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

**目录**

[**Abstract（摘要）** 2](#_Toc29569655)

[**1. Introduction（介绍）** 2](#_Toc29569656)

[**2. Linear Regression Analysis with Generic Learning** 7](#_Toc29569657)

[**3. Experimental Results** 19](#_Toc29569658)

[**4. Conclusion** 32](#_Toc29569659)

**Abstract（摘要）**

我们开发了一种无参数的人脸识别算法，该算法对每个subject使用单个gallery样本，该算法对光照，表情，遮挡和年龄的大变化不敏感。我们利用这样的观察，即等距原型嵌入是一种最佳嵌入，可以最大程度地增加类之间的最小one-against-the-rest余量。 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。 尽管LRA非常简单且没有参数，但它结合了常用的本地描述符（如Gabor表示法和本地二进制模式），在扩展Yale B，CMU PIE，AR和FERET数据库上进行的几种标准实验均优于最新方法。

**1. Introduction（介绍）**

在本文中，我们考虑针对每个subject使用单个gallery样本的人脸识别问题。 这个“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。 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]。

与保留训练数据的全局或局部结构的现有方法不同[14]，我们利用将gallery样本放置在特征空间中等距的位置来解决一个样本问题的方法，而无需考虑观察空间中训练数据的真实结构。

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，这可能与要素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。

从理论上讲，等距嵌入的最佳性可以通过Vapnik的统计学习理论来解释[15]（第353页），这证明了通过放置R可以最大化球体中所包含的k≤K个点的所有二分法的最小余量。 这些点位于规则单形上，其顶点位于球体的表面上。 在一个样本问题中，每个样本代表一个类的唯一原型。 根据Vapnik定理，考虑一类和其余类的所有二分法，等距原型是一种最佳嵌入，可最大程度地增加类之间的最小one-against-the-rest余量。

**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。**

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。

如图1所示，LRA使用最小二乘回归将三个gallery图像映射到，，。

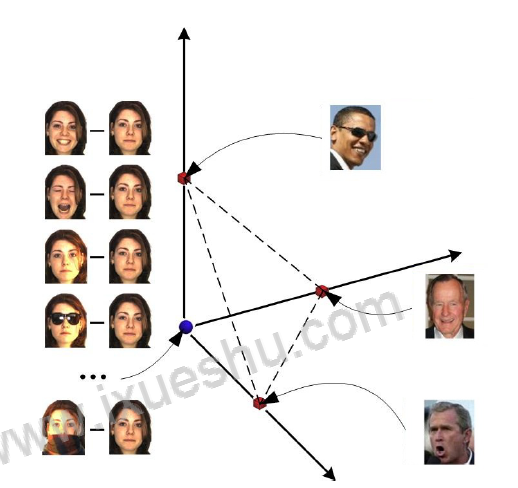


图1：等距原型嵌入空间的几何解释，其中三个gallery图像映射到 ，，和通用面部变化都映射到 。 制定了增量学习算法，以根据新近出现的画廊图像，训练样本或一般变化有效地更新模型。

回归目标的设置旨在避免耗时的搜索最接近的原型，从而可以通过找到响应向量的最大条目来有效地标识测试样本。除了等距原型嵌入之外，我们的模型还有两个其他优点，如下所示。

* 通用学习。LRA将普通人脸的人际差异映射到 ，以便galley脸部上的任何类似变化都不会影响回归结果。实验结果证实，即使从不同的数据库和相机设置收集了通用数据，该通用学习技术也可以在很大程度上改善人脸识别性能。
* 增量学习。LRA的映射矩阵由训练数据矩阵的广义逆来表示，可以通过完善的Greville算法来递增计算[16]。 结果，如果添加了新的画廊类，新的训练样本或新的通用变体，则可以有效地更新LRA模型。我们将显示增量LRA达到与批处理LRA相同的识别结果。

应该指出的是，最小二乘回归是模式识别中的“经典”技术，由William G.Wee于1968年首次应用于多类分类[17]，最近在人脸识别中得到普及[18] [19]。 尽管先前的工作使用线性回归来解决每个类别中有多个样本的问题，但我们提出的LRA方法首次将最小二乘回归与常见的人脸描述符组合使用来解决一个样本问题，并获得了优异的结果与最先进的方法相比。还介绍了LRA和流行的线性回归分类方法之间的理论和实验比较。

本文的其余部分安排如下：第二部分详细介绍LRA模型的数学原理，并说明如何将回归模型应用于人脸识别，第三部分提供实验结果，第四部分总结本文。

**2. Linear Regression Analysis with Generic Learning**

本节首先根据线性回归分析（LRA）制定基于等距原型的识别，然后尝试通过泛型学习，增量学习和不变描述符来增强回归器的泛化能力。 还讨论了LRA与其他线性回归分类模型之间的关系。

2.1. Linear Regression Analysis (LRA)

为了识别K个subjects，通过移动和旋转一个K-1个regular simplex可以实现无数个等距嵌入。为了提高效率，使用类别指标向量来表示第i个对象，其中具有单个1，即其*i*th分量。

等距嵌入的这种设置避免了费时的最近邻搜索以进行识别，因为可以通过向量的最大元素有效地找到最近的原型。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。当将新颖的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：y = Wx。

