

基于马氏距离的稀疏表示分类算法

朱伟冬, 胡剑凌

(苏州大学 电子信息学院, 江苏 苏州 215006)

摘 要:常用分类算法对人脸图像在不同光照条件下的识别效果较不理想。设计了一种新颖的基于马氏距离(Mahalanobis Distance)的人脸识别分类算法(Mahalanobis Distance based Sparse Representation Classification, MSRC)。该算法框架基于稀疏表示原理,通过引入马氏距离和乔里斯基分解(Cholesky decomposition)求出最优稀疏解向量,最终实现人脸特征分类识别。算法首先求解基于马氏距离的最小 L^1 范数,进而对测试样本实现稀疏重构,并通过判断重构样本与原始样本的残差值最终完成分类。与传统稀疏表示分类算法相比,该算法显著降低了光照对人脸图像的影响。在Extended Yale face database B人脸库上的实验结果表明,所提出的基于马氏距离的稀疏表示分类算法能达到97%的分类效率,并且在人脸不同光照情况下仍能得到较好的识别效果。

关键词:稀疏表示;乔里斯基分解;马氏距离;人脸识别

中图分类号:TP391

文献标识码:A

文章编号:1673-629X(2011)11-0027-04

Sparse Representation Classification Algorithm Based on Mahalanobis Distance

ZHU Wei-dong, HU Jian-ling

(School of Electronic Information, Soochow University, Suzhou 215006, China)

Abstract: In this paper a novel Mahalanobis Distance based method for sparse representation classification was designed to improve the recognition efficiency for different illumination condition face images. Mahalanobis Distance and Cholesky decomposition are introduced to solve the sparse solution vector, and Mahalanobis Distance based Sparse Representation Classification (MSRC) is designed to recognize the face image. Firstly, Mahalanobis Distance based L^1 -minimization algorithm is proposed to obtain the sparse representation. Then, reconstruct the test image. Finally, the one that has the minimum reconstruction error is selected as the most matched face. Compared to the traditional SRC algorithms, our algorithm significantly reduces the influence of illumination. Lots of numerical experiments based on ORL face database and Extended Yale face database B are performed. The results show that the proposed Mahalanobis Distance based Sparse Representation Classification algorithm can achieve about 97% recognition rate for normal face images.

Key words: sparse representation; Cholesky decomposition; Mahalanobis Distance; face recognition

0 引言

人脸识别分类算法是人脸识别技术的一个重要环节,与特征提取算法相比,分类算法的研究范围更广泛、复杂度更高。传统的信号分解变换方法是将信号分解在一组完备的正交基上(傅里叶变换、小波变换、Gabor变换等),并且给定信号唯一的表达形式,但如果信号的特性与基函数不完全匹配,所获得的分解结果就不一定是稀疏的。1993年, Mallat 和 Zhang^[1]率先提出了信号基于过完备原子库的分解思想,用来表

示信号的基自适应地根据信号本身的特点自由选取,分解后可以得到原信号非常稀疏的表示。

目前用于人脸识别的分类算法很多,除了运用传统模式分类方法如最近邻法、Bayes 分类算法、欧氏距离、最近邻子空间算法等以外,各种神经网络分类器如径向基函数(Radial Basis Function, RBF)神经网络、支持向量学习机(SVM)等也被广泛采用。这些分类算法对很多数据分类都具有普适性,但当人脸图像存在不同光照情况时,往往识别效果很不理想。

文中设计了一种基于马氏距离的人脸识别分类方法,该算法基于稀疏表示(SRC)算法思想,并在此基础上做进一步改进,提出了基于马氏距离的稀疏表示分类算法(MSRC),求出测试人脸的最稀疏解,通过判断重构样本的最小误差实现人脸正确识别,在一定程度上提高识别精度以及降低样本间相关性的干扰。

收稿日期:2011-03-27;修回日期:2011-07-08

基金项目:苏州市应用基础研究计划(SYG201031)

作者简介:朱伟冬(1984-),男,硕士研究生,研究方向为图像处理与嵌入式系统;胡剑凌,副教授,研究方向为数字信号处理、图像处理等。

在 Extended Yale B 人脸库^[2]上的实验结果表明,该算法能有效提高识别准确率,当测试人脸图像出现不用光照情况时,其识别率优于目前常用人脸识别算法。

