OIsoProjection算法是在IsoProjection算法的基础上，得到一组正交基向量，因此该算法在保留IsoProjection算法线性的特点的同时又能够保持高维数据的流形结构。

OIsoProjection算法希望通过一组正交基改造目标函数，正交基的求法也相类似，都是通过拉格朗日乘子法引入正交基约束条件来推导出一个特征方程，然后求出该特征方程的最大特征值对应的特征向量以求出正交基，再利用该正交基构造特征方程进行特征提取。

### 5.1.1 OIsoProjection算法推导

首先要计算出一组正交基，则目标函数如下：

则向量可以通过求解特征方程的最大特征值所对应的特征向量获得。

若已有d-1个正交基向量。则可以利用如下正交约束条件获得：。

利用拉格朗日乘子法得到如下方程：

为方便求解做出如下定义：

；

；

；

；

经过对目标函数进行一系列运算后（因推导过程冗长繁琐故将详细推导过程放于附录9.1.1），得到如下函数：

可通过求解上式最大特征值对应的特征向量得到正交基。

所以，OIsoProjection算法的正交投影向量就是特征方程的最大特征值所对应的特征向量。

### 5.1.2 OIsoProjection算法流程

OIsoProjection算法的前三步与IsoProjection算法的步骤相同，此处不再赘述。直接从第四步开始介绍：

4．计算正交投影向量。

通过5.2.1中的算法计算出一组正交基，具体流程如下：

1. 通过求解特征方程的最大特征值所对应的特征向量获得计算。
2. 通过求解特征方程的最大特征值所对应的特征向量获得，。

5. 特征提取。

令通过OIsoProjection算法得到的最佳投影矩阵为d维的，通过PCA算法得到的投影矩阵为，则令，则通过线性嵌入：

可以得到对任意x的d维特征。

## 5.2 ONPE

ONPE算法在NPE算法的基础上添加了正交迭代处理。改进的思路与OIsoProjection算法类似，算法的推导过程也与OIsoProjection算法类似。

TODO：基于NPE；改进思路，算法特点

### 5.2.1 ONPE算法推导

定义邻域保护函数为：

最小化该函数将得到NPE算法，而ONPE算法既要找到一组正交向量又要满足上式最小，因此ONPE算法的目标函数如下：

可以通过求解的最小特征对应的特征向量获得。

若已有k-1个正交向量，则第k正交向量满足一下条件：

利用拉格朗日乘子法得到如下方程：

为方便求解做出如下定义：

；

；

；

经过对目标函数进行一系列运算后（因推导过程冗长繁琐故将详细推导过程放于附录9.1.2），得到如下函数：

令：

则：

通过求解这个特征方程就能得到一组正交投影向量。

所以，ONPE算法的正交投影向量就是特征方程的最大特征值所对应的特征向量。

### 5.2.2 ONPE算法流程

ONPE算法的前两步与NPE算法一样，这里从第三步开始介绍：

3.计算正交投影向量。通过5.2.1中的算法计算出一组正交投影向量。

1. 通过求解的最小特征对应的特征向量获得。
2. 通过求解的最小特征对应的特征向量获得。

4. 特征提取。

令通过ONPE算法得到的最佳投影矩阵为d维的，通过PCA算法得到的投影矩阵为，则令，则通过线性嵌入：

可以得到对任意x的d维特征。

## 6.1常用的人脸数据库介绍

使用公开人脸数据库对程序进行测评是评判算法好坏的主要依据。本节将详细介绍目前人脸识别领域常用的几个人脸数据库。

人脸库的图像示例和URL在附录9.2中。

### 6.1.1 FERET人脸数据库

FERET项目从1993年运行至1997年，由美国国防部Counterdrug技术开发项目办公室和美国国防部高级研究计划局(DARPA)发起。该项目的主要任务是研究开发一个可以协助安全、情报和执法人员执行任务的自动的人脸识别系统。

FERET是目前应用最广泛的人脸数据库，该库由14051个正面、左右侧面的人脸头像组成。FERET库为多姿态、不同光照的灰度人脸图像，FERET库收纳人脸的不同光照、不同角度和不同表情的成像，该库包含以下子集：

