Equidistant Prototypes Embedding for Single Sample Based Face Recognition with Generic Learning and Incremental Learning

基于通用学习和增量学习的基于单样本人脸识别的等距原型嵌入

1. 作者想解决什么问题？（摘要）

“一个样本问题”是许多现实世界中有关执法和国土安全的主要挑战之一。 从理论上讲，这是一个极小的样本量问题，使传统的模式识别技术恶化。

2. 作者通过什么理论/模型来解决这个问题？

我们的方法称为linear regression analysis（LRA），而不是保留训练数据的全局或局部结构，它应用最小二乘回归技术将gallery样本映射到等距的位置，而不管训练数据的真实结构如何。 此外，结合了一种新颖的通用学习方法，该方法将通用面部的类内面部差异映射到零矢量，以增强LRA的通用能力。

使用这种新颖的方法，即使仅从不同的数据库和相机设置中收集通用数据，仅基于少数通用类的学习也可以大大提高人脸识别性能。 基于Greville算法的增量学习使映射矩阵能够根据新近出现的gallery类别，训练样本或一般变化而有效地更新。

3. 作者给出的整体实验效果怎么样？（摘要或者实验）

表1：在Extended Yale B，CMU PIE和AR数据库上进行的10个测试的平均识别率和标准偏差（std），每个受试者每个样本具有单个训练样本，分别使用像素强度，Gabor和LBP功能。

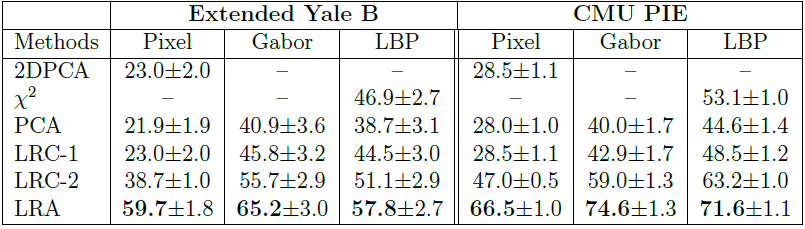


表2：100个AR数据库的受试者的比较识别率，每个受试者有一个训练样本和24个测试样本

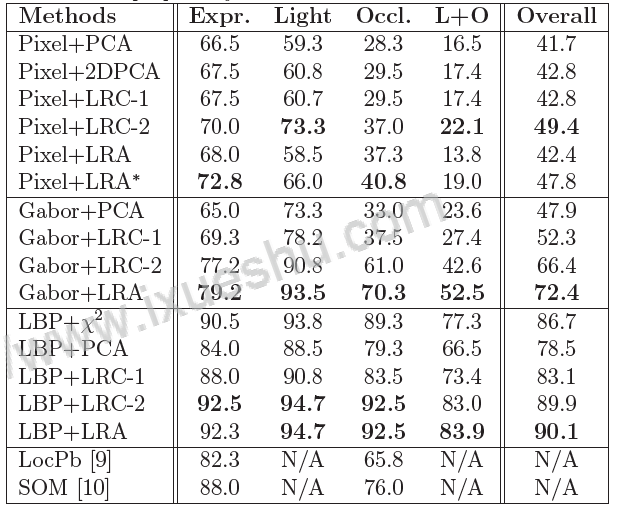


表3：使用单一训练样本对80名AR数据库受试者的LRA和LRA与一般学习（LRAGL）的平均错误率进行比较。Gallery外的20个Subjects用于通用学习。符号↓表示通过从LRA切换到LRA-GL而减少的识别错误百分比。