最后，识别结果由响应向量的最大部分确定：

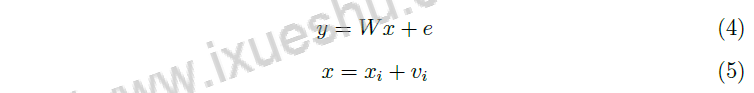


其中表示响应向量y的第i个元素。

应该提到的是，LRA的数学公式与“indicator矩阵的线性回归[20]”相同，但是由于以下两个方面的特点，它被更名为“线性回归分析”。 首先，LRA是等距原型的有效分类器，如Remark 1所示，它对单样本问题具有深刻的理论依据。其次，通过适当地指定回归目标（regression target），LRA得以扩展，以包括更多gallery之外的信息，以进一步提高性能。

2.2. LRA with Generic Learning (LRA-GL)

给定学习的映射矩阵W和subject i的test样本x，LRA的任务是从嘈杂的响应向量y中恢复潜在的。将面部识别问题视为具有测量误差的线性回归模型[21]，如下所示：

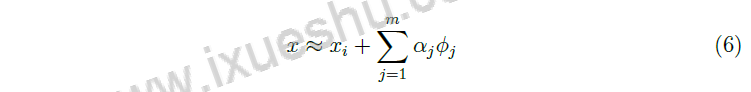


其中test样本x表示为正确的gallery样本xi和相关的类内变化的组合。

在训练阶段，（2）中的最小二乘解假设误差项e为零平均高斯噪声。 然而，在表情，光照和遮挡中存在未观察到的变化时，test图像x可能会与正确的gallery图像大大偏离，并且回归误差实际上与高斯分布相差甚远。 因此，最小平方解W对于具有大变化的识别问题是不可取的。 取代基于鲁棒估计[22] [23]或基于局部分割[9] [18]的技术来解决离群值，我们提出了一种新颖的学习方法，该方法基于以下假设充分利用人脸的相似形状和纹理 ：

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。

给定足够大的m个类内变化基集，Assumption 1保证任何受试者的test图像和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。图2展示了Assumption 1如何在脸部图像上成立的。由太阳镜和侧光引起的gallery脸的面部差异可以用其他脸的一些相似的差异线性近似。事实上，由于类内变异并不是真正独立于人的，线性近似的变异在表示意义上永远不可能是完美的(例如，图2所示的被遮挡眼睛周围的近似误差)。

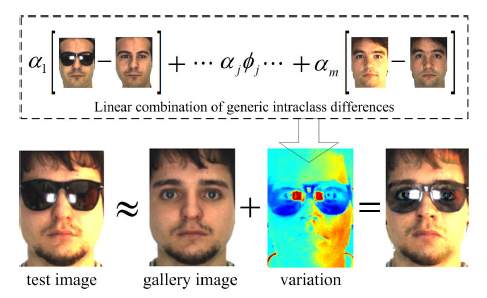
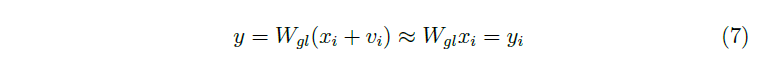
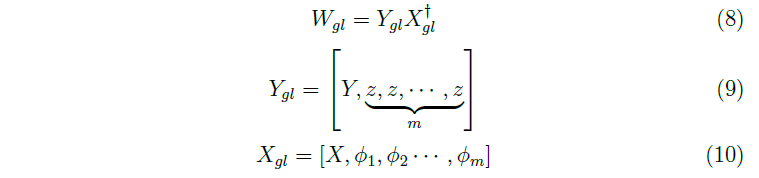


图2:LRA的泛型学习直觉:通过将泛型类内差异映射到零向量，gallery和test图像的响应向量将几乎相同。

然而，出于识别的目的，抵消这样的近似变化足以减轻gallery和test样品之间的大偏差。 此外，由于对齐的人脸的类间和类内变异之间的相关性往往较小[24] [25]，因此该算法有可能在保持gallery样本分离的情况下投射出通用的类内变异。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对应的基本向量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。

为了观察basis vector如何表征人脸，图3显示了从AR和扩展Yale B数据库获得的一些basis vector（图像形式）。

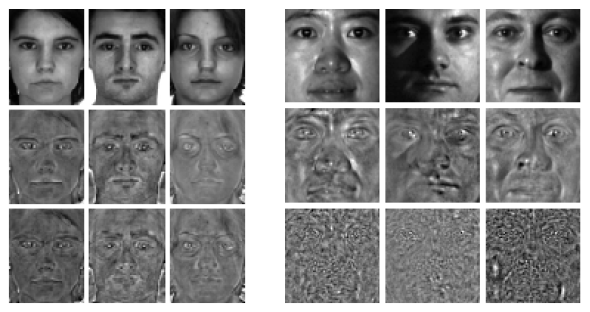


图3：从AR和扩展Yale B数据库获得的一些basis vectors。第一行显示了一些gallery图像，第二行显示了LRA的corresponding basis vectors，最后一行显示了LRA-GL的corresponding basis vectors。

