

# 结合显性与隐性空间光滑的高效二维图像判别特征抽取

朱快快 田 青 陈松灿  
(南京航空航天大学计算机科学与技术学院 南京 210016)  
(zkk1992@nuaa.edu.cn)

## A Fast Discriminant Feature Extraction Framework Combining Implicit Spatial Smoothness with Explicit One for Two-Dimensional Image

Zhu Kuaikuai, Tian Qing, and Chen Songcan  
(College of Computer Science & Technology, Nanjing University of Aeronautics & Astronautics, Nanjing 210016)

**Abstract** Images have two-dimensional inherent spatial structures, and the pixels spatially close to each other have similar gray values, which means images are locally spatially smooth. To extract features, traditional methods usually convert an original image into a vector, resulting in the destruction of spatial structure. Thus 2D image-based feature extraction methods emerge, typically, such as 2DLDA and 2DPCA, which reduce time complexity significantly. However, 2D-based methods manipulate on the whole row (or column) of an image, leading to spatially under-smoothing. To overcome such shortcomings, spatial regularization is proposed by explicitly imposing a Laplacian penalty to constrain the projection coefficients to be spatially smooth and has achieved better performance than 2D-based methods, but sharing the genetic high computing cost with 1D methods. Implicit spatial regularization (ISR) constrains spatial smoothness within each local image region by dividing and reshaping image and then executing 2D-based feature extraction methods, resulting in a performance improvement of the typical bi-side 2DLDA over SSSL (a typical ESR method). However, ISR obtains the spatial smooth implicitly but has lack of explicit spatial constraints such that the feature space obtained by ISR is still not smooth enough. The optimization criteria of bi-side 2DLDA are not jointly convex simultaneously, resulting in high computing cost and globally optimal solution cannot be guaranteed. Inspired by statements above, we introduce a novel linear discriminant model called fast discriminant feature extraction framework combining implicit spatial smoothness with explicit one for two-dimensional image recognition (2D-CISSE). The key step of 2D-CISSE is to preprocess spatial smooth for images, then ISR is executed. 2D-CISSE not only retains spatial smooth explicitly, but also reinforces the explicit spatial constraints. Not only can it achieve globally optimal solution, but it also have generality, i. e. any out-of-shelf image smoothing methods and 2D-based feature extraction methods can be embedded into our framework. Finally, experimental results on four face datasets (Yale, ORL, CMU PIE and AR) and handwritten digit datasets (MNIST and USPS) demonstrate the effectiveness and superiority of our 2D-CISSE.

收稿日期:2016-03-11;修回日期:2016-05-05  
基金项目:国家自然科学基金项目(61472186);高等学校博士学科点专项科研基金项目(20133218110032)  
This work was supported by the National Natural Science Foundation of China (61472186) and the Specialized Research Fund for the Doctoral Program of Higher Education (20133218110032).  
通信作者:陈松灿(s.chen@nuaa.edu.cn)

**Key words** spatial smooth; image Euclidean distance; implicit spatial regularization (ISR); feature extraction; matrix-based pattern

**摘要** 图像具有固有的二维空间结构,空间上邻近的像素点通常具有相近的灰度值,意味着图像具有局部光滑性.为其特征抽取,传统方法常将原始图像拉成向量,造成空间结构的破坏,由此直接基于图像的 2D 特征抽取法应运而生.典型的如 2DLDA,2DPCA,相比向量方法,计算复杂度显著降低,但其操作针对的是图像整行(或整列),导致空间光滑度过粗.为此,空间正则化通过在向量化空间中显式地施加局部空间光滑弥补这一不足,由此获得了比 2D 抽取法更优的分类性能,但其遗传了向量法的高计算代价.最近,隐性空间正则化方法(implicit spatial regularization, ISR)提出利用图像划分与重组隐性地体现图像局部光滑性,而后再利用现有 2D 方法抽取特征,使典型双边 2DLDA 性能优于 SSSL(一种典型的显性空间正则化方法),但是,仅隐性地光滑缺乏显式的强制约束力,其特征空间依然欠光滑,同时双边 2DLDA 由非凸问题获得,计算耗时却不能保证解的全局最优性.鉴于此,提出一种结合显性与隐性空间光滑的高效二维图像判别特征抽取框架(2D-CISSE).其关键步骤是预先对图像显性地全局光滑,紧接着进行 ISR,既继承了 ISR 的隐性光滑又强化了图像局部光滑的显式约束力,不仅可直接获得全局最优投影,同时该框架具有一般性,即现有大部分图像光滑方法与 2D 特征抽取法均可嵌入其中.最后,通过在人脸数据集 Yale, ORL, CMU PIE, AR 以及手写数字数据集 MNIST 和 USPS 上的对比实验验证了 2D-CISSE 框架性能的优越性与计算的高效性.

**关键词** 空间光滑;图像欧氏距离;隐性空间正则化;特征抽取;基于矩阵模式

**中图法分类号** TP391.41; TP751

尽管深度学习已在图像分类<sup>[1-2]</sup>、自然语言理解<sup>[3-4]</sup>等领域取得了几乎无可比拟的进展,其成功的关键在于利用了大量的训练样本,但现实中依然存在仅有少量训练样本的分类问题,而提升此种情况下分类性能的关键之一是如何充分利用数据内在的结构信息.针对图像分类问题,则是如何充分利用图像的二维空间结构信息.传统的特征抽取方法<sup>[5-7]</sup>通常先按一定顺序将  $r \times c$  的图像拉成  $rc$  维向量后用向量方法处理,导致空间结构信息缺失,影响后续学习任务的性能.因此,如何充分利用空间结构信息成为了研究关注点.

针对结构信息利用策略, Tian 等人<sup>[8]</sup>将其归纳为 3 类,即结构嵌入的欧氏度量法、面向估计器学习的结构正则化方法和直接图像操作方法.传统的欧氏距离在用于衡量 2 幅图像的相似度时,包含在图像中的空间信息未能充分体现,对图像分类产生不利影响.鉴于此, Wang 等人<sup>[9]</sup>将空间结构信息嵌入欧氏距离提出一种图像欧氏距离法 (IMage Euclidean distance, IMED),在 手写数字识别与人脸识别的实验中获得了比传统欧氏距离更优的识别效果.而 Li 等人<sup>[10]</sup>拓展了 IMED,将其用于多角度的性别分类并获得了更高的分类精度,随后文献<sup>[11]</sup>又提出了一种自适应的图像欧氏距离度量法.这些结构嵌入的欧氏度量法本质在于通过对图像像素点

施加权重显性地对图像作全局空间光滑,有效补偿了空间信息的损失,但其基于向量模式,不可避免要面对维度过高带来的高计算复杂度和小样本问题.