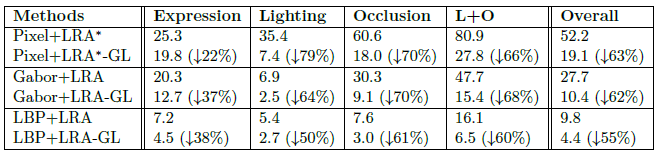
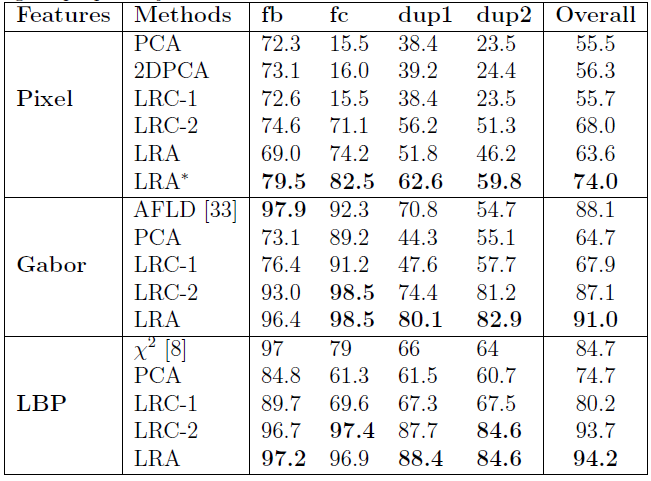


表4：FERET数据库中1196个受试者的相对识别率，每个受试者只有一个训练样本



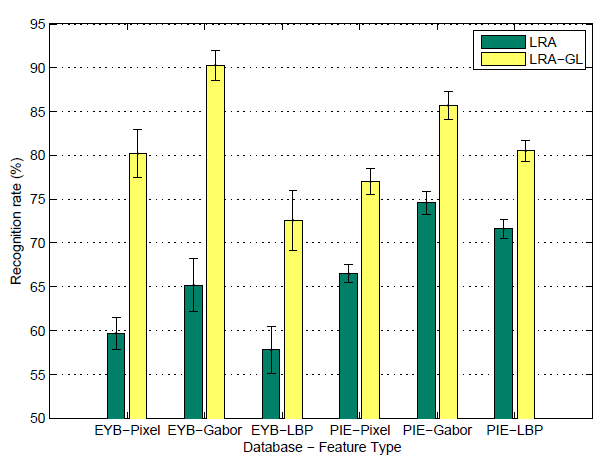


图7：使用三种类型的功能在Extended Yale B（EYB）和CMU PIE数据集上的LRA和LRA-GL的平均识别率，其中LRA-GL方法使用跨数据库训练程序。