我们可以看到，眼睛、鼻子、眉毛以上的前额部分和脸颊在LRA的基向量中占主导地位，强调每个gallery face的这些不变区域可以获得更好的识别性能。通过泛型学习，使被强调的不变区域更加localized，从而进一步提高对sever variations的鲁棒性。

2.3. Incremental Learning of LRA-GL

现实世界中的面部识别系统的训练集会随着时间的推移而变化。 更多的训练样本将被插入到系统中，以扩大gallery或者提高准确性。因此，增量学习对于系统的效率很重要。根据（8），LRA-GL的映射矩阵被公式化为目标矩阵和广义逆矩阵的乘积，它们被有效地更新如下。

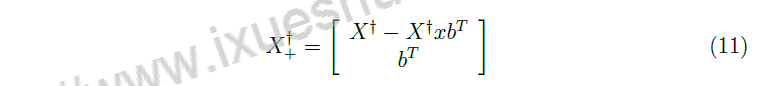
根据三种不同情况，LRA目标矩阵的更新非常简单。

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；

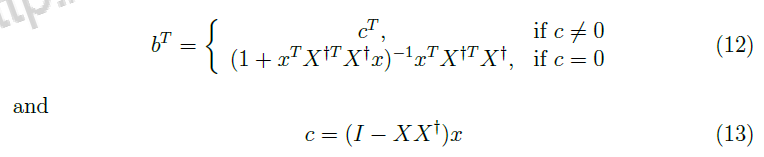
2）如果添加了通用差异矢量，则，其中z是零矢量，维数等于Y的行数；

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。

此外，人脸识别系统可以有效地反馈错误识别的样本，以更新LRA模型，从而避免了对整个系统的重新训练。

2.4. Intrinsic Dimension, Regularization, and Face Descriptors

面部图像中的像素数量通常超过10^4，但是由于人脸的形状和纹理相似，图像空间的固有尺寸低至O（10^2）[26]。低的固有维数导致像素强度特征的特征谱迅速衰减，如图4所示。

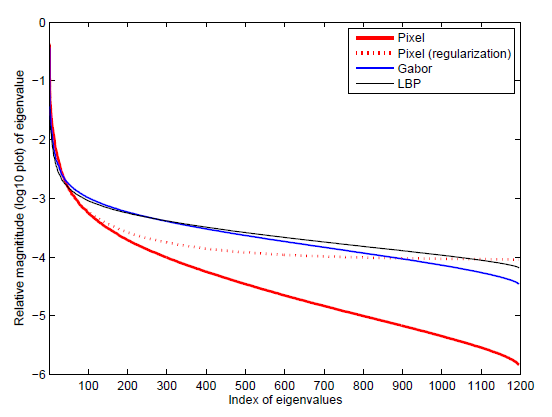


图4：三种流行的面部描述子：像素强度，Gabor小波和LBP的Gram矩阵的特征值的相对幅值（）。 数据来自FERET数据库的1196个库图像。

这可能会降低LRA的识别性能。

具体地，在小样本量条件下，将X的广义逆计算为，其中逆Gram矩阵通过 inverse eigenvalues对面部分量进行加权。

Specifically,under the small sample size condition, the generalized inverse of X is computedas , where the inverse Gram matrix weights the face components by the inverse eigenvalues.

因此，（test图像的）较小的噪声干扰可能会在特征值较小的区域中引起inverse eigenspectrum的较大振动，从而极大地影响回归结果。

Therefore, the small noise disturbance (of the test image) might cause large vibrations of the inverse eigenspectrum in the region of small eigenvalues, and thus largely effects the regression result.

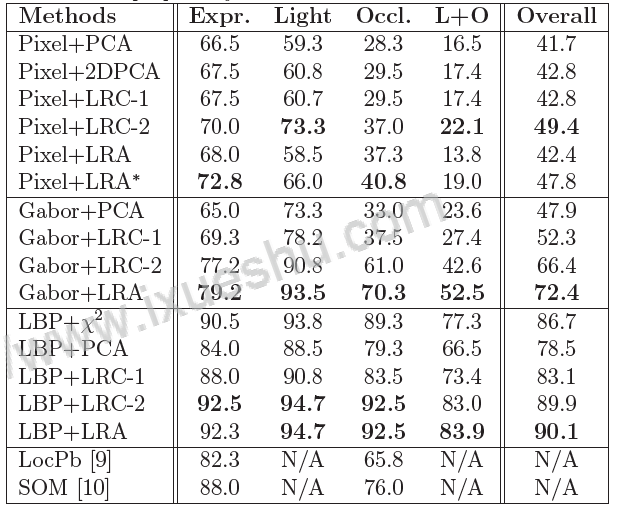
假设，

对于Gabor或者LBP来说，中的元素的大小都会比较平均，不会出现极大和极小的区域

而对于，像素强度来说，就不是那么平均了，于是回归的结果y可能会因为x较小的噪声干扰而被影响！

作为证据，在某些情况下使用像素强度，LRA产生的适度精度与PCA和2DPCA相似（示例请参见表2）。

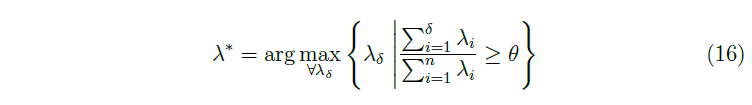
表2：100个AR数据库的受试者的比较识别率，每个受试者有一个训练样本和24个测试样本



