

不同累计贡献率下 PDEs 对 CCA 的影响

雷田田 王 振 范丽亚 杨 静

(聊城大学 数学科学学院, 山东 聊城 252059)

摘 要 针对高维噪声图像(光暗、有遮挡)的分类问题,分类质量的好坏严重依赖于图像降噪方法和降维方法的选择.本文提出基于偏微分方程组(PDEs)的典型相关分析(CCA)作为联合降噪降维方法,并研究了 PDEs 的进化对 CCA 中累计贡献率的影响.分别以支持向量机(SVM)和一对余多类 SVM(OVR-MSVM)为分类器,以分类精度为标准,在 Extended Yale B 人脸数据集和 Palm-print 掌纹数据集上进行了大量的实验.实验结果表明,PDEs 的进化可以弱化累计贡献率的影响,甚至不用选择累积贡献率,通过不超过 5 次的进化即可达到同等条件下的最佳分类精度.

关键词 图像分类;偏微分方程组;典型相关分析;累计贡献率;进化次数

中图分类号 TP391

文献标识码 A

0 引言

图像特征抽取作为模式识别研究的核心问题之一,在机器学习、图像处理以及计算机视觉等相关领域有着广泛的应用背景^[1,2].特征抽取在本质上就是寻找一种映射方式,利用这种方式将数据从高维原始空间映射到低维子空间中,使得降维后的数据仍保持尽可能多的识别特征,以便做进一步处理.对于高维噪声图像(光暗、有遮挡)的分类与识别任务,清洗图像以及抽取图像的有效特征是完成任务的关键.近年来,有关特征抽取方法的研究成果颇丰,其中线性判别分析(Linear Discriminant Analysis, LDA)^[3,4]和主成分分析(Principal Component Analysis, PCA)^[5,6]是最具代表性的两种方法.但是,这两种方法主要是针对数据的一组特征进行降维处理的.

随着信息时代的来临,相同的模式不可避免地会出现多个不同的特征表示方式,每个特征表示都反映了同一模式的不同特性.可以将多个特征表示首尾相连形成一个新的高维特征表示^[7],也可以利用复向量将两组特征合并在一起^[8],然后利用 PCA, LDA 等方法进行特征抽取.这样做在一定程度上可以提高图像的分类或识别质量,但却忽略了特征表示间的内在联系以及表示的多样性.为了改进这一点,本文选用典型相关分析(Canonical Correlation Analysis, CCA)^[9,10]作为图像降维工具,其优势是考虑了多组特征间的线性相关性,但也出现了累计贡献率如何选取的问题.累计贡献率决定了降维子空间的维度,贡献率越高,维度越大,计算成本也越高,但分类精度并不一定越好.目前,针对累计贡献率还没有一个行之有效的选择方法,习惯上选为 100%,这样一来很多时候起不到降维作用.

另一方面,尽管 CCA 通过考虑多组特征间的线性相关性,提高了特征表示的鉴别性与鲁棒性,但其在特征抽取过程中并没有考虑光照、遮挡对图像质量的影响,这限定了 CCA 处理高维噪声图像(光暗、有遮挡)的能力.2010 年和 2013 年, Liu 等人分别提出了将偏微分方程组(Partial Differential Equations, PDEs)应用于图像恢复和计算机视觉的方法^[11,12].2017 年, Fang 等人^[13]提出了一种基于 PDEs 的图像特征提取方法,所提方法既考虑了图像的判别信息,也考虑了图像在平移、旋转下的不变性以及对照光的鲁棒性,但没有考虑图像维数约减问题.2018 年和 2019 年,江等人分别研究了 PDEs 的进化次数对图像特征质量的影响^[14]以及对正则化 SVM 模型参数的影响^[15].

收稿日期:2019-08-01

基金项目:国家自然科学基金项目(11801248);山东省自然科学基金项目(ZR2016AM24, ZR2018BF010)资助

通讯作者:范丽亚,女,汉族,博士,教授,研究方向:机器学习理论与应用, E-mail: fanliya63@126.com.

为了弥补 CCA 的不足,本文选用文献[13]中提出的方法作为图像降噪方法,并借助于文献[14,15]的成果,研究了 PDEs 的进化对累计贡献率的影响.针对 85%,90%,95%,100% 四个累积贡献率,分别以支持向量机 (Support Vector Machine, SVM)^[16,17] 和一对余多类 SVM (One-versus-Rest Multi-class SVM, OVR-MSVM)^[18,19] 为分类器,在 Extended Yale B 数据集和 Palm-print 数据集上进行了大量的实验.

1 支持向量机

本节简要回顾文中要用到的二类分类器 SVM 和多类分类器 OVR-MSVM. 在本文中,分别用 $\mathbf{e}_m = (1, \dots, 1)^\top \in \mathbf{R}^m$ 和 $\mathbf{I}_m \in \mathbf{R}^{m \times m}$ 表示 m 维 1 向量和 $m \times m$ 阶单位矩阵.

1.1 二类支持向量机

给定二分类数据集 $\mathbf{T} = \{(x_i, y_i)\}_{i=1}^m \in \mathbf{R}^d \times \{\pm 1\}$, 其中 $y_i \in \{\pm 1\}$ 表示第 i 个样本 $x_i \in \mathbf{R}^d$ 的类标签. 记 $\mathbf{X} = [x_1, x_2, \dots, x_m] \in \mathbf{R}^{d \times m}$ 和 $\mathbf{D} = \text{diag}(y_1, y_1, \dots, y_m) \in \mathbf{R}^{m \times m}$ 分别为样本矩阵和类标签矩阵, $\mathbf{y} = (y_1, \dots, y_1)^\top \in \mathbf{R}^m$ 为类标签向量.

线性 SVM 通过构建下面的二次规划模型

$$\begin{aligned} \min_{w, b, \xi} \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^m \xi_i \\ \text{s. t. } y_i (\langle \mathbf{w}, \mathbf{x}_i \rangle + b) \geq 1 - \xi_i, \xi_i \geq 0, i = 1, 2, \dots, m, \end{aligned} \quad (1)$$

来寻找分类决策函数 $f(\mathbf{x}) = \mathbf{w}^\top \mathbf{x} + b$, 其中 $\mathbf{w} \in \mathbf{R}^d$ 和 $b \in \mathbf{R}$ 分别是决策函数的法向量和阈值, $C > 0$ 是模型参数, $\{\xi\}_{i=1}^m$ 是松弛变量, $\|\cdot\|$ 和 $\langle \cdot, \cdot \rangle$ 分别表示向量的欧氏范数和内积. 通过求解模型(1)的 Wolfe 对偶模型

$$\begin{aligned} \min_{\alpha} \frac{1}{2} \alpha^\top \mathbf{G} \alpha - \mathbf{e}_m^\top \alpha \\ \text{s. t. } \mathbf{y}^\top \alpha = 0, 0 \leq \alpha \leq C \mathbf{e}_m, \end{aligned} \quad (2)$$

可得最优决策函数, 其中 $\alpha \in \mathbf{R}_+^m$ 是 Lagrange 乘子向量, $\mathbf{G} = \mathbf{D} \mathbf{X}^\top \mathbf{X} \mathbf{D}$ 是对称非负定阵. 具体算法如下

算法 1 (SVM)

- 步 1 给定数据集 $\mathbf{T} = \{(x_i, y_i)\}_{i=1}^m \in \mathbf{R}^d \times \{\pm 1\}$, 选择适当的模型参数 $C > 0$.
- 步 2 求解模型(2), 得最优解 $\alpha^* \in \mathbf{R}_+^m$.
- 步 3 计算 $\mathbf{w}^* = \mathbf{X} \mathbf{D} \alpha^* = \sum_{i=1}^m \alpha_i^* y_i \mathbf{x}_i$.
- 步 4 找 α^* 的一个正分量 $0 < \alpha_j^* < C$, 计算 $b^* = y_j - \langle \mathbf{w}^*, \mathbf{x}_j \rangle = y_j - \sum_{i=1}^m \alpha_i^* y_i \langle \mathbf{x}_i, \mathbf{x}_j \rangle$.
- 步 5 构造分类决策函数 $f(\mathbf{x}) = \langle \mathbf{w}^*, \mathbf{x} \rangle + b^*$.
- 步 6 对任一输入数据 $\tilde{\mathbf{x}} \in \mathbf{R}^d$, 其类标签为 $y_{\tilde{x}} = \text{sign}(f(\tilde{\mathbf{x}}))$.