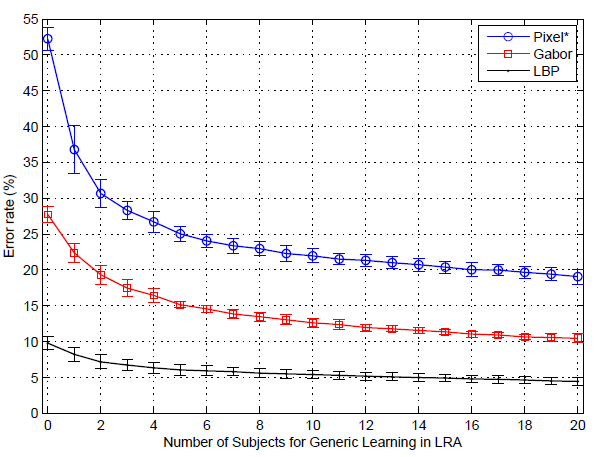


图9：LRA-GL的平均错误率（和标准差）与用于通用学习的subject数成函数关系。

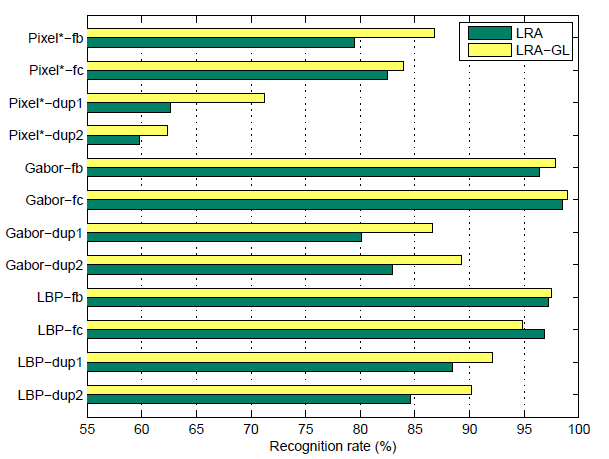


图11：使用三种类型的特征，在四个探针组上LRA和LRA-GL的相对FERET识别率。

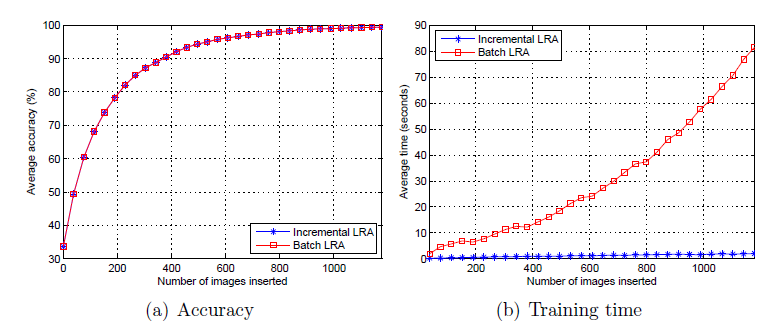


图12：Extended Yale B数据库上的batch-LRA和增量-LRA之间的比较。（a）准确性与插入训练样本数的关系；（b）训练时间与插入训练样本的关系。

4. 作者为什么研究这个课题？（introduction）

受“Remark 1”的启发，我们提出了一种高效且无参数的算法（称为线性回归分析（LRA）），以解决人脸识别中的一个样本问题。

5. 目前这个课题的研究进行到了哪一阶段？（introduction）

由于无监督学习技术不能在没有类内信息的情况下适用，因此无监督技术具有通过ICA [1]，PCA [2]或其变体[3] [4] [5]查找gallery数据的低维嵌入的功能。虽然这些方法被广泛应用，但是这些方法仅对于人脸表示(face representation)是最佳方法，并且仅对微小变化下的识别有效。

通过不变的特征（例如Gabor表示[6] [7]和局部二进制模式[8]）识别人脸可提高对人脸变化的鲁棒性。

但是，由于基于整体图像特征的方法会丢弃有关人脸3D布局的所有信息，因此它们对未观察到的变化的描述能力有限。 不幸的是，事实证明克服这些限制以建立有效的3D面部模型非常具有挑战性，尤其是当必须使识别系统在不受控制的光照和遮挡的情况下工作时。 基于可变形模型和虚拟人脸渲染的方法以显着的计算开销实现了鲁棒性。

一种更有效的方法是通过一组不重叠的块来识别人脸[9] [10]，但是该想法的现有实现产生了不确定的结果[11]。

正如MPEG-7标准在人脸检索上所提出的那样，提高人脸变化的鲁棒性的另一种策略是从通用数据集中学习判别统计，但是尚未考虑优化gallery脸部之间的区别。最近有关稀疏表示的一些工作提出通过新的字典设计来整合gallery集和外部通用数据集的信息[12] [？]。 有关一个样本问题的详细调查，请参见[13]。

6. 作者使用理论基于哪些假设？ （introduction）

此过程（以下称为“等距嵌入equidistant embedding”）利用了两个观察结果：

1）许多gallery面孔看起来非常相似，这可能与特征空间中的近距离类相对应。 通过将近距离类映射为远距离，等距嵌入会放大微妙的差异，这对于区分相似的面孔至关重要，同时还减少了在光照和遮挡变化的情况下身份混淆的风险。

2）如果gallery图像是在不同条件下收集的，则在特征空间中，它们可能会根据采集条件紧密聚类，并且测试样本可能会向条件相似的错误gallery样本倾斜。 通过消除基于条件的聚类，等距嵌入可减少因采集条件变化而引起的偏差。

