

基于 L1 范数的二维局部保留映射

邢红杰 赵浩鑫

(河北大学数学与计算机学院 河北省机器学习与计算智能重点实验室 保定 071002)

摘要 提出了一种基于 L1 范数的二维局部保留映射(two-dimensional locality preserving projections based on L1-norm, 2DLPP-L1)特征提取方法。与传统的基于 L2 范数的二维局部保留映射(2DLPP)相比,所提方法有两个优点。首先,由于 L1 范数对噪声不敏感,因此它具有更强的抗噪声能力;其次,它不需要进行特征值分解。在两个人脸数据库和一个手写数字数据集上的实验结果表明,当训练集中有噪声时,所提的 2DLPP-L1 能够取得优于传统 2DLPP 的分类性能。

关键词 特征提取, L1 范数, 局部保留映射, 人脸识别

中图分类号 TP391.4 **文献标识码** A

Two-dimensional Locality Preserving Projections Based on L1-norm

XING Hong-Jie ZHAO Hao-Xin

(Key Laboratory of Machine Learning and Computation Intelligence, College of Mathematics and Computer Science, Hebei University, Baoding 071002, China)

Abstract This paper presented a method of two-dimensional locality preserving projection based on L1-norm(2DLPP-L1). The proposed approach has two advantages compared with the conventional L2-norm based two-dimensional locality preserving projection(2DLPP). Firstly, it is more robust against outliers because L1-norm is insensitive to noises. Moreover, it does not require the eigenvalue decomposition. Experiments on two face databases and one hand-written digit dataset illustrate that compared with 2DLPP, the proposed method exhibits better performance when there are outliers in training sets.

Keywords Feature extraction, L1-norm, Locality preserving projection, Face recognition

1 引言

特征提取被认为是模式识别和机器学习领域中一个关键的研究课题,它通常作为一种预处理方法来克服维数灾难难题。主成分分析(principal component analysis, PCA)^[1]和线性判别分析(linear discriminant analysis, LDA)^[2]是两种常用的特征提取方法。PCA 的目标是寻找一个最优的正交线性变换,使数据在低维的特征空间中被进一步处理,同时保留尽可能多的数据信息(例如最大方差)。然而,传统的 PCA 是一种无监督的学习方法,没有考虑类别信息,无法使用训练数据中的判别信息。LDA 通过同时最小化类内散布矩阵和最大化类间散布矩阵来保留判别信息,使得 LDA 在取得的低维空间中能够获得最具判别能力的投影。Martínez 和 Kak 指出,在处理实际问题时, LDA 通常比 PCA 更具优势^[3]。例如,在某些人脸识别任务中,当训练数据充分时, LDA 通常能够取得比 PCA 更高的分类准确率。尽管如此,他们还指出当训练集数据很少时, LDA 的性能会劣于 PCA^[3]。最近, He 和 Niyogi 提出了一种新的特征提取方法,即局部保留映射(locality preserving projections, LPP)。LPP 能够保留数据中暗

含的局部结构。文献[4]指出,与 PCA 和 LDA 相比, LPP 表现出了更优的判别能力。

以上提到的这些特征提取方法都是基于数据的向量空间表示形式,所以当训练数据很少时,经常会遇到不适定问题,例如,当训练集中的训练样本非常少时, PCA、LDA 和 LPP 需要的特征值分解就会产生奇异值问题。为了解决这个问题,目前许多的工作将上述方法推广到它们对应的矩阵表示形式,如二维主成分分析(two-dimensional PCA, 2DPCA)^[5,6]、二维线性判别分析(two-dimensional LDA, 2DLDA)^[7,8]和二维局部保留映射(two-dimensional LPP, 2DLPP)^[9-11]均已提出。与基于向量形式的方法相比,这些基于矩阵形式的二维方法直接从给定的图像矩阵中提取图像特征,而不是在图像矩阵转化成的向量中提取。这些二维特征提取方法比它们的一维版本更具吸引力,因为它们能够取得更优的性能。此外,它们的计算复杂度要比相应的一维版本低得多。在这 3 种基于矩阵形式的方法中,与 2DPCA 和 2DLDA 相比, 2DLPP 能够取得更高的识别准确率和更快的特征提取速度。

上述一维和二维特征提取方法都是基于 L2 范数的。然

到稿日期:2011-06-20 返修日期:2011-09-16 本文受国家自然科学基金(60903089, 61073121), 中国博士后科学基金(20080440820), 河北省自然科学基金(F2009000231)资助。

邢红杰(1976—),男,博士,副教授,主要研究方向为模式识别、机器学习, E-mail: hixing@hbu.edu.cn.

