

整合原始人脸图像和其虚拟样本的人脸分类算法

刘 梓¹ 宋晓宁^{1,2} 唐振民¹

(南京理工大学计算机科学与工程学院 南京 210094)¹ (江南大学物联网学院 无锡 214122)²

摘要 人脸识别作为最具吸引力的生物识别技术之一,由于会受到不同的照明条件、面部表情、姿态和环境的影响,仍然是一个具有挑战性的任务。众所周知,一幅人脸图像是对人脸的一次采样,它不应该被看作是脸部的绝对精确表示。然而在实际应用中很难获得足够多的人脸样本。随着稀疏表示方法在图像重建问题中的成功应用,研究人员提出了一种特殊的分类方法,即基于稀疏表示的分类方法。受此启发,提出了在稀疏表示框架下的整合原始人脸图像和虚拟样本的人脸分类算法。首先,通过合成虚拟训练样本来减少面部表示的不确定性。然后,在原始训练样本和虚拟样本组成的混合样本中通过计算来消除对分类影响较小的类别和单个样本,在系数分解的过程中采用最小误差正交匹配追踪(Error-Constrained Orthogonal Matching Pursuit, OMP)方法,进而选出贡献程度大的类别样本并进行分类。实验结果表明,提出的方法不仅能获得较高的人脸识别的精度,而且还具有更低的计算复杂性。

关键词 稀疏表示,贪婪算法,人脸识别,分类

中图分类号 TP391.4 文献标识码 A DOI 10.11896/j.issn.1002-137X.2015.5.059

Integrating Original Images and its Virtual Samples for Face Recognition

LIU Zi¹ SONG Xiao-ning^{1,2} TANG Zhen-min¹

(School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing 210094, China)¹

(School of Internet of Things Engineering, Jiangnan University, Wuxi 214122, China)²

Abstract As one of the most attractive biometric techniques, face recognition is still a challenging task. This is mainly owing to the varying lighting, facial expression, pose, and environment. In this sense, a face image is just an observation and it should not be considered as the absolutely accurate representation of the face. However, even in a real world face recognition system, it is difficult to obtain enough samples. The great success of sparse representation in image reconstruction triggers the research on sparse representation based pattern classification. Inspired by this, a sparse representation based classification method using category elimination and greedy search strategy was proposed for face recognition. First, we reduced the uncertainty of the face representation by synthesizing the virtual training samples. We applied an error-constrained orthogonal matching pursuit algorithm to exploit an optimal representation result of training samples from the classes by eliminating the category and the specific training samples. The final remaining training samples are used to produce a best representation of the test sample and to classify it. Then, we selected useful training samples that are similar to the test sample from the set of all the original and synthesized virtual training samples. Finally, we devised a representation approach based on the selected useful training samples to perform face recognition. Experimental results on five widely used face databases demonstrate that our proposed approach can not only obtain higher face recognition accuracy, but also has a lower computational complexity than the other state-of-the-art approaches.

Keywords Sparse representation, Greedy algorithm, Face recognition, Classifications

生物识别技术是模式识别领域的一个重要分支,人脸识别则是生物识别技术中最具挑战性和吸引力的一个研究方向。目前尽管人脸识别技术已经取得很多成就,但是这仍然是一项艰巨的任务。不同的光照条件、面部表情、姿势和遮挡^[1-3]等因素都会影响面部识别的效果和效率。为了解决这些问题,学者们做出了很多努力,比如 Chen^[4]等提出了光照

补偿的方法来进行人脸识别, Sharma^[5]等提出了姿态不变的虚拟分类来实现人脸识别。仔细观察可以发现,在人脸图像集中如果包含了足够多的不同的位姿、光照条件和不同表情的训练样本,那么将有利于分类进而会得到更高的识别率。但是不幸的是在实际应用中,由于采样条件的限制,不可能得到大量的训练样本。一个训练样本只是人脸图像的一个采

到稿日期:2014-05-29 返修日期:2014-09-19 本文受国家自然科学基金重点项目(90820306),国家自然科学基金(61100116, 61125305),江苏省自然科学基金(BK2011492),中国博士后科学基金(2011M500926),江苏省博士后科学基金(1102063C)资助。

刘 梓(1985-),男,博士生,主要研究方向为模式识别与智能控制、图象识别、计算机视觉, E-mail: sandylaublue@163.com; 宋晓宁(1975-),男,博士,副教授,主要研究方向为模式识别与智能控制、图象识别、计算机视觉; 唐振民(1961-),男,教授,博士生导师, CCF 高级会员,主要研究方向为智能机器人与目标识别、图象处理与模式识别、智能信息系统、故障诊断与容错技术等。

样,并不能反映其可能的变化。为了解决这个问题,基于小样本的人脸识别方法被相继提出^[6-8]。与此同时,构造虚拟样本的方法从另一个角度解决了训练样本匮乏的问题。

构造的虚拟训练样本,提供了更多的面部观测信息,减少了在人脸识别中的人脸数据的不确定性。例如,Tang^[9]等在原始的人脸图像中利用光流法获得不同表情的虚拟样本,Jung^[10]等人利用噪声来生成新的训练样本,Thian^[11]等使用简单的几何变换来构造新的人脸样本,Liu^[12]等人将每个单独的图像表示成平移样本张成的子空间。此外,构建虚拟样本的思路也被用于手写体数字识别^[13]。