Remark 1. 在每个类只有一个样本的特征空间中，将gallery样本放在等距的位置等效于最大化任何类与其余类之间的最小余量。

Assumption 1. 任何gallery面孔的类内变化都可以通过大量的通用面孔的类内差异的线性组合来近似得出。

7. 这篇文章存在哪些缺陷（如果作者没说，就要自己总结）？（conclusion）

待补充！

8. 作者关于这个课题的发展有哪些思考？（conclusion）

计划设计新的人脸图像描述子来捕捉更多identity-specific的面部特征（也就是从特征提取方法继续发展）

9. 研究的数据从哪里来？（实验）

使用四个标准数据库，即Extended Yale B [27]，CMU PIE [28]，AR [29]和FERET [30]，进行了广泛的实验来说明所提出方法的有效性。

* Extended Yale B（EYB）数据库包含27个人的正面图像，每个人都有64种不同的照明。 在我们的实验中，选择了两个子集（子集2和3），它们具有从13°到50°的24个光源方向。
* CMU PIE数据库[31]包含68个subject，共41368张图像。对于每个subject，在我们的实验中使用了包含21种不同光源的照明子集（C27）。

对于这两个数据集，首先通过相似度转换对所有图像进行归一化，该相似度转换将居中的眼间水平线设置为水平且相距70个像素，然后将其裁剪为128×128的大小，并使眼睛的中心位于（29，34 ）和（99，34）提取纯脸区域。 在我们的实验中没有进行进一步的预处理程序。 裁剪后的图像直接用于[32]中所述的Gabor和LBP特征提取，并调整大小为64×64以进行像素强度特征提取。 图6显示了三个数据集的一些裁剪图像。

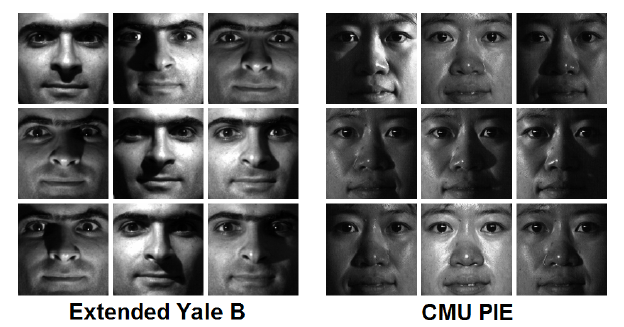


图6：扩展Yale B，CMU PIE和AR数据集的裁剪图像。 对于每个数据集，每个subject随机选择一个训练样本进行训练，其余样本用于测试。

* AR数据库的较大子集，其中包含来自100个受试者的2500张图像。 对于每个subject，使用Session 1的自然图像进行训练，并使用其他24张图像（Session 1的其余12张图像和Session 2的12张图像）进行测试。

如[9]中所述，以165×120的尺寸裁剪图像并转换为灰度。 对裁剪图像上提取LBP特征（具有11×8的cells）。裁剪后的图像将尺寸调整为128×128以进行Gabor特征提取，并将尺寸调整为82×60以进行像素特征处理。 图8显示了此实验中使用的一个人的25张裁剪图像，从图中可以看到测试图像的表情，照明和遮挡（太阳镜和围巾）以及照明+遮挡（L + O）的表情不同 ）。



图8：AR数据库中一个人的裁剪图像。 单个自然图像用于训练，而其他24个严重变化的图像用于测试。

* 实验遵循FERET数据库的标准数据分区：

•通用训练集包含429人的1,002张图像，这些图像列在FERET标准训练CD中。

•gallery训练集包含1,196人的1,196张图像。

•fb probe组包含使用其他面部表情拍摄的1,195张图像。

•fc probe组包含在不同光照条件下拍摄的194张图像。

•dup1 probe组包含在不同时间拍摄的722张图像。

•dup2 probe组包含至少一年后拍摄的234张图像，是dup1集的子集。

首先根据眼睛坐标对所有图像进行归一化和裁剪，然后按照第一组实验中的描述提取像素，Gabor和LBP特征。 图10显示了我们的实验中使用的一些裁剪后的图像，从图中可以看出，该数据库的类内probe-gallery变异性很复杂，因为它是在几年中的许多次会议中收集的。