而, L2 范数能够放大训练数据中噪声的不良影响, 因此基于 L2 范数的方法往往对噪声非常敏感。为了解决该问题, 最近出现了基于 L1 范数的 PCA (PCA-L1)^[12], 并与基于 L2 范数的 PCA 以及基于旋转不变 L1 范数的 PCA (R1-PCA)^[13] 进行了比较。实验结果表明, 在一些基准数据集和人脸重构任务中, PCA-L1 更具优势。受到 PCA-L1 的启发, Pang 等人提出了基于 L1 范数的张量分析方法 (TPCA-L1)^[14] 和基于 L1 范数的 LPP (LPP-L1)^[15]。他们通过实验表明, 所提出的两种基于 L1 范数的方法与相应的基于 L2 范数的方法相比, 都取得了更优的性能。此外, Li 等人提出了基于 L1 范数的 2DPCA (2DPCA-L1)^[16], 其与 PCA-L1 以及传统的基于 L2 范数的 2DPCA (2DPCA-L2) 相比, 具有诸多优点。通过使用旋转不变性 L1 范数, Li 等人^[17] 提出了 R1-LDA, R1-2DLDA 和 R1-TLDA。这些方法的一个局限性是, 它们都需要通过特征值分解来获得特征投影, 这就使得上述算法经常会遇到奇异值问题, 从而变得不稳定。

受 2DPCA-L1^[16] 的启发, 使用 L1 范数代替传统 2DLPP 中的 L2 范数, 从而提出基于 L1 范数的 2DLPP (2DLPP-L1)。与 2DLPP 相比, 所提方法有两个优点:

鲁棒的抗噪声能力: 因为 L1 范数对噪声不敏感, 所以所提的 2DLPP-L1 比基于 L2 范式的 2DLPP 具有更强的抗噪声能力;

避免了特征值分解: 2DLPP 必须用特征值分解来提取特征向量, 这可能会引发奇异值问题。提出的 2DLPP-L1 则不同, 它不需要进行特征值分解, 而是使用一个迭代算法来获得最优映射。

提出的 2DLPP-L1 也比它的一维版本 LPP-L1 更具优势。LPP-L1 尽管也具有鲁棒的抗噪声能力且不需要进行特征值分解, 但是不能充分利用图像的空间信息, 因为它将图片的二维几何信息退化到高维的一维向量中。相比而言, 2DLPP-L1 将图片自然地表示成矩阵形式, 取得了更多的空间信息。

本文第 2 节简要地回顾了传统的 2DLPP 算法; 第 3 节详细阐述了所提的 2DLPP-L1 方法; 对于一个有监督学习任务, 第 4 节描述了特征提取和分类的过程; 第 5 节使用相关实验验证了所提方法在带有噪声的数据集上能够取得优于另外 3 个相关模型的性能; 最后给出了概括性总结和对未来研究工作的建议。

2 2DLPP

2DLPP 是一种基于矩阵的特征提取方法, 可以看成是 LPP 的二维扩展模型。在 2DLPP 中, 存在于图像矩阵中的局部信息可以保留到获得的二维子空间中。为了得到最优的投影向量, 最小化下面的目标函数:

$$\sum_{i=1}^N \sum_{j=1}^N \|y_i - y_j\|_2^2 s_{ij} \quad (1)$$

式中, $\|\cdot\|_2$ 是 L2 范数; $y_i = X_i w$ 是图像 X_i 投影后的特征向量, s_{ij} 是图像 X_i 和 X_j 之间的相似度, 它的具体定义形式可参见文献[9]。通过数学推导, 式(1)中最小化问题的解可以通过求解下面的特征值分解问题获得^[9]:

$$X^T (L \otimes I_m) X w = \lambda X^T (D \otimes I_m) X w \quad (2)$$

式中, $X^T = [X_1^T, X_2^T, \dots, X_N^T]$, D 是一个对角阵, 其元素为

$D_{ii} = \sum_{j=1}^N s_{ij}$, $L = D - S$ 为拉普拉斯矩阵, $S_{ij} = s_{ij}$, \otimes 是 Kronecker 积。 I_m 是一个 $m \times m$ 的单位矩阵。详细的推导可参见文献[9-11]。