稀疏表示方法在图像重建中取得了巨大成功^[14,15],基于稀疏表示的分类方法(Sparse Representation based Classification, SRC)从一个新颖的视角来解决分类问题^[7],在人脸识别和验证中取得了良好的效果。这些方法利用训练样本的线性组合来稀疏地表示测试图像,通过评估每类测试样本的表示能力来进行分类。文献^[16-19]率先使用了用训练样本稀疏地表示测试样本的方法。稀疏度是由线性组合的系数的 l_1 范数来评估的,一个较小的 l_1 范数意味着更高层次的稀疏。这里的稀疏指的是线性组合的系数等于或接近于零。文献^[16,17]中采用了多目标规划的方法获得线性组合的表示系数,但是具有非常高的时间复杂度。此外,同时获得最小偏差和最小 l_1 范数的算法几乎是不可能的。Zhang等人^[20]讨论了SRC模型中的协作表示(Collaborative Representation, CR)特性,并验证了在SRC模型中并不是稀疏性约束而是协作表示机制在分类中起到了重要作用。有了这个有趣的结论,可以适当弱化SRC模型中的强稀疏性约束,采用基于贪婪搜索策略的稀疏系数分解方法,在降低了复杂性的同时进一步提高了识别的精确度。

基于SRC的方法中,字典的优化问题是系数分解的基本问题,例如文献^[21,22]强化了字典间的鉴别性并用于分类,此外,人脸图像的全局与局部特性也是可以相互融合的,如文献^[23,24]。合理的虚拟样本不仅要求可以反映人脸图像表情和姿态的变化,也应该有助于获得较高的识别准确度。在本文中,提出的算法试图根据已有的训练样本来构造和合理利用新的虚拟训练样本来完成人脸识别。为此,本文提出了在SRC框架下的人脸图像虚拟样本分类算法。首先根据已有的训练样本构造虚拟训练样本,其次在所有的训练样本中选择若干最具有表示能力的候选样本来协作表示测试样本,最后通过评估候选测试样本的表示能力来进行分类。

本文第1节概述了基于稀疏表示分类方法(SRC)的算法的基本框架思路;第2、3节详细阐述了文中提出的整合原始样本和虚拟样本的人脸图像分类算法;第4节是实验结果和分析;最后为结论。

1 SRC 算法概述

假设存在 L 类共计 n 个训练样本,为方便运算,训练样本记作 n 个列向量 x_1, \dots, x_n 。测试样本 y 可以近似地表示为所有训练样本的线性组合,即:

$$y = X\alpha \quad (1)$$

式中, $X = (x_1, \dots, x_n)$, $\alpha = (\alpha_1, \dots, \alpha_n)^T$ 。对于式(1),在SRC

框架下更为一般的解向量可以表示为:

$$\hat{\alpha} = \arg \min_{\alpha} \|\alpha\|_0 \quad \text{s. t.} \quad \|y - X\alpha\|_2 < \epsilon \quad (2)$$

式中,正常数 ϵ 用来平衡信号 y 的编码误差和稀疏度量 α 。由于 l_0 问题是NP问题,因此更常用的是 l_1 范数约束,如下所示:

$$\hat{\alpha} = \arg \min_{\alpha} \|\alpha\|_1 \quad \text{s. t.} \quad \|y - X\alpha\|_2 < \epsilon \quad (3)$$

从式(1)中可以看出,在表示测试样本时,每一个训练样本都有了自己的贡献,而且第 i 个训练样本对于表示测试样本 y 的贡献可记为 $\alpha_i x_i$ 。得到了每一个训练样本的贡献后,可以计算出每一类训练样本的贡献。例如,所有来自第 k 类的训练样本的集合为 x_1, \dots, x_k ,则第 k 类训练样本的贡献总和为 $g_k = \alpha_1 x_1 + \dots + \alpha_k x_k$ 。全局的算法认为,偏离度 $e_k = \|y - g_k\|_2^2$ 越小,则第 k 类训练样本的贡献越大,进而测试样本 y 归为使 e_k 最小的那一类。

2 提出的算法描述

学者们一直在研究SRC模型的鲁棒性,为什么能在分类问题中取得良好的效果。然而在第一节所述的分类问题中,训练样本 X 的各列是高度相关的,即各列都是人脸图像的向量排列,不能保证原子之间的相互独立性。这也就意味着在欠定线性方程(1)中,式(2)、式(3)不具备充分的条件来保证得到的解空间是 l_0 或 l_1 最优的。对此,文献^[16]提出了一个Cross-and-Bouquet模型来解释式(3)在分类问题中是可以被 l_1 恢复的,这就保证了SRC模型的良好性能。此外,在SRC模型分类过程中,协作表示(CR)^[20]机制发挥着重要的作用。因此,可以从两个方面来提升SRC模型分类性能,即字典优化和协作表示。

2.1 虚拟样本的构造

每个训练样本是对真实人脸数据的一次采样,但是在采样的过程中,由于受到光照、位姿和表情的不同的影响,导致训练样本并不能反映真实人脸的各种变化信息。也就是说有限的训练样本并不能真实地反映面部不同的变化。一种合理的应对这种数据不确定性问题的方式就是利用现成的有限个训练样本来合成虚拟的样本,从而能够反映出面部更多可能的变化趋势。

但是在这里,要讨论如何构造合理的虚拟样本。首先,合成出来的虚拟样本如果与已有的训练样本太过相似,则不能体现出虚拟样本的作用,也失去了增加人脸图像可能变化趋势的意义。另一方面,如果合成出来的虚拟样本与已有的训练样本差别太大,可能会导致合成出来的虚拟样本与其余类别的样本相似度很高,从而对随后的分类产生副作用。综合这两方面的考虑,一个好的虚拟样本应该和对应的实际训练样本有一个适当的差异。

