

# 改进 SRC 算法在人脸识别中的应用

雷明军, 彭进业, 冯晓毅

(西北工业大学 电子信息学院, 陕西 西安 710129)

摘要: 稀疏表示分类算法 (Sparse Representation-based Classification, SRC) 在人脸数据库上有很高的识别性能。然而, 对于姿态变化, SRC 的识别效果并不理想。针对 SRC 算法不能解决测试样本与训练样本存在偏移误差的问题, 本文提出了基于 SRC 的改进算法。该算法将每一类的训练样本单独作为训练字典, 利用迭代校正和基于金字塔分层机构的运动偏移估计方法得到最终的偏移量, 最后对校正后的测试样本使用 SRC 算法实现分类。实验结果表明该方法对于有偏移误差的人脸图像具有较好的鲁棒性及识别率。

关键词: SRC; 人脸识别; 迭代校正; 运动偏移估计

中图分类号: TP312

文献标识码: A

文章编号: 1674-6236(2013)04-0137-03

## Application of a kind of improved SRC algorithm in face recognition

LEI Ming-jun, PENG Jin-ye, FENG Xiao-yi

(Dept. of Electronics and Information, Northwestern Polytechnical University, Xi'an 710129, China)

Abstract: Sparserepresentation of classification algorithm has high identification performance on the face database. However, because of the attitude changes, SRC recognition effect is not ideal. Because the SRC algorithm can not solve the offset error problem between test samples and the training samples, this paper puts forward the improved algorithm based on SRC. The algorithm regards each type of training samples as training dictionary individually and get the final offset using the iterative correction and based on the Pyramid layered motion offset estimation method, finally using SRC algorithm classified the corrected test samples. The experimental results show that this method for face images with offset error performs better robustness and the identification rate.

Key words: SRC; face recognition; iterative correction; motion estimation

基于图像的人脸识别技术是近年来计算机视觉领域研究热点之一, 人脸识别技术利用从人脸图像中提取出有效特征信息, 与数据库中的已知人脸信息进行匹配, 计算出待测试人脸的特征与数据库中哪一类人脸最相近, 从而鉴别个人身份。现有大多数人脸识别算法主要基于统计学习理论, 利用大量训练样本来学习人脸分类器, 根据得到的分类器模型进行人脸识别。

稀疏表示人脸识别算法 (简称“SRC 方法”) 是近年来新出现的一个比较有效的人脸识别算法<sup>[1]</sup>。它假设任一张人脸图像都可以用同一个人训练样本的线性组合来表示, 通过寻找测试样本相对于整个训练集的稀疏表示系数来发现测试图像所属的用户身份。借助于先进的高维凸优化技术 (如 L1 范数最小化<sup>[2]</sup>) 稀疏表示系数可以被精确稳定地恢复出来, 解的精度和鲁棒性都有理论上的保证。与现有大多数方法相比, SRC 方法直接利用了高维数据分布的基本特性 (即“稀疏性”) 进行统计推断, 可以有效地应对维数灾难问题。同时, SRC 方法直接利用原始像素进行人脸识别, 可以有效地避免

各种预处理过程。通过选择合适的训练图像 (不同光照和表情) 并利用稀疏信号对损毁像素建模, SRC 方法可以有效地应对光照变化、表情变化、局部遮挡等问题, 具有良好的鲁棒性。但是, 现有 SRC 方法要求测试图像和训练图像严格对齐, 姿态变化会引起人脸图像产生对齐误差, 降低 SRC 方法的识别性能。姿态问题已经成为 SRC 方法走向实用必须解决的一个主要问题。

## 1 SRC 算法基本理论

第  $i$  类训练样本按列排列表示成矩阵  $A_i, A_i = [v_{i,1}, v_{i,2}, \dots, v_{i,n_i}] \in R^{m \times n_i}$ 。其中  $v_{i,j}$  表示第  $i$  类目标的第  $j$  个样本, 对于一幅  $w \times h$  的灰度图像  $v_{i,j} \in R^m (m=wh)$ 。根据线性子空间原理, 如果第  $i$  类样本足够多, 那么第  $i$  类测试样本  $y$  可以写成:

$$y = v_{i,1} \alpha_1 + v_{i,2} \alpha_2 + \dots + v_{i,n_i} \alpha_{n_i}, \alpha_j \in R, j=1, 2, \dots, n_i$$

$$A = [A_1, A_2, \dots, A_k] = [v_{1,1}, v_{1,2}, \dots, v_{k,n_k}] \in R^{m \times \sum_{i=1}^k n_i}$$

$A$  表示  $k$  类目标的所有训练样本所组成的矩阵。将测试样本  $y$  扩展到用所有训练样本进行线性表示, 则表示式如下:

