

基于稀疏编码的人脸识别算法

徐东勋, 李超, 袁昌斌

(中国海洋大学 信息科学与工程学院, 山东 青岛 266100)

摘要: 介绍了一种基于稀疏编码的人脸识别算法。先对 10 副自然图像应用稀疏编码, 学习到基函数和图像稀疏表示的拟合分布的参数。在人脸识别中, 用稀疏编码和已得到的基函数表示图像的稀疏, 再经过拟合分布函数得到人脸图像的最终表示, 然后应用多分类线性支持向量机(SVM)来完成识别算法。通过在人脸数据库上的实验表明, 该算法具有很高的识别正确率。

关键词: 人脸识别; 稀疏编码; 多类线性 SVM; 循环 ICA

中图分类号: TP391.41

文献标识码: A

文章编号: 1674-7720(2012)22-0035-03

Face recognition via sparse coding

Xu Dongxun, Li Chao, Yuan Changbin

(School of Information Science and Engineering, Ocean University of China, Qingdao 266100, China)

Abstract: This paper mainly presents an approach to face recognition via sparse coding. Learn the basis function and parameters of the sparse representation's fitted distribution after applying sparse coding to ten natural images. In the process of face recognition, the basis function is used to get the images' sparse representation via sparse coding, followed by the fitted distribution function to get final image presentation. Multi-class linear SVM is chosen as the classifier to finish the recognition. The algorithm is applied to face datasets and the results show that it has high recognition correct rate.

Key words: face recognition; sparse coding; multi-class linear SVM; recursive ICA

人脸识别是生物特征识别的重要组成部分, 相对于其他的生物特征识别技术, 人脸识别具有操作简单、结果直观、隐蔽性好等优点。从 100 年多前被提出来后, 人脸识别的研究热度一直没有降低。随着互联网等信息技术的发展, 安全识别变得更加重要, 人脸识别的应用面也逐渐扩大, 在商业、安全以及司法等行业内都有非常广泛的应用。人脸识别就是通过计算机完成人脸智能判别任务。图像识别是人脸识别的主要方法。

由于自然条件下不同光照条件、拍摄角度、人脸表情等因素的影响, 人脸图像有较强的不稳定性。找到对这些不利因素不敏感而又能有效鉴别特征的算法是提高识别性能的关键。主成分分析(PCA)是目前应用最多的人脸识别方法。但是 PCA 方法只考虑了图像的二阶统计信息, 没有考虑更高阶的信息。基于稀疏编码的人脸识别则可以得到更高阶的信息, 但目前的稀疏编码的识别算法中基函数的学习大多依赖于人脸数据库, 基函数会随着识别任务不同而需要重新学习得到^[1]。

Shan^[2]探索了一种模拟人类视觉过程的循环 ICA(独立成分分析)算法。提出通过对自然图像的应用循环 ICA

学习, 得到一种通用的视觉特征, 这种视觉特征能应用在视觉的所有识别任务中。

本文基于稀疏编码学习到的通用视觉特征, 把学习到的通用视觉特征应用到基于稀疏编码的人脸识别任务中, 通过对两种通用人脸数据库的实验, 均得到了较高的识别正确率。

1 稀疏编码理论

稀疏编码理论由 OLSHAUSEN B A 和 FIELD D J^[3]于 1996 年提出, 这种概率模型可以用于捕获自然图像的高阶相关结构。标准的算式表示为:

$$\min \|a\|_1 \text{ s.t. } X=Da$$

X 为训练样本, $X=\{x_1, x_2, \dots, x_n\}$, x_n 代表图像的第 n 个图像子块, D 为图像基函数, a 为图像的稀疏表示。

经拉格朗日乘法后的通用算式为:

$$(D, a) = \arg \min_{D, a} \|Da - X\|_2^2 + \lambda \|a\|_1$$

$$\text{s.t. } \|D_i\|_2^2 \leq 1, i=1, 2, \dots, k$$

其中, 一阶范数 $\|a\|_1$ 来保证稀疏性, 有关 D 的二阶范数是为了标准化。研究表明: 对式中 D, a 都不是凸优化函

《微型机与应用》2012 年 第 31 卷 第 22 期

图形、图像与多媒体

Image Processing and Multimedia Technology

数。但是当其中一个已知时,另一个则是凸优化函数,可解。

求解算法过程如下:

(1) 初始化 D : 用高斯随机矩阵初始化 D , 保证 D 的每一列都标准化。

(2) D 为已知, 更新 a

$$a = \arg \min_a \|Da - X\|_2^2 + \lambda \|a\|_1 \quad (1)$$

通过线性编码可有效解决。

(3) a 已知, 更新 D

$$D = \arg \min_D \|Da - X\|_2^2$$

$$\text{s.t. } \|D_i\|_2 \leq 1, i=1, 2, \dots, k$$

这是个有限制的凸优化问题, 可通过优化函数包解决。

(4) 在步骤(2)和步骤(3)间循环直到收敛。即可得到图像的基 D 和稀疏系数 a 。

本文中用到的稀疏编码算法为参考文献[3]的主页提供的程序包^[4]。

2 通用视觉特征的学习

依照参考文献[2]的模拟人类视觉系统理论, 人类在小时候学习到通用的视觉特征, 并用到随后的所有的视觉任务中, 识别算法分为两部分: 学习通用的视觉特征和应用到具体的识别任务。