为尽可能保持像素点间的空间结构信息,一个自然的做法是直接对图像(或重组图像)矩阵操作.因而很多 2D 特征抽取法相继被提出,代表性的有 Chen 等人<sup>[12-16]</sup>通过对图像双线性投影所发展出的一系列分类器,例如 MatMHKS<sup>[12]</sup> 和 MatFE + MatCD<sup>[16]</sup>,并在人脸识别与手写数字识别的任务中获得比基于向量模式更优的识别性能.迄今为止,大部分基于向量子空间的方法已被成功地拓展到 2D 或张量的版本,如主成分分析 (PCA) 的 2D 版本 2DPCA<sup>[17]</sup> 和广义低秩近似矩阵 (GLRAM)<sup>[18]</sup> 等,线性判别分析 (LDA) 的 2D 版本单边 2DLDA<sup>[19]</sup> 和双边 2DLDA<sup>[20]</sup> 等,支持向量机的 2D 版本,即支持矩阵机<sup>[21]</sup> (support matrix machines, SMM). 与传统向量法相比,2D 特征抽取法具有 3 点优势:

- 1) 直接基于图像矩阵定义的离散矩阵维度很低,这能有效缓和小样本问题,也避免了原有类内散度矩阵的奇异问题;

- 2) 能显著降低特征抽取的时间和空间计算复杂度;

- 3) 可部分利用像素点间的空间结构信息.

众多实验结果表明 2D 特征抽取法比其对应的

1D 向量法更优,但文献[22]中指出 2D 方法仅部分利用了图像在同一行或列的空间信息,并没有充分利用整个图像中的空间信息,其空间光滑度仍较粗.

鉴于此,有学者尝试将空间光滑正则化用于惩罚相关目标函数,从而使优化结果尽可能空间光滑<sup>[22-24]</sup>.代表性方法为 Cai 等人<sup>[22]</sup>提出的空间光滑子空间学习模型(spatial smooth subspace learning, SSSL).在 SSSL 中,通过引入拉普拉斯惩罚使投影系数在空间上保持光滑. Zuo 等人<sup>[23]</sup>进一步对拉普拉斯惩罚函数做高斯加权来实现多尺度图像光滑. Lotte 等人<sup>[24]</sup>利用拉普拉斯惩罚对欠光滑滤波器光滑约束提出空间正则化的共同空间模式,在脑电波滤波分类实验中获得更优分类性能.通过显性地光滑图像空间的投影向量,空间正则化使向量特征抽取法获得了比 2D 特征抽取法更优的分类性能.但是,这种空间正则化策略依然存在 2 点不足:1)它遗传了原向量法固有的高计算代价;2)选取正则化参数的交叉验证耗时且参数选取依然是一个开放性问题.

在图像分类领域,随着空间信息利用策略的不断深入研究,很多综合策略相继被提出.如 2013 年, Gao 等人<sup>[25]</sup>将结构嵌入的欧氏度量法结合到图像直接操作法中,通过将传统欧氏距离替换成 IMED 来实现像素点间空间关系的融合利用,提出了一种基于图像欧氏距离的 2 维最大局部变化(two-dimensional maximum local variation, 2DMLV)人脸识别方法,并在人脸识别实验中获得比 2DLDA, 2DLPP, 2DPCA 等更高的分类精度.然而,2DMLV 缺失了与 SSSL 的性能比较,且由于 Kronecker 积的计算导致较高的空间复杂度.2016 年, Zhu 等人<sup>[26]</sup>提出了一种不同于 SSSL 显性空间正则化的隐性空间正则化策略(implicit spatial regularization,

ISR),基于 SSSL 的假设:局部图像区域(也称作空间窗口,如  $3 \times 3$ )通常是光滑的,以重叠和非重叠方式将原图像划分成同大小的空间窗口,再依次将各空间窗口拉成向量作为新矩阵的列,从而重新组装成新的矩阵作为后续特征抽取的对象,具体操作如 3.2 节图 2 所示.重组的本意在于使新矩阵每列能保持自然的光滑和局部空间关系.由此结合现成的 2D 特征抽取法所提出的 ISR+双边 2DLDA (ISR-b)方法获得了优于 ISR+单边 2DLDA (ISR-s)和 SSSL 的分类精度.但其仍然存在 2 点不足:

1) 尽管其中的双边 2DLDA 达到了同类方法当前最好的结果,但其目标函数非凸,只能采用交替迭代优化算法求解,计算代价大,同时也不能保证解的全局最优性;

2) 划分重组的隐性方法尽管直觉上较好地保持了图像的局部空间光滑,但因其缺乏显性光滑的强制,其空间仍欠光滑.

为克服上述方法的不足,提出了一种结合显性与隐性空间光滑的高效二维图像判别特征抽取框架(2D-CISSE). 2D-CISSE 框架主要包含 3 个步骤,如图 1 所示,其关键步骤体现在步骤 1 中,即对原图像预先显性空间光滑,而后 2 个 Step 等同上述 ISR. 尽管 2D-CISSE 仅简单地在非重叠 ISR 基础上施加显性空间光滑约束,但令人惊讶的是其仅需结合单边 2DLDA 即可获得比 ISR-b(非重叠 ISR 中最优)显著更优的分类性能,由此不仅大大降低了计算复杂度,同时还保证了解的全局最优性.最后利用最近邻分类器分别在人脸数据集 ORL, Yale, CMU PIE, AR 以及手写数字数据集 MNIST 和 USPS 上的分类实验结果显示 2D-CISSE 框架获得了比 SSSL, 2DMLV, ISR 更优的分类性能和更高的计算效率.

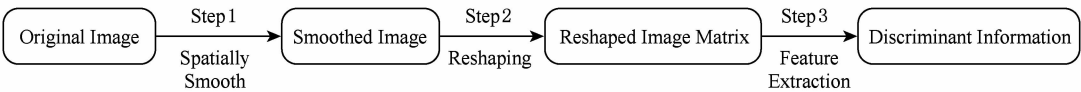


Fig. 1 2D-CISSE framework

图 1 结合显性与隐性空间光滑的 2D 图像判别特征抽取框架

1 相关工作

$\{A_i\}_{i=1}^N$  表示一组训练图像集,  $\{x_i\}_{i=1}^N$  是其相应的向量化的训练集,其中  $N$  为训练样本数,  $A_i \in \mathbb{R}^{r \times c}$  表示第  $i$  张图像,  $x_i \in \mathbb{R}^{r \times 1}$  是  $A_i$  的向量表达,令  $X = (x_1, x_2, \dots, x_N) \in \mathbb{R}^{r \times N}$ . 假设  $\mathcal{G}$  为以训练样本阵  $X$ (或  $x$ ) 为顶点集构造的邻接图,  $W$  与  $L$  分别为图  $\mathcal{G}$  对应的边权重矩阵和图 Laplacian 矩阵,通过这些

符号,简单回顾 SSSL 与 2DMLV.