收稿日期: 2012-10-07

稿件编号: 201210015

基金项目: 国家自然科学基金资助 (61075014)

作者简介: 雷明军 (1986—) 男, 湖南常德人, 硕士。研究方向: 图像信息处理。

$$y = Ax_0 \quad R^m$$

其中  $x_0 = [0, \dots, 0, i_1, i_2, \dots, i_{n_i}, 0, \dots, 0]^T$  是系数向量,  $x_0$  中只有和  $y$  同一类的训练样本所对应的系数是非零值, 而其他系数均为零值。当训练集的样本总数远大于第  $i$  类训练样本的个数, 则是稀疏的。由于  $y = Ax$  通常是欠定系统, 因此, 方程  $y = Ax$  的解是不唯一的, 由于要求所求的  $x$  是稀疏的, 由压缩感知原理可知, 问题可以转换为求解  $l_0$  范数最小化问题, 模型如下:

$$\hat{x}_0 = \arg \min_x \|x\|_0 \text{ s.t. } y = Ax$$

其中,  $\|\cdot\|_0$  表示  $l_0$  范数, 表示向量中非零元素的个数。然而此模型属于非凸优化模型, 求解是一个 NP 问题, 近年来在稀疏表示和压缩感知领域的研究表明, 如果  $x$  足够稀疏, 那么  $l_0$  范数最优化问题等效于下面的  $l_1$  范数最小化问题:

$$\hat{x}_1 = \arg \min_x \|x\|_1 \text{ s.t. } y = Ax$$

如果测试样本包含轻微污损 (含噪声或局部遮挡) 可以通过求解  $x$  和  $e$  的最小范数问题, 实现正确的分类, 该算法的模型为下式:

$$\min \|x\|_1 + \|e\|_1 \text{ subj } y = Ax + e$$

## 2 基于 SRC 的改进算法

在测试图像与训练图像没有对齐的情况下, 测试样本不能用训练样本稀疏表示。但是, 可以对测试样本先校正, 再对校正后的测试样本求解稀疏表示。算法模型如式 (1):

$$\hat{x} = \arg \min \|x\|_1 + \|e\|_1 \text{ subj } T(y, \theta) = Ax + e \quad (1)$$

其中  $y$  是原始图像,  $\theta$  是仿射变换参数,  $T(\cdot)$  是仿射变换算子。  $T(y, \theta)$  表示测试样本  $y$  经过仿射变换与训练样本对齐后的图像。当未知时, 要同时最优化  $x$ ,  $e$  和  $\theta$  是一个非凸的最优化问题。如果用户自己设定的初始值求解问题 (1) 测试图像  $y$  可能会与非同类的训练样本形成校正的趋势, 从而陷入局部最小值, 最终得到的  $x$  和  $e$  不是全局最优的。为了避免陷入局部最小值, 本算法将每一类用户的训练图像单独用作训练字典, 算法模型变为式 (2):

$$\hat{x} = \arg \min \|e\|_1 \text{ subj } T(y, \theta) = A_i x + e \quad i=1, 2, \dots, k \quad (2)$$

其中  $k$  表示训练图像中的种类。在模型 (2) 中, 由于只使用一类用户的图像作为训练样本, 所以只需使  $e$  最稀疏化, 而不必使  $x$  稀疏化。

假设测试样本与训练样本只存在微小的平移偏差, 则可以通过一阶泰勒展开式来近似校正后的测试图像:

$$T(y, \theta) = y + x \cdot Y_x + y \cdot Y_y \quad (3)$$

其中  $Y_x$  是图像  $y$  水平方向的一阶导数,  $Y_y$  表示图像  $y$  竖直方向的一阶导数。所以对于只有较小偏移的测试样本可以将模型 (3) 改写成 (4)。偏移量较大时, 可以不断迭代将上述过程, 同时更新校正后的测试样本, 每次迭代都能得到相对偏移量, 直到收敛停止迭代。再将每次得到的相对偏移量累加, 最终得到原始的测试样本相对于同类训练样本偏移量。

$$\hat{x} = \arg \min \|e\|_1 \text{ subj } y + x \cdot Y_x + y \cdot Y_y = A_i x + e \quad (4)$$

将模型 (4) 写成如下形式:

$$\hat{x} = \arg \min \|e\|_1 \text{ subj } y = A_i x - x \cdot Y_x - y \cdot Y_y + e = Bw \quad (5)$$