图10：一些gallery图像和相应probe图像的裁剪图像。

10. 研究中用到的重要指标有哪些？（实验）

average recognition rate：平均识别率

standard deviation：标准差

错误减少率（ERR）（用符号↓表示），以衡量通用学习的有效性。

average error rate：平均错误率

11. 方法实现具体分哪几步？每一步分别得出了什么结论？（实验）

1）LRA基础

使用类别指标向量来表示第i个subject，

其中具有单个1，即其*i*th分量。

使用这些目标作为gallery样本的多元输出，我们可以用矩阵符号表示回归模型



其中是一个K×K目标矩阵，W是一个K×p映射矩阵，而是一个K×K误差（errors）矩阵。 在不失一般性的前提下，我们假设gallery样本是根据它们所在的类排序的，因此我们可以将Y写为K×K单元矩阵。

为了最小化平方误差之和，即，可以如下计算最佳变换矩阵。



其中X†表示X的广义逆。

当将新的test图像提供给基于LRA的分类器时，首先提取图像的特征向量（用x表示），然后将其归一化为零均值和单位长度。response vector响应向量 通过线性变换得出：

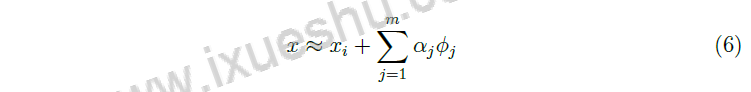
最后，识别结果由响应向量的最大部分确定：



其中表示响应向量y的第i个元素。

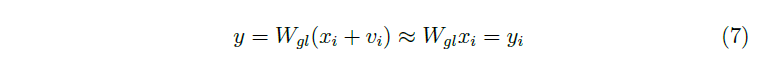
2）LRA with Generic Learning (LRA-GL)

给定足够大的m个类内变化基集，Assumption 1保证任何subject的test图像x和gallery图像之间未观察到的差别都可以通过“person-independent”基的线性组合来近似。

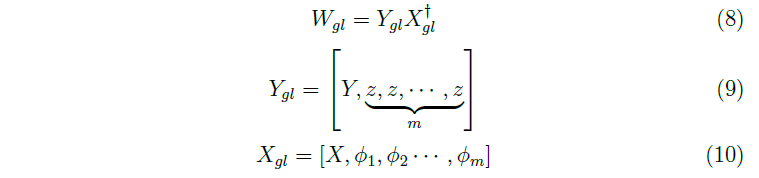


在我们的实验中，类内变量基的计算方法很简单，就是在一个通用数据集中从同类的其他样本中减去一个样本。

令是通用学习的理想映射矩阵。通过将所有类内变异基映射到零向量，即，j = 1，...，m，LRA-GL通过 减少未观察到的类内变异。 这样，底层的就可以被准确地恢复。



为了达到这个目的，LRA-GL的映射矩阵如下：



其中是p×（K + m）扩展数据矩阵，而是K×（K + m）扩展目标矩阵。 在一个样本问题的典型条件下，的秩最大为K + m，的秩为K，因此的秩为K。

LRA和LRA-GL都得出一个秩为K的K×p映射矩阵。

K个galley图像中的每个图像都有一个corresponding basis vector（映射矩阵的行向量），该向量spans等距嵌入空间。

test样本会被识别出来是通过选择最好的correlated basis vector。

3）Incremental Learning of LRA-GL

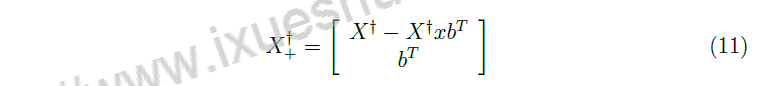
根据三种不同情况，LRA目标矩阵的更新非常简单。

（1）如果添加了第ith类的样本，则根据更新目标矩阵Y，其中y是第i个元素为1的稀疏向量；

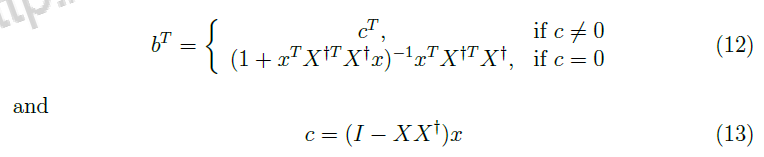
（2）如果添加了generic variant矢量，则，其中z是零矢量，维数等于Y的行数；