1.1 空间光滑的子空间学习模型 SSSL

空间光滑的子空间学习(SSSL)<sup>[22]</sup>是一种典型的基于向量的显性空间正则化方法,其利用离散的 Laplacian 光滑算子对投影向量显式地约束使其局部光滑,在图嵌入框架<sup>[27]</sup>中其目标函数为

$$\arg \max_a \frac{a^T X W X^T a}{(1-\beta) a^T X L X^T a + \beta f(a)}, \quad (1)$$

其中,  $a$  是待优化的投影向量,正则化因子  $0 \leq \beta \leq 1$

控制光滑度,  $J$  是一个离散的 Laplacian 正则化函数:

$$J(\mathbf{a}) = \|\Delta \cdot \mathbf{a}\|^2 = \mathbf{a}^\top \Delta^\top \Delta \mathbf{a}, \quad (2)$$

其中二维 Laplacian 算子的离散近似  $\Delta \in \mathbb{R}^{r \times r}$  可以表示为

$$\Delta = \mathbf{D}_1 \otimes \mathbf{I}_2 + \mathbf{I}_1 \otimes \mathbf{D}_2, \quad (3)$$

其中,  $\mathbf{I}_1$  与  $\mathbf{I}_2$  分别是  $r \times r$  和  $c \times c$  的单位矩阵,  $\otimes$  表示 Kronecker 积,  $\mathbf{D}_1$  ( $\mathbf{D}_2$ ) 是在图像行(列)方向一个  $r \times r$  ( $c \times c$ ) 二阶梯度光滑算子或矩阵:

$$\mathbf{D}_j = \frac{1}{h_j^2} \begin{bmatrix} -1 & 1 & & & 0 \\ & 1 & -2 & 1 & \\ & & \ddots & \ddots & \ddots \\ & & & 1 & -2 & 1 \\ 0 & & & & 1 & -1 \end{bmatrix}, \quad (4)$$

其中,  $h_1$  ( $h_2$ ) 是样本矩阵在水平(竖直)方向的宽度.

## 1.2 二维最大局部变化法 2DMLV

基于图像欧氏距离的二维最大局部变化法 (two-dimensional maximum local variation, 2DMLV)<sup>[25]</sup> 将包含图像的差异性与判别性信息的局部变化混合到目标函数中来降维, 利用图像欧氏距离替换传统欧氏距离, 从而将图像像素点间的空间关系运用到 2D 特征抽取法中. 对于任意图像矩阵  $\mathbf{A}_j \in \mathbb{R}^{r \times c}$ , 目标是通过线性变换  $\mathbf{y} = \mathbf{A}\mathbf{a}$  得到一个低维的特征向量  $\mathbf{y} \in \mathbb{R}^c$ ,  $\mathbf{a}$  为待优化的投影向量.

首先, 定义邻接关系矩阵  $\mathbf{V}$ , 当  $\mathbf{A}_j$  是  $\mathbf{A}_l$  的  $k$  近邻或者  $\mathbf{A}_l$  是  $\mathbf{A}_j$  的  $k$  近邻时,  $V_{jl} = 1$ , 否则  $V_{jl} = 0$ . 实际上, 近邻空间的图像可能来自不同的类别, 为了保持图像多样性, 保证距离较大的 2 张图像在降维后依然保持较大的距离, 可分别定义类内邻接关系矩阵  $\mathbf{V}^w$  与类间邻接关系矩阵  $\mathbf{V}^b$ :

$$V_{jl}^w = \begin{cases} V_{jl} \times \exp(-t/\|\mathbf{A}_j - \mathbf{A}_l\|_F^2), & \tau_j = \tau_l, \\ 0, & \text{otherwise}, \end{cases} \quad (5)$$

$$V_{jl}^b = \begin{cases} V_{jl} \times \exp(-t/\|\mathbf{A}_j - \mathbf{A}_l\|_F^2), & \tau_j \neq \tau_l, \\ 0, & \text{otherwise}, \end{cases} \quad (6)$$

其中, 参数  $t \geq 0$ ,  $\tau_j$  表示图像  $\mathbf{A}_j$  的类标号,  $\|\cdot\|_F$  表示 Frobenius 范数.

定义矩阵版本的图像欧氏距离<sup>[25]</sup>:

$$d_{\text{IMED}}(\mathbf{A}_j - \mathbf{A}_l) = \|\hat{\mathbf{A}}_j - \hat{\mathbf{A}}_l\|_F, \quad (7)$$

其中,  $\hat{\mathbf{A}}_j = \mathbf{G}_1^{1/2} \mathbf{A}_j (\mathbf{G}_2^{1/2})^\top$ , 对称矩阵  $\mathbf{G}_1 \in \mathbb{R}^{r \times r}$  与  $\mathbf{G}_2 \in \mathbb{R}^{c \times c}$  定义如下:

$$\mathbf{G}_1(p, p') = \left(\frac{1}{2\pi}\right)^{\frac{1}{2}} \exp\left\{-\frac{(p-p')^2}{2\sigma^2}\right\}, \quad (8)$$

$$\mathbf{G}_2(q, q') = \left(\frac{1}{2\pi}\right)^{\frac{1}{2}} \exp\left\{-\frac{(q-q')^2}{2\sigma^2}\right\}, \quad (9)$$

其中,  $p, p' = 1, 2, \dots, r$ ;  $q, q' = 1, 2, \dots, c$ ;  $\sigma$  为宽度参数, 用于调整光滑度.

分别最大化类内和类间的局部方差, 目标函数可以定义为

$$\begin{aligned} \arg \max_{\mathbf{a}} \sum_{l=1}^N \sum_{j=1}^N ((\hat{\mathbf{A}}_l \mathbf{a} - \hat{\mathbf{A}}_j \mathbf{a})^\top (\hat{\mathbf{A}}_l \mathbf{a} - \hat{\mathbf{A}}_j \mathbf{a}) V_{jl}^w), \\ \arg \max_{\mathbf{a}} \sum_{l=1}^N \sum_{j=1}^N ((\hat{\mathbf{A}}_l \mathbf{a} - \hat{\mathbf{A}}_j \mathbf{a})^\top (\hat{\mathbf{A}}_l \mathbf{a} - \hat{\mathbf{A}}_j \mathbf{a}) V_{jl}^b). \end{aligned} \quad (10)$$

之后可利用线性和将其转换为单目标的优化问题, 可以通过特征分解求出优化目标.

## 2 2D-CISSE 框架

针对结合显性与隐性的空间光滑的二维图像判别特征抽取框架 (2D-CISSE) 所包含如图 1 所示的 3 个步骤 (Step1 ~ Step3), 本节将分别对其详细地展开介绍.