其中,

$$B_i = [A_i, Y_x, Y_y, I] \quad R^{m \times (n+m+2)}$$

$$w = \begin{bmatrix} x \\ -x \\ -y \\ e \end{bmatrix} \quad R^{n_1+m+2}$$

因此问题 (5) 转成经典的范数最小化问题:

$$\arg \min \|e\|_1 \text{ subj } y = B_i w \quad (6)$$

由于图像的维度很大, 计算复杂度很高。因此, 对式 (6) 等号左右两边同时乘以随机投影矩阵  $R$ , 以降低维度。因此问题 (6) 转换为:

$$\arg \min \|e\|_1 \text{ subj } R_y = R B_i w \quad (7)$$

其中  $R \in R^{d \times m}$ ,  $d < m$ 。通过模型 (7) 求解得到  $w = [x - y e]^T$ 。

测试样本  $y$  是校正的版本  $y_{est}$  和残差  $y$  之和。即  $y = y_{est} + y$ 。其中  $y_{est} = A_i x_i$ 。得到校正估计  $y_{est}$  之后, 利用基于金字塔分层机构的运动偏移估计方法<sup>[11]</sup>估计当前测试样本  $y$  与校正版本估计  $y_{est}$  的平移量。即求解下式:

$$\hat{\theta} = \arg \min \|T(y, \theta) - y_{est}\|_2 \quad (8)$$

求出测试样本相对于不同的训练样本的仿射变换参数  $\theta$  后, 利用式 (9) 求解原始测试样本相对于校正后的不同类型的训练样本的误差  $e$ 。当训练样本矩阵  $A_i$  与测试样本同类时, 测试样本  $y$  可以用同类校正后的训练样本线性表示, 则  $e$  主要是噪声所引起的误差, 所以  $e$  的模值最小。但是, 当训练样本矩阵  $A_i$  与测试样本不同类时, 测试样本不会与非同类的经仿射变换后的训练样本对齐, 测试样本不能用这样的训练样本线性表示,  $e$  不仅包含噪声, 还包含由于未对齐所导致的误差, 所以  $e$  的模值较大。因此可以通过比较不同类的训练样本矩阵  $A_i$  所对应的  $e$  的模值的大小来对测试样本进行分类, 如式 (10)。

$$\hat{x}_i = \arg \min \|x_0\|_1 + \|e\|_1 \text{ subj } y = T(A_i, \theta) x + e \quad (i=1, 2, \dots, k) \quad (9)$$

$$\text{Identify}(y) = \arg \min \|e\|_2 \quad (i=1, 2, \dots, k) \quad (10)$$

总结文中改进算法步骤如表 1 所示。

## 3 实验设计及仿真结果

实验选择目前常用的 Extended Yale B 人脸数据库作为实验数据集。然后将本文提出的改进方法与 SRC 方法, 文献 [3] 的算法在该数据库下进行比较。本实验是在 MATLAB R2007b 平台下完成, PC 机配置为 2.0 GHz, 1 G 内存。

本实验随机选取了 10 个人的人脸图像作为实验数据, 并且对每类人脸随机选取 30 幅图像作为训练样本, 再在剩余的图像中随机选取 15 幅图像作为测试样本, 并采用双线性插值法将所选的样本图像都归一化至 9684 (单位: 像素) 的尺寸。因此, 本实验中一共有 300 幅训练图像和 150 幅测试图像。

为了公平比较, 将本文的算法、SRC 算法和文献 [3] 的算



表 1 改进的 SRC 算法步骤  
Tab. 1 Improved SRC algorithm steps

```

改进的 SRC算法步骤
Input :测试样本  $y \in R^m$ , k 类训练样本  $A_1, A_2, \dots, A_k \in R^{m \times n_i}$ 
for subject=1 :k
     $i = 1, \dots, k$ 
    while  $\epsilon > \text{thd}$ 
         $y = T(y, \epsilon), Yx = \frac{Yy}{Yx}, Yy = \frac{Yy}{Yx}$ 
         $\epsilon = \arg \min \|e\|_1, \text{subj } y + xY_x + yY_y = A_i x + e$ 
         $i = i + 1$ 
    end
    Solve L1-minimization problem :
     $\arg \min \|x\|_1 + \epsilon_i, \text{subj } y = T(A_i, \epsilon_i) x + e_i (i=1, 2, \dots, k)$ 
end
Output :identify(y) = arg min(r(i)) i=1, 2, ... k
    
```

法都采用随机脸特征,用随机矩阵将人脸图像投影至 150 维 (d=150) 的低维特征空间。对于特征脸识别方法也将人脸图像降维到 150 维的特征脸空间。

实验结果如表 2 所示,为了使比较更加直观,采用折线图 2 展现出来。

表 2 识别结果比较  
Tab. 2 Comparison of recognition results

偏移值	0	2	5	7	9
SRC	98.67	77.33	14.67	12.00	10.00
文献	94.00	91.33	89.33	73.33	70.67
本文方法	94.67	93.33	90.00	74.67	69.33