1 基于马氏距离的稀疏表示分类算法

1.1 基本思想

稀疏表示(Sparse Representation)理论又被称为压缩感知(Compressed Sensing)理论,由于其在理论和实践中的较大突破,尤其表现在计算复杂度的下降和分类效果的显著提升,使得该理论被广泛应用到信号处理领域的各个方面。人脸图像信号作为一种特殊的信号,其本身及其所代表的含义在很多时候都具备“稀疏”这一特点,这就使得将稀疏表示理论引入到人脸识别技术成为了一种可能。

基于过完备字典的稀疏表示理论是信号处理领域的研究热点,字典中的元素即为原子,信号由原子的基本线性组合构成。不同于常规的字典构造方式^[3,4],将人脸基图像作为字典的基本元素。稀疏表示就是在过完备字典 A 中选择少量的原子(列向量)来表示信号 y ,使得解向量 x 最稀疏^[5]。

用第 i 类人脸样本构成样本矩阵 $A_i = [V_{i,1}, V_{i,2}, \dots, V_{i,n_i}] \in R^{m \times n_i}$,其中 m 为特征维数, n_i 为 i 类样本数量。那么属于 i 类的测试样本 $y \in R^m$ 可表示为:

$$y = \varphi_{i,1}v_{i,1} + \varphi_{i,2}v_{i,2} + \dots + \varphi_{i,n_i}v_{i,n_i} \quad (1)$$

其中系数 $\varphi_{i,j} \in R, j=1,2,\dots,n_i$ 。由式(1)进而构造一个完备的矩阵 A :

$$A = [A_1, A_2, \dots, A_k] = [v_{1,1}, v_{1,2}, \dots, v_{k,n_k}] \quad (2)$$

此时属于 i 类样本的测试样本 y 可重新表示为:

$$y = Ax \in R^m \quad (3)$$

其中 $x = [0, \dots, 0, \varphi_{i,1}, \varphi_{i,2}, \dots, \varphi_{i,n_i}, 0, \dots, 0]^T \in R^n, n$ 为第 i 个人脸的样本数量。

通过大量实验发现:当式(3)中的字典 A 样本数量足够多时,就能通过最小 L^1 范数求得 y 的一个稀疏解,并且样本越多解向量 x 就越稀疏。如样本种类 $k=20$,并且每类样本数量 n 相同,那么向量 x 中的非零系数仅有 5% (1/20)。

1.2 通过最小 L^1 范数求稀疏解

1999 年, S. Chen 和 D. Donoho 等人提出了基追踪(Basis Pursuit)算法^[6],文中指出,找到最稀疏的解向量 x ,相当于求解最小 L^0 范数:

$$L^0: x_0 = \arg \min \|x\|_0, \text{ 满足 } Ax=y \quad (4)$$

这是 NP(non-deterministic polynomial)问题^[7],可以通过不同的方法逼近,其中 $\|x\|_0$ 等于向量 x 非零系数的个数。

压缩感知理论指出^[8],当向量 x 高度稀疏时,求解最小 L^0 范数近似于求解最小 L^1 范数:

$$L^1: x_0 = \arg \min \|x\|_1, \text{ 满足 } Ax=y \quad (5)$$

通常用于人脸识别的图像数据是含噪的,因此式(5)并不能准确得出稀疏解 x ,可以通过求解如下最小稳态 L^1 范数来近似得到稀疏解 x :

$$L^1: x_0 = \arg \min \|x\|_1, \text{ 满足 } \|Ax-y\|_2 \leq \varepsilon \quad (6)$$

其中 ε 是重建图像所能容忍的最大误差。

目前解决上述条件最优化问题的常用方法是对优化问题做正规化处理(Regularization),典型的如支持向量机(SVM)理论。因此式(6)可表示为:

$$L^1: x_1 = \arg \min \|x\|_1 + \mu \|Ax-y\|_2^2 \quad (7)$$

其中 μ 是图像重建后对 ε 的补偿系数^[9]。

1.3 马氏距离

马氏距离是一种有效地计算两个样本集相似度的算法(数据间协方差距离),它考虑到了各种特性之间的联系,并且是与尺度无关的(scale-invariant)。通常情况下,对于服从同一分布且协方差矩阵为 S 的两个随机向量 $x, y \in R^n$,其马氏距离可以表示为:

