



论 文

# 基于光照补偿空间的鲁棒人脸识别

李月龙<sup>①\*</sup>, 孟丽<sup>②</sup>, 封举富<sup>③</sup>, 武继刚<sup>①</sup>

① 天津工业大学计算机科学与软件学院, 天津 300387

② 军事交通学院汽车指挥系, 天津 300161

③ 北京大学机器感知与智能教育部重点实验室, 北京 100871

\* 通信作者. E-mail: liyuelong@pku.edu.cn

收稿日期: 2013-01-20; 接受日期: 2013-03-17

国家自然科学基金 (批准号: 61302127、61173032)、天津市高等学校科技发展基金计划项目 (批准号: 20120805) 和国家重点基础研究发展计划 (批准号: 2011CB302400) 资助项目

**摘要** 如何应付光照条件变化是人脸识别的主要研究课题, 文中提出了一种基于单幅注册图像的光照不变人脸识别方法. 文中首先分析证明了同一个主体的不同光照图像之间的差别, 可以通过一个与身份和光照独立的子空间进行线性描述, 我们称之为光照补偿空间; 然后介绍了光照补偿空间的构建, 并且提出了一种简单有效的基于光照补偿的人脸识别算法, 将图像光照差异补偿与人脸识别过程统一在一起, 通过一次  $\ell_1$  优化就可以完成人脸身份的识别. 文中方法在没有明显增加优化复杂度的前提下, 有效降低了光照差异对于人脸识别算法精度的影响. 文中提出的方法在 Extended Yale B 图像库上进行了验证.

**关键词** 人脸识别 光照补偿 稀疏表示 线性子空间  $\ell_1$  优化

## 1 引言

识别出图像、视频中人脸的身份对于人类而言是与生俱来的基本技能, 除了极少数神经疾病患者, 绝大多数正常人从很小的年龄开始都具备这种能力, 这是人类作为高智慧生命的重要标志之一. 因此, 使人类所创造出的机器具备自动识别人脸的能力, 对于人工智能、人机交互研究而言就显得极为重要. 但这一对人类而言最基本的能力, 却是极难用机器进行模拟实现的. 迄今为止, 一方面, 在生物、神经学科等领域, 对于人类的神经系统是如何实现人脸识别仍无可靠的研究成果; 另一方面, 计算机相关领域对于自动人脸识别问题的研究涉及图像处理、计算机视觉、模式识别、人工智能等多个领域技术的交叉综合运用, 有无数理论和实际问题需要进行探索, 仍有许多科研难题虽然经过众多优秀科研人员几十年的共同努力, 仍然没有可靠的解决方案.

目前, 人脸识别中公认的主要难题是人脸图像的光照和姿态变化所造成的识别效果的降低<sup>[1]</sup>, 因为这两个因素的变化会直接导致获取到的图像间存在极大差异. 巨大的图像差异会使得测试样本与注册样本间的类内距离超过类间距离, 从而最终导致错误的分类识别结果 (模式识别研究中, 能进行正确分类的前提就是类内距离小于类间距离). 本文将主要针对如何减小光照对人脸识别的影响进行研究.

应对光照差异, 最直接的思路是应用各种图像处理技术减小图像差别, 例如, 直方图均衡化 (HE)<sup>[2]</sup>、

**引用格式:** 李月龙, 孟丽, 封举富, 等. 基于光照补偿空间的鲁棒人脸识别. 中国科学: 信息科学, 2013, 43: 1398-1409, doi: 10.1360/112012-380

均值为零方差为一的归一化处理<sup>[3]</sup>、离散余弦变换<sup>[4]</sup>等。但受光源位置、形状差异的影响,不同光照条件下获取的人脸图像千差万别,仅通过对单幅图像进行图像处理实际上很难有效减小光照变化的影响。从信息论角度,不同光照图像所包含的信息有很大差异,所以实际上不可能实现单幅图像上完美的归一化。因此,许多能够综合利用多幅图像上的有效信息的方法策略被设计了出来,基于子空间的方法就是其中最典型的代表。子空间的方法通过寻找高维人脸图像数据所分布的低维子空间,在有效降低数据处理规模的同时,揭示区别不同身份人脸的本质性因素,从而有效降低包括光照差异在内的各种非身份信息所造成的影响。依据分布形式(即流形)的不同,基于子空间的方法大致可以分为线性与非线性两大类。线性子空间方法假设原始高维数据存在线性低维分布结构,实现简单、速度快,可以有效降低数据的处理规模,在各种模式识别问题中被广泛采用。PCA(principal component analysis)<sup>[5]</sup>是最早也是最典型的线性子空间方法。PCA以最能保存原数据分布形式的特征向量作为基构建子空间,但PCA的设计初衷是降低数据的维数而不是探索发现数据内部的低维分布结构,因此,虽然PCA有良好的降维效果,但由于没有揭示能有效区分身份的信息,这一方法对人脸识别效果的改进十分有限。LDA(linear discriminant analysis)<sup>[6]</sup>则更为强调寻找能有效决定分类的子空间,因此,对人脸识别性能有更好的提升,是十分有效的线性子空间算法。相对线性子空间,非线性子空间方法假设数据的低维分布结构是非线性的,对数据分布形式有着更为灵活且精确的描述,此类方法也被称作流行学习<sup>[6]</sup>。LPP(locality preserving projections)<sup>[7,8]</sup>是非线性子空间方法在人脸识别问题上的一个有效应用,它以线性计算方式进行非线性的流形学习,大幅降低了算法复杂度。另一方面,将PCA、LDA等线性子空间方法利用核方法诱导到超高维数据空间<sup>[9]</sup>,可以使线性方法具有处理非线性数据的能力。基于稀疏表示的人脸识别(SRC)<sup>[10]</sup>,以不同光照条件下获得的Gallery图像为基向量直接构建线性子空间,虽然这种子空间的形式极为简单,但由于光照空间的结构本身较为简单,空间维数也较低,在Gallery图像充足的条件下, SRC对光照变化仍然具有一定适应能力。但对Gallery样本数量的依赖会限制该算法的实用性,因为在实际应用中,很难为每一个人注册时都进行多幅不同光照图像的摄取,而且算法的运算规模也会因此变大。