1.2 多类支持向量机

给定 $K (K \geq 3)$ 类数据集 $\mathbf{T} = \{(\mathbf{x}_k, y_k)\}_{k=1}^m \in \mathbf{R}^d \times \{1, 2, \dots, K\}$, 其中第 i 类有 m_i 个样本且 $m_1 + \dots + m_K = m$. 用 $\mathbf{X} = [x_1, \dots, x_m] \in \mathbf{R}^{d \times m}$ 和 $\mathbf{X}_i = [x_1^{(i)}, \dots, x_{m_i}^{(i)}] \in \mathbf{R}^{d \times m_i}$ 分别表示整体样本矩阵和第 i 类样本矩阵, 则 \mathbf{X} 可表示为 $\mathbf{X} = [\mathbf{X}_1, \dots, \mathbf{X}_K]$. 用 $\mathbf{X}_{-i} = [\mathbf{X}_1, \dots, \mathbf{X}_{i-1}, \mathbf{X}_{i+1}, \dots, \mathbf{X}_K] \in \mathbf{R}^{d \times (m - m_i)}$ 表示去掉第 i 类的样本矩阵.

线性 OVR-MSVM 的基本思想是以 \mathbf{X}_i 为正类, 以 \mathbf{X}_{-i} 为负类 (即 $y_k^i = 1, \mathbf{x}_k^i \in \mathbf{X}_i; y_k^i = -1, \mathbf{x}_k^i \in \mathbf{X}_{-i}$), 利用线性 SVM 思想, 构建如下二次规划模型

$$\begin{aligned} \min_{w_i, b_i, \xi_k^i} \frac{1}{2} \|\mathbf{w}_i\|^2 + C_i \sum_{k=1}^m \xi_k^i \\ \text{s. t. } y_k^i (\mathbf{w}_i^\top \mathbf{x}_k^i + b_i) \geq 1 - \xi_k^i, \xi_k^i \geq 0, k = 1, 2, \dots, m, \end{aligned} \quad (3)$$

来寻找分类决策函数 $f_i(\mathbf{x}) = (\mathbf{w}_i^*)^\top \mathbf{x} + b_i^*$, 其中 $\mathbf{w}_i \in \mathbf{R}^d$ 和 $b_i \in \mathbf{R}$ 分别是超平面 $f_i(\mathbf{x}) = 0$ 的法向量和阈值, $C_i > 0$ 是模型参数, $\{\xi_k^i\}_{k=1}^m$ 是松弛变量. 通过求解模型(3)的 Wolfe 对偶形式

$$\begin{aligned} \min_{\alpha^i} \frac{1}{2} \alpha^{i\top} \mathbf{G}_i \alpha^i - \mathbf{e}_m^{\top} \alpha^i \\ \text{s. t. } \mathbf{y}^{i\top} \alpha^i = 0, 0 \leq \alpha^i \leq C_i \mathbf{e}_m, \end{aligned} \quad (4)$$

可得最优决策函数, 其中 $\alpha^i \in \mathbf{R}_+^m$ 是 Lagrange 乘子向量, $\mathbf{G}_i = \mathbf{D}_i \mathbf{X}_i^\top \mathbf{X}_i \mathbf{D}_i$, $\mathbf{D}_i = \text{diag}(y_1^i, y_2^i, \dots, y_m^i) \in \mathbf{R}^{m \times m}$. 下

面给出具体算法.

算法 2 (OVR-MSVM)

- 步 1 给定数据集 $T = \{(\mathbf{x}_i, y_k)\}_{k=1}^m \in \mathbf{R}^d \times \{1, 2, \dots, K\}$, 取 $i=1$.
- 步 2 令 \mathbf{X}_i 为正类, \mathbf{X}_{-i} 为负类, 选择适当的模型参数 $C_i > 0$.
- 步 3 求解模型(4), 得最优解 $\boldsymbol{\alpha}^i, \boldsymbol{\alpha} \in \mathbf{R}_+^m$.
- 步 4 计算 $\mathbf{w}_i = \mathbf{X}\mathbf{D}_i\boldsymbol{\alpha}^i = \sum_{k=1}^m \alpha_k^i y_k^i \mathbf{x}_k$.
- 步 5 找 $\boldsymbol{\alpha}^i$ 的一个正分量 $0 < \alpha_j^i < C_j$, 计算 $b_i = y_j^i - \langle \mathbf{w}_i, \mathbf{x}_j \rangle = y_j^i - \sum_{k=1}^m \alpha_k^i y_k^i \langle \mathbf{x}_k, \mathbf{x}_j \rangle$.
- 步 6 构造分类决策函数 $f_i(\mathbf{x}) = \langle \mathbf{w}_i, \mathbf{x} \rangle + b_i$.
- 步 7 若 $i < K$, 置 $i = i + 1$, 转步 2, 否则转步 8.
- 步 8 对任一输入数据 $\tilde{\mathbf{x}} \in \mathbf{R}^d$, 其类标签为 $y_{\tilde{\mathbf{x}}} = \operatorname{argmax}(f_i(\tilde{\mathbf{x}}) / \|\mathbf{w}_i\|)$.

从算法 2 中可以看出, OVR-MSVM 需要学习 K 个 SVM, 相比于其他多类分类器, OVR-MSVM 的学习时间短, 速度快, 但会产生数据不平衡的问题.