3 基于 L1 范数的 2DLPP

本节首先对所提的 2DLPP-L1 方法加以公式化描述, 然后证明其收敛性, 最后描述 2DLPP-L1 具体的算法实现。

3.1 问题的数学描述

设 X_i 表示一个 $m \times n$ 像素的图片, 其中 $i=1, 2, \dots, N$, w 表示一个 n 维的单位列向量。每张图片 X_i 在 w 上的投影定义为:

$$y_i = X_i w = \begin{pmatrix} x_{i1} \\ \vdots \\ x_{im} \end{pmatrix} w \quad (3)$$

式中, $y_i \in \mathcal{R}^{n \times 1}$ 称为投影特征, $x_{il} \in \mathcal{R}^{1 \times n}$ 是 X_i 的第 l 个行向量。与文献[15]中的定义类似, X_i 和 X_j 之间的不相似度 d_{ij} 定义为:

$$d_{ij} = 1 - s_{ij} \quad (4)$$

式中, s_{ij} 为 X_i 和 X_j 的相似度, 其定义如下:

$$s_{ij} = \begin{cases} \exp\left(\frac{-\|X_i - X_j\|_2}{\gamma}\right), & \text{如果 } X_i \text{ 是 } X_j \text{ 的 } k \text{ 近邻点之一,} \\ & \text{或者 } X_j \text{ 是 } X_i \text{ 的 } k \text{ 近邻点之一} \\ 0, & \text{其他} \end{cases} \quad (5)$$

注意, 这里定义的相似度还有其它的定义形式, 详细描述可参见文献[9]中的有关内容。式(5)中的参数 γ 可预定义为:

$$\gamma = C \times \max(\|X_i - X_j\|_2) \quad (i, j=1, 2, \dots, N) \quad (6)$$

式中, C 是一个合适的常数。

因此, 2DLPP-L1 中需要被最大化的目标函数可以表示为:

$$g(w) = \sum_{i=1}^N \sum_{j=1}^N \|y_i - y_j\|_1 d_{ij} \quad (7)$$

式中, $\|\cdot\|_1$ 是 L1 范数。将式(3)代入式(7), 可得:

$$\begin{aligned} g(w) &= \sum_{i=1}^N \sum_{j=1}^N \|X_i w - X_j w\|_1 d_{ij} \\ &= \sum_{i=1}^N \sum_{j=1}^N \sum_{l=1}^m |x_{il} w - x_{jl} w| d_{ij} \\ &= \sum_{i=1}^N \sum_{j=1}^N \sum_{l=1}^m |(x_{il} - x_{jl}) w| d_{ij} \end{aligned} \quad (8)$$

为了去掉式(8)中的绝对值符号, 定义了如下一个极性函数 p_{ijl} :

$$p_{ijl} = \begin{cases} 1, & \text{如果 } (x_{il} - x_{jl}) w > 0 \\ -1, & \text{如果 } (x_{il} - x_{jl}) w \leq 0 \end{cases} \quad (9)$$

将式(9)代入式(8), 并且考虑到 $w^T w = 1$, 2DLPP-L1 的最优化问题为:

$$\begin{cases} \max_w g(w) = \max_w \sum_{i=1}^N \sum_{j=1}^N \sum_{l=1}^m p_{ijl} (x_{il} - x_{jl}) w d_{ij} \\ \text{s. t. } w^T w = 1 \end{cases} \quad (10)$$

如文献[12]中的描述和证明, 在第 $t+1$ 次迭代中最优的 $w(t+1)$ 的更新公式也可类似地给出:

$$w(t+1) = \frac{\sum_{i=1}^N \sum_{j=1}^N \sum_{l=1}^m p_{ijl}(t) (x_{il} - x_{jl})^T d_{ij}}{\left\| \sum_{i=1}^N \sum_{j=1}^N \sum_{l=1}^m p_{ijl}(t) (x_{il} - x_{jl})^T d_{ij} \right\|_2} \quad (11)$$