到底什么是适当的差异,在不断尝试中发现通过对现有训练样本之间的加权平均构造出来的虚拟样本,不仅可以提供人脸样本可能的变化趋势,还可以保证这些变化不太剧烈。

假设来自某个类别的 l 个训练样本集合为 X_l ,即 $X_l = [x_l^1, \dots, x_l^l]$ 。如果 x_l^i 和 x_l^j 是来自 i 类别的两个真实训练样本,那么对应的虚拟样本可以表示为

$$x_l^v = \omega_j (x_l^i + x_l^j) \quad (4)$$

式中, ω_j 为加权系数。如果每两个训练样本合成一个新的虚拟样本,且加权系数的组合个数为 N ,则可以产生 $N \times C_2^2$ 个新的虚拟样本。在本文的实验中, ω_j 的权值选择为 $\omega_j = [1/4, 1/2, 3/4, 1]$,所有的真实训练样本和合成的虚拟样本构成了新的训练样本集合。

2.2 协作表示能力优化

虽然来自所有类别的训练样本 X 都可以用来设计分类算法,但是所有训练样本对最后的分类决策并不提供相同程度的影响。此外,从数值计算的角度来看,增加了虚拟的样本并没有增加训练矩阵的秩,因此没有必要把所有的训练样本都用来表示测试样本。

某些与测试样本相似度高的类别对分类决策具有更高的影响,那些距离测试样本很远的样本则具有很少的影响,甚至会对分类的决策产生副作用。所以,在分类决策之前,第一步就是去除若干对测试样本贡献度较少的类别或样本,然后在剩余的类别或样本中再做分类决策。从而,可能会获得更高的分类精确度。在通常的分类问题中,类别越多,可能的最大分类精确度就会越低。

在本文提出的算法中,预先剔除了距离测试样本较远的若干样本。由于候选类样本减少,因此分类算法的效率和精度将会增加。从协作表示的角度来看,当剔除了若干离群样本后,真正类别的表示权值将会增大,保证了后续分类的准确性。为此,在得到真实样本和虚拟样本的混合样本集合以后,首先会剔除距离测试样本较远的样本,为了简便起见,以欧氏距离为准则,一次性剔除若干个训练样本,仅保留 K 个训练样本作为候选类别。这样做的好处是,既能快速剔除掉对分类影响不大的类别,也能避免由于训练类别中某些单个样本由于受到光照、姿态和遮挡等情况的影响,减少了对测试样本的贡献度的这种情况。

如果剩余类别所有的训练样本记作 x_1, \dots, x_t ,则第 k 个类别表示测试信号 y 的贡献程度为 $g_k = b_1 x_1 + \dots + b_t x_t$,最小的 $e_k = \|y - \sum_{i=1}^t b_i x_i\|^2$ 表示第 k 个训练样本在表示测试样本 y 时贡献最大,其中 p 表示第 k 个类别中样本的个数。因此,将测试样本 y 归为偏差最小的那一类。

2.3 稀疏系数的求解

许多文献,比如文献[20],已经验证了在 SRC 模型中是协作表示机制在分类中起到了重要作用,但是稀疏性约束并不能简单地被移除。从式(3)中可以看出,测试样本 y 不仅可以由训练样本精确地表示出来,而且在基于字典 X 的分解系数也是稀疏的。比如,类别 i 和类别 j 是很相似的,对于测试样本 y 来说重建误差可以分别表示为 $e_i = y - X_i a_i$ 和 $e_j = y - X_j a_j$ 。由于类别 i 和类别 j 是很相似的,因此 e_i 和 e_j 的距离也很近,这就给分类带来了困难。如果在表示信号 y 的时候加入了稀疏性度量,假如 y 是 i 类别的,那么在 i 类别中应当可以用较少的样本个数来表示信号 y 。同样,假如 y 是 i 类别的,在相同的稀疏约束的情况下, i 类别的重建误差应该更小。

从前面的分析可知,本文算法的目标是在训练样本集合中先找到若干个最接近于测试样本的近邻,然后用这些候选

训练样本来表示测试样本。对于给定的测试样本 y 和训练集合样本 X ,某个训练样本的贡献越大,则说明测试样本在该训练样本的基空间中的投影越大。因此,计算每个训练样本与测试样本 y 之间的内积,选择最大的内积值 a_i 作为对应训练样本 x_i 的表示系数,将残差 $r = y - a_i x_i$ 作为下一次迭代的输入,继续计算残差在训练样本集下面的表示系数,直到满足预先设定的阈值为止。这种搜索计算策略可以将贡献较大的训练样本逐个找出来。因此,本文引入了误差最小的贪婪搜索算法(Error-Constrained Greedy Search Method)来解决式(3)的问题。众所周知,解决式(3)的问题通常有两种方式,一种是贪婪算法,另一种是松弛算法。松弛法将 l_0 约束松弛为 l_1 约束来求解,而贪婪算法在每次迭代的时候都会对某一个非零向量求解。目前来说最常用的稀疏分解包括 Lasso(Least Absolute Shrinkage and Selection Operator)^[25]、Lars(Least Angle Regression)^[26]、匹配追踪(Matching Pursuit, MP)^[27]和正交匹配追踪(Orthogonal Matching Pursuit, OMP)^[28]等方法。Lasso、Lars 这类方法将稀疏逼近问题转化为凸优化问题,凸优化理论中有大量的算法来完成求解。MP 和 OMP 属于贪婪搜索算法,这种方法的特点是通过不停地迭代来直接逼近原始信号,期望通过局部最优解的累加来得到全局最优解。然而, l_1 范数只关注向量的累加值,而不关心非零向量的个数,所以一个向量如果有很多非零位置且很小的系数,虽然有较小的 l_1 范数,也会有很大的 l_0 范数。因此,求解式(3)时,对于分解系数可以不用那么强的 l_1 约束。到目前为止,如何平衡稀疏性和协作表示的关系依然是研究热点。