学习通用视觉特征用的自然图像来自参考文献[3]的主页, 是 10 副 512×512 大小的自然图像。

2.1 图像预处理

首先将 10 副图像进行白化处理, 然后标准化每副图像成零平均值和单位方差。白化后, 靠近边界的 6 个像素被切除以避免边界影响。每个图像的大小是 500×500 。然后提取所有的 2430490 个 8×8 的图像子块, 每个子块再减去它的局部均值。

2.2 计算稀疏基函数

经过上面的处理得到图像子块集合为 $\{x_1, x_2, \dots\}$ 。

首先用高斯随机变量初始化基函数, 接着在每次循环中随机挑选 100 副图像子块 $\{x_1, x_2, \dots, x_{100}\}$ 。混合矩阵 D 用下式进行迭代更新:

$$D = D + \frac{R}{100} \sum_i (x_i - Da_i) a_i'$$

其中, $R=0.01$ 代表学习的速率, a_i 是在给定输入 x_i 和现有的混合矩阵 D 得到的最可能的基本信号源。每次更新后, D 的列被标准化成单位长度来加速学习过程。重复这个过程 100000 次, 如此可得到图像基函数 D 。图 1 为当 a 的维数是 64 时学习到的图像基函数。

2.3 估计 a 的累积分布函数

使用参考文献[2]提到的累积分布函数方法, 可以将图像稀疏表示 a_i 的密集区域拉开距离以便于识别, 而非密集的区域则改变不大。得到累积分布函数的方法是使用拟合的方式: 首先计算得到所有 2430490 提取图像子

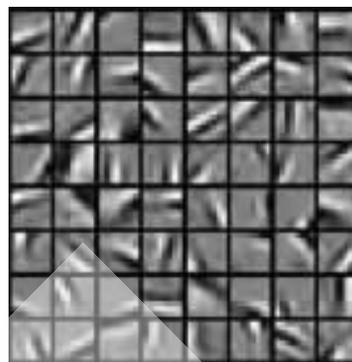


图 1 64 维图像基函数

块的基本信号表示 a_i , 接着在 0~15 之间以每格 0.001 的长度来计算 $|a_i|$ 的柱状图。实际 $|a_i|$ 的累积分布函数 (cdf) 从这个柱状图中产生 (如图 2 所示), 接着通过下面的函数调整:

$$F_i(|a_i|) = \Gamma((|a_i|/\tau)^\theta, 1/\theta)$$

其中, Γ 表示不完整的伽马函数。

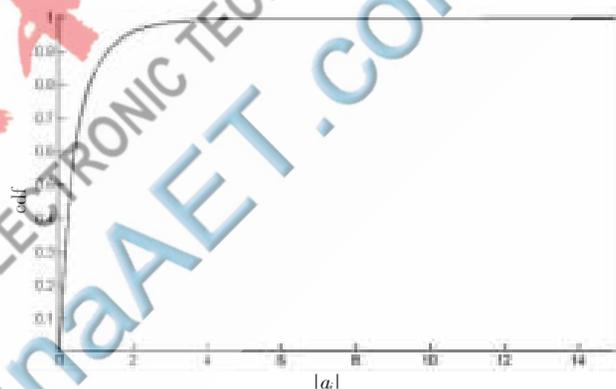


图 2 其中一个 $|a_i|$ 的实际 cdf 函数和调整后的 cdf 的函数 (a 的维数是 64)

从 10 副自然图像中学习得到线性混合矩阵 D 和非线性激励函数 $G_i = g(F_i)$, g 表示标准正态分布的逆函数。这就是所需要的通用的视觉特征, 可以不作任何调整直接应用到人脸识别中。

3 具体的人脸识别实验

3.1 基于 Yale 脸部数据库的实验

Yale 数据库包含了 15 个人, 每人 11 副脸部图像, 总共 165 副灰度图像, 其中一个人的 11 副图像如图 3 所示。每个人人脸图像都有很大的光线差异, 人的脸部表



图 3 Yale 人脸数据库其中一人的 11 副图像

图形、图像与多媒体

情和位置也有一定程度的变化。本文选择直接从参考文献[5]的主页下载已经手动剪切和处理过的 64×64 大小的图像。

(1) 用 Matlab 中的 `imresize` 函数把所有人脸图像缩小为 32×32 。使用 2.1 节中的白化滤波器白化每个图像,接着标准化每副图像成零平均值和单位方差。对每副图像提取所有的 625 个 8×8 大小的图像子块。

(2) 利用从自然图像已学习到的基函数 D ,应用稀疏编码的求解过程的式(1)得到每个图像子块的最可能的稀疏表示 a_i 。

(3) 对 a_i 的每一维应用非线性激励函数 G_i 。当 a 的维数是 64 时,现在每个人脸图像被一个 $625 \times 64 = 40\,000$ 维向量表示。