### 2.1 原始图像显式光滑

#### 2.1.1 全局图像欧氏距离光滑

图像欧氏距离 (IMED)<sup>[9]</sup> 巧妙地将像素点间的空间关系嵌入欧氏距离成为有效的图像相似性度量.

**定义 1**<sup>[9]</sup>. 图像欧氏距离 IMED. 假设有 2 幅图像  $\mathbf{A}_j \in \mathbb{R}^{r \times c}$ ,  $\mathbf{A}_l \in \mathbb{R}^{r \times c}$ , 其距离可表示为

$$d_{\text{IMED}}^2(\mathbf{A}_j - \mathbf{A}_l) = \overline{\text{Vec}}(\mathbf{A}_j - \mathbf{A}_l)^\top \mathbf{G} \overline{\text{Vec}}(\mathbf{A}_j - \mathbf{A}_l), \quad (11)$$

其中,  $\overline{\text{Vec}}(\mathbf{A}_j - \mathbf{A}_l) = \mathbf{x}_j - \mathbf{x}_l$ , 对称矩阵  $\mathbf{G} \in \mathbb{R}^{n \times n}$  中元素  $G_{ii,jj}$  可以定义为

$$G_{ii,jj} = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{1}{2}((p-p')^2 + (q-q')^2)\right), \quad (12)$$

$$ii = (p-1)c + q,$$

$$jj = (p'-1)c + q',$$

$$p, p' = 1, 2, \dots, r, \quad q, q' = 1, 2, \dots, c.$$

由于矩阵  $\mathbf{G}$  维度过高, 通常计算困难, 幸运的是能利用简单变换将其改写为

$$\mathbf{G} = \mathbf{G}_1 \otimes \mathbf{G}_2, \quad (13)$$

对称矩阵  $\mathbf{G}_1 \in \mathbb{R}^{r \times r}$  与  $\mathbf{G}_2 \in \mathbb{R}^{c \times c}$  同式 (8) (9). 利用式 (13), 有如下定理:

**定理 1**<sup>[28]</sup>. 设有 2 个矩阵  $\mathbf{B}_1 \in \mathbb{R}^{m \times p}$ ,  $\mathbf{B}_2 \in \mathbb{R}^{n \times q}$ ,  $\mathbf{h}_i^\top$  ( $i = 1, 2, \dots, p$ ) 表示矩阵  $\mathbf{H} \in \mathbb{R}^{p \times q}$  的第  $i$  行, 那么矩阵  $\mathbf{H}$  的向量化:  $\overline{\text{Vec}}(\mathbf{H}) = (\mathbf{h}_1^\top \quad \mathbf{h}_2^\top \quad \dots \quad \mathbf{h}_p^\top)^\top$ . 如果存在  $\mathbf{z} = \overline{\text{Vec}}(\mathbf{Z}) = (\mathbf{B}_1 \otimes \mathbf{B}_2) \overline{\text{Vec}}(\mathbf{H})$ , 则有  $\mathbf{Z} = \mathbf{B}_1 \mathbf{H} \mathbf{B}_2^\top$ . 其中  $\otimes$  表示 Kronecker 积,  $\overline{\text{Vec}}$  代表矩阵的向量化操作.

根据定理 1 可变换得到图像欧氏距离的矩阵版本如 1.2 节式(7)所示. 因此即得图像欧氏距离对原始图像的光滑操作:

$$\hat{\mathbf{A}}_j = \mathbf{G}_1^{1/2} \mathbf{A}_j (\mathbf{G}_2^{1/2})^T, \quad (14)$$

$\hat{\mathbf{A}}_j$  称作对原图像矩阵  $\mathbf{A}_j$  的全局图像欧氏距离光滑 (GIMED), 具体操作见算法 1. 而后将其运用到图像样本集  $\{\mathbf{A}_i\}_{i=1}^N$  得到光滑的样本集  $\{\hat{\mathbf{A}}_i\}_{i=1}^N$ .

**算法 1.** 全局图像欧氏距离光滑算法 GIMED.

输入: 训练图像集  $\{\mathbf{A}_i\}_{i=1}^N, \mathbf{A}_i \in \mathbb{R}^{r \times c}$ , 参数因子  $\sigma$ ;

输出: 光滑的图像矩阵  $\hat{\mathbf{A}} = (\hat{\mathbf{A}}_1, \hat{\mathbf{A}}_2, \dots, \hat{\mathbf{A}}_N)$ .

① 分别利用式(8)(9)计算  $\mathbf{G}_1, \mathbf{G}_2$ ;

② for  $i=1$  to  $N$  do / \* 对每个图像矩阵 \* /

③ 用式(14)计算图像矩阵  $\mathbf{A}_i$  对应光滑矩阵

$$\hat{\mathbf{A}}_i;$$

④ end for

### 2.1.2 局部图像欧氏距离光滑

除了对图像矩阵全局空间光滑, 本文将 IMED 结合到第 2 阶段的图像矩阵划分重组 (详见 3.2 节) 之中, 提出局部图像欧氏距离光滑算法 (LIMED).

假设有 2 个原始图像  $\mathbf{A}_j \in \mathbb{R}^{r \times c}, \mathbf{A}_l \in \mathbb{R}^{r \times c}$ , 选择尺寸为  $m \times n$  的空间窗口, 利用如图 2 方式将其划分重组为

$$\mathbf{A}'_j = (\mathbf{a}'_j{}^1, \mathbf{a}'_j{}^2, \dots, \mathbf{a}'_j{}^K), \mathbf{A}'_l = (\mathbf{a}'_l{}^1, \mathbf{a}'_l{}^2, \dots, \mathbf{a}'_l{}^K),$$

其中,  $K = (rc)/(mn)$ .

重组后矩阵  $\mathbf{A}'_j$  的第  $i$  列为原始图像矩阵  $\mathbf{A}_j$  划分后第  $i$  个空间窗口  $\mathbf{A}_j^i \in \mathbb{R}^{m \times n}$  的向量化, 利用传统欧氏距离可得:

$$d^2(\mathbf{A}'_j - \mathbf{A}'_l) = \sum_{i=1}^K (\mathbf{a}'_j{}^i - \mathbf{a}'_l{}^i)^T (\mathbf{a}'_j{}^i - \mathbf{a}'_l{}^i) = \sum_{i=1}^K d^2(\mathbf{A}_j^i - \mathbf{A}_l^i), \quad (15)$$

用式(7)中矩阵图像欧氏距离替换传统欧氏距离:

$$d_{\text{IMED}}^2(\mathbf{A}'_j - \mathbf{A}'_l) =$$

$$\sum_{i=1}^K d_{\text{IMED}}^2(\mathbf{A}_j^i - \mathbf{A}_l^i) = \sum_{i=1}^K \|\hat{\mathbf{A}}_j^i - \hat{\mathbf{A}}_l^i\|,$$

$$\hat{\mathbf{A}}_j^i = \mathbf{G}_1^{1/2} \mathbf{A}_j^i (\mathbf{G}_2^{1/2})^T, \mathbf{G}_1 \in \mathbb{R}^{m \times m}, \mathbf{G}_2 \in \mathbb{R}^{n \times n}. \quad (16)$$

对于每个原始图像矩阵  $\mathbf{A}_j$ , 式(16)即为对其每一个空间窗口进行空间光滑操作, 称为局部图像欧氏距离光滑算法 LIMED.