（3）如果添加了新类的样本，则，其中z是零向量，维数等于Y的列数，y是稀疏向量，最后一个元素为1 。

数据矩阵的广义逆的更新是通过完善的Greville算法进行的[16]。 将更新后的数据矩阵表示为，其中是旧数据矩阵，是相加项（样本或通用变体矢量）。Greville算法计算矩阵的广义逆，如下所示：



其中被定义为，



将项目添加到训练数据时，增量LRA（ILRA）首先将Y和X†分别更新为和†，然后将更新的映射矩阵计算为



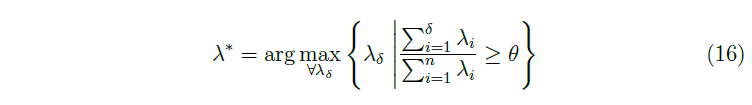
此更新过程由少量矩阵和向量的乘法和加法组成，这比直接计算矩阵逆的效率要高得多。通过增量学习，可以在线方式进行LRA的gallery训练和通用训练。

4）LRA \*方法

为了稳定inverse eigenspectrum，常用的方法是通过添加一个倍数单元矩阵，来规范化Gram矩阵的eigenspectrum。



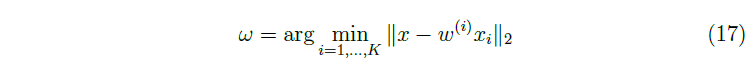
请注意，交叉验证不适用于选择一个样本问题的参数。相反，我们提出了一种LRA \*方法，以通过eigenvalue spectrum本身的统计量来控制正则化的强度。



其中是降序的第i个特征值，n是特征值的总数，因此阈值θ表示在第一个δ特征向量中捕获的能量的理想比例。为简单起见，我们在实验中为像素特征设置θ= 0.98。

对照组方法：

LRC通过找到test样本与特定类别gallery的线性组合之间的最小残差对面孔进行分类。在一个样本问题中，我们发现LRC简化为加权最近邻分类器，



基于计算标量权重的不同方式，提出了两种算法。

1）LRC-1

计算权重作为x在第i个gallery样本上的投影。



如图5所示，对于单位长度特征向量，残差仅由的正弦确定。LRC-1实际上是针对一个样本问题的基于角度的最近邻分类器。 此外，如果将所有样本标准化为均值和单位方差为零，则此过程等效于选择与测试样本最相关的gallery样本。

2）LRC-2

通过x在整个gallery集上的投影来协作计算权重。



其中是提取与第i个gallery样本相对应的系数的向量。 如图5所示，LRC2中使用的残差涉及权重和角度。

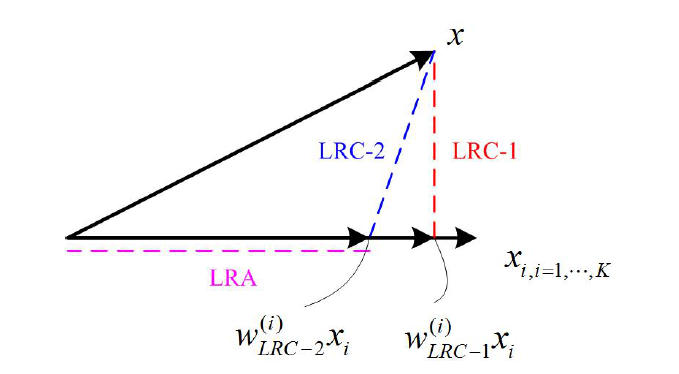


图5：流行的基于线性回归的分类方法之间的关系。红线的长度表示LRC-1中使用的残差。 蓝线代表LRC-2中使用的残差。 品红色线的长度表示LRA中使用的系数。