2 基于 PDEs 的图像降噪方法

本节简要回顾基于 PDEs 的图像降噪方法, 详细推导见文献[13, 14]. 考虑 c 类噪声(光暗、有遮挡)图像集 $\{(\mathbf{I}_m, \mathbf{h}_m)\}_{m=1}^M \in \mathbf{R}^{d_1 \times d_2} \times \mathbf{R}^c$, 其中 \mathbf{I}_m 为原始图像, \mathbf{h}_m 为 \mathbf{I}_m 对应的标签向量, 如果 \mathbf{I}_m 属于第 i 类, 则第 i 个分量 $h_{mi} = 1$; 反之 $h_{mi} = 0$, 记 $\mathbf{H} = [\mathbf{h}_1, \dots, \mathbf{h}_M] \in \mathbf{R}^{c \times M}$, 则基于 PDEs 的图像降噪模型为

$$\begin{aligned} \min_{\mathbf{W}, \{\mathbf{a}_i(t)\}} \mathbf{E} &= \mathbf{M}^{-1} \|\mathbf{H} - \mathbf{W} \cdot \mathbf{U}|_{t=T}\|_F^2 + \lambda \|\mathbf{W}\|_F^2 \\ \text{s. t. } \mathbf{u}_m^{n+1} &= \mathbf{u}_m^n + \Delta t \sum_{i=0}^5 a_i^n g(\operatorname{inv}_i(\mathbf{u}_m^n)) \in \mathbf{R}^{d_1 \times d_2}, \\ \mathbf{u}_m^0 &= \mathbf{I}_m, n=0, 1, \dots, N-1, m=1, 2, \dots, M, \end{aligned} \quad (5)$$

其中 $\mathbf{A} = [a_i^n]_{N \times 6}$ 为控制系数矩阵, $\mathbf{W} \in \mathbf{R}^{c \times d_1 d_2}$ 为权矩阵, $\mathbf{u}|_{t=T}$ 表示时刻 $t=T$ 的特征矩阵, $\operatorname{vec}(\mathbf{u})$ 表示矩阵 \mathbf{u} 的向量化, $\mathbf{U}|_{t=T} = [\operatorname{vec}(\mathbf{u}_1|_{t=T}), \dots, \operatorname{vec}(\mathbf{u}_M|_{t=T})] \in \mathbf{R}^{d_1 d_2 \times M}$, $\lambda > 0$ 是调节参数, $\{\operatorname{inv}_i(\mathbf{u}_m^n)\}_{i=0}^5$ 是直至二阶的基本微分不变量, $g(x)$ 为非线性压缩映射, N 为进化次数.

控制系数 a_i^n 可通过下面的迭代公式进行更新

$$(a_i^n)^{k+1} = (a_i^n)^k - \eta (\partial \mathbf{E} / \partial a_i^n)^k, i=0, 1, \dots, 5, n=0, 1, \dots, N-1, \quad (6)$$

其中 $\eta > 0$ 是迭代步长,

$$\partial \mathbf{E} / \partial a_i^n = \Delta t \cdot \operatorname{Tr}[(\partial \mathbf{E} / \partial \mathbf{U}^{n+1})^T \mathbf{B}_i^n], \quad (7)$$

且 $\mathbf{B}_i^n = [\operatorname{vec}(g(\operatorname{inv}_i(\mathbf{u}_1^n))), \dots, \operatorname{vec}(g(\operatorname{inv}_i(\mathbf{u}_M^n)))] \in \mathbf{R}^{d_1 d_2 \times M}$, $\mathbf{C}_i^n = [\operatorname{vec}(\mathbf{Z}(i, 1, n)), \dots, \operatorname{vec}(\mathbf{Z}(i, M, n))] \in \mathbf{R}^{d_1 d_2 \times M}$, $\partial \mathbf{E} / \partial \mathbf{U}^N = 2\mathbf{M}^{-1} \mathbf{W}^T (\mathbf{W} \mathbf{U}^N - \mathbf{H})$, $\partial \mathbf{E} / \partial \mathbf{U}^n = \partial \mathbf{E} / \partial \mathbf{U}^{n+1} + \Delta t \sum_{i=0}^5 a_i^n \mathbf{C}_i^n$,

$$\begin{aligned} \mathbf{Z}(i, m, n)(p, q) &= \frac{\partial \mathbf{E}}{\partial \mathbf{u}_m^{n+1}}(p+1, q) \cdot \frac{\partial g(\operatorname{inv}_i(\mathbf{u}_m^n)(p+1, q))}{\partial \mathbf{u}_m^n(p, q)} + \frac{\partial \mathbf{E}}{\partial \mathbf{u}_m^{n+1}}(p-1, q) \cdot \frac{\partial g(\operatorname{inv}_i(\mathbf{u}_m^n)(p-1, q))}{\partial \mathbf{u}_m^n(p, q)} \\ &+ \frac{\partial \mathbf{E}}{\partial \mathbf{u}_m^{n+1}}(p, q+1) \cdot \frac{\partial g(\operatorname{inv}_i(\mathbf{u}_m^n)(p, q+1))}{\partial \mathbf{u}_m^n(p, q)} + \frac{\partial \mathbf{E}}{\partial \mathbf{u}_m^{n+1}}(p, q-1) \cdot \frac{\partial g(\operatorname{inv}_i(\mathbf{u}_m^n)(p, q-1))}{\partial \mathbf{u}_m^n(p, q)} \\ &+ \frac{\partial \mathbf{E}}{\partial \mathbf{u}_m^{n+1}}(p, q) \cdot \frac{\partial g(\operatorname{inv}_i(\mathbf{u}_m^n)(p, q))}{\partial \mathbf{u}_m^n(p, q)}, \end{aligned}$$

$$\mathbf{Z}(i, m, n) = [\mathbf{Z}(i, m, n)(p, q)] \in \mathbf{R}^{d_1 \times d_2}, i=0, 1, \dots, 5, n=0, 1, \dots, N-1, m=0, 1, \dots, M.$$

下面给出具体算法.

算法 3 (基于 PDEs 的图像降噪方法)

- 步 1 初始化. 设 $k=0, \epsilon > 0, \rho = 0.95, N=5, \eta=1, \Delta t=0.5$. 在 $[-1, 1]$ 中均匀采样得矩阵 $\mathbf{A}^k \in \mathbf{R}^{N \times 6}$.
- 步 2 利用迭代公式 $\mathbf{u}_m^{n+1} = \mathbf{u}_m^n + \Delta t \sum_{i=0}^5 a_i^n g(\operatorname{inv}_i(\mathbf{u}_m^n))$ 计算 $(\mathbf{U}^n)^k = [\operatorname{vec}((\mathbf{u}_1^n)^k), \dots, \operatorname{vec}((\mathbf{u}_M^n)^k)] \in \mathbf{R}^{d_1 d_2 \times M}, n=0, 1, \dots, N-1$, 其中 $\mathbf{A} = \mathbf{A}^k$.
- 步 3 求解模型(5), 得 \mathbf{W}^k , 其中 $\mathbf{U}^N = (\mathbf{U}^N)^k$.
- 步 4 计算 $\mathbf{E} = \mathbf{M}^{-1} \|\mathbf{H} - \mathbf{W}^k (\mathbf{U}^N)^k\|_F^2 + \lambda \|\mathbf{W}^k\|_F^2$.
- 步 5 利用(7)式计算 $(\partial \mathbf{E} / \partial a_i^n)^k, i=0, 1, \dots, N-1$, 其中 $\mathbf{A} = \mathbf{A}^k, \mathbf{W} = \mathbf{W}^k$.
- 步 6 更新 $\eta \leftarrow \rho \eta$, 并利用(6)式更新 \mathbf{A}^k 至 \mathbf{A}^{k+1} .

步 7 利用 $\mathbf{u}_m^{k+1} = \mathbf{u}_m^k + \Delta t \sum_{i=0}^5 a_i^n g(\text{inv}_i(\mathbf{u}_m^k))$ 更新 $(\mathbf{U}^n)^k$ 至 $(\mathbf{U}^n)^{k+1}$, 其中 $\mathbf{A} = \mathbf{A}^{k+1}$.

步 8 求解模型(5), 得 \mathbf{W}^{k+1} , 其中 $\mathbf{U}^N = (\mathbf{U}^n)^{k+1}$.

步 9 计算 $\mathbf{E}^{k+1} = \mathbf{M}^{-1} \|\mathbf{H} - \mathbf{W}^{k+1}(\mathbf{U}^N)^{k+1}\|_F^2 + \lambda \|\mathbf{W}^{k+1}\|_F^2$.