此外, Shashua 和 Riklin-Raviv<sup>[11]</sup>提出了利用熵图像合成各种光照下人脸图像的方法; Tan 和 Triggs<sup>[12]</sup>设计了一个综合光照预处理、局部纹理特征、核方法、多特征融合等各种策略识别不同光照人脸图像的方法; Chen 等<sup>[13]</sup>提出了基于加权最小二乘滤波器的图像光照渲染策略,该方法可以依照单幅目标图像的光照来渲染任意新的人脸图像,其核心思想是首先利用滤波器将源图像和目标光照图像都拆分为光照相关图和光照无关图两部分,然后依据目标图像的光照相关图修正源图像的光照相关图,最后将修正后的源图像光照相关图与其光照无关图进行组合获得新光照下的图像; Han 等<sup>[14]</sup>设计了基于光照分布估计的自适应图像光照预处理算法; Xie 和 Lam<sup>[15]</sup>将人脸图像的光源划分成 65 类,并按照源图像的光照所属类别进行光照补偿。

本文提出了一种基于光照补偿子空间的人脸识别方法,与以往大多数光照不变人脸识别算法不同,我们只需要在 Gallery 集合中为每个主体保存一幅标准光照图像,就可以进行任意光照条件下的人脸识别。本文内容的主要组织方式如下:首先证明通过一个与身份和光照独立的子空间可以弥补任意光照人脸图像与其对应标准光照图之间的差异;然后介绍这一光照补偿子空间的学习构建;最后提供一个简单有效的识别策略,使得通过一次  $\ell_1$  优化就可以直接识别出人脸图像的身份。

## 2 图像表示模型

本文所用人脸图像模型以 Lambertian 假设<sup>[16]</sup>为基础。按照 Lambertian 假设,图像是曲面反射

率 (由纹理决定)  $\rho(x, y)$  和曲面法向量  $\mathbf{n}(x, y)$  与光源方向  $\mathbf{s}$  的内积的乘积, 即

$$I(x, y) = \rho(x, y) \mathbf{n}(x, y)^T \mathbf{s}, \quad (1)$$

其中,  $0 \leq \rho(x, y) \leq 1$ ,  $\mathbf{n}(x, y)$  和  $\mathbf{s}$  均是三维方向列向量.

此外, 本文沿用文献 [11] 提出的理想物体的概念, 假定人脸是一类形状相同但曲面反照率不同的三维物体, 即, 公式 (1) 中的形状相关项  $\mathbf{n}(x, y)$  对于每一个人脸而言都是相同的.

在实际数据处理过程中, 通常会将二维图像数据按列相连转化为一维列向量进行处理, 即  $I(x, y) \rightarrow \mathbf{I}$ , 同样也可以将图像的反照率转化为列向量  $\rho(x, y) \rightarrow \boldsymbol{\rho}$ , 按照这种数据表示方式, 参照公式 (1), 任一主体  $p$  在任一光照  $\mathbf{s}_{\text{arb}}$  下摄取的人脸图像可以表示为

$$\mathbf{I}_{p, \text{arb}} = \boldsymbol{\rho}_p \circ (\mathbf{N} \cdot \mathbf{s}_{\text{arb}}), \quad \mathbf{N} = \begin{bmatrix} \mathbf{n}_1^T \\ \vdots \\ \mathbf{n}_W^T \end{bmatrix} \in \mathbb{R}^{W \times 3}, \quad (2)$$

其中,  $\circ$  表示 Hadamard 积运算,  $\mathbf{n}_i$  表示图像上第  $i$  个像素的曲面法向量方向,  $W$  为图像像素数目.

令标准正面光照为  $\mathbf{s}_{\text{nor}}$ , 则对于任意光照  $\mathbf{s}_{\text{arb}}$ , 必存在一个光照偏差  $\mathbf{s}_{\text{left}}$ , 使得有  $\mathbf{s}_{\text{arb}} = \mathbf{s}_{\text{nor}} + \mathbf{s}_{\text{left}}$ , 从而公式 (2) 可转化为

$$\begin{aligned} \mathbf{I}_{p, \text{arb}} &= \boldsymbol{\rho}_p \circ [\mathbf{N} \cdot (\mathbf{s}_{\text{nor}} + \mathbf{s}_{\text{left}})] = \boldsymbol{\rho}_p \circ (\mathbf{N} \cdot \mathbf{s}_{\text{nor}}) + \boldsymbol{\rho}_p \circ (\mathbf{N} \cdot \mathbf{s}_{\text{left}}) \\ &= \mathbf{I}_{p, \text{nor}} + \boldsymbol{\rho}_p \circ (\mathbf{N} \cdot \mathbf{s}_{\text{left}}), \end{aligned} \quad (3)$$

其中,  $\mathbf{I}_{p, \text{nor}}$  表示人脸  $p$  在标准正面光照下摄取的图像.

令  $\boldsymbol{\rho}_1, \boldsymbol{\rho}_2, \dots, \boldsymbol{\rho}_M$  为可以描述所有人脸反照率的子空间的基 (假设子空间的维数为  $M$ ), 从而对任一  $\boldsymbol{\rho}$ , 都存在一组系数  $\alpha_i$ , 使得  $\boldsymbol{\rho} = \sum_{i=1}^M \alpha_i \boldsymbol{\rho}_i$ ; 令  $\mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_3$  为光源方向向量描述空间的基, 从而任一光源  $\mathbf{s}$  都可以表示为  $\mathbf{s} = \sum_{j=1}^3 \beta_j \mathbf{s}_j$ ,  $\beta_j$  为对应系数. 引入这种线性子空间描述形式之后, 公式 (3) 的右边第二项变为

$$\begin{aligned} \boldsymbol{\rho}_p \circ (\mathbf{N} \cdot \mathbf{s}_{\text{left}}) &= \left( \sum_{i=1}^M \alpha_i \boldsymbol{\rho}_i \right) \circ \left[ \mathbf{N} \cdot \left( \sum_{j=1}^3 \beta_j \mathbf{s}_j \right) \right] = \sum_{i=1}^M \sum_{j=1}^3 \alpha_i \beta_j \boldsymbol{\rho}_i \circ (\mathbf{N} \cdot \mathbf{s}_j) \\ &= \sum_{i=1}^M \sum_{j=1}^3 \alpha_i \beta_j \boldsymbol{\rho}_i \circ (\mathbf{N} \cdot \mathbf{s}_j) = \sum_{i=1}^M \sum_{j=1}^3 (\alpha_i \beta_j) [\boldsymbol{\rho}_i \circ (\mathbf{N} \cdot \mathbf{s}_j)]. \end{aligned} \quad (4)$$