延续参考文献[5]中的方法,随机地将图像分成训练和测试样本集,每次试验时从每个人的图像中随机选取 $M=2,3, \dots, 8$ 个图像作为训练图像,剩余的作为测试图像。对应每个 M ,选择测试 50 个随机队列,最后的识别率是所有队列的平均识别率。

当训练样本和测试样本选定后,每副图像由 40 000 维的向量表示,维数太高对识别的计算时间和准确度都有很大影响,接着选择使用 PCA 来降低它的维数。主成分数目的选择要使 95% 的方差被捕获。例如,当 M 从 2~8 时,主成分的数目为 27~105。接着使用映射的 PCA 矩阵来得到训练样本和测试样本。

(4) 在分类器的选择上,选用支持向量机 SVM,考虑到非线性 SVM 的随维数增加计算时间指数级地增长,而选用线性 SVM。采用的参考文献[4]中用到的多类识别的线性 SVM,其中的 λ 设置为 0.01。

表 1 列出了在测试图像上识别效果,随着训练图像数目的增加,图像的识别率逐渐增高,使用参考文献[5]在其主页最新更新的结果作为对比,得到的识别效果要明显地优于参考文献[5]的结果。即使是其最近更新的结果依旧不如本文的识别率高。

表 1 Yale 数据库上的识别率对比

| 算法 | $M=2$ | $M=3$ | $M=4$ | $M=5$ | $M=6$ | $M=7$ | $M=8$ |
|-----------|-------|-------|-------|-------|-------|-------|-------|
| 本文 | 0.69 | 0.793 | 0.855 | 0.888 | 0.903 | 0.920 | 0.928 |
| 参考文献[5] | 0.576 | 0.723 | 0.778 | 0.817 | — | — | — |
| 参考文献[5]更新 | 0.625 | 0.745 | 0.807 | 0.853 | 0.877 | 0.897 | 0.913 |

3.2 在 ORL 脸部数据库上的实验

ORL 数据库总共包含 400 个脸部图像,分属 40 个人,每人 10 副。这些图像拍摄在不同的时间、光照条件、脸部表情下,如图 4 所示。如 Yale 人脸识别的实验一样,从参考文献[5]作者的主页下载已手动对齐和切割的 64×64 大小的图像,接着用 `imresize` 函数缩小成 32×32 图像。具体的实验步骤与测试 Yale 数据库时一样。从每个人的图像中随机挑选 $M=2,3, \dots, 8$ 个图像作为训练图像,

表 2 ORL 数据库上识别率的对比

| 算法 | $M=2$ | $M=3$ | $M=4$ | $M=5$ | $M=6$ | $M=7$ | $M=8$ |
|-----------------------|-------|-------|-------|-------|-------|-------|-------|
| 本文 | 0.830 | 0.903 | 0.948 | 0.971 | 0.981 | 0.988 | 0.99 |
| Cai ^[4] | 0.851 | 0.923 | 0.958 | 0.972 | — | — | — |
| Cai ^[4] 更新 | 0.851 | 0.927 | 0.96 | 0.977 | 0.983 | 0.987 | 0.991 |

剩余作为测试。平均识别正确率如表 2 所示,本文依然选择参考文献[5]和其主页最近更新的识别率作为比较。



图 4 ORL 人脸数据库的一人的 10 副图像

本文算法在 ORL 人脸数据库的表现也是随着训练样本数目的增加而增加,并且维数越大越接近参考文献[5]的结果,如维数为 7 或 8 时,得到的结果几乎没有差距。

根据参考文献[2]中的通用视觉特征理论,利用稀疏编码从自然图像中学习基函数和稀疏系数的累积分布函数参数,并利用稀疏编码应用到人脸识别中。通过选用两种通用的人脸数据库进行测试,均得到了比较高的识别率。并与别人的结果进行了对比。算法还有很多可以改进的地方:如分类器的其他选择、预处理方法的改进等。

参考文献
[1] 杨荣根,任明武,杨静宇. 基于稀疏表示的人脸识别方法[J]. 计算机科学,2010,37(9):267-269.

[2] SHAN H H, ZHANG L Y, COTTRELL G W. Recursive ICA[C]. In Advances in Neural Information Processing Systems, 2007:1273-1280.

[3] OLSHAUSEN B A, FIELD D J. Emergence of simple-cell receptive field properties by learning a sparse code for natural images[J]. Nature,1996(381):607-609.

[4] YANG J, YU K, GONG Y, et al. Linear spatial pyramid matching using sparse coding for image classification[C]. IEEE Confenece on CVPR, 2009:1794-1801.

[5] CAI D, HE X F, HE Y X, et al. Learning a spatially smooth subspace for face recognition[C]. IEEE Confenece on CVPR, 2007:1-7.

(收稿日期:2012-08-25)

作者简介:

徐东勋,男,1987年生,硕士研究生,主要研究方向:机器学习,图像识别,稀疏编码。