步 10 若 $\|\mathbf{E}^{k+1} - \mathbf{E}^k\|_F < \varepsilon$, 停止迭代, 置 $\mathbf{A}^* \leftarrow \mathbf{A}^{k+1}, \mathbf{W}^* \leftarrow \mathbf{W}^{k+1}$; 否则, 置 $k \leftarrow k+1$, 转步 5.

在文献[15]中, 江等人研究发现通过 PDEs 的进化可以弱化模型参数的影响, 甚至不需要考虑调参. 为此, 本文在实验中选取了固定参数.

3 基于 PDEs 的 CCA

作为数据降维工具, CCA 考虑了多组特征间的线性相关性, 提高了特征表示的鉴别性与鲁棒性, 但其在特征抽取过程中并没有考虑光照、遮挡等噪声对图像质量的影响, 同时还出现了累计贡献率如何选取的问题. 为了提高图像的分类精度, 本节将 PDEs 融入 CCA 中, 提出了基于 PDEs 的 CCA 作为图像的降噪降维方法.

3.1 CCA 与累计贡献率

给出两个高维随机矢量 $\mathbf{x} \in \mathbf{R}^p$ 和 $\mathbf{y} \in \mathbf{R}^q$, 它们的取值分别为 $\{\mathbf{x}_i\}_{i=1}^m$ 和 $\{\mathbf{y}_i\}_{i=1}^m \in \mathbf{R}^q$. CCA 的目标是寻找典型方向 $\mathbf{w}_x \in \mathbf{R}^p$ 和 $\mathbf{w}_y \in \mathbf{R}^q$, 使得投影后的随机变量 $\mathbf{w}_x^T \mathbf{x}$ 和 $\mathbf{w}_y^T \mathbf{y}$ 间具有极大相关性, 即满足如下准则

$$\max_{\mathbf{w}_x, \mathbf{w}_y} \rho = \frac{\text{cov}(\mathbf{w}_x^T \mathbf{x}, \mathbf{w}_y^T \mathbf{y})}{\sqrt{\text{var}(\mathbf{w}_x^T \mathbf{x}) \text{var}(\mathbf{w}_y^T \mathbf{y})}} \quad (8)$$

记

$$\begin{cases} \mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_m] \in \mathbf{R}^{p \times m}, \mathbf{Y} = [\mathbf{y}_1, \dots, \mathbf{y}_m] \in \mathbf{R}^{q \times m}, \bar{\mathbf{x}} = m^{-1} \mathbf{X} \mathbf{e}_m \in \mathbf{R}^p, \bar{\mathbf{y}} = m^{-1} \mathbf{Y} \mathbf{e}_m \in \mathbf{R}^q, \\ \mathbf{C}_{xx} = \sum_{i=1}^m (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^T \in \mathbf{R}^{p \times p}, \mathbf{C}_{yy} = \sum_{i=1}^m (\mathbf{y}_i - \bar{\mathbf{y}})(\mathbf{y}_i - \bar{\mathbf{y}})^T \in \mathbf{R}^{q \times q}, \\ \mathbf{C}_{xy} = \sum_{i=1}^m (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{y}_i - \bar{\mathbf{y}})^T \in \mathbf{R}^{p \times q}, \end{cases}$$

则 $\text{cov}(\mathbf{w}_x^T \mathbf{x}, \mathbf{w}_y^T \mathbf{y}) = m^{-1} \mathbf{w}_x^T \mathbf{C}_{xy} \mathbf{w}_y$, $\text{var}(\mathbf{w}_x^T \mathbf{x}) = m^{-1} \mathbf{w}_x^T \mathbf{C}_{xx} \mathbf{w}_x$, $\text{var}(\mathbf{w}_y^T \mathbf{y}) = m^{-1} \mathbf{w}_y^T \mathbf{C}_{yy} \mathbf{w}_y$. 于是, 模型(8)可表示为

$$\max_{\mathbf{w}_x, \mathbf{w}_y} \rho = \frac{\mathbf{w}_x^T \mathbf{C}_{xy} \mathbf{w}_y}{\sqrt{\mathbf{w}_x^T \mathbf{C}_{xx} \mathbf{w}_x \cdot \mathbf{w}_y^T \mathbf{C}_{yy} \mathbf{w}_y}} \quad (9)$$

从 \mathbf{C}_{xx} 和 \mathbf{C}_{yy} 的定义可以看出, 模型 $\min \mathbf{w}_x^T \mathbf{C}_{xx} \mathbf{w}_x$ 和 $\min \mathbf{w}_y^T \mathbf{C}_{yy} \mathbf{w}_y$ 分别表示降维后的数据集 $\{\mathbf{w}_x^T \mathbf{x}_i\}_{i=1}^m$ 和 $\{\mathbf{w}_y^T \mathbf{y}_i\}_{i=1}^m$ 越紧致越好. 由于模型(9)与典型方向 $\mathbf{w}_x, \mathbf{w}_y$ 的范数 $\|\mathbf{w}_x\|$, $\|\mathbf{w}_y\|$ 无关, 所以可等价地表示为

$$\begin{aligned} \max_{\mathbf{w}_x, \mathbf{w}_y} & \mathbf{w}_x^T \mathbf{C}_{xy} \mathbf{w}_y \\ \text{s. t. } & \mathbf{w}_x^T \mathbf{C}_{xx} \mathbf{w}_x = 1, \mathbf{w}_y^T \mathbf{C}_{yy} \mathbf{w}_y = 1. \end{aligned} \quad (10)$$

考虑模型(10)的 Lagrange 函数, 并令其对 $\mathbf{w}_x, \mathbf{w}_y$ 的偏导数为 0, 可将模型(10)转化为广义特征方程

$$\begin{bmatrix} 0 & \mathbf{C}_{xy} \\ \mathbf{C}_{xy} & 0 \end{bmatrix} \begin{bmatrix} \mathbf{w}_x \\ \mathbf{w}_y \end{bmatrix} = \lambda \begin{bmatrix} \mathbf{C}_{xx} & 0 \\ 0 & \mathbf{C}_{yy} \end{bmatrix} \begin{bmatrix} \mathbf{w}_x \\ \mathbf{w}_y \end{bmatrix} \quad (11)$$

其中 $\mathbf{C}_{xy}^T = \mathbf{C}_{yx}$. 利用矩阵的特征值分解和奇异值分解可求解方程(11), 下面给出具体算法.

算法 4 (CCA)

步 1 给定高维数据集 $\{\mathbf{x}_i\}_{i=1}^m \in \mathbf{R}^p$ 和 $\{\mathbf{y}_i\}_{i=1}^m \in \mathbf{R}^q$, 计算矩阵 $\mathbf{C}_{xx} \in \mathbf{R}^{p \times p}$, $\mathbf{C}_{yy} \in \mathbf{R}^{q \times q}$ 和 $\mathbf{C}_{xy} \in \mathbf{R}^{p \times q}$.

步 2 为了避免矩阵的奇异性, 置 $\mathbf{C}_{xx} \leftarrow \mathbf{C}_{xx} + t\mathbf{I}_p$, $\mathbf{C}_{yy} \leftarrow \mathbf{C}_{yy} + t\mathbf{I}_q$, 其中 $t > 0$ 是正则化参数.