令  $a_{i,j} = \alpha_i \beta_j$ ,  $\phi_{i,j} = \boldsymbol{\rho}_i \circ (\mathbf{N} \cdot \mathbf{s}_j)$ , 则上式可表示为

$$\boldsymbol{\rho}_p \circ (\mathbf{N} \cdot \mathbf{s}_{\text{left}}) = \sum_{i=1}^M \sum_{j=1}^3 a_{i,j} \phi_{i,j}. \quad (5)$$

将上式由双下标表示等价转化为单下标表示:

$$\boldsymbol{\rho}_p \circ (\mathbf{N} \cdot \mathbf{s}_{\text{left}}) = \sum_t^{3M} a_t \phi_t. \quad (6)$$

由此可以得出结论, 任意光照下人脸图像与对应标准光照图像之间的差别可以通过一个线性子空间进行表示. 由于这一差别是由光照差异所造成的, 本文将这一子空间命名为光照描述空间. 由于

$\phi_t$  ( $t = 1, 2, \dots, 3M$ ) 只与人脸反照率基  $\rho_i$  ( $i = 1, 2, \dots, M$ ) 和光照基  $s_j$  ( $j = 1, 2, 3$ ) 相关, 而与具体人脸主体和光照条件无关, 因此, 这一光照描述空间的组成并不依赖于具体人脸和光照, 而是适用于所有人脸和光照情形, 这是我们能够通过学习的办法获取光照描述空间的理论依据.

将公式 (6) 代入公式 (3), 得到

$$\mathbf{I}_{p,\text{arb}} = \mathbf{I}_{p,\text{nor}} + \sum_t^{3M} a_t \phi_t. \quad (7)$$

由上式可以看出, 只要能够构建出光照描述空间  $\phi_t$  ( $t = 1, 2, \dots, 3M$ ), 在借助一组描述系数  $a_t$  ( $t = 1, 2, \dots, 3M$ ) 的情况下, 就可以利用标准光照下的人脸图像来重建对应人脸在任意光照条件下的图像. 这一性质可以被直接用来进行基于  $\ell_2$  的光照不变人脸识别: 输入图像  $\mathbf{I}_{q,\text{arb}}$  可以通过下述方法判断所属类别:

$$p^o = \arg \min_p \left( \min_{a_t, \phi_t} \left\| \mathbf{I}_{q,\text{arb}} - \mathbf{I}_{p,\text{nor}} - \sum_t^{3M} a_t \phi_t \right\|_2 \right). \quad (8)$$

### 3 光照不变人脸识别

在上一节我们已经获得了一种光照不变人脸识别的计算方法, 但由于实际上光照描述空间是未知的, 公式 (8) 的真正形式应该为

$$p^o = \arg \min_p \left( \min_{a_t, \phi_t} \left\| \mathbf{I}_{q,\text{arb}} - \mathbf{I}_{p,\text{nor}} - \sum_t^{3M} a_t \phi_t \right\|_2 \right). \quad (9)$$

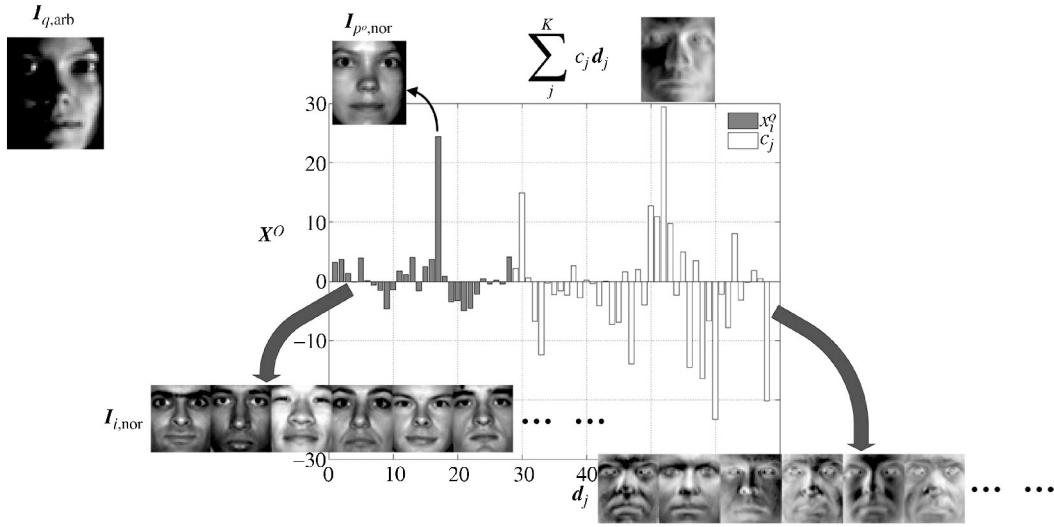
若令图像维数为  $N$ , 则上式中仅括弧内的优化, 就涉及  $(N+1) \times 3M$  个变量. 这是一个几乎不可能完成的任务, 因为变量的数目众多, 不可能得到唯一稳定解. 所以我们在原优化问题中加入了光照稀疏性, 即  $\ell_0$  限制, 使优化解趋于稳定,

$$p^o = \arg \min_p \left[ \min_{a_t} \left( \min_{\phi_t} \left\| \mathbf{I}_{q,\text{arb}} - \mathbf{I}_{p,\text{nor}} - \sum_t^{3M} a_t \phi_t \right\|_2 + \lambda \|\mathbf{a}\|_0 \right) \right], \quad (10)$$

其中,  $\mathbf{a} = [a_1, a_2, \dots, a_{3M}]^T$ ,  $\lambda$  是平衡  $\ell_2$  与  $\ell_0$  的权重系数.