完成 $w(t+1)$ 的更新后,第 $t+1$ 次迭代的 p_{ijl} 可由下式给出:

$$p_{ijl}(t+1) = \begin{cases} 1, & \text{如果 } (x_{il} - x_{jl})w(t+1) > 0 \\ -1, & \text{如果 } (x_{il} - x_{jl})w(t+1) \leq 0 \end{cases} \quad (12)$$

3.2 收敛性性质

受 Kwak^[12] 提出的 PCA-L1 收敛性定理以及 Pang 等人关于 LPP-L1 的收敛性定理和证明^[15] 的启发,将 2DLPP-L1 的收敛性性质描述如下。

定理 1 设投影向量和极性函数的更新公式分别由式(11)和式(12)定义,则投影向量 w 收敛到式(8)的一个局部最大值点。

证明:根据 Kwak 在文献^[12]中给出的证明,只需要证明 $g(w(t+1)) \geq g(w(t))$ 。从式(12)中极性函数的定义以及 $d_{ij} \geq 0$,可知:

$$p_{ijl}(t+1)(x_{il} - x_{jl})w(t+1)d_{ij} \geq 0 \quad (13)$$

然而,可有下式成立:

$$p_{ijl}(t)(x_{il} - x_{jl})w(t+1)d_{ij} \leq 0 \quad (14)$$

进而可得:

$$g(w(t+1)) = \sum_{i,j,l} p_{ijl}(t+1)(x_{il} - x_{jl})w(t+1)d_{ij} \geq \sum_{i,j,l} p_{ijl}(t)(x_{il} - x_{jl})w(t+1)d_{ij} \quad (15)$$

用式(11)代替不等号右面的 $w(t+1)$,则有:

$$\sum_{i,j,l} p_{ijl}(t)(x_{il} - x_{jl})w(t+1)d_{ij} = \left\| \sum_{i,j,l} p_{ijl}(t)(x_{il} - x_{jl})^T d_{ij} \right\|_2 \quad (16)$$

并且,

$$\begin{aligned} & \left\| \sum_{i,j,l} p_{ijl}(t)(x_{il} - x_{jl})^T d_{ij} \right\|_2 \geq \\ & \left\| \sum_{i,j,l} p_{ijl}(t)(x_{il} - x_{jl})^T d_{ij} \right\|_2 \times \cos \langle \sum_{i,j,l} p_{ijl}(t)(x_{il} - x_{jl})^T d_{ij}, \sum_{i,j,l} p_{ijl}(t-1)(x_{il} - x_{jl})^T d_{ij} \rangle \\ & = \left[\sum_{i,j,l} p_{ijl}(t)(x_{il} - x_{jl})^T d_{ij} \right] \\ & \frac{\sum_{i,j,l} p_{ijl}(t-1)(x_{il} - x_{jl})^T d_{ij}}{\left\| \sum_{i,j,l} p_{ijl}(t-1)(x_{il} - x_{jl})^T d_{ij} \right\|_2} \\ & = \sum_{i,j,l} p_{ijl}(t)(x_{il} - x_{jl})d_{ij}w(t) = g(w(t)) \end{aligned} \quad (17)$$

因此,可得 $g(w(t+1)) \geq g(w(t))$,证毕。

3.3 2DLPP-L1 算法

使用 2DLPP-L1 提取单个投影特征向量的过程概括在了算法 1 中。这里需要强调的是,算法 1 仅能提取第一个投影向量 w 。因此,对于某个图像 X_i ,只能得到一个向量形式的特征向量 y_i 。

算法 1 用于提取单个投影特征向量的 2DLPP-L1

输入:二维图像因素 $\{x_i \in \mathcal{R}^{m \times n}\}_{i=1}^N$,不相似度矩阵 $(d_{ij})_{N \times N}$
 输出:最优投影向量 $w^* \in \mathcal{R}^{n \times 1}$
 初始化:投影向量 $w(0)$ 中的元素随机地从 $(0, 1)$ 中选取,设 $w(0) = w(0) / \|w(0)\|_2$, $t=0$ 且 $flag=0$
 while ($flag=0$) do
 Step 1 if $(x_{il} - x_{jl})w(t) > 0$ then
 $p_{ijl}(t) = 1$;
 else
 $p_{ijl}(t) = -1$ 。
 endif
 Step 2 令 $t=t+1$, $w(t) = \sum_{i,j,l} p_{ijl}(x_{il} - x_{jl})^T d_{ij}$ 且 $w(t) = w(t) / \|w(t)\|_2$ 。
 Step 3 if $w(t) = w(t-1)$ then
 if 存在 $(x_{il} - x_{jl})w(t) = 0$ then