为了便于标记,文中将贪婪搜索策略记作 $\alpha = EcOMP(D, y, \epsilon)$,其中 y 为某个测试样本, D 为训练样本构成的字典。 $EcOMP(D, y, \epsilon)$ 的返回结果为测试样本信号 y 在训练样本 D (即字典 D)下的稀疏系数,且重建误差为 ϵ 。

2.4 算法框架示意图及详细流程

综上所述,算法的详细运算流程如下。

目标任务:对于给定的测试样本 y ,通过构造虚拟的训练样本及剔除离群样本来更好地计算表示 y 的系数,解决分类问题。

步骤1:初始化。假定来自 L 类共计 n 个训练样本,将每个训练样本的初始二维矩阵图像转化为一维列向量 $x_i \in R^{q \times 1}$ (q 为训练样本的维数),并利用式(4)来构造新的 m 个虚拟样本。每个列向量为一个原子,则由 $n+m$ 个原子组成的矩阵 $X \in R^{q \times (m+n)}$ 即是要构建的混合字典。

步骤2:标准化。将字典 X 的每一列原子标准化。

步骤3:计算测试样本和训练样本之间的欧氏距离,距离较大的若干类先被剔除,保留 K 个候选训练样本,完成原子矩阵的更新。

步骤4:使用贪婪搜索的策略完成稀疏系数的分解。在更新后的字典 X' 下计算测试样本 y 的分解系数 $\hat{\alpha} = EcOMP(y, X', \epsilon)$ 。

步骤5:计算残差。如果剩余的所有类别记作 x_1, \dots, x_t ,则可以计算每个类别 r 对于测试样本 y 的贡献程度 $g_r = b_1 \hat{\alpha}_1 x_1 + \dots + b_t \hat{\alpha}_t x_t$, g_r 与测试样本 y 的偏差可由 $D_r = \|y - \sum_{i=1}^t b_i \hat{\alpha}_i x_i\|^2$

($r \in C$) 计算得到, 其中 p 表示第 k 个类别中样本个数。

步骤 6: 结果输出。测试样本 y 归为偏差最小的那一类, 即 $Identify(y) = \arg \min_r \{D_r\}$ 。

算法框架示意图如图 1 所示。

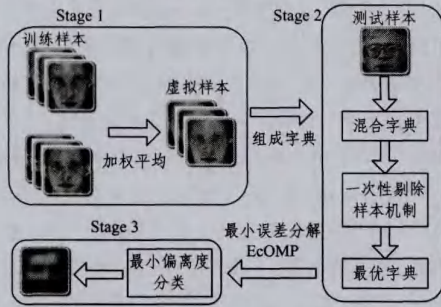


图 1 算法框架示意图

3 算法分析

本节将讨论前文所提出算法的特点、基本原理和潜在的优点。此方法的基本特征是用训练样本的子集来线性表示测试样本, 并利用最优表示结果来执行分类。这里的“最优”是指测试样本与线性表示结果的误差最小。下文将做进一步的分析 and 广泛讨论。

3.1 算法的基本原理

在 SRC 模型中, 系数分解采用迭代求解方式获得, 大部分的非负系数等于或趋近于零。但是在实际的人脸识别场景中, 由于一些遮挡残缺或是干扰像素的影响, 导致得到理想的分解系数是很困难的。因此, 本文提出的算法遵循以下两个关键点: 一是协作表示的机制对分类的重要影响, 二是系数分解时的稀疏约束条件是必要的。

为此, 首先利用所有的训练样本构造虚拟样本, 来尽可能地描述样本可能存在的变化趋势。其次, 为了增强协作表示的能力, 一次性剔除了若干对分类影响较小的训练样本。最后在求解稀疏系数的过程中使用了基于贪婪搜索策略的弱 l_1 优化求解算法。

3.2 算法的稀疏性描述

算法的基本思路是将测试样本表示为训练样本的线性组合方式, 如式(1)所示。训练样本的集合则为稀疏表示的字典, 线性表示的系数则为测试样本在字典下的分解系数。分解系数的稀疏性描述来自两个方面: 一方面, 测试样本在字典的表示下, 也就是用所有训练本来精确表示测试样本的线性组合中, 一些训练样本的系数等于或接近于零; 另一方面, 在预处理的过程中, 剔除距离测试样本较远的类别或样本。被剔除的类别和样本, 在最后的系数分解中并不参与运算, 所以在重建误差计算结果中, 被剔除的训练样本的系数为零。

3.3 剔除策略的概率说明

SRC 算法的分类规则可以描述如下: c_1, \dots, c_L 分别表示 L 个类别, 则概率 $p(c_i | y)$ 表示测试样本 y 和第 i 类训练样本来自同一类的概率。如果测试样本和第 i 类训练样本的距离记作 $dist_i$, 那么后验概率 $p(c_i | y)$ 可以被定义为 $p(c_i | y) \propto (1 - dist_i / \sum_{j=1}^L dist_j), i=1, 2, \dots, L$ 。距离 $dist_i$ 可以用 $e_i = \|y - A_i b_i\|^2$ 描述, 那么后验概率 $p(c_i | y)$ 可以被计算得到 $p(c_i | y) \propto D_{\max} - \|y - g_i\|^2, c_i \in C, D_{\max}$ 表示一个数量大于最大的