### 2.1.3 双边滤波光滑

Tomasi 等人<sup>[29]</sup>提出的双边滤波 (bilateral filter, BLF) 在处理相邻像素点的灰度值时, 兼顾了几何空间上的临近关系和灰度值上的相似性, 对两者非线性组合, 自适应滤波后得到光滑图像, 优点在

于图像光滑的同时模糊的边缘信息能够得以保持. 这里将其用作原图像全局光滑.

原始图像  $\mathbf{A}_l$  经过双边滤波光滑得到图像  $\hat{\mathbf{A}}_l$ , 其空间坐标  $(i, j)$  处像素点的灰度值  $\hat{\mathbf{A}}_{ij}$  的变换如下:

$$\hat{\mathbf{A}}_{ij} = \frac{\sum_{p, q \in S_{i, j}} w(p, q) g(p, q)}{\sum_{p, q \in S_{i, j}} w(p, q)}, \quad (17)$$

其中,  $S_{i, j}$  表示以空间坐标  $(i, j)$  为中心点大小为  $(2M+1) \times (2M+1)$  的邻域 (或窗口),  $M$  为滤波器的半宽,  $M$  值越大, 平滑作用越强, 实验中取  $M=5$ . 式(17)实际就是对中心像素点邻域内像素点灰度值的加权平均,  $w(p, q)$  为空间坐标  $(i, j)$  处的加权系数, 由 2 部分因子的乘积组成:

$$w(p, q) = w_s(p, q) w_r(p, q), \quad (18)$$

其中, 空间临邻度因子  $w_s(p, q)$  与灰度值相似度因子  $w_r(p, q)$  分别表示为

$$w_s(p, q) = \exp\left(-\frac{|p-i|^2 + |q-j|^2}{2\sigma_s^2}\right), \quad (19)$$

$$w_r(p, q) = \exp\left(-\frac{A(p, q) - A(i, j)}{2\sigma_r^2}\right), \quad (20)$$

其中, 参数  $\sigma_s$  与  $\sigma_r$  分别控制空间邻近度因子与灰度值相似度因子的衰减程度.

双边滤波光滑算法详细过程如下.

**算法 2.** 双边滤波光滑算法 BLF.

输入: 训练图像集  $\{\mathbf{A}_i\}_{i=1}^N, \mathbf{A}_i \in \mathbb{R}^{r \times c}$ , 参数因子  $\sigma_r, \sigma_s$ , 滤波器半宽  $M$ ;

输出: 光滑的图像矩阵  $\hat{\mathbf{A}} = (\hat{\mathbf{A}}_1, \hat{\mathbf{A}}_2, \dots, \hat{\mathbf{A}}_N)$ .

① for  $i=1$  to  $N$  do / \* 每个训练图像样本 \* /

② for  $(p, q) = (1, 1)$  to  $(r, c)$  do / \* 对每个像素点 \* /

③ 利用式(19)(20)计算  $(p, q)$  处的  $w_s, w_r$ ;

④ 用式(17)计算图像矩阵在空间位置  $(p, q)$  处的灰度值  $\hat{\mathbf{A}}_{pq}$ ;

⑤ end for

⑥ end for

## 2.2 图像划分重组

图像重组技术在人脸识别中已有广泛应用, 如我们此前的工作 MatPCA 与 MatFLDA<sup>[30]</sup> 将向量模式重组为相应矩阵后再利用 2DPCA 与 2DLDA 特征抽取. SSSL<sup>[22]</sup> 中分析了 2D 特征抽取方法只能整行或整列利用空间信息. 实际上图像整行 (或列) 并不光滑, 故这些 2D 嵌入函数的空间光滑度仍较粗. 以此为动机的隐式正则化<sup>[19]</sup> 将原始图像划分重

组以体现列向量的光滑度,隐式地利用图像空间结构信息.

本文借鉴了 ISR<sup>[19]</sup> 的非重叠的划分策略,如图 2 所示,将给定的大小为  $r \times c$  的光滑图像划分为  $K$  个相同尺寸( $m \times n$ )的空间窗口,  $K = (rc)/(mn)$ , 依次将每个空间窗口拉成向量从而组成维度为  $mn \times K$  的新矩阵. 直观上,通过划分重组,新矩阵的列(相应空间窗口)通常是光滑的,隐含的局部空间关系得以高度充分地保留.

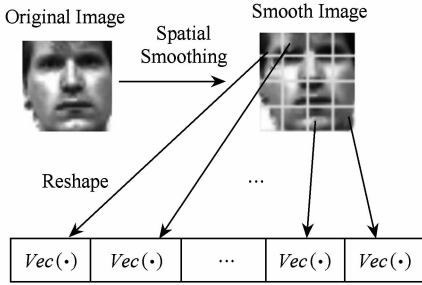


Fig. 2 An illustration of smoothed image reshaping in non-overlapping way

图 2 光滑图像的非重叠重组图示

### 2.3 单边 2DLDA 特征抽取

对重组后的图像矩阵,可采用任何现有 2D 特征抽取法对特征进行抽取. 令  $\hat{\mathbf{A}}' = (\hat{\mathbf{A}}'_1, \hat{\mathbf{A}}'_2, \dots, \hat{\mathbf{A}}'_N)$  为一组重组后的图像矩阵,分为  $C$  类,第  $i$  类含有  $N_i$  个图像样本. 以  $\hat{\mathbf{A}}'$  为顶点集构造无向加权图  $\mathcal{G} = \{\hat{\mathbf{A}}', \mathbf{W}'\}$ , 其中权重矩阵  $\mathbf{W}'$  的元素  $W'_{i,j}$  为定点  $\hat{\mathbf{A}}'_i$  与  $\hat{\mathbf{A}}'_j$  之间的相似度,  $\mathbf{D}$  为对角矩阵,其对角元素  $D_{i,i} = \sum_{j \neq i} W'_{i,j}$ . 令  $\mathbf{Y} = (\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_N)$  为抽取得到的特征集,在一般图嵌入框架中目标函数可表示为

$$\min \sum_{i,j} \|\mathbf{Y}_i - \mathbf{Y}_j\|^2 W'_{i,j}. \quad (21)$$

本文采用基于线性判别分析(LDA)的 2D 版本,单边 2DLDA 抽取特征. LDA 通过最小化类内散度的同时最大化类间散度以将不同类别尽可能分开,在图嵌入框架中定义 LDA 准则,权重矩阵  $\mathbf{W}'$  为

$$W_{i,j} = \begin{cases} 1/N_i, & \tau_i = \tau_j = t; \\ 0, & \text{otherwise.} \end{cases} \quad (22)$$