$$d(x, y) = \|x-y\|_M = \sqrt{(x-y)^T S^{-1} (x-y)} \quad (8)$$

其中 $S = R^{m \times n}$ 为随机向量的正定协方差矩阵。由上述马氏距离的定义,进而得到了原始人脸样本 y 和重建人脸样本 Ax 间的马氏距离:

$$d(Ax, y) = \|Ax-y\|_M = \sqrt{(Ax-y)^T S^{-1} (Ax-y)} \quad (9)$$

于是公式(9)可重新表示为:

$$L^1: x_1 = \arg \min \|x\|_1 + \mu \|Ax-y\|_M^2 \quad (10)$$

而协方差矩阵 S 由特征向量中的主要成分决定^[10],如果协方差矩阵 S 为单位矩阵,那么马氏距离就简化为欧氏距离;如果协方差矩阵 S 为对角矩阵,则可以称其为正规化的欧氏距离。

乔里斯基分解求解式(10)中基于马氏距离的解向量 x 。因为协方差矩阵 S 正定,由乔里斯基分解理论:

$$S^{-1} = L^* \cdot L \quad (11)$$

其中 L 为上三角矩阵且对角线元素正定, L^* 为 L 的共轭转置矩阵。于是式(10)可重新表示为:

$$\begin{aligned} x &= \arg \min \|x\|_1 + \mu ((Ax-y)) L^* L (Ax-y)) \\ &= \arg \min \|x\|_1 + \mu (L(Ax-y)) L^* (Ax-y)) \\ &= \arg \min \|x\|_1 + \mu \|L(Ax-y)\|_2 \end{aligned} \quad (12)$$

式(12)是一种凸优化问题,可通过二阶锥规划方法得到有效解决^[11,12]。

1.4 分类算法步骤

文中 SRC 分类算法主要目的是为了找到重建图像后的最小残差值 $r_i(y)$, 进而完成人脸识别工作。对于任意类样本 i , 存在稀疏向量 $\delta_i(x) \in R^n$ (其非零元素与 i 类样本相关) 使得测试样本 y 可近似的表示为, $\hat{y}_i$ 完成 y 样本的近似重建。通过马氏距离求出 y 和 $\hat{y}_i$ 的残差值:

$$r_i(y) = \|y - A\delta_i(\hat{x}_i)\|_M \quad (13)$$

当残差 $r_i(y)$ 最小时, 测试样本 y 即被认为属于这个 i 类样本。稀疏表示分类算法步骤如表 1 所示。

表 1 稀疏表示分类算法表

SRC 分类算法步骤
· 输入: 过完备字典 A , 补偿系数 μ , 测试样本 y
· 求解基于马氏距离的最小 L^1 范数, 得到稀疏解 $\hat{x}$
$\hat{x}_i = \operatorname{argmin} \ x\ _1 + \mu \ Ax - y\ _M^2$
· 计算残差 $r_i(y)$, $r_i(y) = \ y - A\delta_i(\hat{x}_i)\ _M$ 其中 $i = 1, \dots, k$
· 输出: $y = \operatorname{argmin} r_i(y)$, 得到识别结果

2 仿真实验与结果分析

实验共分 3 组, 实验 1 构建了基于 MSRC 的人脸分类算法框架; 第 2 组实验验证 MSRC 分类器对三维人工数据的分类效果; 第 3 组实验对不同光照条件下人脸样本做分类识别, 并与普通 SRC 算法做对比。

实验 1, 设计基于稀疏表示的人脸分类算法框架。本实验人脸样本来自 Extend Yale B 人脸库, 该人脸库包含 38 个人的 2427 张人脸图像 (192×168), 选取其中 1216 张人脸图像做训练 (人均 32 张), 其余用于测试。当输入第一个人的人脸图像做测试样本, 通过求解基于马氏距离的最小 L^1 范数, 得到稀疏系数 $\hat{x}$ 的分布图如图 1 所示, 横坐标为训练样本数量, 纵坐标对应稀疏系数。