$\|y - g_i\|^2$ 的值。

在算法中, 预先剔除了若干距离测试样本比较远的类别和样本, 也就是说这些被剔除类别的 $p(c_i | y) = 0$ 。因 $\sum_i p(c_i | y)$ 是一个常量, 故在剩余的 M 个类别中重新计算的后验概率 $p(c_i | y)$ 将会增加。距离测试样本越近的类别, 其后验概率值将会越大。通过剔除类别和重新计算后验概率的迭代过程, 一方面去除了距离比较远的类别的干扰, 另一方面加强了剩余类别的权值。这种方式可以很好地处理初始时各类别的后验概率相差不大的情况。随着迭代的进行, 重新计算的后验概率会拉开差距。最后, 将测试样本的类别标记为具有最大后验概率的类。

3.4 贪婪搜索策略

MP 是一种常用的贪婪搜索算法, 它每次都会寻找一个最佳匹配的原子, 在计算完残差后继续寻找下一个最匹配的原子, 直到满足预先设定的重建阈值。在 SRC 模型和 MP 算法中, 可以看出相似点和互补的地方。首先, 它们都要寻找最能有效表示信号的原子, 此外, 垂直投影是 MP 的主要步骤, 而 SRC 使用了类似的思路。鉴于此, 本文算法中并没有使用经典的松弛优化方法, 而是使用了贪婪的方法, OMP 是对 MP 方法的改进, OMP 在分解的每一步对所选择的全部原子进行正交化处理, 这使得在精度要求相同的情况下, OMP 算法的收敛速度更快。在第 4 节的实验中, 将会在 ORL 数据库上对比 Lasso、Lars、MP 和 EcOMP 几种方法的识别率和计算时间。

4 实验结果

在本节中, 将本文提出的算法在 ORL^[29]、FERET^[30] 和 AR^[31] 人脸图像数据库上进行了大量的实验。这 3 个人脸图像数据库中的照片是在不同的姿态、面部表情和照明条件下采集的。在 AR 数据库中甚至还包含面部被遮挡的人脸样本。ORL 数据库共计 400 张图片, 分别来自 40 个类别, 每个类别提供 10 张样本图像。ORL 人脸数据库中的部分样本见图 2。



图 2 ORL 数据库中部分的标准图像

FERET 人脸数据库已成为检验面部识别技术的一个标准数据库, 在文中只使用了 FERET 数据库的一部分图像, 共计 1400 张图片, 分别来自 200 个类别, 每个类别提供 7 张样本图像。在文中使用的 FERET 数据库的每个类别的 7 张样本中, 分别包含了 1 张标准的正面照片, 1 张表情变化的正面照片, 1 张光线变化的正面照, 2 张小幅度姿态变化的样本 ($\pm 15^\circ$), 2 张大角度姿态变化的样本 ($\pm 25^\circ$)。FERET 人脸数据库中的部分样本见图 3。



图 3 FERET 数据库中部分的标准图像

AR 人脸数据库中共计使用了 3120 张图片,分别来自 120 个类别,每个类别提供 26 张样本图像。AR 数据库中的人脸样本表情变化丰富,有墨镜和围巾的遮挡。AR 人脸数据库中的部分样本见图 4。



图 4 AR 数据库中的部分的标准图像

在实验部分,将文中提出的算法与目前主流的基于整幅人脸图像的识别算法做了对比,包括经典的主成分分析算法、线性鉴别分析、K 近邻、协作表示的分类算法(Collaborative Representation Classification, CRC)、传统的 SRC 分类、线性回归分类(Linear Regression Classification, LRC)。首先将来自 ORL 数据库中的人脸图像降采样为 46×56 的大小,同样地,FERET 和 AR 人脸数据库图像分别降采样为 40×40 和 40×50 的大小。在 AR 数据库的实验中,选取每个类别的前 8 个图像作为训练样本,剩余的作为测试样本。因为 AR 数据库的样本太多,所以只做了一种组合方式。同样地,在 ORL 数据库中做了 3 组实验。每个类别中分别选取 4、5 和 6 个图像作为训练样本,剩余的图片作为测试样本。在 10 个样本中任意选择 5 个样本共有 C_{10}^5 种可能,在实验中随机选择了 10 种排列组合方式。在 FERET 数据库的实验中,每个类别中分别选取 4 个和 5 个图像作为训练样本,剩余的图片作为测试样本。在实验中也随机选择了 10 种排列组合方式。测试数据集的构成如表 1 所列。文中所有的实验均使用 Matlab 来实现,并在英特尔 I7 3770 (3.4GHz) 的 PC 机上运行。

表 1 3 个人脸测试数据集组成

数据库	类别数	类别样本数	训练样本数		实验次数(总数)
			样本总数	训练样本个数	
ORL	40	10	400	4,5,6	10(210), 10(252),10(210)
FERET	200	7	1400	4,5	10(35),10(21)
AR	120	26	3120	8	1

如图 5 所示,现有训练样本之间的加权平均构造出来的虚拟样本,一方面可以提供人脸样本可能的变化趋势,另一方面可以保证这些变化不太剧烈。在式(4)中 ω_j 为加权系数,在实验中, ω_j 的权值选择为 $\omega_j = [1/4, 1/2, 3/4, 1]$,所有的真实训练样本和合成的虚拟样本构成了新的训练样本集合。