单边 2DLDA 旨在找到投影矩阵  $\mathbf{U}$  或  $\mathbf{V}$ , 通过  $\mathbf{Y}_i = \mathbf{U}^T \hat{\mathbf{A}}'_i$  或  $\mathbf{Y}_i = \hat{\mathbf{A}}'_i \mathbf{V}$  获取相应特征. 经过前两个步骤得到的矩阵  $\hat{\mathbf{A}}'$  在列方向是光滑的,保持其光滑特性的同时在行方向利用  $\mathbf{Y}_i = \hat{\mathbf{A}}'_i \mathbf{V}$  抽取特征,带入式(21)得到单边 2DLDA 目标函数:

$$\begin{aligned} \min \sum_{i,j} \|\mathbf{Y}_i - \mathbf{Y}_j\|^2 W'_{i,j} = \\ \min \sum_{i,j} \|\hat{\mathbf{A}}'_i \mathbf{V} - \hat{\mathbf{A}}'_j \mathbf{V}\|^2 W'_{i,j}, \end{aligned} \quad (23)$$

通过简单的公式变换,式(23)重写为

$$\max \text{tr}(\mathbf{V}^T (\sum_{i,j} W'_{i,j} \hat{\mathbf{A}}'_i \hat{\mathbf{A}}'^T_j) \mathbf{V}), \quad (24)$$

$$\text{s. t. } \text{tr}(\mathbf{V}^T (\sum_{i,i} D_{i,i} \hat{\mathbf{A}}'_i \hat{\mathbf{A}}'^T_i) \mathbf{V}) = 1.$$

投影矩阵  $\mathbf{V}$  的优化可通过解决下式(25)的特征分解问题求得,具体算法参考算法 3 中 Step3.

$$\sum_{i,j} W'_{i,j} \hat{\mathbf{A}}'_i \hat{\mathbf{A}}'^T_j \mathbf{V} = \lambda \sum_i (D_{i,i} \hat{\mathbf{A}}'_i \hat{\mathbf{A}}'^T_i) \mathbf{V}. \quad (25)$$

### 2.4 分类

给定测试图像矩阵  $\mathbf{P}$ , 利用 2.1 节与 2.2 节对其空间光滑与重组后得到新的矩阵  $\hat{\mathbf{P}}'$ , 之后采用单边 2DLDA 对其抽取特征:  $\mathbf{F} = \hat{\mathbf{P}}' \mathbf{V}$ . 假设  $\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_N$  是 2D-CISSE 框架得到的相应训练样本的特征矩阵, 利用最近邻分类器算法, 对于所有的  $i = 1, 2, \dots, N$ , 如果  $\text{dis}(\mathbf{Y}_i, \mathbf{F}) = \min_i (\text{dis}(\mathbf{Y}_i, \mathbf{F}))$ , 则将图像  $\mathbf{P}$  判别到训练样本图像  $\mathbf{Y}_i$  所属的类.

定义特征矩阵  $\mathbf{Y}_1 = (\mathbf{y}_1^1, \mathbf{y}_1^2, \dots, \mathbf{y}_1^d)$  与特征矩阵  $\mathbf{Y}_2 = (\mathbf{y}_2^1, \mathbf{y}_2^2, \dots, \mathbf{y}_2^d)$  的距离如下:

$$\text{dis}(\mathbf{Y}_1, \mathbf{Y}_2) = \sum_{i=1}^d \|\mathbf{y}_1^i - \mathbf{y}_2^i\|^2. \quad (26)$$

#### 算法 3. 2D-CISSE.

输入: 训练图像集  $\{\mathbf{A}_i\}_{i=1}^N, \mathbf{A}_i \in \mathbb{R}^{r \times c}$ , 空间窗口尺寸  $m \times n$ , 光滑算法选择项 *choice*;

输出: 投影矩阵  $\mathbf{V}$ .

Step1. 原始图像空间光滑.

- ① if *choice* = GIMED, 执行算法 1;
- ② 转到 Step2;
- ③ end if
- ④ if *choice* = BLF, 执行算法 2;
- ⑤ 转到 Step2;
- ⑥ end if

Step2. 划分重组.

- ① for  $i = 1$  to  $N$  do /\* 对每张图像 \*/
- ② 将  $\mathbf{A}_i$  或  $\hat{\mathbf{A}}_i$  按照  $m \times n$  尺寸子块划分;
- ③ if *choice* = LIMED;
- /\* 若选择 LIMED 算法 \*/
- ④ for  $j = 1$  to  $K$  do /\*  $K = rc/mn$  \*/
- ⑤ 用式(16) 对每个空间窗口空间光滑;
- ⑥ end for

- ⑦ end if
- ⑧ 将每个空间窗口(矩阵块)拉成向量,作为列向量重组为新的矩阵  $\hat{\mathbf{A}}'_i$ ;
- ⑨ end for
- Step3. 单边 2DLDA.

- ① 计算权重矩阵  $\mathbf{W}'$  与  $D_{i,i} = \sum_{i \neq j} \mathbf{W}'_{i,j}$ ;
- ② 求  $\sum_{i,j} \mathbf{W}'_{i,j} \hat{\mathbf{A}}'_i{}^T \hat{\mathbf{A}}'_j \boldsymbol{\varphi} = \lambda \sum_i (D_{i,i} \hat{\mathbf{A}}'_i \hat{\mathbf{A}}'_i{}^T) \boldsymbol{\varphi}$   
前  $d$  个最大特征值的特征向量  $\boldsymbol{\varphi}_1, \boldsymbol{\varphi}_2, \dots, \boldsymbol{\varphi}_d$ ;
- ③  $\mathbf{V} \leftarrow (\boldsymbol{\varphi}_1, \boldsymbol{\varphi}_2, \dots, \boldsymbol{\varphi}_d)$ .

算法 3 分别将全局图像欧氏距离光滑算法 GIMED、局部图像欧氏距离光滑算法 LIMED 和双边滤波算法 BLF 嵌入 2D-CISSE 框架中,得到 3 种算法:GIMED-ISR-S,LIMED-ISR-S,BLF-ISR-S.