引入光照稀疏性的合理性在于: 如第 2 节所述, 光照补偿空间  $\phi_t$  ( $t = 1, 2, \dots, 3M$ ) 由人脸反照率空间  $\rho_i$  ( $i = 1, 2, \dots, M$ ) 和光源方向空间  $s_j$  ( $j = 1, 2, 3$ ) 综合作用决定, 即光照描述空间实际包含可以描述所有人脸的信息, 因此, 此空间的维数不应该很低; 但如公式 (3) 所示, 在描述某一具体人脸图像时, 光照补偿空间的作用是补偿这个人的当前图像与其对应标准光照图之间的光照差异, 显然决定这一差异的决定性因素是光照而不是人脸身份 (从图 2 第二行列出的光照补偿图可以明显看到这一点), 而描述光照的只需要三维子空间, 因此光照补偿空间在进行单幅图像差异补偿时所真正用到实际子空间的维数可以很低, 文献 [17] 指出的, 进行单个人脸描述只需要很低维的子空间的结论, 也从另一个侧面证明了上述结论的正确性. 基于以上两点原因, 在利用过完备字典 [18] 表示光照补偿空间的前提下, 在原优化问题中加入光照稀疏性限制是合理的. 在具体计算时, 由于  $\ell_0$  优化是 NP 难问题, 优化速度缓慢, 按照惯例 [10] 我们利用  $\ell_1$  优化对其进行近似代替求解.

但如果直接依赖公式 (10), 在进行识别的同时, 优化得到过完备光照字典, 由于一次优化只涉及一个训练样本, 显然无法得到有效可靠的字典; 另一方面, 在每次进行识别时, 都同时进行光照描述空

图 1 某一输入样本对应的最优系数  $x^o$ Figure 1  $x^o$ , the optimal coefficient of an input sample

间的优化, 既增加了问题的复杂性又耗费了更多计算时间, 效率较低. 因此, 我们设计将光照描述空间的学习构建过程与人脸的识别过程分开, 首先进行光照空间的离线构建, 而后再进行人脸的在线识别. 具体方法: 令  $D_{\text{ill}} = [d_1, d_2, \dots, d_K]$  为先期训练得到的描述光照空间的过完备字典<sup>[19]1)</sup>, 字典维数为  $K$ , 则公式 (10) 描述的识别优化问题被转化为

$$p^o = \arg \min_p \left[ \min_{a_t} \left( \|I_{q,\text{arb}} - I_{p,\text{nor}} - D_{\text{ill}} a\|_2 + \lambda \|a\|_1 \right) \right]. \quad (11)$$

上式所描述的优化中, 主要的缺陷在于需要针对每一个库中标准脸分别进行一次优化, 时间复杂度较高. 另一方面, 同类样本必然更为相似, 所以在所有样本中, 使用同类样本进行相互表示时对应系数的  $\ell_1$  模最小, 将这一特点用于进行识别可以将所有库中标准脸综合在一次优化之中, 从而有效减少优化问题的规模. 因此, 我们将所有标准库中的人脸图像与光照描述过完备字典放在一起, 组成整体描述字典, 通过一次  $\ell_1$  优化<sup>[20]2)</sup> 确定输入样本的所属类别, 即, 求最优描述系数  $x^o$ ,

$$x^o = \arg \min_x (\|I_{q,\text{arb}} - Ux\|_2 + \lambda \|x\|_1), \quad (12)$$

其中,  $U = [I_{1,\text{nor}}, I_{2,\text{nor}}, \dots, I_{G,\text{nor}}, D_{\text{ill}}]$  为整体过描述字典;  $I_{i,\text{nor}}$  ( $i = 1, 2, \dots, G$ ) 表示已注册的每一个人所对应的标准光照图像;  $G$  为注册库中人脸数目; 令  $x^o$  中前  $G$  个对应于库中人脸的系数项为  $x_i^o$ , 后  $K$  个对应光照字典的系数为  $c_j$ . 依据公式 (12) 的优化结果, 输入样本的所属类别为所有  $x_i^o$  中最大项所对应的类别, 即

$$p^o = \arg \max_i (x_i^o). \quad (13)$$

图 1 展示了利用上述方法进行人脸识别时, 针对一个输入训练样本求得的最优系数  $x^o$  (图中一并列出发光照描述空间补充的光照  $\sum_j^K c_j d_j$ , 以及整体描述字典中某些列所对应的图像  $I_{i,\text{nor}}$  和  $d_j$ ).

1) 本文算法中, 光照描述字典  $D_{\text{ill}}$  的训练学习采用文献 [19] 的基于 K-SVD 的学习算法.

2) 本文算法的  $\ell_1$  优化, 采用文献 [20] 的 Large-Scale  $\ell_1$  优化算法.



图 2 测试样本图像与对应光照描述, 图中第一行为测试样本图像; 第二行为光照描述空间补充的测试样本与标准光照图像之间的差异  $\sum_j^K c_j d_j$

**Figure 2** Probe face images and corresponding illumination descriptions, where the first row consists of probe images, while the second one shows  $\sum_j^K c_j d_j$ , the differences between probe and corresponding normal illumination images complemented by illumination description space

#### 4 实验

我们在 Extended Yale B 人脸图像数据库<sup>[21,22]</sup>上实验本文提出的光照不变人脸识别算法. 我们的实验数据集包含 38 个人的 2414 幅正面人脸图像, 图像大小  $192 \times 168 = 32256$ , 分别拍摄于 64 种不同光照条件下, 其中摄像机镜头与光源的水平角度偏差由左至右从  $-130^\circ$  渐变至  $+130^\circ$ , 垂直角度偏差由上至下从  $-40^\circ$  渐变至  $+90^\circ$ . 由于原始图像维数较高, 直接进行处理优化的时间复杂度极高, 通常的解决方法是首先对图像进行下采样, 因此是否在下采样后仍然能达到较好的识别效果, 也是关系到本文算法是否实用的重要因素. 我们实验中使用的下采样率有  $1/4$ ,  $1/8$ ,  $1/16$ , 对应的图像数据维数分别为 2016, 504, 126.