 令 $w(t) = [w(t) + \Delta w] / \|w(t) + \Delta w\|_2$, 返回 Step 1。

 endif
 令 $flag=1$ 。
 endif

end while

$w^* = w(t)$

然而,在许多特征提取任务中,仅用一个主投影特征向量来重构原始数据集或者图像集的信息是不够的。因此,为了使 2DLPP-L1 能够提取多个投影特征向量,将算法 1 加以扩展,并将特征提取过程概括在算法 2 中。

算法 2 用于提取多个投影特征向量的 2DLPP-L1

输入:二维图像集 $\{X_i \in \mathcal{R}^{m \times n}\}_{i=1}^N$,不相似度矩阵 $(d_{ij})_{N \times N}$

输出:最优投影矩阵 $W^* = [w_1^*, w_2^*, \dots, w_d^*] \in \mathcal{R}^{n \times d}$

初始化:令 $w_0 = 0 \in \mathcal{R}^{n \times 1}$, $Z_j^0 = [X_i - X_j]_{\mathcal{R}^{m \times n}}$

for $r=1$ to d do

$$Z_j^r = Z_j^{(r-1)} - Z_j^{(r-1)} w_{r-1} w_{r-1}^T$$

 通过执行算法 1 中的步骤寻找最优的投影向量 w_r^* , 令 $w_r = w_r^*$

end for

$$W^* = [w_1^*, w_2^*, \dots, w_d^*]$$

与 LPP-L1^[15] 类似,从算法 2 中可以观察到 2DLPP-L1 需将原始图像集 $\{X_i\}_{i=1}^N$ 转化成它们对应的差异矩阵,即 $\{X_i - X_j\}_{i,j=1}^N$ 。此外,当使用 2DLPP-L1 提取多个投影特征向量时,需将差异矩阵看作一个整体来处理。

4 特征提取和分类

给定一个训练图像集 $\{X_i\}_{i=1}^N$, 2DLPP-L1 寻找 d 个最佳的投影向量 $w_1^*, w_2^*, \dots, w_d^*$, 它们是矩阵 W^* 的列向量。对于一个给定的图像 X_k , 可通过以下方式获得其投影特征矩阵 Y_k :

$$Y_k = X_k W^* \quad (18)$$

式中, $Y_k = [y_k^1, y_k^2, \dots, y_k^d] \in \mathcal{R}^{m \times d}$ 称作图像 X_k 的特征矩阵。

因此,对于一个给定的训练图像集 $\{X_i\}_{i=1}^N$, 可以获得 N 个特征矩阵,即 Y_1, Y_2, \dots, Y_N 。本文中最近邻分类器(1-nearest neighbor classifier, 1NN)来分类。给定两个特征矩阵 Y_i 和 Y_j , 它们之间的距离定义如下:

$$d(Y_i, Y_j) = \sum_{r=1}^d \|y_i^r - y_j^r\|_2 = \sum_{r=1}^d \sqrt{\sum_{p=1}^m (y_i^{(rp)} - y_j^{(rp)})^2} \quad (19)$$

式中, $y_i^r = [y_i^{(r1)}, y_i^{(r2)}, \dots, y_i^{(rm)}]^T$ 。因此,假设 $\{X_i\}_{i=1}^N$ 的特征矩阵为 $\{Y_i\}_{i=1}^N$, 其中每个矩阵对应的类标为 ω_i 。对于一个测试图片 X , 它的特征矩阵为 Y 。如果 $d(Y, Y_i) = \min_i d(Y, Y_i)$ 并且训练图片 X_i 的类标是 ω_i , 该测试图片 X 就被划分到 ω_i 类。