2.5 算法复杂度分析

假设有  $N$  个训练样本,每个图像样本维度为  $r \times c$ ,空间窗口大小为  $m \times n$ . 本文利用的 GIMED, LIMED,BLF 三个光滑算法的计算复杂度分别为  $O(Nrc(r+c)), O(Nrc(m+n)), O(NrcM^2)$ , 其中  $M$  为 BLF 的窗口半宽. 本文 2D-CISSE 框架中图像划分重组与特征抽取 2 个步骤的计算复杂度同 ISR-s,均为  $O((r^3c^3)/(m^3n^3))$ ,而 ISR-b 是由非凸问题获得,若其交替迭代优化选取的迭代次数为  $Ite$ ,则其计算时间复杂度为  $O((m^3n^3 + (r^3c^3)/(m^3n^3)) \times Ite)$ ,SSSL 与 2DMLV 的计算时间复杂度分别为  $O(r^3c^3)$  和  $O(N^2rc + r^3c^3)$ . 因此可以发现:尽管需要付出 Step1 中对原始图像空间光滑的计算代价,但 2D-CISSE 框架在整体上依然具有最低的计算复杂度.

3 实 验

为验证 2D-CISSE 框架在现实数据集上的分类性能,将利用其所提出的 GIMED-ISR-S,LIMED-ISR-S,BLF-ISR-S 三个算法分别在 4 个人脸数据集(Yale,ORL,CMU PIE,AR)、3 个手写数字数据集(MNIST350,MNIST2k2k,USPS)上与 SSSL、非重叠 ISR+双边 2DLDA(简记为 ISR-b)、非重叠 ISR+单边 2DLDA(简记为 ISR-s)和 2DMLV 四个代表性的空间信息利用算法进行比较. 表 1 为本文所用 7 个数据集的详细介绍.

Table 1 Introduction of Datasets

表 1 数据集介绍

| Data Sets | Number of Samples | Number of Classes | Number of Each Class | Original Dimension | Resizing Dimension |
|-----------|-------------------|-------------------|----------------------|--------------------|--------------------|
| Yale      | 165               | 11                | 11                   | 32×32              | 33×33              |
| ORL       | 400               | 40                | 10                   | 32×32              | 33×33              |
| CMU PIE   | 11 560            | 68                | 170                  | 32×32              | 33×33              |
| AR        | 2 600             | 100               | 26                   | 33×33              | 33×33              |
| MNIST350  | 3 500             | 10                | 350                  | 28×28              | 27×27              |
| MNIST2k2k | 4 000             | 10                | *                    | 28×28              | 27×27              |
| USPS      | 9 298             | 10                | *                    | 16×16              | 15×15              |

Note: “\*” means each class has different numbers.

SSSL 与 2DMLV 均采用 5 折(5-fold)交叉验证选取参数. 本文算法和 ISR 在图像划分重组时均将空间窗口大小设为  $3 \times 3$ (文献[26]中验证为最优),为了对图像完整地非重叠划分,利用近邻差值法将一些数据集的维度进行了调整,如表 1 所示. 本文的 BLF-ISR-S 算法通过 5 折交叉验证获得不同训练样本量的参数  $\sigma_r$  与  $\sigma_s$ ,基于图像欧氏距离的 GIMED 与 LIMED 的参数  $\sigma$  对训练样本数目弱敏感,因此每个数据集交叉验证获得的  $\sigma$  可运用于对其不同的训练划分. 实验中,每个训练集均是从完整数据集随机采样获得.

3.1 人脸数据集

表 2 为在图 3 所示的人脸数据集 Yale,CMU PIE,ORL 上的实验结果,表 2 中结果为 20 次随机采样训练的均值与标准差. 可以发现:1)本文的 3 种算法均优于 ISR-s,意味着对原始图像的光滑策略(Step1)确实提升了分类性能;2)2 种全局光滑算法 GIMED-ISR-S 与 BLF-ISR-S 均获得比 ISR-b,ISR-s,SSSL,2DMLV 更优的分类性能;3)局部光滑策略 LIMED-ISR-S 的分类性能比起 2 种全局策略稍低,但在大部分训练集中能得到不低于 4 种对比算法的分类与稳定性能.