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。 图4示出了正则化eigenspectrum的示例。

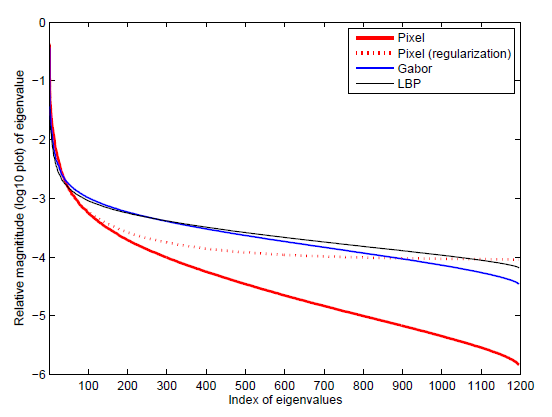


图4：三种流行的面部描述子：像素强度，Gabor小波和LBP的Gram矩阵的特征值的相对幅值（）。 数据来自FERET数据库的1196个库图像。

在实践中，一个更可取的解决方案是应用局部描述符（例如Gabor和LBP）来表示人脸而不是像素强度。 在人脸识别社区中，分类算法主要是在像素域中实现和评估的，如[2] [4] [18] [19]所示，因此它们忽略了相邻像素点之间的潜在关系。 相反，局部描述符通过预定义的滤波器卷积或图案匹配来考虑图像像素的局部关系，因此倾向于具有更高的固有尺寸。 如图4所示，Gabor和LBP的特征谱的衰减与像素强度的正规化之一一样慢，并且在倒数形式中会更加稳定。

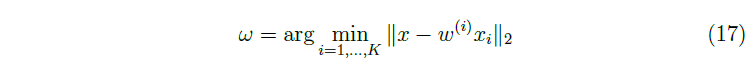
因此，基于局部特征的表示不仅由于其不变性而且还具有稳定的inverse eigenspectrums，因此对人脸识别具有鲁棒性。

与像素强度相比，在局部特征上实施LRA可能会产生更可靠的性能。 作为证据，在我们实验的所有测试条件下，使用LRA进行的基于局部特征的分类始终均比对照方法具有更高的准确性。

2.5. Relations to Linear Regression Classification

线性回归分类（LRC），例如[18] [19]，最近在人脸识别中越来越受欢迎。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。

尽管它适用于一个样本问题，但据我们所知，尚未对LRC进行过每类单个样本的研究或测试。 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。

第一种算法[18]（其后称为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。

应该意识到，特征向量之间的相关性对光照和遮挡的变化非常敏感。

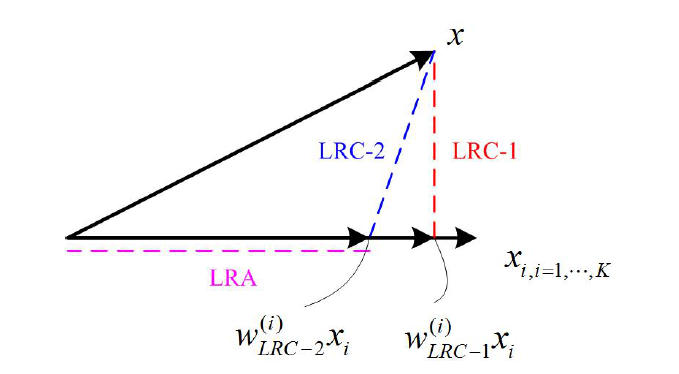


图5：流行的基于线性回归的分类方法之间的关系。红线的长度表示LRC-1中使用的残差。 蓝线代表LRC-2中使用的残差。 品红色线的长度表示LRA中使用的系数。

第二种算法[19]（以下称为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。 如果test图像和正确的gallery图像之间存在一些严重变化，则本质上代表特征相关性的角度可能会成为识别中的误导因素。

例如，如果在左照明下捕获测试图像，则所有左照明的gallery图像都倾向于与它成小角度。 一旦正确的gallery图像不在左光下，而其他一些gallery图像在左光下，则LRC-2的误识别可能会发生。

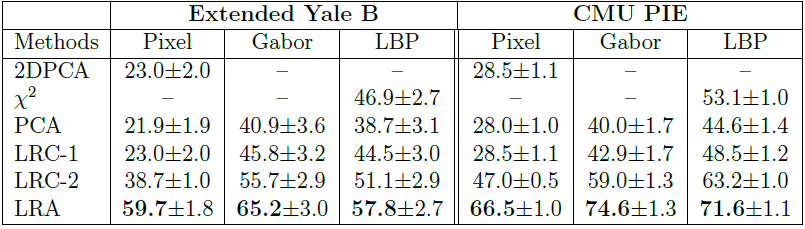
有趣的是，（19）中的LRC-2的权重与（3）中的LRA的系数相同。也就是

LRA中的判别公式（3）为



换句话说，LRA消除了LRC-2中的误导因素，仅使用gallery样本的权重进行识别。 从这个角度来看，如果gallery集合中的图像是在不同条件下采集的，则LRA比LRC-2更为健壮（有关证据，请参见表1）。