图 5 FERET 数据库中的 5 个原始样本(第一行)和构造的虚拟样本

图 6 为来自 FERET 数据库的测试样本图像以及对应的

4 个类别的重构图像。其中,图(I)为测试样本,图(II)~图(V)为所有训练样本中贡献最多的前 4 类样本的 2 维重构图像。(a)部分图像是采用传统的 SRC 算法得到的实验结果,(b)部分是由文中提出的合成虚拟样本算法得到的实验结果。可以看出,在传统的 SRC 方法中,正确类别的样本提供了较多的贡献,但是由于所有的测试样本都参与了运算,因此正确类别的样本的贡献度没有特别体现出来。但是在文中提出的方法中,通过构造虚拟样本,并去除距离较远的类别和单个测试样本,使得正确的类别样本的贡献程度在迭代运算中得到加强,所以容易得到正确的分类结果。

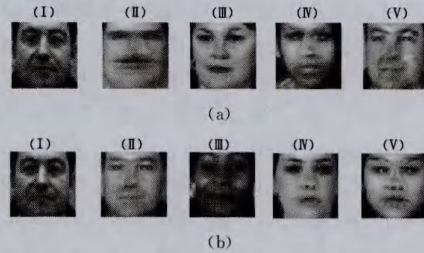


图 6 给定测试样本的 4 类最大贡献样本重构图像

图 7 示出在 ORL 数据库中不同训练样本的情况下, EcOMP、MP、Lasso 及 Lars 算法做稀疏分解时分类算法的最终识别率,其他参数都相同。图 8 为 EcOMP、MP、Lasso 及 Lars 算法做系数分解时的速度。从图中可以看出误差最小的贪婪搜索算法的有效性。

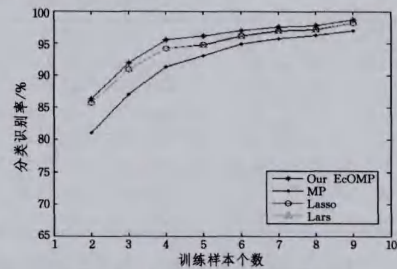


图 7 ORL 库下不同的稀疏分解方法在不同训练样本下的识别率

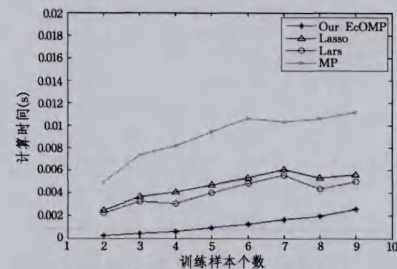


图 8 ORL 库下不同的稀疏分解方法在不同训练样本下的计算速度

根据已有的训练样本,可以构造出若干虚拟样本。将虚拟样本和原始样本组合成新的混合字典后,以欧氏距离为准则,一次性剔除若干个训练样本,仅保留 K 个训练样本作为候选类别。这样做的好处是,既能快速剔除掉对分类影响不大的类别,也能避免训练类别中某些单个样本由于受到光照、姿态和遮挡等的影响而减少了对测试样本的贡献度的情况。图 9 为在 FERET、ORL 和 AR 库上针对不同的 K 值做出的识别曲线图,其中 Feret4 表示在 FERET 数据库上 4 个原始训练样本。

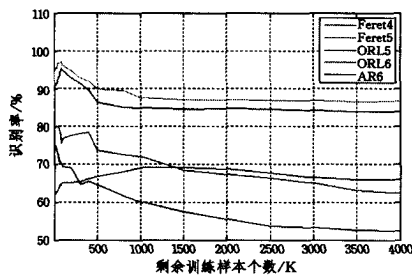


图9 FERET、ORL和AR库下不同的K值的分类识别率

此外,分别在ORL、FERERT和AR数据库上,将文中提出的算法与主流的基于整幅人脸图像的识别算法做了对比,其中包括经典的主成分分析算法(PCA)、线性判别分析(LDA)、K近邻、协作表示的分类算法(Collaborative Representation Classification, CRC)、传统的SRC分类、线性回归分类(Linear Regression Classification, LRC),如表2—表4所列。其中PCA(100)、PCA(150)表示PCA在做特征提取的时候分别用了100和150个变换轴;LDA(39)表示LDA在做特征提取的时候用了39个变换轴。Our为本文提出的算法。ORL数据库下的各种算法的分类识别率见表2, FERET数据库下的各种算法的分类识别率见表3, AR数据库下的各种算法的分类识别率见表4。

表2 ORL数据库上不同训练样本个数在不同方法下的分类识别率(%)的比较

方法	原始训练样本个数		
	3	4	5
PCA(100)	88.41±0.08	91.22±0.07	92.82±0.06
PCA(150)	86.72±0.06	90.63±0.06	92.14±0.07
LDA(39)	87.68±0.05	90.87±0.08	92.28±0.08
SRC	88.12±0.05	91.24±0.05	92.72±0.06
CRC	89.76±0.04	90.76±0.05	91.92±0.05
KNN(K=5)	89.83±0.03	90.83±0.03	92.05±0.02
LRC	90.27±0.07	91.27±0.06	91.76±0.06
Our	91.83±0.05	94.83±0.04	95.45±0.04

表3 FERET数据库上不同算法的识别结果比较

方法	原始训练样本个数	
	4	5
PCA(100)	69.54±0.08	72.35±0.07
PCA(150)	68.37±0.06	71.56±0.06
LDA(39)	67.68±0.05	71.87±0.08
SRC	70.65±0.08	75.24±0.05
CRC	70.81±0.06	76.14±0.05
KNN(K=5)	69.83±0.05	74.83±0.06
LRC	72.46±0.06	76.38±0.06
Our	74.06±0.04	77.08±0.05