在实验中, 为了减小图像之间由于整体亮度不同所带来的差异, 所有图像被首先标准化为均值为 0, 方差为 1. 所有实验被重复 10 次, 每次实验的训练集和测试集都是随机进行划分, 所有下文所涉及的识别率, 均为这 10 次随机实验的结果的平均值.

在光照描述字典构建阶段, 我们分别随机选取 3, 5, 8, 10 个人的所有光照图像进行训练, 其余图像全部用于进行识别实验. 利用 K-SVD 算法学习构建字典, 需要指定两个参数,  $L$ : 进行数据描述时所允许使用的最大字典元素数 (列数);  $K$ : 所构建字典的元素数目. 本文实验中, 使用的  $L$ 、 $K$  值组合共 8 种:  $[10\ 20]$ ,  $[10\ 30]$ ,  $[10\ 40]$ ,  $[20\ 20]$ ,  $[20\ 30]$ ,  $[20\ 40]$ ,  $[30\ 30]$ ,  $[30\ 40]$ , 下文中涉及的识别率均是这 8 种情况的均值. 所有实验中参数  $\lambda$  的取值统一设置为 0.01.

在组成整体描述字典  $\mathbf{U}$  时, 为使每一幅库中标准图像在进行  $\ell_1$  选择时有相同的出现权重, 各个  $\mathbf{I}_{i,\text{nor}}$  的模必须一致; 另一方面,  $\mathbf{U}$  中各图像列  $\mathbf{I}_{i,\text{nor}}$  和光照描述列  $\mathbf{d}_j$  的模的相互关系, 并不影响算法的整体表现. 因此, 不失一般性, 所有  $\mathbf{U}$  中各列的模都被标准化为 1.

如第 3 节所述, 本文算法进行识别时, 只需要每一个人的一幅标准光照图像, 其余 63 种光照条件下的图像全部用于测试. 识别过程中, 输入样本图像与标准光照图像之间的光照差异由光照描述字典弥补. 图 2 展示了一部分测试样本图像及通过本文算法计算出的光照描述字典所表示的光照. 从图中可以看出, 本文算法计算出的光照, 可以在一定程度上反映测试样本的光照条件, 从而能够有效减小测试样本与库中标准光照图像之间的差别, 进而提高识别精度.

为了便于展示分析结果, 本文将所有 63 种光照按光源位置分成 6 组. 由于每幅图像的光照取决于拍摄时摄像机与光源的水平和垂直夹角, 本文的划分标准是两个夹角的绝对值中较大一个所落入的区间, 从组 1 至组 6 角度偏差逐渐增大, 相应识别难度也逐渐增加. 表 1 列出了各分组配置, 图 3 展示了各组中的图像示例.

图 4 至图 6 展示了本文算法在各光照组上的识别表现 (从左至右、从上至下四幅图像分别对应于由 3, 5, 8, 10 个人的图像组成训练集学习光照描述空间). 其中, SID(sparse illumination dictionary)

表 1 光照分组情况

	Table 1 The configurations of illumination grouping					
	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6
Light orientation scope	$0^{\circ}\sim 30^{\circ}$	$30^{\circ}\sim 40^{\circ}$	$40^{\circ}\sim 50^{\circ}$	$50^{\circ}\sim 70^{\circ}$	$70^{\circ}\sim 110^{\circ}$	$110^{\circ}\sim 130^{\circ}$
Image number	18	5	5	10	13	12



图 3 光照分组中的样本示例, 其中第 1 行至第 6 行分别对应于组 1 至组 6

Figure 3 Samples coming from each illumination group, where row 1 to 6 correspond to group 1 to 6 respectively