表1：在扩展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。



**3. Experimental Results**

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。 已采用了面部识别文献中报道的几种标准评估协议，并使用三种广泛使用的面部描述符（像素强度，Gabor和LBP）对建议的方法和最新技术进行了全面比较。

3.1. Extended Yale B, CMU PIE: Recognizing Faces Under Variable Illuminations

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。

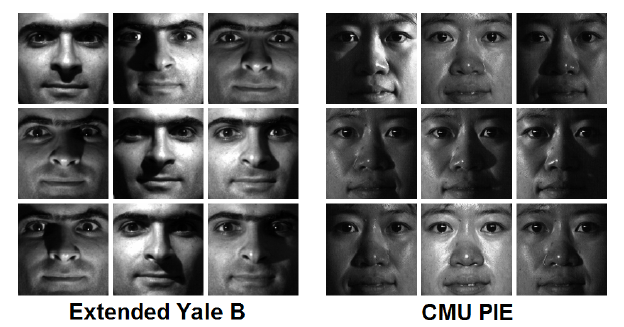
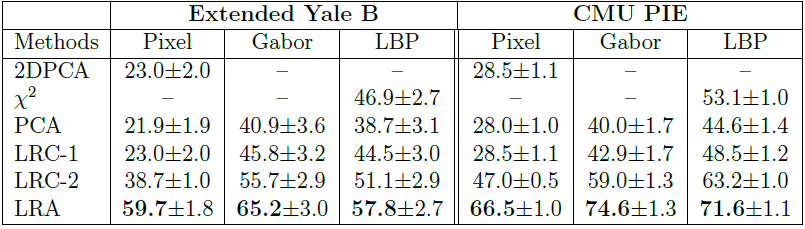


图6：扩展Yale B，CMU PIE和AR数据集的裁剪图像。 对于每个数据集，每个subject随机选择一个训练样本进行训练，其余样本用于测试。

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。 为了使报告的结果有意义，我们报告了10个随机训练/测试分区的平均识别率和标准差。

表1列出了分别使用三种不同特征的三种数据集上的六种识别方法的比较性能。不出所料，PCA在所有数据集上的准确性都非常低，这表明在可变光照下测试图像和单个图库图像之间的特征相关性对于识别而言并不可靠。 2DPCA [4]和LRC-1 [18]都可以提高PCA基线性能，但是这种改进是微不足道的，因为它们主要基于特征相关性（在此实验中它们产生相同的准确性）。

表1：在扩展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。



利用直方图匹配的优势，[8]准确性比基于相关的方法要好得多，但其性能仍低于LRC-2 [19]和LRA。 在所有测试的数据集和特征类型中，拟议的LRA方法始终优于LRC-2，其次是LRC-1，这清楚地表明等距嵌入到一个样本问题上的优越性。 通过平衡gallery面孔之间的距离，可以强调每个面孔不变区域周围的特征，如图3所示，这使得LRA明显优于其他方法。

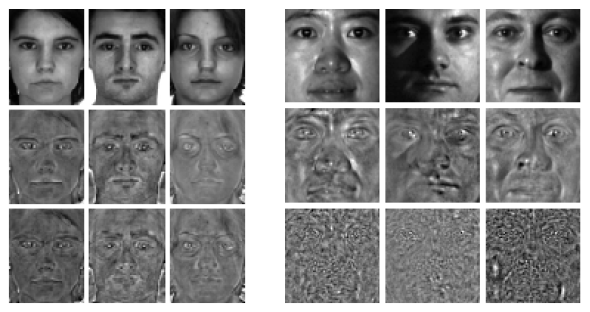


图3：从AR和扩展Yale B数据库获得的一些基本向量。第一行显示了一些gallery图像，第二行显示了LRA的对应基本向量，最后一行显示了LRA-GL的对应基本向量。

LRC-1和LRC-2的相对较低的准确性可能是因为它们在光照变化的情况下参与了不可靠的特征相关性。 特别是，在像素强度的相关性最容易引起误解的情况下，LRA比LRC-2的性能高出20％至30％，比其他方法的性能高出30％至60％。

由于EYB和PIE数据集的性能相对较低，我们评估通过跨数据库的通用学习来进一步改善LRA。 具体而言，我们将PIE数据集用作识别EYB数据库的通用训练集，并将EYB数据集用作识别PIE数据库的通用训练集。 图7示出了LRA-GL方法的结果。

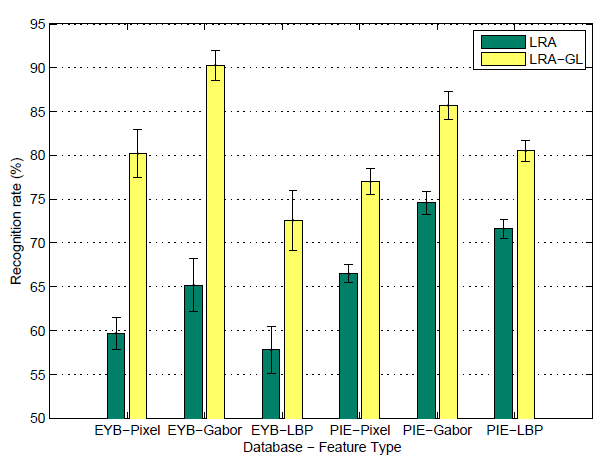


图7：使用三种类型的功能在Extended Yale B（EYB）和CMU PIE数据集上的LRA和LRA-GL的平均识别率，其中LRA-GL方法使用跨数据库训练程序。