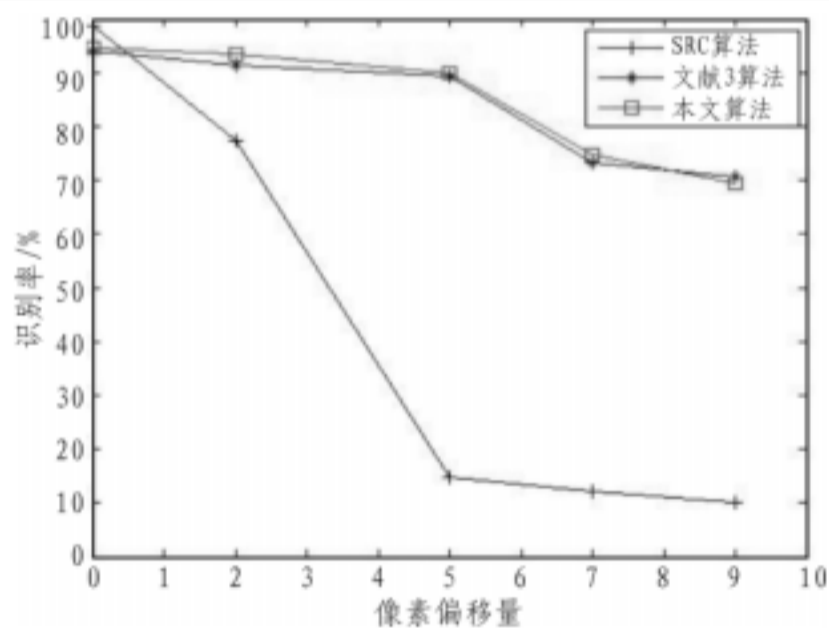


图 1 识别结果图  
Fig. 1 Recognition results map

从图 1 实验结果可以看出,在测试样本和训练样本对齐的情况下(偏移值等于零) SRC 的识别率最高,但是随着偏移误差的增大,识别率急剧下降。说明了 SRC 对图像平移十分敏感,鲁棒性差。本文方法与文献 [3] 的方法,明显优于 SRC 方法,识别率和鲁棒性得到明显的改善。但是,本文提出的方法的识别率比起文献 [3] 的方法又略高一筹。

#### 4 结论

针对 SRC 算法的不能解决测试样本与训练样本存在偏

移误差的问题,提出了基于 SRC 的改进算法。该算法利用迭代校正的方法得到每一次校正的偏移量,直到迭代停止,将每次的偏移结果累加就是最终的偏移量。估计当前测试样本与校正样本的偏移量的过程中,使用了基于金字塔分层机构的运动偏移估计方法。为了避免求出的稀疏解偏向于非同类的训练样本,而收敛到局部最小值的情况,这里只将每一类的训练样本单独作为训练字典。最后对校正后的测试样本使用 SRC 算法实现分类。实验结果表明该方法对于有偏移误差的图像具有较好的鲁棒性,识别性能良好。

参考文献:

- [1] 周激流,张晔. 人脸识别理论研究进展[J]. 计算机辅助设计与图形学学报,1999,11(2):180-184.
- [2] ZHOU Ji-liu,ZHANG Ye. Face recognition theory research progress[J]. Computer Aided Design and Computer Graphics, 1999,11(2):180-184.
- [3] Hong Z Q. Algebraic feature extraction of image for recognition[J]. Pattern Recognition,1991,24(3):211-219.
- [4] Huang J,Huang X, Metaxas D. Simultaneous image transformation and sparse representation recovery[J]. In CVPR,2008.
- [5] Lades M,Vorbuggen J,Buhmann J C et al. Distortion invariant object recognition in the dynamic link architecture[J]. IEEE Transactions on Computers,1993,42(3):300-311.
- [6] Samaria F. Face recognition using hidden markov models[M]. Ph.Dthesis, University of Cambridge,1994.
- [7] Yuile A L. Deformable templates for face recognition [J]. Journal of Cognitive Neuroscience,1991,3(1):59-70.
- [8] Liu C J,Wechsler H. Comparative assessment of independent component analysis (ICA) for face recognition[C]// Int. Conf. on Audio and Video Based Biometric Person Authentication, 1999.
- [9] Moghaddam B. Principal manifolds and Bayesian subspaces for visual recognition [C]//Tech Report 99-35, A Mitsuishi Electric Research Lab,1999.
- [10] Wright,Yang A,Ganesh A et al. Robust face recognition via sparse representation [C]//IEEE Transactions on Pattern Analysis and Machine Intelligence,2007.
- [11]Wagner A,Wright J,Ganesh A et al. Towards a practical face recognition system:Robust registration and illumination via sparse representation [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Miami,USA:IEEE Computer Society,2009:597-604.
- [11]Bergen J,Anandan P,Hanna K,et al. Hingorani. Hierarchical model-based motion estimation[J]. In Proceedings of ECCV,1992:237-252.