步 3 对 \mathbf{C}_{xx} 做特征值分解: $\mathbf{C}_{xx} = \mathbf{U} \Sigma_x \mathbf{U}^T$, 其中 $\mathbf{U} \in \mathbf{R}^{p \times p}$ 是正交矩阵, $\sigma_1 \geq \dots \geq \sigma_p > 0$ 是 \mathbf{C}_{xx} 的全部非零特征值且 $\Sigma_x = \text{diag}(\sigma_1, 1, \dots, \sigma_p)$.

步 4 令 $\mathbf{B} = \Sigma_x^{-1/2} \mathbf{U}^T \mathbf{C}_{xy} \mathbf{C}_{yy}^{-1/2} \in \mathbf{R}^{p \times q}$ 且 $\text{rank}(\mathbf{B}) = r_B \leq \min\{p, q\}$.

步 5 对矩阵 \mathbf{B} 做奇异值分解: $\mathbf{B} = [\mathbf{P}_1 \Sigma_B, 0] \mathbf{Q}^T$, 其中 $\mathbf{P}_1 = [\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_{r_B}] \in \mathbf{R}^{p \times r_B}$ 是列正交阵, $\Sigma_B = \text{diag}(\lambda_1, \dots, \lambda_{r_B})$, $\lambda_1 \geq \dots \geq \lambda_{r_B} > 0$ 是 \mathbf{B} 的全部非零奇异值.

步 6 若累积贡献率满足 $(\lambda_1^2 + \dots + \lambda_d^2) / (\lambda_1^2 + \dots + \lambda_{r_B}^2) \geq \lambda$, 则取 $d \leq r_B$ 对典型方向

$$\mathbf{w}_{xi} = \mathbf{U} \Sigma_x^{-1/2} \mathbf{p}_i, \mathbf{w}_{yi} = \lambda_i^{-1} \mathbf{C}_{yy}^{-1} \mathbf{C}_{yx} \mathbf{U} \Sigma_x^{-1/2} \mathbf{p}_i, i = 1, 2, \dots, d.$$

步 7 利用典型方向矩阵 $\mathbf{W}_x = [\mathbf{w}_{x1}, \dots, \mathbf{w}_{xd}] \in \mathbf{R}^{p \times d}$, $\mathbf{W}_y = [\mathbf{w}_{y1}, \mathbf{w}_{y2}, \dots, \mathbf{w}_{yd}] \in \mathbf{R}^{q \times d}$ 进行降维, 得降维数据集

$\{W_x^T x_i\}_{i=1}^m \in \mathbf{R}^d$ 和 $\{W_y^T y_i\}_{i=1}^m \in \mathbf{R}^d$.

从算法 4 中可以看出, 累积贡献率取的不同, 低维子空间的维度 d 也会不同. 习惯上将累积贡献率选为 100%, 这样一来很多时候起不到降维的作用. 若累积贡献率取的过低, 将会导致数据判别信息的丢失, 影响数据的分类精度. 目前还没有一个行之有效的方法选择累积贡献率.

3.2 基于 PDEs 的 CCA

为了弥补 CCA 没有考虑光照、遮挡等噪声对图像质量影响的不足, 本节提出一个基于 PDEs 的图像降噪降维方法——基于 PDEs 的 CCA (CCA\ PDEs). 对给定的 $c(c \geq 2)$ 类高维噪声(光暗、有遮挡)图像集 $\{(I_m, h_m)\}_{m=1}^M \in \mathbf{R}^{d_1 \times d_2} \times \mathbf{R}^c$, CCA\ PDEs 的基本思路如下

算法 5 (CCA\ PDEs)

步 1 利用 PDEs 的进化对原始图像 $I_m \in \mathbf{R}^{d_1 \times d_2}$ 进行降噪处理, 得特征矩阵 $u_m \in \mathbf{R}^{d_1 \times d_2}$.

步 2 将特征矩阵分为左右两部分, 得矩阵集 $\{(u_{xm}, h_m)\}_{m=1}^M \in \mathbf{R}^{d_1 \times d_3} \times \mathbf{R}^c$ 和 $\{(u_{ym}, h_m)\}_{m=1}^M \in \mathbf{R}^{d_1 \times (d_2 - d_3)} \times \mathbf{R}^c$, 其中 $d_3 < d_2$.

步 3 对向量集 $\{(\text{vec}(u_{xm}), h_m)\}_{m=1}^M \in \mathbf{R}^{d_1 d_3} \times \mathbf{R}^c$ 和 $\{(\text{vec}(u_{ym}), h_m)\}_{m=1}^M \in \mathbf{R}^{d_1 (d_2 - d_3)} \times \mathbf{R}^c$, 利用 CCA 进行降维处理, 得降维向量集 $\{(x_m, h_m)\}_{m=1}^M \in \mathbf{R}^d \times \mathbf{R}^c$ 和 $\{(y_m, h_m)\}_{m=1}^M \in \mathbf{R}^d \times \mathbf{R}^c$.

步 4 融合两个降维向量集, 得 $\left\{ \begin{bmatrix} x_m \\ y_m \end{bmatrix}, h_m \right\}_{m=1}^M \in \mathbf{R}^{2d} \times \mathbf{R}^c$, 选择分类器对其进行分类.

步 5 通过调整 PDEs 的进化次数, 得最佳分类精度.

4 PDEs 的进化对累积贡献率的影响

本节针对 $c(c \geq 2)$ 类高维噪声(光暗、有遮挡)图像集, 以分类精度为标准, 通过实验研究 PDEs 的进化对累计贡献率的影响. 对二类图像集, 选用 SVM 作为分类器; 对多类图像集(以三类图像集为例), 选用 OVR-MSVM 作为分类器. 分类器均采用五折交叉验证法. 借助文献[14, 15]的研究成果, 分类器的模型参数取为 0.5 或 0.01 且 PDEs 的进化不超过 7 次. 所有试验都在 Extended Yale B 人脸数据集和 Palm-print 掌纹数据集上进行, 且考虑累积贡献率为 85%, 90%, 95%, 100% 四种情况.

4.1 数据集描述

Palm-print 掌纹数据集包含 400 个人的 8000 个掌纹, 每人 20 个, 前 10 个和后 10 个分别在不同时间获取, 每个掌纹尺寸为 32×32 . 随机选取 12 人的掌纹(见图 1)组成 6 个二类图像集以及随机选取 18 人的掌纹(见图 2)组成 6 个三类图像集进行试验