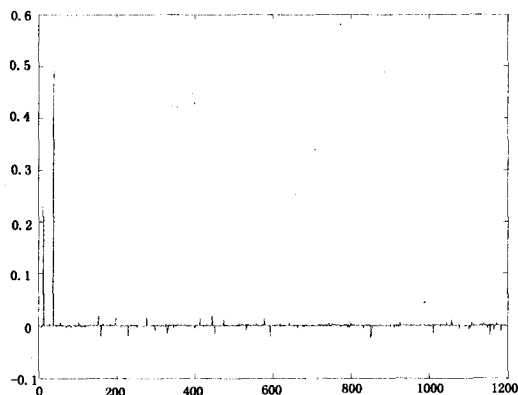


图 1 稀疏系数的分布情况

实验 2, 将 MSRC 算法对三维数据做分类测试, 验证分类效果。图 2 显示的是三维数据训练样本分布情况, 每种数据有 50 个样本; 图 3 显示的是测试样本的分布情况, 每种数据有 20 个样本; 图 4 显示的 MSRC 分类算法的分类结果。

况, 每种数据有 50 个样本; 图 3 显示的是测试样本的分布情况, 每种数据有 20 个样本; 图 4 显示的 MSRC 分类算法的分类结果。

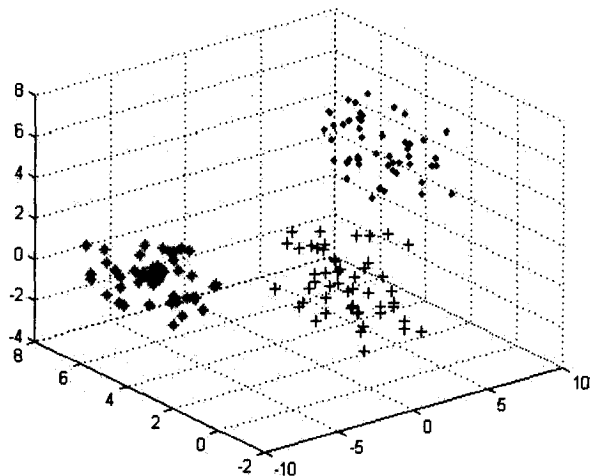


图 2 三维训练样本

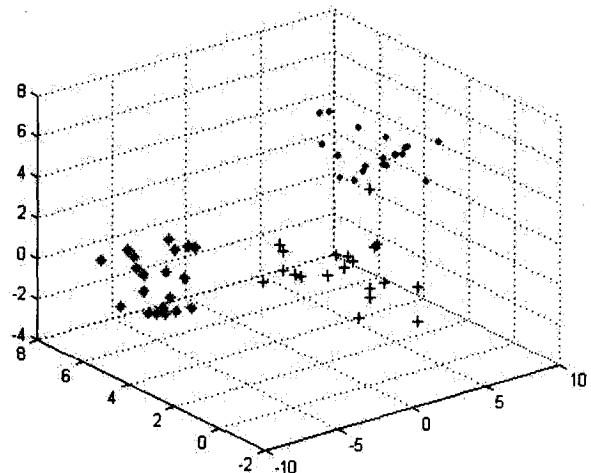


图 3 三维测试样本

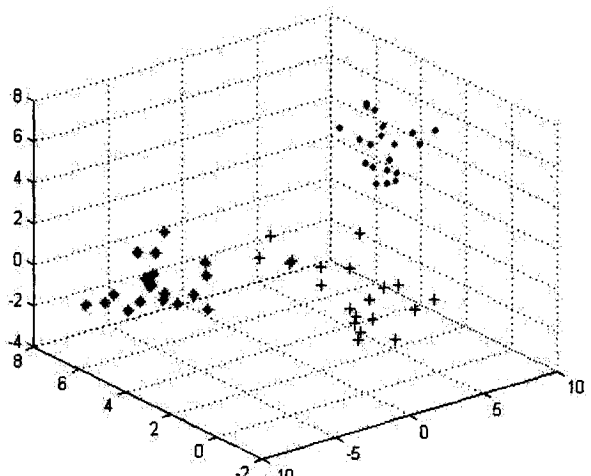


图 4 MSRC 算法分类效果

从实验结果明显看出, MSRC 分类算法能较准确地完成三维数据的分类。

实验 3, 不同光照条件是影响常用人脸分类算法

主要因素之一,为了考察文中方法对不同光照人脸图像分类识别的鲁棒性,使用 Extend Yale B 人脸库中不同光照情况的人脸图像作为测试样本。随机选择每个人的一半照片作为测试样本,另一半作为训练样本,特征维数分别选择 30、56、72、90、110、140 六种,以此来对比 SRC 算法与文中 MSRC 算法的分类效果。图 5 为不同光照条件下的四组人脸测试样本。