尽管这两个数据库是在不同的光照和相机设置下收集的，但是它们之间的通用学习可以大大提高人脸识别性能。 通常，从PIE数据集进行通用学习可以将EYB数据集的准确性提高15％到25％。 从EYB数据集学习可以将PIE数据集的准确性提高大约10％。该结果表明，即使从不同的数据库和相机设置中收集了学习集，LRA-GL在大范围的实际应用中也是可行的。

3.2. AR Database: Recognizing Partially Occluded, Expression, and Lighting

Variant Faces

第二组实验使用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。

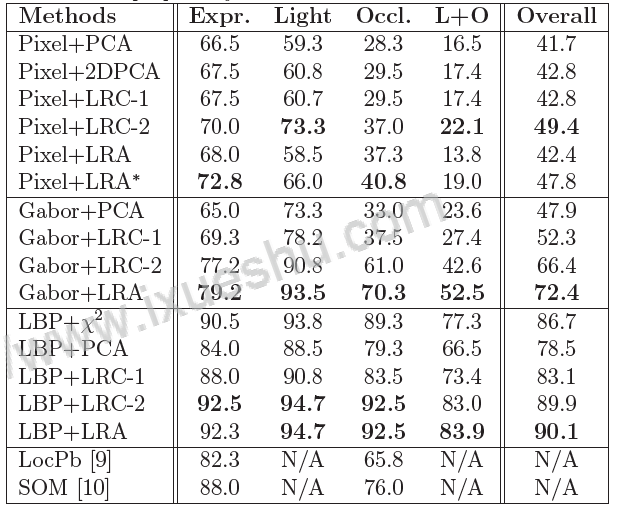
如[9]中所述，以165×120的尺寸裁剪图像并转换为灰度。 在具有11×8像元的裁剪图像上提取LBP特征。裁剪后的图像将尺寸调整为128×128以进行Gabor特征提取，并将尺寸调整为82×60以进行像素特征处理。 图8显示了此实验中使用的一个人的25张裁剪图像，从图中可以看到测试图像的表情，照明和遮挡（太阳镜和围巾）以及照明+遮挡（L + O）的表情不同 ）。



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。

表2报告了六种不同方法的比较准确性，并且LRA使用不同类型的功能执行不同的操作。 具体来说，结合Gabor功能，LRA的准确性要比其他方法高得多。 与LBP功能一起使用时，LRA的性能比LRC-2略好，但仍比其他方法好得多。 然而，使用像素特征，精度是中等的，这可能是由Gram矩阵的不稳定inverse eigenspectrum引起的。因此，采用LRA \*方法对特征谱进行正则化，可以观察到性能显着提高。

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



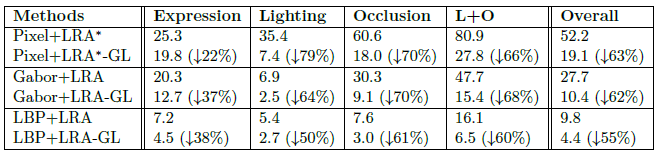
通常，与表1中的结果相比，LRA在此实验中的优势较小。 这可能是因为该实验的gallery集是在统一条件下收集的，而对照方法（如LRC-2）受不可靠关联的影响较小。

比较不同功能集的性能也很有趣。 通常，LBP功能的性能要比Gabor好得多，其次是像素强度。 令人惊讶的是，在存在遮挡的情况下，测试集上LBP与像素之间的精度差异可能高达70％。 在被遮挡的面部上的出色性能可能是因为在脸部的被遮挡部分很少出现均匀的LBP，并且被遮挡的部分会自动打折以供识别。 我们还将我们的方法与最新方法[10] [9]进行比较，该方法通过先验信息通过局部分区解决表达和遮挡问题。 SOM-face方法需要事先知道图像的遮挡区域，而LocPb方法使用通用图像来学习局部区域的权重。 与这些方法相比，使用LRA的基于LBP的分类提供了一种相当简单但更可靠的解决方案，可以识别部分遮挡，表情和光照的变脸。

接下来，我们研究了通用学习对提高LRA准确性的有效性。在此实验中，我们从AR数据库中选择了80个subjects的随机子集。 按照先前的设置，将Session 1中每个subject的自然图像用于训练，而其他24个图像用于测试。 为了构建用于通用学习的类内变体基础，选择了另外20个subjects（与80个gallery subjects不重叠），每个subject有13个Session 1的图像。 为了使报告的结果有意义，我们报告了10个随机gallery/通用分区的平均错误率。 表3列举了该实验的识别错误率，其中我们定义了一个错误减少率（ERR）（用符号↓表示），以衡量通用学习的有效性。