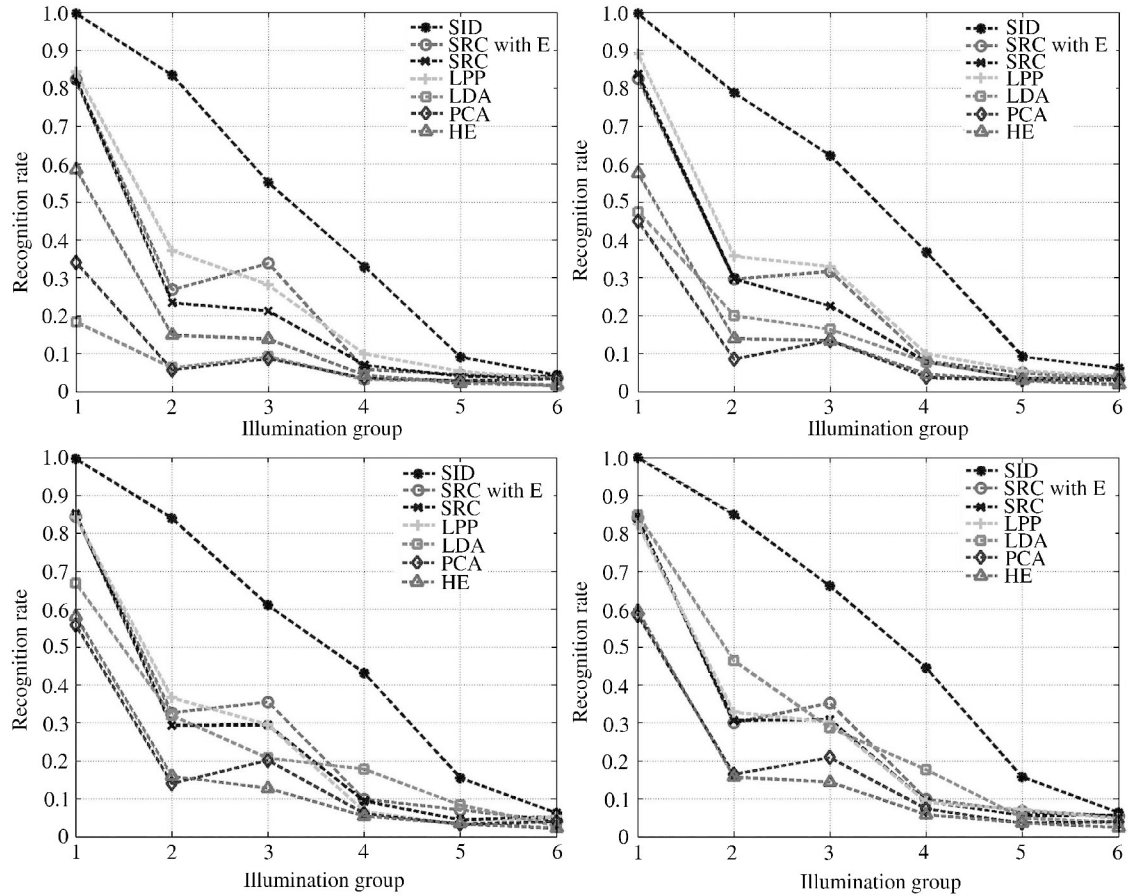


图 4 下采样率为 1/16 时本文算法的识别效果

Figure 4 The recognition performance under downsampling rate 1/16

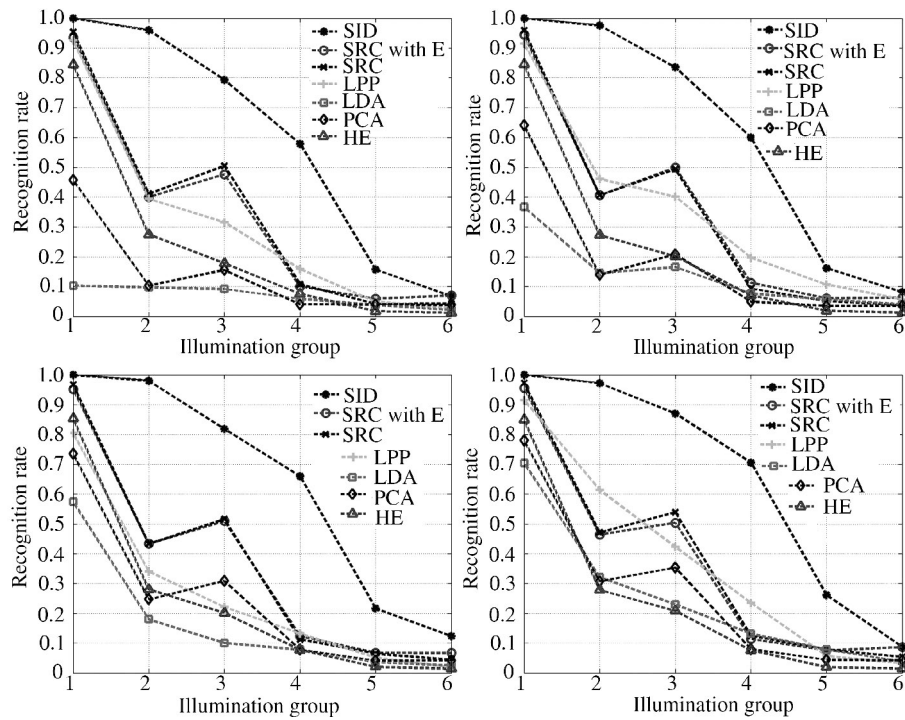


图 5 下采样率为 1/8 时本文算法的识别效果

Figure 5 The recognition performance under downsampling rate 1/8

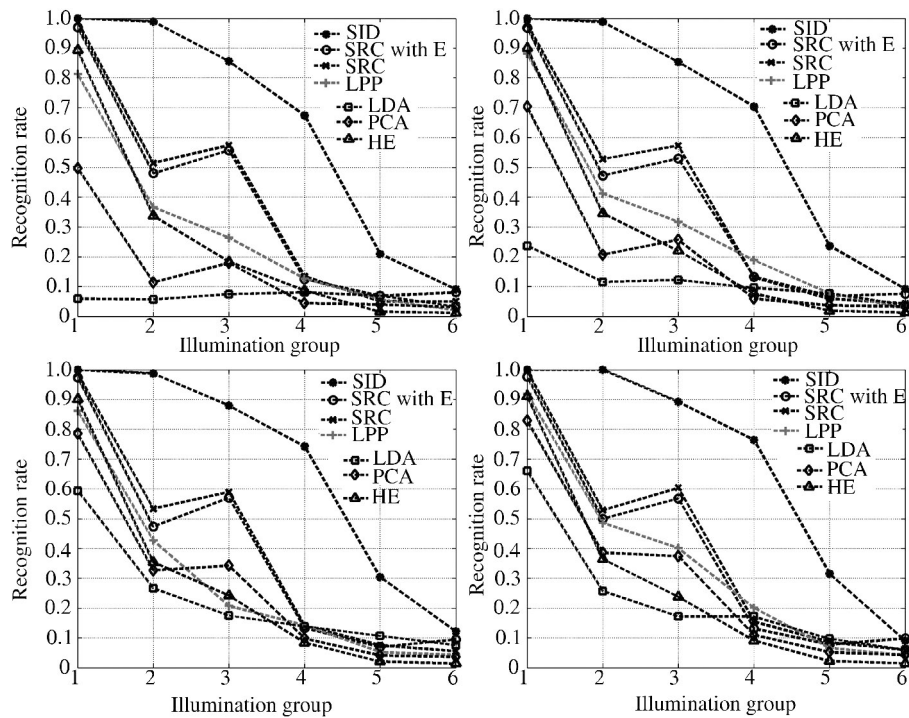


图 6 下采样率为 1/4 时本文算法的识别效果

Figure 6 The recognition performance under downsampling rate 1/4



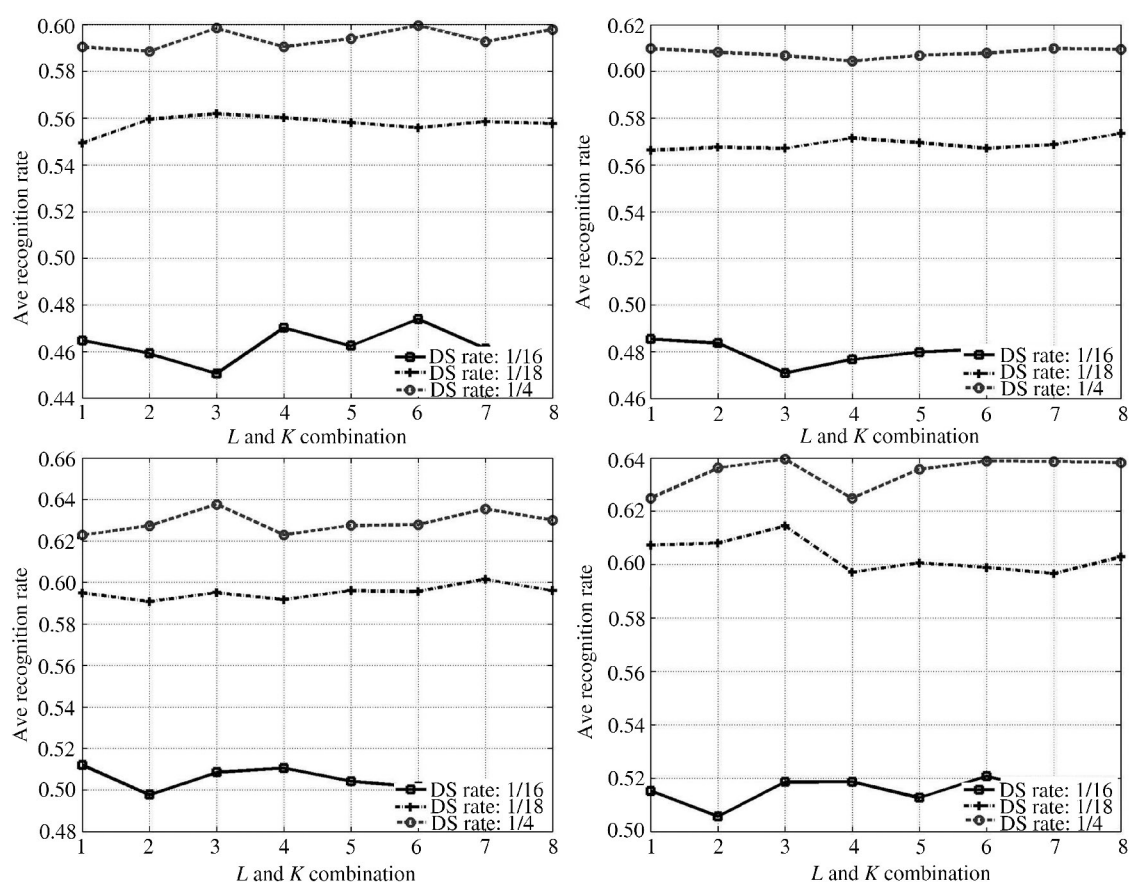
图 7 K-SVD 算法的  $L$ 、 $K$  参数取值对本文算法平均识别率的影响Figure 7 The influence of  $L$  and  $K$  parameters of K-SVD algorithm to the average recognition rate of our proposed strategy

表 2 识别错误率对比

Table 2 The comparison of false recognition rate

Algorithm	错误率 (%)					
	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6
DCT[4]+SID	0.51	0.89	1.14	1.43	2.03	2.31
DCT[4]	2.71	3.01	3.57	4.27	5.46	5.99
BHE <sup>[15]</sup>	3.52	5.71	6.31	7.23	8.41	9.26
Quotient image <sup>[11]</sup>	3.74	5.89	7.21	9.38	16.26	20.28