5 实验结果

以下实验中所有的代码都是使用 MATLAB 语言编写的。为了验证所提 2DLPP-L1 方法的有效性,在 3 个图像数据库,即 ORL(AT&T)人脸数据库、Yale 人脸数据库和手写数字数据集上进行了实验,并与现有的方法进行了比较。在下面的 3 个实验中,部分训练图片被加入噪声,以测试 2DLPP-L1 的鲁棒性。对于 2DLPP-L2 (基于 L2 范数的 2DLPP,即传统 2DLPP)和 2DLPP-L1,式(6)中的常数 C 决定着热核中的参数 γ 。在实验中将常数 C 设置为 2000。值得一提的是,实验中参与比较的方法都是基于矩阵形式的二维特

征提取方法。我们没有将所提方法与基于L2范数的2DLDA进行对比,原因是2DLDA使用了类别信息,而2DLPP-L1没有使用任何判别信息。

5.1 在ORL人脸数据库上的实验结果

ORL(AT&T)人脸数据库¹⁾由10个人的400张图片组成,每个人40张图片。这400张图片都是像素为112×92的灰度图像。为了降低计算复杂度,所有的图片都裁剪为56×46像素的图片。实验中,选取每个人的前2张图片进行训练,剩余图片用来测试,则在训练集中共有80张图片,在测试集中有320张图片。

为了测试2DLPP-L1的鲁棒性,从训练集中随机选取10%的图片加入噪声。与文献[12]中使用的噪声生成方法相同,在每张带有噪声的图片中,长方形噪声的大小是随机的,且噪声在随机选取的训练图片中的位置也是随机的。长方形噪声的大小被随机地设置在14×12到56×46像素之间。长方形噪声中的点被随机地设置为黑白点。图1展示了8张噪声图片。对于2DLPP-L2和2DLPP-L1,在构建相似矩阵时的近邻个数被设置为1。

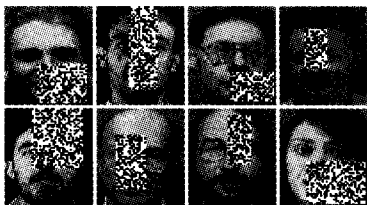


图1 ORL人脸数据库训练集中的噪声图片

当投影向量的数目不同时,4种不同方法在测试集上的识别误差率如图2所示。从图2中可以观察到,2DLPP-L1、2DPCA-L2(基于L2范式的2DPCA,即2DPCA)、2DPCA-L1和2DLPP-L2的最低错误率分别为15.63%、15.94%、16.88%和46.88%。与2DLPP-L2相比,所提2DLPP-L1的错误率更低,抗噪声能力更强。此外,与2DPCA-L2和2DPCA-L1相比,2DLPP-L1取得了更优的识别性能。

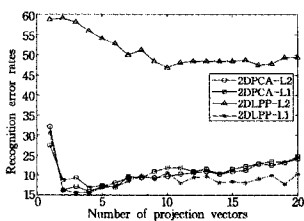


图2 在ORL人脸数据库测试集上不同方法的识别误差率

5.2 在Yale人脸数据库上的实验结果

Yale人脸数据库²⁾由15个人的165张图片组成,每个人各有11张不同的照片。Yale数据库中这些图片都是灰度图像,图像的原始大小为243×320像素。在预处理步骤中,人工地剪裁这些图片,并调整为40×50像素。然后,使用5折交叉验证来测试2DLPP-L1的性能。也就是说,这个图片集被平均分为5组,其中4组图片用来训练,余下的一组用来测试。最后的测试误差率是这5次实验的平均测试误差率。为

了检验2DLPP-L1的鲁棒性,在每次实验中,随机选择训练集中的12张图片来生成噪声图片。生成噪声的方法与5.1节中的方法类似。图3展示了在最后一次实验中训练集的12张噪声图片。在2DLPP-L2和2DLPP-L1中,构建相似矩阵时的近邻个数均设置为3。

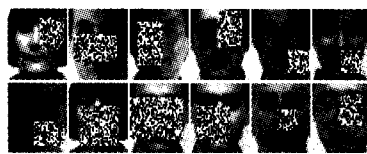


图3 Yale人脸数据库训练集中的噪声图片

2DPCA-L2、2DPCA-L1、2DLPP-L2和2DLPP-L1 4种方法在Yale人脸数据库上的实验结果概括在表1中。在表1中对比了投影矩阵维数均取50×4时的分类误差率。从表1的结果中可以发现,与其他3种方法相比,2DLPP-L1取得了最好的泛化性能。表1中的试验结果同时展示了参数相同同时,2DLPP-L1比2DLPP-L2具有更强的抗噪声能力。