例如，由于基于LBP特征的LRA-GL将总错误率从52.2％降低到19.1％，所以ERR为↓63％[（52.2-19.1）/52.2]，这表明通过切换LRA到LRA-GL可以避免63％的识别错误。

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。



对于这三种类型的测试功能，总体ERR约为↓55-↓63％，有力地证明了LRA-GL方法对各种功能的普遍有效性。 特别是，使用LRA-GL的基于LBP特征的分类仅可实现4.4％的总错误率。 为了更好地理解LRA-GL的影响，表2还分别列举了四个测试变量的错误率。 在所有三种经过测试的特征类型上，LAR-GL在所有类型的变体上都是可行的，但有效性似乎有所不同。 照明和遮挡的错误减少率（↓50–↓79％）显着高于表情的错误减少率（↓22–↓38％）。 用于表达的相对较低的ERRs表示，表达的变化对单个面孔的特定面部形状最敏感，并且最难进行线性近似。

即使具有出色的表现，仍然存在一个有趣的问题：构建通用学习集需要多少个通用subjects。图9显示了平均错误率（和标准偏差）与 用于学习的通用subjects的数量。

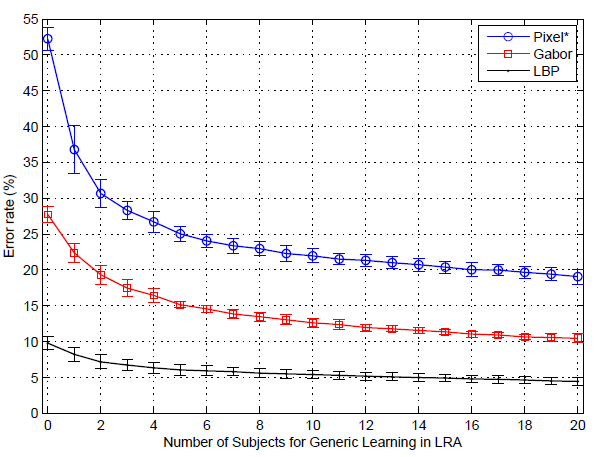


图9：LRA-GL的平均错误率（和标准差）与用于通用学习的subject数成函数关系。

对于所有类型的特征，少量subjects的类内变异库足以大大降低错误率。 例如，使用基于像素特征的LRA-GL，从五个subjects中倾斜可以将错误率从52％降低到26％。

这一发现表明，一旦根据测试条件适当地设计了类内变异基，从少数学科中进行的通用学习就足以显着提高面部识别性能。

3.3. FERET Database: Large-Scale Face Recognition Under Complex Variations

第三组实验旨在测试LRA在大规模面部数据库上针对复杂面部变化的鲁棒性。 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。

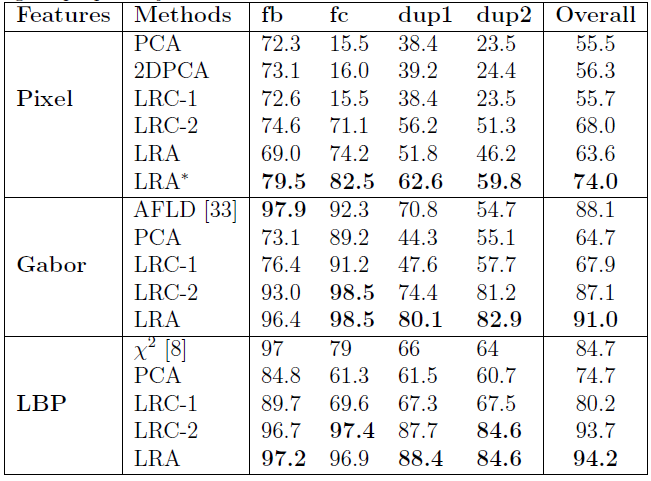
首先根据眼睛坐标对所有图像进行归一化和裁剪，然后按照第一组实验中的描述提取像素，Gabor和LBP特征。 图10显示了我们的实验中使用的一些裁剪后的图像，从图中可以看出，该数据库的类内探针-画廊变异性很复杂，因为它是在几年中的许多次会议中收集的。



图10：一些gallery图像和相应probe图像的裁剪图像。

表4列出了FERET数据库上的比较性能，结果再次表明LRA方法最适合同时处理表情，光照和年龄的变化，尽管使用LBP功能时LRC-2紧随其后。 此外，我们将LRA与其他一样本方法进行了比较，即自适应LDA [33]和加权LBP匹配[8]。 自适应LDA使用通用学习集构造类内散布矩阵，使用gallery集构造类间散布矩阵，这使其性能优于标准PCA。

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



加权方法表征了有关不同面部区域重要性的先验知识。 但是，它们的表现仍低于仅使用gallery进行训练的LRA。 该比较表明，LRA的等距嵌入为利用gallery集中的识别信息提供了优化的方法。

尽管LRA的准确性已经很高，但是将信息纳入通用训练集中可以进一步提高性能。 图11显示LRA-GL在12个测试用例（3个特征×4个探针组）中的11个病例中提高了LRA的识别率。 特别是，基于Gabor特征的LRA-GL在fc集合上具有99％的精度，是处理光照变化的最佳选择，而基于LBP特征的LRA-GL在dup1和dup2集合上具有90％的精度，是解决衰老影响的专家。