表4 AR数据库上不同算法的识别结果比较

方法	PCA(200)	LDA(119)	SRC	CRC
识别率(%)	50.00	56.47	64.83	67.82
方法	KNN(K=5)	LRC	Our	
识别率(%)	56.38	57.59	69.81	

以上实验在稀疏系数的分解过程中,分别从识别率和计算速度方面对比了OMP、MP、Lasso和Lars算法。基于贪婪搜索的EcOMP算法由于其计算原理和快速收敛的性质,在各种方法中脱颖而出。构造的虚拟样本在一定程度上预测了样本可能存在的变化,将会对分类识别率起到积极的影响。表2—表4的数据表明,文中提出的方法在上述的方法中是最有效的。

结束语 本文提出了一种构造虚拟样本的方法,其在基于稀疏表示的分类框架下完成了人脸识别。首先,根据已有的训练样本通过加权平均的方式合成新的虚拟样本,并将虚拟样本和原始的样本组成新的训练字典;然后剔除若干个距离测试样本较远的训练样本;最后把测试样本表示成剩余训练样本线性组合的方式,旨在所有训练样本中消除对分类影响较小的样本,进而突出贡献大的样本在分类策略中的权值。相对于传统的PCA、LDA及其改进方法,本文工作可以说是一种新的思路和分类模式,具有较高的研究价值。

参考文献

- [1] Pishchulin L, Gass T, Dreuw P, et al. Image warping for face recognition: From local optimality towards global optimization [J]. *Pattern Recognition*, 2012, 45(9): 3131-3140
- [2] Kautkar S N, Atkinson G A, Smith M L. Face recognition in 2D and 2.5 D using ridgelets and photometric stereo [J]. *Pattern Recognition*, 2012, 45(9): 3317-3327
- [3] Wang J, You J, Li Q, et al. Orthogonal discriminant vector for face recognition across pose [J]. *Pattern Recognition*, 2012, 45(12): 4069-4079
- [4] Chen W, Er M J, Wu S. Illumination compensation and normalization for robust face recognition using discrete cosine transform in logarithm domain [J]. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 2006, 36(2): 458-466
- [5] Sharma A, Dubey A, Tripathi P, et al. Pose invariant virtual classifiers from single training image using novel hybrid-eigenfaces [J]. *Neurocomputing*, 2010, 73(10): 1868-1880
- [6] Beymer D, Poggio T. Face recognition from one example view [C] // *Proceedings of the Fifth International Conference on Computer Vision*, 1995. IEEE, 1995: 500-507
- [7] Vetter T. Synthesis of novel views from a single face image [J]. *International Journal of Computer Vision*, 1998, 28(2): 103-116
- [8] Tan X, Chen S, Zhou Z H, et al. Face recognition from a single image per person: A survey [J]. *Pattern Recognition*, 2006, 39(9): 1725-1745
- [9] Bin T, Siwei L, Hua H. High performance face recognition system by creating virtual sample [C] // *Proceedings of the 2003 International Conference on Neural Networks and Signal Processing*, 2003. IEEE, 2003, 2: 972-975
- [10] Jung H C, Hwang B W, Lee S W. Authenticating corrupted face image based on noise model [C] // *Proceedings of the Sixth IEEE International Conference on Automatic Face and Gesture Recognition*, 2004. IEEE, 2004: 272-277
- [11] Thian N P H, Marcel S, Bengio S. Improving face authentication using virtual samples [C] // *Proceedings of the 2003 IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2003 (ICASSP'03). IEEE, 2003, 3: III-233-6
- [12] Liu J, Chen S, Zhou Z H, et al. Single image subspace for face recognition [M] // *Analysis and Modeling of Faces and Gestures*. Springer Berlin Heidelberg, 2007: 205-219
- [13] Dahmen J, Keysers D, Ney H. Combined classification of handwritten digits using the 'virtual test sample method' [M] // *Multiple Classifier Systems*. Springer Berlin Heidelberg, 2001: 109-118
- [14] 孙玉宝. 图像稀疏表示模型及其在图像处理反问题中的应用 [D]. 南京: 南京理工大学, 2010

(下转封三)

方法,与 NSCT 变换的方法差别非常小。因此本文提出的融合算法在红外与可见光图像融合中是有效果的。

表 1 客观评价实验结果

图像	实验方法	AvG	IE	PSNR	MI
第一组图像	DWT	3.5996	5.9178	38.8659	5.0796
	CL	4.2706	5.9277	38.8458	5.0796
	NSCT	5.3668	5.9351	42.7811	5.6661
	本文算法	5.4814	6.8965	47.8865	5.6658
第二组图像	DWT	5.2610	6.2469	29.7307	4.7849
	CL	5.4595	6.2471	29.7309	4.7849
	NSCT	6.4154	6.3779	33.7544	4.9398
	本文算法	6.5902	6.8489	38.3352	5.0299

结束语 本文对 Tetrolet 变换进行了有效改进,使其更适用于图像融合。针对红外和可见光的特点,提出一种新的低频融合规则。相对于其他融合方法,本文算法的融合图像目标物体更明显,红外与可见光的优点都得到了体现;图像信息量增加,对比度也增强,可视效果更好;对于在不同背景环境下的源图像,取得了较好的融合效果。其不足之处在于运行时间稍长,如何优化程序、进一步改进算法是下一步的研究工作。