EPF <sup>[13]</sup>	2.63	3.00	3.41	4.03	5.12	5.21

代表本文算法, SRC with E 表示带错误描述矩阵的 SRC 算法, 其他英文简写含义参见第 1 节。

从各组图像中可以看出, 用于训练光照描述空间的训练样本图像数目对于最终识别效果没有明显的影响, 这充分说明光照描述空间的维数本身较低, 用较少训练图像就可以学习得到比较精确的光照空间 (这与被广泛接受的人脸身份和光照子空间维数本身都比较低的结论<sup>[17]</sup>是一致的)。各光照组

中, 识别率随下采样率改变的方差分别为 0.15%, 8.92%, 14.0%, 17.0%, 7.19%, 2.22%, 属于较低水平, 说明本文算法受下采样率影响较小, 在较高下采样率下运行仍可保持一定精度.

本文算法中需要人工指定 2 个参数: K-SVD 算法的  $L$  和  $K$ . 这两个参数取值是否直接影响识别性能是决定本文算法应用价值的重要因素 (过于依赖参数配置会限制算法的可移植性). 图 7 展示了  $L, K$  参数取值对于本文算法平均识别率的影响 (从左至右、从上至下 4 幅图像分别对应于以 3, 5, 8, 10 个人为光照描述训练集). 从图中可以看出, 在所有训练集和下采样率情况下,  $L, K$  参数的取值对于本文算法识别性能的影响不超过 3%, 充分反映了本文算法对这 2 个参数的鲁棒性.

通过与基于频域变换的图像预处理策略相结合, 本文算法的识别精度可以得到进一步提升. 表 2 展示了本文提出的 SID 策略在经过 DCT 光照预处理<sup>[4]</sup>后的图像上的识别错误率. 从中可以看出 DCT+SID 方法可以达到良好的识别精度.

## 5 总结

本文的主要创新点有两个: 一、提出并证明了, 同一个人在不同光照条件下获取的面部图像之间的差异可以通过一个子空间进行线性描述, 而且此子空间的构成与此人的身份和当前图像光照条件无关; 二、提出了一种基于光照补偿的鲁棒人脸识别算法, 此算法只需要每个注册人的一幅标准光照图像就可以识别任意光照条件下的人脸图像.