表1 在Yale人脸数据库上4种不同方法的测试误差率

方法	投影矩阵的维数	测试误差率
2DPCA-L2	50×4	20.00%
2DPCA-L1	50×4	19.11%
2DLPP-L2	50×4	20.67%
2DLPP-L1	50×4	17.78%

5.3 在手写数字数据库上的实验结果

手写数字数据库³⁾包含1404张、数字0~9及大写字母A~Z的二值图片。其中,每个数字或大写字母都有39张图片。在这个实验中,只使用数字图片,则共有10类样本和390张图片。这390张图片均为20×16像素的二值图片(即每个像素值非0则1)。我们仍然使用5折交叉验证法来测试所有方法的性能。此外,在每次实验中,随机地从训练集中选择30张图片来生成噪声图片。为了生成噪声图片,在每张图片中随机选择32个像素点并将它们的像素值取反,即如果像素值是1就置成0,如果像素值为0就置为1。图4展示了在第一次实验中训练集生成的30张噪声图片。在这个实验中,2DLPP-L2和2DLPP-L1构建相似矩阵时的近邻个数均设为8。2DPCA-L2、2DPCA-L1、2DLPP-L2和2DLPP-L1 4种方法在手写数字数据库上的实验结果如表2所列。



图4 手写数字数据库训练集中的噪声图片

表2 在手写数字数据库上4种不同方法的测试误差率

方法	投影矩阵的维数	测试误差率
2DPCA-L2	16×4	8.07%
2DPCA-L1	16×4	8.61%
2DLPP-L2	16×4	8.29%
2DLPP-L1	16×4	7.79%

(下转第238页)

¹⁾ <http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html>

²⁾ <http://cvc.yale.edu/projects/yalefaces/yalefaces.html>

³⁾ <http://www.cs.nyu.edu/~roweis/data/>

到的 CSS 的过程中,仍需要根据实际情况进行具体分析。然而,提出完备的 SFS 形式化推演规则是基于 CSS 的 DFT 定性分析方法的最终目标,因为基于该推演规则,可以利用计算机辅助方法自动完成 CSS 的定性分析,进而快速给出动态系统中导致顶事件失效的基本事件动态失效模式。

参 考 文 献

- [1] Dugan J B, Bavuso S, Boyd M. Dynamic fault tree models for fault tolerant computer systems[J]. IEEE transaction on reliability, 1992, 41(3): 363-377
- [2] 李堂经, 王新阁, 杨哲. 动态故障树的综合分析方法[J]. 装备制造技术, 2009(8): 22-23, 49
- [3] Zhang X, Miao Q, Fan X, et al. Dynamic fault tree analysis based on Petri nets[C]//Proceedings of 8th international conference on reliability, maintainability and safety. Chengdu, China, 2009: 138-142
- [4] Marquez D, Neil M, Fenton N. Solving dynamic fault trees using a new hybrid bayesian network inference algorithm[C]//Proceedings of 16th Mediterranean Conference on Control and Automation. Ajaccio, France, 2008: 609-614

- [5] Boudali H, Crouzen P, Stoelinga M. Dynamic fault tree analysis using input/output interactive markov chains[C]//Proceedings of 37th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN'07). Edinburgh, UK, 2007: 708-717
- [6] Sullivan K J, Dugan J B, Coppit D. The Galileo fault tree analysis tool[C]//Proceedings of the 29th Annual International Symposium on Fault-Tolerant Computing. Madison, Wisconsin, 1999: 232-235
- [7] Relex[EB/OL]. <http://www.ptc.com/products/relex/>, 2010-12-16
- [8] Liu D, Xing W, Zhang C, et al. Cut sequence set generation for fault tree analysis[C]//Proceedings of 2007 International Conference on Embedded Software and Systems(ICESS 2007). Dae-gu, South Korea, 2007: 58-69
- [9] 陈越洲, 谭琳, 邢维艳, 等. 一种新的故障树定性分析方法[J]. 计算机工程, 2008, 34(13): 67-69
- [10] Liu D, Zhang C, Xing W, et al. Quantification of cut sequence set for fault tree analysis[C]//Perrott R, et al., eds. Proceedings of the 3rd international conference on high performance computing and communication(HPCC 2007). 2007: 755-765