1. fa: 正面的人脸图像
2. fb: 是对应于fa中的图像数秒后获取的。
3. ba: 类似于fa中的正面人脸图像。
4. bj: 是对应于ba中的图像数秒后获取的。
5. bk: 在不用灯光的条件下，对应于ba中的图像数秒后获取的。
6. ra: 各种角度的姿态库，角度是连续的，但是角度是没有确切统计过的。

### 6.1.2 Yale人脸数据库

Yale人脸数据库是耶鲁大学计算视觉与控制中心创建的。该库由15个人的人脸图像组成，共有165张图像，每个人有11张图像，分别基于不同的光照变化、表情变化和姿态变化，如眨眼、正常、瞌睡、惊奇、悲伤和高兴，还有佩戴眼镜的人脸图像。

### 6.1.3 Yale B人脸数据库

Yale B人脸库是来自不同光照和不同姿态下的单光源人脸图像，该库由10个人的5850张图像组成。该库主要用于在光照和姿态问题上的建模与分析。

该库包含了在64种光照条件下的9种不同姿态共576种不同的观察条件。该库还包括每种姿态下获得的包含背景的图像，所以，该库中共有5760+90=5850张图像。该库压缩后大约1GB。

### 6.1.4 ORL人脸数据库

ORL人脸库由剑桥大学AT&T实验室创建，该库由40个人，每人10张图像，共400张图像组成。图像均在不同光照、不同角度和不同表情下获得。每张图片的分辨率均为像素。该库的面部表情表情变化主要是笑/不笑和睁眼/闭眼，还有一些其他的面部细节。拍摄的是受试者的正面、垂直头像，允许倾斜或旋转。

### 6.1.5 AR人脸数据库

AR人脸库由西班牙巴塞罗那计算机视觉中心创建，该库包含126人4000多张彩色图像，图像分两次拍摄获得，并且严格控制采集环境，如摄像机参数、摄像机距离和光照环境等。

AR库中的图像都是白种人图像，该库的控制条件是：中性表情；笑；愤怒；尖叫；左侧光照；右侧光照；两侧光照；墨镜；墨镜和左侧光照；墨镜和右侧光照；围巾；围巾和左侧光照；围巾和右侧光照。

### 6.1.6 XM2VTS人脸数据库

英国的XM2VTS人脸库是受欧洲ACTS项目的研究计划M2VTS资助建立的身份认证资料数据库，是一个大范围的多模数据库。该库数据是由英国Surrey大学的295名志愿者在4个不同的时间段的图像和语音视频片段。该数据库是商业收费数据库。

XM2VTS有2360张图像，每张图片的分辨率均为像素，控制条件为：低头；抬头；戴眼镜；不戴眼镜；愤怒；平和；侧脸；正脸；有妆；无妆等等。

该库还有293人的三维模型，以VRML格式存储。

### 6.1.7 CMU PIE人脸数据库

CMU PIE是卡耐基梅隆大学在2000年10月到12月收集的，有68个人的41368张图像，包含大量关于位置、表情和光照变化的图像。该库对基于位置的人脸识别产生很大影响。

## 6.2实验结果

分别运行NPE、ONPE、IsoProjection和OIsoProjection算法，并在下一节对比这些算法，分析这些算法的优缺点。

# 9附录

## 9.1 算法详细推导流程

### 9.1.1 OIsoProjection算法详细推导过程

条件不再赘述，详见5.1.1。

对于以下方程：

令：，则：

方程两边同时左乘：

得：

将(9.1.1.2)式方程两边同时左乘，得到d-1个式子：

5.1.1中定义过：

；

；

；

；

所以就可以用矩阵形式表示这个方程了：

得：

将(9.1.1.2)两边同乘，得：

矩阵表示：

由和得：

得：

得特征方程：

TODO：基于流形学习算法的新生儿疼痛表情识别[D].

### 9.1.2 ONPE算法详细推导过程

条件不再赘述，详见5.2.1。

对于以下方程：

令，则：

得：

式两边同时左乘，得到k-1个式子：

5.2.1中定义过：

；

；

；

所以就可以用矩阵形式表示这个方程了：

得：

式两边同时左乘，得：

通过式和式得：

得：

得特征方程：