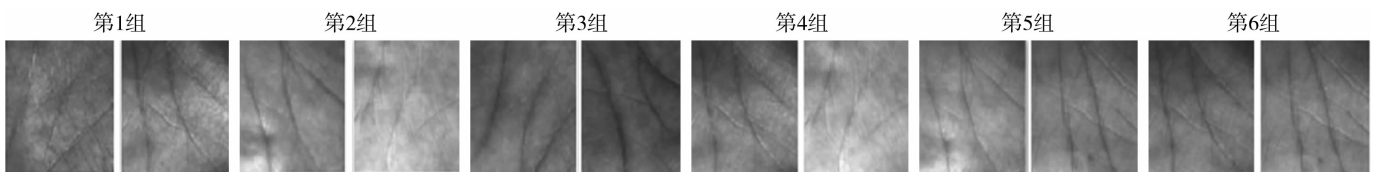


图 1 Palm-print 数据集中的 12 人掌纹图像

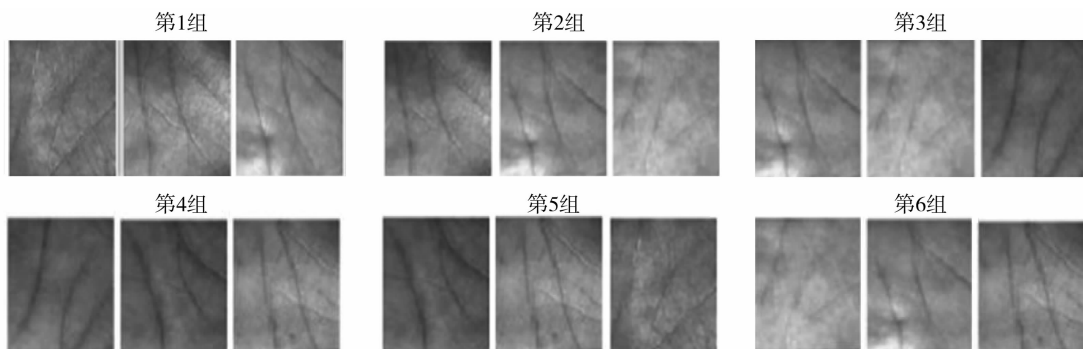


图 2 Palm-print 数据集中的 18 人掌纹图像

Extended Yale B 人脸数据集包含 38 个人分别在 $12^\circ, 25^\circ, 50^\circ, 77^\circ, 90^\circ$ 的不同光照角度下的 2414 张面部图像, 每张尺寸为 32×28 . 随机选取 8 人(见图 3), 每人受光照影响较大的 30 张面部图像组成四个二类图

像集(从左至右编号分别为 1 至 8)以及随机选取 12 人(见图 4),每人受光照影响较大的 30 张面部图像组成四个三类图像集进行试验.



图 3 Extended Yale B 数据集中的 8 人面部图像



图 4 Extended Yale B 数据集中的 12 人面部图像

4.2 PDEs 对累积贡献率的影响

本节只考虑 PDEs 的进化对累积贡献率的影响,不考虑压缩函数的影响,故取 $g(x)=x$. 用 λ 表示累积贡献率, n 表示 PDEs 的进化次数. $n=0$ 表示只是利用 CCA 进行降维处理, $n=1,2,\dots,7$ 表示利用 CCA \ PDEs 同时进行降噪降维处理,只是 PDEs 的进化次数不同.表 1 和表 2 给出了对二类图像集的影响,表 3 和表 4 给出了对三类图像集的影响.

表 1 Palm-print 数据集下 PDEs 对累积贡献率的影响 (二类, $C=0.5$)

	n	0	1	2	3	4	5	6	7
第 1 组	85%	0.5500	0.5250	1.0000	0.5000	1.0000	1.0000	1.0000	0.9750
	90%	0.5750	0.5250	1.0000	0.5000	1.0000	1.0000	1.0000	1.0000
	95%	0.5750	0.5250	1.0000	0.5000	1.0000	1.0000	1.0000	1.0000
	100%	0.8500	0.6750	1.0000	0.6500	1.0000	1.0000	1.0000	1.0000
第 2 组	85%	0.5000	0.5000	1.0000	1.0000	0.5250	0.9500	0.9250	0.9250
	90%	0.5500	0.5250	1.0000	1.0000	0.5250	1.0000	1.0000	0.9500
	95%	0.5500	0.5250	1.0000	1.0000	0.5250	1.0000	0.9750	0.9500
	100%	0.8500	0.6000	1.0000	1.0000	0.6250	1.0000	1.0000	1.0000
第 3 组	85%	0.7000	0.5000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9750
	90%	0.6250	1.0000	1.0000	0.9750	0.9500	0.9500	0.9250	0.8250
	95%	0.7000	0.5000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9750
	100%	0.7750	0.9000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
第 4 组	85%	0.6000	0.5250	0.9500	1.0000	0.5750	0.9750	0.9500	0.9250
	90%	0.6000	0.5750	0.9750	1.0000	0.8000	1.0000	0.9750	0.9500
	95%	0.5500	0.5750	0.8250	1.0000	0.9000	1.0000	0.9750	0.9500
	100%	0.7250	0.6250	0.8500	1.0000	0.7250	1.0000	1.0000	1.0000
第 5 组	85%	0.6250	1.0000	1.0000	1.0000	0.9500	0.9500	0.9250	0.8500
	90%	0.6250	1.0000	1.0000	0.9750	0.9500	0.9500	0.9250	0.8250
	95%	0.6250	1.0000	0.9750	0.9750	0.9500	0.9750	0.9750	0.9500
	100%	0.8000	1.0000	1.0000	1.0000	0.9750	0.9750	0.9750	0.9500
第 6 组	85%	0.5250	0.5250	0.5000	0.5750	1.0000	0.9000	0.9250	0.9250
	90%	0.5250	0.5250	0.5000	0.5750	1.0000	0.9500	0.9500	0.9250
	95%	0.5250	0.5000	0.5000	0.5750	0.9750	0.8250	0.9000	0.9000
	100%	0.7750	0.5000	0.5750	0.5750	0.9750	0.7250	0.9750	0.9500

表 2 Extended Yale B 数据集下 PDEs 对累积贡献率的影响(二类图像, $C=0.5$)

	n	0	1	2	3	4	5	6	7
第 1 组	85%	0.8000	0.8833	0.8667	0.9167	0.9000	0.8333	0.8333	0.8167
	90%	0.7833	0.9000	0.8667	0.9167	0.8833	0.8667	0.8333	0.7667
	95%	0.9000	0.9167	0.9167	0.9500	0.9333	0.9000	0.8167	0.7667
	100%	0.8833	0.8833	0.8833	0.9167	0.9333	0.9333	0.9000	0.9167
第 2 组	85%	0.8167	0.9000	0.7833	0.7500	0.8500	0.8833	0.8167	0.8167
	90%	0.8167	0.8000	0.8333	0.9333	0.8167	0.8833	0.8500	0.9333
	95%	0.8667	0.9167	0.8500	0.8833	0.8667	0.8667	0.8000	0.8000
	100%	0.8167	0.8333	0.7833	0.8000	0.8167	0.8167	0.8167	0.8667
第 3 组	85%	0.7667	0.7500	0.7667	0.7333	0.7500	0.7500	0.7000	0.6833
	90%	0.7833	0.8000	0.7667	0.8500	0.7500	0.7000	0.7333	0.6333
	95%	0.7833	0.7833	0.8667	0.6833	0.7500	0.7500	0.7167	0.6667
	100%	0.8167	0.9167	0.8500	0.8000	0.7500	0.8500	0.8667	0.8333
第 4 组	85%	0.8667	0.9000	0.8667	0.9000	0.9000	0.8833	0.9167	0.8833
	90%	0.9167	0.9000	0.9000	0.9000	0.9333	0.9333	0.9333	0.9000
	95%	0.9167	0.9333	0.9500	0.9333	0.9000	0.9167	0.9000	0.8833
	100%	0.9167	0.9333	0.9500	0.9500	0.9167	0.9167	0.9167	0.9000