图 5 不同光照情况的人脸测试图像

在人脸图像不同光照条件下,随着光照强度的不断减弱, SRC 和 MSRC 分类结果如图 6 所示,可以明显看出所提分类算法明显优于文献[5]所提的 SRC 人脸识别算法。当特征维数为 140 时,文中算法的最高识别率达到了 97.1%,而 SRC 算法的识别率最高为 94.77%。从基于光照变化的实验可以发现,基于马氏距离的分类算法在分类准确率上要明显高于普通 SRC 算法,同时大大降低计算复杂度,从整体上提高了分类算法的性能。

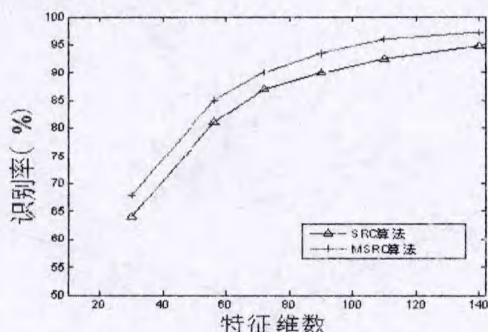


图 6 基于不同光照条件的分类效果对比

3 结束语

设计了一种基于马氏距离的稀疏表示人脸分类器

算法——MSRC 算法,算法首先构建基于训练样本的完备字典,然后求解人脸基图像在此字典上的稀疏表示,通过重构的人脸图像,求得原始图像和重构图像的残差实现分类。在 Extend Yale B 人脸库上的实验结果表明, MSRC 人脸分类方法具有较优异的分类效果。当人脸样本数量足够的情况下,分类效果好于传统 SRC 人脸识别分类算法,并且克服了光照条件对人脸图像分类识别的影响。

参考文献:

- [1] Mallat S G, Zhang Z F. Matching pursuits with time-frequency dictionaries [J]. IEEE Trans. on Signal Processing, 1993, 41(12): 3397-3415.
- [2] Lee K C, Ho J, Kriegman D. Acquiring Linear Subspaces for Face Recognition under Variable Lighting [J]. IEEE Trans. Pattern Anal. Mach. Intelligence, 2005, 27(5): 684-698.
- [3] 李武军, 王崇骏, 张 炜, 等. 人脸识别研究综述 [J]. 模式识别与人工智能, 2006, 19(1): 58-66.
- [4] 尹 飞, 冯大政. 基于 PCA 算法的人脸识别 [J]. 计算机技术与发展, 2008, 18(10): 31-33.
- [5] 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.
- [6] Donoho D. For most large underdetermined systems of linear equations the minimal ℓ_1 -norm solution is also the sparsest solution [J]. Comm. on Pure and Applied Math, 2006, 59(6): 797-829.
- [7] Amaldi E, Kann V. On the approximability of minimizing non-zero variables or unsatisfied relations in linear systems [J]. Theoretical Computer Science, 1998, 209: 237-260.
- [8] Ji Yangfeng, Lin Tong, Zha Hongbin. Mahalanobis Distance Based Non-negative Sparse Representation for Face Recognition [C]// IEEE International Conference on Machine Learning and Applications. [s. l.]: [s. n.], 2009.
- [9] He X, Yan S, Hu Y, et al. Face recognition using Laplacian-faces [J]. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2005, 27(3): 328-340.
- [10] Belhumeur P, Hespanda J, Kriegman D. Eigen faces vs. Fisher faces: recognition using class specific linear projection [J]. IEEE Trans. on Pattern Analysis and Machine Intelligence, 1997, 19(7): 711-720.
- [11] 黄国宏, 刘 刚. 一种新的基于 Fisher 准则的线性特征提取方法 [J]. 计算机技术与发展, 2008, 18(5): 227-230.
- [12] Chen S, Donoho D. Atomic decomposition by basis pursuit [J]. SIAM Journal on Scientific Computing, 1999, 20(1): 33-61.