(上接第 204 页)

结束语 受 2DPCA-L1 的启发, 本文提出了 2DLPLP-L1。2DLPP-L1 使用的是 L1 范数而非 L2 范数, 这样 2DLPP-L1 就具有了更强的抗噪声能力。另外, 2DLPP-L1 使用迭代方法来寻找最优的投影向量, 避免了复杂的特征值分解, 从而不会出现奇异值问题。与传统的 2DLPP-L2 相比, 所提出的 2DLPP-L1 抗噪声能力更强。与 2DPCA-L2 和 2DPCA-L1 相比, 2DLPP-L1 也表现出了更优的识别性能。

为了使所提方法更具吸引力, 在以后的工作中, 我们会从以下两个方面对 2DLPP-L1 加以研究和探讨。首先, 2DLPP-L1 需要将原始的数据集转化成由差异向量构成的矩阵。而基于差异向量的迭代过程非常耗时, 下一步我们会减小 2DLPP-L1 的时间复杂度并将之应用于大规模数据集; 第二, 2DLPP-L1 没有考虑类别信息。在以后的工作中, 我们将对利用判别信息的监督 2DLPP-L1 加以研究。

参 考 文 献

- [1] Jolliffe I T. Principal Component Analysis(2nd Edition)[M]. New York: Springer, 2002
- [2] Duda R O, Hart P E, Stork D H. Pattern Classification(2nd Edition)[M]. Wiley Interscience, 2000
- [3] Martínez A M, Kak A C. PCA Versue LDA[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2001, 23(2): 228-233
- [4] He X, Niyogi P. Locality Preserving Projections[C]//Advance in Neural Information Processing System. 2003, 16: 152-160
- [5] Yang J, Zhang D, Frangi A, et al. Two-dimensional PCA: A New Approach to Appearance-based Face Representation and Recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2004, 26(1): 131-137
- [6] Meng J, Zhang W. Volume Measure in 2DPCA-based Face Recognition[J]. Pattern Recognition Letters, 2007, 28(10): 1203-1208

- [7] Ye J, Janardan R, Li Q. Two-dimensional Linear Discriminant Analysis[C]//Advance in Neural Information Processing Systems. 2004, 2: 1569-1576
- [8] Liang Z, Li Y, Shi P. A Note on Two-dimensional Linear Discriminant Analysis[J]. Pattern Recognition Letters, 2008, 29: 2122-2128
- [9] Chen S, Zhao H, Kong M, et al. A Two-Dimensional Extension of Locality Preserving Projections[J]. Neurocomputing, 2007, 70: 912-921
- [10] Hu D, Feng G, Zhou Z. Two-dimensional Locality Preserving Projections(2DLPP) with Its Application to Palmprint Recognition[J]. Pattern Recognition, 2007, 40: 339-342
- [11] Yu W. Two-dimensional Discriminant Locality Preserving Projections for Face Recognition[J]. Pattern Recognition Letters, 2009, 30(15): 1378-1383
- [12] Kwak N. Principal Component Analysis Based on L1-norm Maximization[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2008, 30(9): 1672-1680
- [13] Ding C H Q, Zhou D, He X, et al. R1-PCA: Rotational Invariant L1-norm Principal Component Analysis for Robust Subspace Factorization[C]//Proceedings of the 23rd International Conference on Machine Learning. 2006: 281-288
- [14] Pang Y, Li X, Yuan Y. Robust Tensor Analysis with L1-norm[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2010, 20(2): 172-178
- [15] Pang Y, Yuan Y. Outlier-resisting Graph Embedding[J]. Neurocomputing, 2010, 73: 968-974
- [16] Li X, Pang Y, Yuan Y. L1-norm-based DPCA[J]. IEEE Transactions on Systems, Man, Cybernetics-Part B: Cybernetics, 2009, 40(4): 1170-1175
- [17] Li X, Hu W, Wang H, et al. Linear Discriminant Analysis Using Rotational Invariant L1-norm[J]. Neurocomputing, 2010, 73: 2571-2579