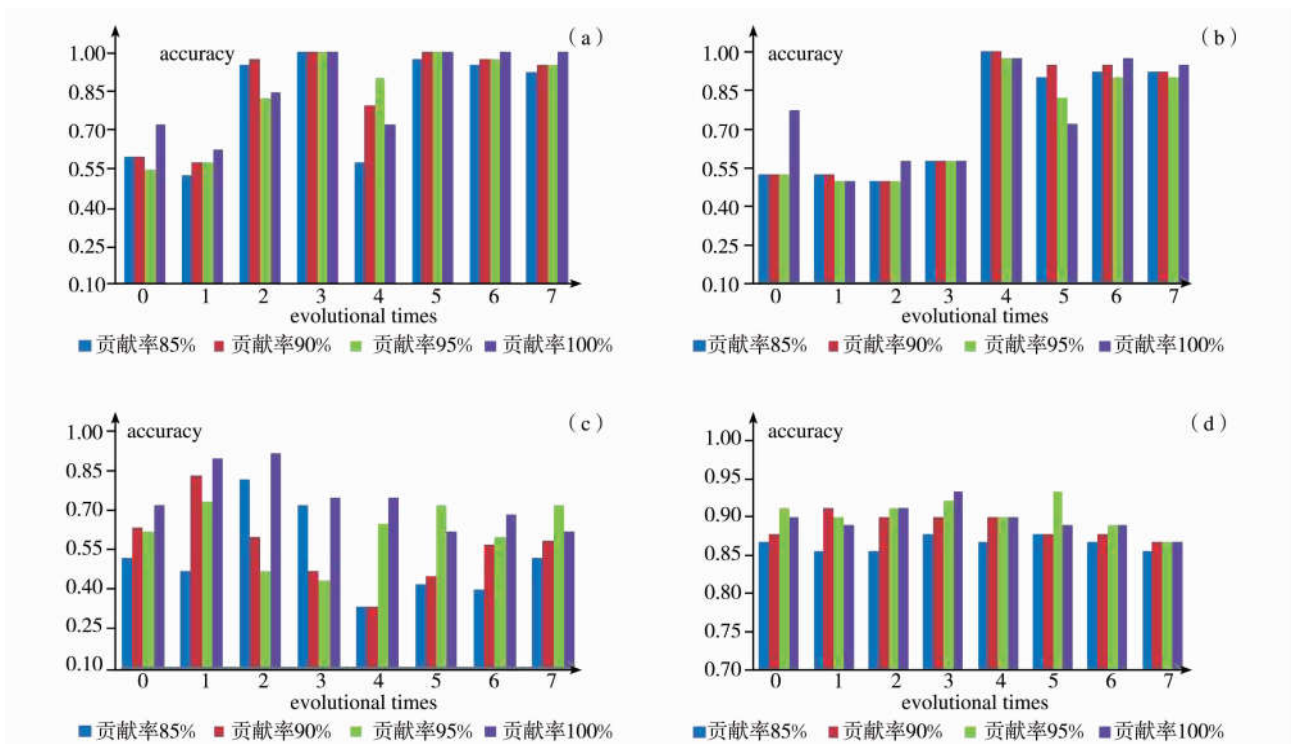
表 3 在 Palm-print 数据集下 PDEs 对累积贡献率的影响(三类, $C_i=0.01$)

	n	0	1	2	3	4	5	6	7
第 1 组	85%	0.3333	0.5500	0.6667	0.8500	0.3500	0.4500	0.3667	0.4667
	90%	0.3333	0.8500	0.4833	0.3333	0.3333	0.3833	0.5167	0.5333
	95%	0.3333	0.5167	0.4667	0.4000	0.4000	0.7167	0.5833	0.5000
	100%	0.4000	0.7333	0.8000	0.6667	0.5833	0.6500	0.7500	0.6333
第 2 组	85%	0.4667	0.5333	0.7667	0.5833	0.3333	0.5333	0.5167	0.4667
	90%	0.4667	0.5667	0.6000	0.4500	0.3333	0.6500	0.6167	0.5333
	95%	0.6000	0.5833	0.3333	0.4000	0.6500	0.3500	0.5833	0.5000
	100%	0.5333	0.8000	0.6167	0.6667	0.6667	0.7167	0.7333	0.7000
第 3 组	85%	0.6667	0.4667	0.6667	0.8833	0.3333	0.3833	0.4500	0.3667
	90%	0.6667	0.5333	0.7500	0.4000	0.3333	0.5000	0.7167	0.5000
	95%	0.7500	0.8000	0.4000	0.3333	0.7667	0.6500	0.4500	0.5667
	100%	0.8000	0.8167	0.7333	0.8667	0.9000	0.9333	0.8667	0.8333
第 4 组	85%	0.5333	0.6000	0.5333	0.9833	0.3333	0.6167	0.5167	0.5167
	90%	0.4667	0.6667	0.8667	0.7333	0.4000	0.6833	0.6000	0.6333
	95%	0.8667	0.9333	0.4667	0.5833	0.7000	0.7667	0.6000	0.5333
	100%	0.8500	0.7833	0.8167	0.8667	0.8667	1.0000	0.9500	0.8833
第 5 组	85%	0.5167	0.4667	0.8167	0.7167	0.3333	0.4167	0.4000	0.5167
	90%	0.6333	0.8333	0.6000	0.4667	0.3333	0.4500	0.5667	0.5833
	95%	0.6167	0.7333	0.4667	0.4333	0.6500	0.7167	0.6000	0.7167
	100%	0.7167	0.9000	0.9167	0.7500	0.7500	0.6167	0.6833	0.6167
第 6 组	85%	0.5333	0.6667	0.6000	0.7667	0.9000	0.8167	0.4000	0.3333
	90%	0.8000	0.7333	0.8667	0.9000	0.8667	0.4500	0.3333	0.4000
	95%	0.8500	0.9500	0.4333	0.4667	0.4000	0.5500	0.4667	0.6667
	100%	0.9000	0.7333	1.0000	0.7167	0.8667	0.7333	0.7333	0.7333

表 4 在 Extended Yale B 数据集下 PDEs 对累积贡献率的影响 (三类, $C_i = 0.5$)

	n	0	1	2	3	4	5	6	7
第 1 组	85%	0.9111	0.9000	0.9444	0.9556	0.9222	0.9778	0.9667	0.9000
	90%	0.9556	0.9444	0.9222	0.9778	0.9000	0.9000	0.9667	0.9444
	95%	0.9222	0.9667	0.9556	0.9556	0.9556	0.9111	0.9222	0.9111
	100%	0.7778	0.7889	0.7556	0.7556	0.8000	0.8000	0.7778	0.7778
第 2 组	85%	0.9000	0.9444	0.9556	0.8556	0.9000	0.8889	0.8556	0.8222
	90%	0.8556	0.9556	0.8889	0.8667	0.8667	0.9222	0.9222	0.8556
	95%	0.9444	0.9333	0.9000	0.9667	0.9667	0.9222	0.9667	0.9444
	100%	0.9222	0.9444	0.9444	0.8889	0.9333	0.8778	0.9222	0.9222
第 3 组	85%	0.8778	0.8889	0.8667	0.8778	0.8111	0.8778	0.8667	0.8667
	90%	0.9000	0.8667	0.9111	0.8778	0.8222	0.8778	0.8778	0.8889
	95%	0.8889	0.8778	0.9111	0.9000	0.8444	0.8667	0.8667	0.9111
	100%	0.7667	0.7667	0.7444	0.7111	0.8111	0.8667	0.7778	0.7778
第 4 组	85%	0.8667	0.8556	0.8556	0.8778	0.8667	0.8778	0.8667	0.8556
	90%	0.8778	0.9111	0.9000	0.9000	0.9000	0.8778	0.8778	0.8667
	95%	0.9111	0.9000	0.9111	0.9222	0.9000	0.9333	0.8889	0.8667
	100%	0.9000	0.8889	0.9111	0.9333	0.9000	0.8889	0.8889	0.8667

为直观起见,以每组实验中的最后一组数据为例,给出了影响柱形图(见图 5)



注:(a) Palm-print 第 6 组;(b) Extended Yale B 第 4 组 (二类);(c) Palm-print 第 6 组;(d) Extended Yale B 第 4 组 (三类).