## 参考文献

- 1 Zhao W, Chellappa R, Phillips P J, et al. Face recognition: a literature survey. *ACM Comput Surv*, 2003, 35: 399–458
- 2 Gonzalez R C, Woods R E. *Digital Image Processing*. 3rd ed. New Jersey: Prentice Hall, 2006
- 3 Viola P, Jones M J. Robust real-time face detection. *Int J Comput Vision*, 2004, 57: 137–154
- 4 Chen W, Er M J, Wu S. Illumination compensation and normalization for robust face recognition using discrete cosine transform in logarithm domain. *IEEE Trans Syst Man Cy B*, 2006, 36: 458–466
- 5 Belhumeur P N, Hespanha J P, Kriegman D J. Eigenfaces vs. fisherfaces: recognition using class specific linear projection. *IEEE Trans Pattern Anal*, 1997, 19: 711–720
- 6 Lee J A, Verleysen M. *Nonlinear Dimensionality Reduction*. New York: Springer, 2007
- 7 He X, Yan S, Hu Y, et al. Face recognition using Laplacianfaces. *IEEE Trans Pattern Anal*, 2005, 27: 328–340
- 8 He X, Niyogi P. Locality preserving projections. In: *Proceedings of Advances in Neural Information Processing Systems*. Cambridge: MIT Press, 2003
- 9 Gärtner T. *Kernels for Structured Data*. Singapore: World Scientific, 2009
- 10 Wright J, Yang A Y, Ganesh A, et al. Robust face recognition via sparse representation. *IEEE Trans Pattern Anal*, 2009, 31: 210–227
- 11 Shashua A, Riklin-raviv T. The quotient image: class-based re-rendering and recognition with varying illuminations. *IEEE Trans Pattern Anal*, 2001, 23: 129–139
- 12 Tan X, Triggs B. Enhanced local texture feature sets for face recognition under difficult lighting conditions. *IEEE Trans Image Process*, 2010, 19: 1635–1650
- 13 Chen X, Chen M, Jin X, et al. Face illumination transfer through edge-preserving filters. In: *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Providence, 2011. 281–287
- 14 Han H, Shan S, Qing L, et al. Lighting aware preprocessing for face recognition across varying illumination. In: *Proceedings of the European Conference on Computer Vision*. Berlin: Springer, 2010. 308–321
- 15 Xie X, Lam K M. Face recognition under varying illumination based on a 2D face shape model. *Pattern Recogn*, 2005, 38: 221–230

- 16 Szeliski R. Computer Vision: Algorithms and Applications. London: Springer, 2010
- 17 Li S Z, Jain A K. Handbook of Face Recognition. 2nd ed. London: Springer, 2011
- 18 Candès E J, Eldar Y C, Needell D, et al. Compressed sensing with coherent and redundant dictionaries. Appl Comput Harmon A, 2010, 31: 1–21
- 19 Aharon M, Elad M, Bruckstein A. K-svd: an algorithm for designing overcomplete dictionaries for sparse representation. IEEE Trans Signal Proces, 2006, 54: 4311–4322
- 20 Kim S J, Koh K, Lustig M, et al. An interior-point method for large-scale  $\ell_1$ -regularized least squares. IEEE J Sel Top Sig Proc, 2007, 1: 606–617
- 21 Georgiades A S, Belhumeur P N, Kriegman D J. From few to many: illumination cone models for face recognition under variable lighting and pose. IEEE Trans Pattern Anal, 2001, 23: 643–660
- 22 Lee K C, Ho J, Kriegman D. Acquiring linear subspaces for face recognition under variable lighting. IEEE Trans Pattern Anal, 2005, 27: 684–698

## Illumination compensation subspace based robust face recognition

LI YueLong<sup>1\*</sup>, MENG Li<sup>2</sup>, FENG JuFu<sup>3</sup> & WU JiGang<sup>1</sup>

1 School of Computer Science and Software Engineering, Tianjin Polytechnic University, Tianjin 300387, China;

2 Automobile Transport Command Department, Military Transportation University, Tianjin 300161, China;

3 Key Laboratory of Machine Perception (MOE), School of Electronics Engineering and Computer Science, Peking University, Beijing 100871, China

\*E-mail: liyuelong@pku.edu.cn

**Abstract** How to deal with illumination variation is an important research topic of face recognition. In this paper, an illumination invariant face recognition strategy is put forward to deal with single gallery recognition. We give a proof that the illumination distinction among face images of  $s$  specific people could be linearly described through a carefully constructed subspace that is independent of both appearance and illumination, which is called the illumination compensation subspace. And then the subspace construction approach is introduced. Finally, a straightforward yet effective illumination compensation based face recognition algorithm is proposed which unifies both the calculation of illumination compensation and face recognition process into once  $\ell_1$  optimization. In a word, our approach effectively reduces the influence of illumination diversity on face recognition accuracy without incurring evident increase in optimization complexity. The proposed method has been evaluated on the Extended Yale B dataset.

**Keywords** face recognition, illumination compensation, sparse representation, line subspace,  $\ell_1$  optimization



**LI YueLong** received his Ph.D. degree in computer science from Key Laboratory of Machine Perception (MOE), Peking University. He was once a visiting researcher in the Department of Computer Science, University of Otago, New Zealand, and a research intern in the Real 3D Group, Samsung Advanced Institute of Technology. He joined Tianjin Polytechnic University as a Lecturer, in 2012. His research interest includes face recognition, image synthesizing, shape modeling and analysis, and compressed sensing. He is a reviewer for a great deal of journals and conferences, such as TIP and Neurocomputing.



**FENG JuFu** received his B.S. degree in 1989 and Ph.D. degree in 1997 both in mathematics from Peking University. In 1992, he joined the Center for Information Science and State Key Laboratory of Machine Perception at Peking University. He is currently a Professor in Key Laboratory of Machine Perception (MOE), Department of Machine Intelligence, School of Electronics Engineering and Computer Science, Peking University. His research interest includes image processing, pattern recognition, machine learning and biometrics.



**MENG Li** received her B.E. and M.S. degrees from University of Jinan. She is currently a Lecture at Military Transportation University. Her research interest includes image processing, pattern recognition, and video analysis.



**WU JiGang** received the B.S. degree from Lanzhou University, China in 1983, and doctoral degree from University of Science and Technology of China (USTC) in 2000. He worked in the Center for High Performance Embedded Systems, School of Computer Engineering, Nanyang Technological University, Singapore from 2000 to 2009, as a research fellow. He joined Tianjin Polytechnic University in 2009, as Dean, professor of School of Computer Science and Software Engineering. His research interest includes reconfigurable VLSI design, hardware/software codesign, and combinatorial search.