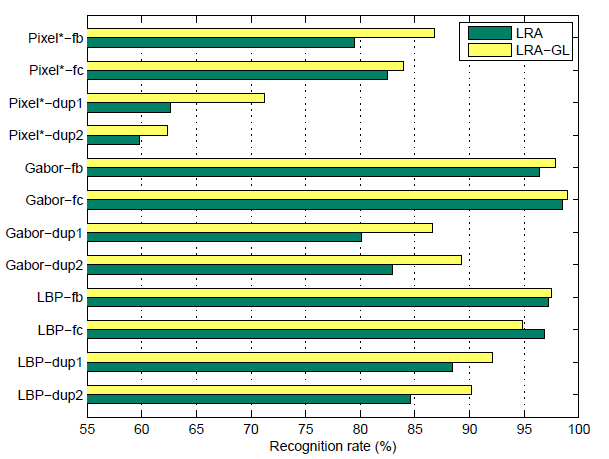


图11：使用三种类型的特征，在四个探针组上LRA和LRA-GL的相对FERET识别率。

在所有测试的功能中，dup1和dup2集的识别精度都有显着提高，这是在接近现实条件的不受控制的设置中获取的，这表明即使在复杂情况下人脸的类内变异性也是可以共享的。

3.4. Incremental Learning from Sufficient Samples

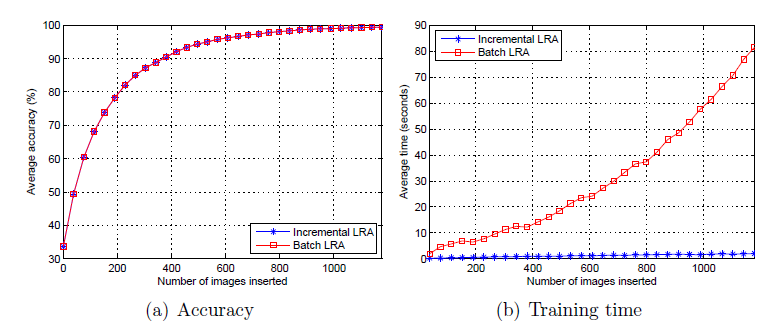
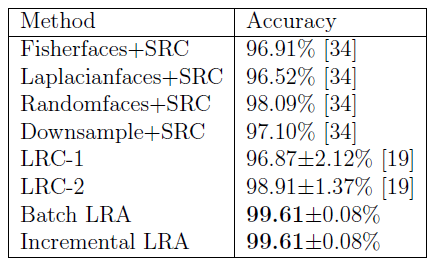


图12：Extended Yale B数据库上的batch-LRA和增量-LRA之间的比较。（a）准确性与插入训练样本数的关系；（b）训练时间与插入训练样本的关系。

表5:SRC和LRA在Extended Yale B数据库中的比较识别率，每个受试者使用32张训练图像。



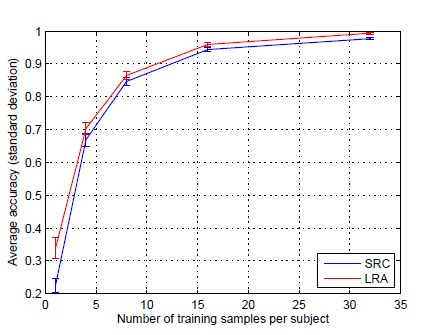


图13：使用每个subject不同数量的训练样本，在Extended Yale B数据库上比较SRC和LRA的识别率。

**4. Conclusion**

实验就人脸识别中的一个样本问题提出了若干结论：

1）PCA、2DPCA和LRC-1在不同的数据集上表现相似，因为它们本质上是基于受光照和遮挡变化影响的test样本和gallery样本之间的相关性。

2） LRA方法似乎是同时处理光照、表达式、遮挡和年龄变化的最佳方法，尽管LRC-2有时是紧随其后的第二种方法。从判别的角度看，等距原型是一个样本问题的最优嵌入。

3） 在具有挑战性的应用中，由于LRA不涉及test样本与gallery样本之间不可靠的相关性，因此在不同条件下采集gallery图像时，LRA的性能大大优于其他方法。

4） 基于局部特征的LRA比基于像素强度的LRA更有效，部分原因是Gram矩阵具有稳定的inverse eigenspectrum。

5） 在图像充分覆盖测试条件的情况下，无论使用何种特征，从5-10名受试者的类内变异基中进行泛型学习都能显著提高LRA的识别率。

此外，当从不同的数据库和相机设置收集通用数据时，LRA-GL仍然有效。

此外，虽然等距离嵌入的思想对各种类型的特征都是有效的，但是人脸描述符的选择对于解决复杂的人脸可变性是至关重要的。例如，LBP描述符在同时处理光照、表达式、遮挡和年龄变化方面是最好的。沃尔夫等人也做了类似的观察。[36]，他将几种新颖的LBP-like和Gabor-like描述子结合起来，在“野外有标签的人脸”（LFW）的非受控人脸验证问题上取得了有史以来最好的结果。虽然在多个基准数据集上取得了最佳结果，但LRA本身可能不足以解决现实世界中的一个样本问题。因此，我们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。