图 5 PDEs 对累积贡献率的影响

从表 1 和表 2 中可以看出:针对二类图像集,不论累积贡献率取为多少,最多经过 4 次 PDEs 进化就可达到同等条件下的最佳分类精度.对 Palm-print 数据集,分类精度至少提高了 15%,尤其对 85%的累积贡献率,至少提高了 30%.对 Extended Yale B 数据集,除极个别情况外,分类精度至少提高了 3.3%.从表 3 和表 4 中可以看出:针对三类图像集,不论累积贡献率取为多少,最多经过 5 次 PDEs 进化就可达到同等条件下的最佳分类精度.对 Palm-print 数据集,分类精度至少提高了 5%,对 85%的累积贡献率,影响更为明显,分

类精度至少提高了 21.7%。对 Extended Yale B 数据集,分类精度最少提高了 1.1%,对 100% 的累积贡献率,精度至少提高了 2.2%。

综上所述,针对高维噪声图像(光暗、有遮挡)的分类问题,本文所提的 CCA\PDEs 是一个有效和高效的降噪降维方法,且通过 PDEs 的进化可以弱化累计贡献率的影响,甚至不用调优累积贡献率,只需进行不超过 5 次的 PDEs 进化即可达到同等条件下的最佳分类精度。

5 结论

众所周知,对高维噪声图像(光暗、有遮挡)的分类问题来说,分类精度的高低严重依赖于图像降噪方法和降维方法的选择。目前,大部分的降噪方法不具备降维作用,同样地,大部分的降维方法也不具备降噪作用。本文提出的 CCA\PDEs 算法可以同时降噪降维,但出现的问题是如何选择 CCA 中的累计贡献率。累计贡献率的大小不仅决定着降维子空间的维度,而且还影响着图像的分类精度,一般常选为 100%,但这样一来很多时候起不到降维的作用。为此,本文研究了 PDEs 的进化对累计贡献率的影响。通过在 Extended Yale B 人脸数据集和 Palm-print 掌纹数据集上的实验表明,PDEs 的进化可以弱化累计贡献率的影响,甚至不用调优累积贡献率,通过不超过 5 次的 PDEs 进化即可达到同等条件下的最佳分类精度。下一步我们将研究 CCA\PDEs 算法对分类器的影响,尤其是对分类器模型参数的影响。

参 考 文 献

- [1] Affonso C, Rossi A L, Vieira F H, et al. Deep learning for biological image classification[J]. *Expert Systems with Applications*, 2017, 85: 114-122.
- [2] Zhao Zhongqiu, Xu Shoutao, Liu Dian, et al. A review of image set classification[J]. *Neurocomputing*, 2019, 335: 251-260.
- [3] Gyanfi K S, Brusey J, Hunt A, et al. A dynamic linear model for heteroscedastic LDA under class imbalance[J]. *Neurocomputing*, 2019, 343: 65-75.
- [4] Dend Ping, Wang Hongjun, Li Tianrui, et al. Linear discriminant analysis guided by unsupervised ensemble learning[J]. *Information Sciences*, 2019, 480: 211-221.
- [5] 杨静, 范丽亚. 基于模糊隶属度的加权广义不定核判别分析[J]. *山东大学学报(工学版)*, 2012, 42(3): 31-38.
- [6] 崔燕, 范丽亚. 高维数据正定核与不定核的 KPCA 变换阵比较[J]. *山东大学学报(工学版)*, 2011, 41(1): 17-23.
- [7] Liu C J, Wechsler H. A shape and texture based enhanced fisher classifier for face recognition[J]. *IEEE Transactions on Image Processing*, 2001, 10(4): 598-608.
- [8] Yang J, Yang J Y, Zhang D, et al. Feature fusion: parallel strategy vs. serial strategy[J]. *Pattern Recognition*, 2003, 36(6): 1369-1381.
- [9] Gao Xizhan, Sun Quansen, Xu Haitao. Multiple-rank supervised canonical correlation analysis for feature extraction, fusion and recognition [J]. *Expert Systems with Applications*, 2017, 84: 171-185.
- [10] 杨静, 高希占. 边界监督多重集典型相关分析[J]. *聊城大学学报(自然科学版)*, 2019, 32(3): 13-22.
- [11] Liu R S, Lin Z C, Zhang W, et al. Learning PDEs for image restoration via optimal control[C]. //The 11th European Conference on Computer Vision, 2010.
- [12] Liu R, Lin Z, Zhang W, et al. Toward designing intelligent PDEs for computer vision: An optimal control approach[J]. *Image and Vision Computing*, 2013, 31(1): 43-56.
- [13] Fang C, Zhao Z, Zhou P, et al. Feature learning via partial differential equation with applications to face recognition[J]. *Pattern Recognition*, 2017, 69: 14-25.
- [14] 江珊珊, 杨静, 范丽亚. 基于 PDEs 的图像特征提取方法[J]. *山东大学学报(工学版)*, 2018(4): 27-36.
- [15] 江珊珊, 范丽亚. PDEs 对 RSVM 中模型参数的影响[J]. *聊城大学学报(自然科学版)*, 2019, 32(3): 36-45.
- [16] Gaudio M, Gorgone E. Lagrangian relaxation for SVM feature selection[J]. *Computers & Operations Research*, 2017, 87: 137-145.
- [17] Yuan Y, Zhang M, Luo P, et al. SVM-based Detection in Visible Light Communications[J]. *Optik*, 2017, 151: 55-64.
- [18] Liu Y, Zheng Y F. One-against-all multi-class SVM classification using reliability measures[C]. //IEEE International Joint Conference on Neural Networks, 2005.
- [19] Lee H, Song J, Park D. Intrusion Detection System Based on Multi-class SVM[J]. *Rough Sets Fuzzy Sets Data Mining & Granular Computing*, 2005, 3642(3): 511-519.

Effect of PDEs on the Cumulative Contribution Rate of CCA

LEI Tian-tian WANG Zhen FAN Li-ya YANG Jing

(School of Mathematical Sciences, Liaocheng University, Liaocheng 252059, China)

Abstract The classification qualities of multi-class high-dimensional noisy (dark and shaded) images heavily depend on the choices of denoising methods and dimension reduction methods for images. In this paper, a canonical correlation analysis (CCA) based on partial differential equations (PDEs) is proposed as a joint denoising and dimension reduction algorithm, and the influence of PDEs evolution on the cumulative contribution rate of CCA is studied. Taking support vector machine (SVM) and one-versus-rest SVM (OVR-MSVM) as classifiers and based on classification accuracy, a lot of experiments are carried out on Extended Yale B face data set and Palm-print data set. The experimental results indicate that the influence of cumulative contribution rates can be weakened by the evolutions of PDEs, even without selecting cumulative contribution rates. The optimal classification accuracy under the same conditions can be achieved by no more than five evolutions of PDEs.

Key words image classification; partial differential equations; canonical correlation analysis; cumulative contribution rate; evolutionary times