参 考 文 献

- [1] Hu Qian, Du Jun-ping, Han Peng-cheng, et al. Multi-sensor Image fusion with SCDPT Transform[C]//IEEE International Conference on Communication Technology. 2013
- [2] 何国栋,石建平,冯友宏,等. 一种新的红外与可见光图像融合算法[J]. 传感器与微系统,2014,33(4):139-141
- [3] Liu Huan-xi, Zhu Tian-hong, Zhao Jia-jia. Infrared and Visible Image fusion Based on Region of Interest Detection and Nonsub-sampled Contourlet Transform[J]. Journal of Shanghai Jiaotong University; Science, 2013, 18(5):526-534
- [4] 江平,张强,李静,等. 基于 NSST 和自适应 PCNN 的图像融合算法[J]. 激光与红外,2014,44(1):108-113
- [5] Do M N, Vetterli M. The Contourlet Transform: An Efficient Directional Multi-resolution Image Representation [J]. IEEE Transaction on Image Processing, 2005, 14(12):1-16
- [6] Xu T Y, Fang Y. Remote sensing image interpolation via the Nonsub-sampled Contourlet Transform[C]// The 2nd IEEE/IET International Conference on Image Analysis and Signal Processing. 2010:695-698
- [7] 李财莲,孙即祥,等. 利用偏微分方程的 Tetrolet 变换图像去噪[J]. 海南大学学报:自然科学版,2011,29(2):165-170
- [8] Chen Yuan, Zhang Rong, Yin Don. SAR Image Sparse Representation Based on Tetrolet Packet Transform[J]. Journal of Electronics & Information Technology, 2012, 34(2):261-267
- [9] 沈瑜,党建武,冯鑫,等. 基于 Tetrolet 变换的红外与可见光图像融合[J]. 光谱学与光谱分析, 2013, 33(6):1506-1511
- [10] Krommweh J. Tetrolet Transform: A New Adaptive Haar Wavelet Algorithm for Sparse Image Representation[J]. Vis. Commun. Image R., 2010, 21(4):364-374
- [11] 延翔,秦翰林,等. 基于 Tetrolet 变换的图像融合[J]. 光电子·激光, 2013, 24(8):1629-1633
- [12] 孙红进. Haar 小波在图像多尺度分解与重构中的应用[J]. 煤炭技术, 2010, 29(11):157-159
- [13] 王春华,马国超,马苗. 基于目标提取的红外与可见光图像融合算法[J]. 计算机工程, 2010, 36(2):197-200
- [14] Ellmauthaler A, Pagliari C L. Multiscale Image Fusion Using the Undecimated Wavelet Transform With Spectral Facrorization and Nonorthogonal Filter Banks[J]. IEEE Transactions on Image Processing, 2013, 22(3):1005-1017
- [15] 童涛,杨桃,等. 基于 NSCT 变换的多传感器图像融合算法[J]. 地理与地理信息科学, 2013, 29(2):22-25
- [16] 王振宇,秦立龙,刁俊良. 基于 K-SVD 和稀疏表示的数字调制模式识别[J]. 计算机科学, 2013, 40(10):65-67
- [17] Wright J, Ma Y, Mairal J, et al. Sparse representation for computer vision and pattern recognition [J]. Proceedings of the IEEE, 2010, 98(6):1031-1044
- [18] Wright J, Yang A Y, Ganesh A, et al. Robust face recognition via sparse representation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2009, 31(2):210-227
- [19] Shi Y, Dai D, Liu C, et al. Sparse discriminant analysis for breast cancer biomarker identification and classification [J]. Progress in Natural Science, 2009, 19(11):1635-1641
- [20] Elhamifar E, Vidal R. Sparse subspace clustering[C]// IEEE Conference on Computer Vision and Pattern Recognition, 2009 (CVPR 2009). IEEE, 2009:2790-2797
- [21] Zhang L, Yang M, Feng X, et al. Collaborative representation based classification for face recognition [J]. arXiv preprint arXiv:1204.2358, 2012
- [22] 朱杰,杨万扣,唐振民. 基于字典学习的核稀疏表示人脸识别方法[J]. 模式识别与人工智能, 2012, 25(5):859-864
- [23] Xu Y, Zhu X J, Li Z M. Using the original and symmetrical face training samples to perform representation based two-step face recognition [J]. Pattern Recognition, 2013, 46:1151-1158
- [24] 胡正平,宋淑芬. 基于全局和分离部件融合的双 L1 稀疏表示人脸图像识别算法[J]. 模式识别与人工智能, 2012, 25(2):78-83
- [25] Fang X Z, Xu Y, Li X L, et al. Locality and similarity preserving embedding for feature selection [J]. Neurocomputing, 2014, 128:304-315
- [26] Tibshirani R. Regression shrinkage and selection via the lasso [J]. Journal of the Royal Statistical Society. Series B (Methodological), 1996:267-288
- [27] Efron B, Hastie T, Johnstone I, et al. Least angle regression [J]. The Annals of statistics, 2004, 32(2):407-499
- [28] Moghadam A E, Shirani S. Matching pursuit-based region-of-interest image coding [J]. IEEE Transactions on Image Processing, 2007, 16(2):406-415
- [29] Needell D, Tropp J A. CoSaMP: Iterative signal recovery from incomplete and inaccurate samples [J]. Applied and Computational Harmonic Analysis, 2009, 26(3):301-321
- [30] ORL database[OL]. <http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html>
- [31] FERET database [OL]. <http://www.itl.nist.gov/iad/humanid/feret/feret?master.html>
- [32] AR database[OL]. <http://cobweb.ecn.purdue.edu/~aleix/al-eix?face?DB.html>

(上接第 294 页)