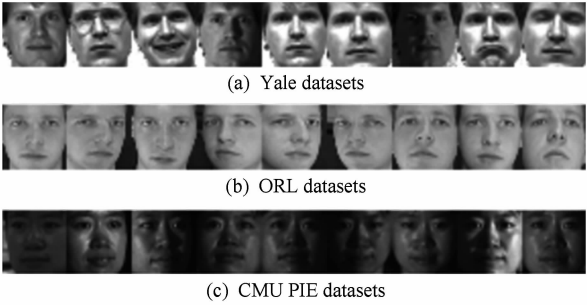


Fig. 3 Sample images from Yale, ORL and CMU PIE datasets

图 3 来自 Yale, ORL, CMU PIE 数据集的样本图像

Table 2 Recognition Accuracy on Yale, CMU PIE, ORL Datasets

表 2 在 Yale,CMU PIE,ORL 数据集上分类识别率

| Data Sets | # $N_i$ | Mean±Std               |                 |                        |                 |                 |                 |                 |
|-----------|---------|------------------------|-----------------|------------------------|-----------------|-----------------|-----------------|-----------------|
|           |         | GIMED-ISR-S            | LIMED-ISR-S     | BLF-ISR-S              | ISR-s           | ISR-b           | SSSL            | 2DMLV           |
| Yale      | 4       | <b>0.782 9±0.048 4</b> | 0.772 9±0.041 9 | 0.767 1±0.043 1        | 0.766 7±0.046 7 | 0.796 4±0.046 7 | 0.764 3±0.036 4 | 0.655 2±0.031 9 |
|           | 5       | <b>0.836 7±0.036 8</b> | 0.822 8±0.033 9 | 0.831 1±0.032 2        | 0.813 3±0.032 8 | 0.828 9±0.024 8 | 0.806 1±0.027 6 | 0.700 0±0.040 3 |
|           | 6       | 0.860 7±0.034 5        | 0.845 3±0.044 4 | <b>0.865 3±0.036 7</b> | 0.845 3±0.038 5 | 0.860 0±0.037 8 | 0.848 0±0.037 3 | 0.731 3±0.047 6 |
|           | 7       | <b>0.885 0±0.025 9</b> | 0.875 8±0.031 7 | 0.869 2±0.030 7        | 0.860 8±0.025 5 | 0.865 0±0.037 8 | 0.858 3±0.027 3 | 0.752 5±0.080 1 |
|           | 8       | <b>0.905 6±0.041 3</b> | 0.895 6±0.040 2 | 0.903 3±0.049 0        | 0.886 7±0.040 1 | 0.895 6±0.044 5 | 0.898 9±0.032 6 | 0.778 9±0.073 7 |
| CMU PIE   | 5       | 0.775 7±0.013 4        | 0.766 7±0.012 0 | <b>0.778 4±0.007 2</b> | 0.764 2±0.012 3 | 0.755 1±0.016 0 | 0.675 5±0.012 7 | *               |
|           | 10      | 0.880 7±0.007 5        | 0.876 8±0.006 8 | <b>0.881 3±0.006 2</b> | 0.876 4±0.006 9 | 0.871 1±0.007 0 | 0.829 5±0.010 4 | *               |
|           | 20      | 0.935 7±0.003 6        | 0.933 8±0.004 1 | <b>0.936 5±0.004 0</b> | 0.930 8±0.004 7 | 0.925 5±0.003 8 | 0.914 9±0.006 7 | *               |
|           | 30      | 0.953 5±0.002 5        | 0.951 5±0.002 5 | <b>0.954 5±0.002 4</b> | 0.949 7±0.002 3 | 0.947 5±0.002 0 | 0.942 4±0.003 8 | *               |
| ORL       | 2       | 0.831 7±0.028 3        | 0.805 5±0.034 1 | <b>0.833 8±0.021 2</b> | 0.802 5±0.038 0 | 0.823 3±0.031 7 | 0.826 1±0.028 1 | 0.695 0±0.035 5 |
|           | 4       | <b>0.961 9±0.011 4</b> | 0.954 8±0.014 7 | 0.958 5±0.014 4        | 0.947 9±0.023 2 | 0.959 6±0.013 5 | 0.941 7±0.019 1 | 0.843 1±0.022 4 |
|           | 5       | <b>0.978 0±0.012 3</b> | 0.976 3±0.011 9 | 0.977 8±0.009 9        | 0.972 5±0.012 7 | 0.975 0±0.013 1 | 0.962 3±0.016 4 | 0.874 5±0.026 6 |
|           | 6       | 0.983 8±0.010 4        | 0.983 4±0.011 9 | <b>0.987 8±0.009 8</b> | 0.978 1±0.009 8 | 0.983 4±0.010 2 | 0.970 3±0.013 4 | 0.905 6±0.022 5 |
|           | 8       | <b>0.990 0±0.011 2</b> | 0.988 8±0.010 7 | 0.989 4±0.010 9        | 0.984 4±0.009 0 | 0.986 3±0.012 8 | 0.978 8±0.012 9 | 0.926 3±0.019 0 |

Notes: The best performance in each line has been bolded; “#  $N_i$ ” means number of samples for training in each class; “\*” means out of memory (similarly hereinafter).

图 4 是 AR 人脸数据集,用于检测分类算法在有遮挡人脸集上的性能.图 5 中 50 次随机采样实验的结果显示:1)本文框架下 3 个算法均获得比 ISR-b

等方法更优的性能;2)局部光滑的 LIMED-ISR-S 所有训练划分中获得最优分类性能.

3.2 手写字母数据集

手写数字数据集 MNIST 中常用的 MNIST2k2k 含有 2000 训练样本与 2000 测试样本,包含手写数字 0~9 共 10 个类别,一些图像样本如图 6 所示.为便于划分训练样本,本文从 MNIST 中每类随机抽取 350 个样本组成数据集 MNIST350,并对其 50 次随机采样训练.如图 7 所示的 USPS 数据集将前 7 291 个样本作为训练集,余下用于测试.表 3 的实验结果表明:1)GIMED-ISR-S 与 BLF-ISR-S 均能获得比 ISR 与 SSSL 更优的分类性能,且识别率比 ISR-b 分别提高 10 和 15 个百分点,前者在 USPS



Fig. 4 Sample images from AR dataset

图 4 AR 数据集的样本图像

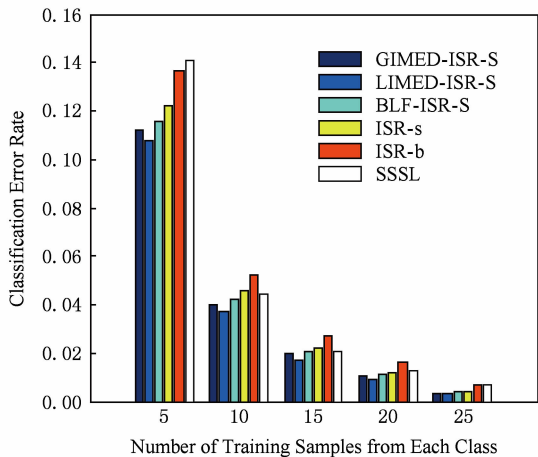


Fig. 5 Error rate (err) comparison on AR dataset

图 5 AR 数据集上的错误率比较



Fig. 6 Sample images from MNIST

图 6 MNIST 数据集的图像示例

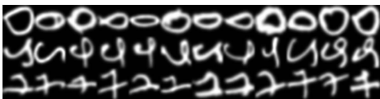


Fig. 7 Sample images from USPS

图 7 USPS 数据集的图像示例



Table 3 Recognition Accuracy on MNIST350, MNIST2k2k, USPS Datasets

表 3 在 MNIST350, MNIST2k2k, USPS 数据集上分类识别率

| Data Sets | # $N_i$ | Mean±Std        |                 |                        |                 |                 |                 |       |
|-----------|---------|-----------------|-----------------|------------------------|-----------------|-----------------|-----------------|-------|
|           |         | GIMED-ISR-S     | LIMED-ISR-S     | BLF-ISR-S              | ISR-s           | ISR-b           | SSSL            | 2DMLV |
| MNIST350  | 100     | 0.784 5±0.150 7 | 0.720 1±0.128 4 | <b>0.885 5±0.007 2</b> | 0.702 6±0.150 7 | 0.704 6±0.160 4 | 0.772 8±0.007 5 | *     |
|           | 150     | 0.789 6±0.078 8 | 0.726 9±0.101 5 | <b>0.899 2±0.005 2</b> | 0.716 4±0.117 3 | 0.745 3±0.107 1 | 0.798 0±0.007 2 | *     |
|           | 200     | 0.826 2±0.051 5 | 0.765 4±0.094 4 | <b>0.909 5±0.006 7</b> | 0.751 8±0.090 5 | 0.750 5±0.081 6 | 0.817 2±0.008 7 | *     |
|           | 250     | 0.830 4±0.045 4 | 0.778 3±0.087 5 | <b>0.913 3±0.008 7</b> | 0.755 7±0.093 7 | 0.759 4±0.093 7 | 0.824 1±0.011 5 | *     |
|           | 300     | 0.855 2±0.039 5 | 0.822 6±0.072 0 | <b>0.919 4±0.010 5</b> | 0.754 1±0.104 5 | 0.774 6±0.073 9 | 0.833 5±0.015 1 | *     |
| MNIST2k2k |         | 0.880 0         | 0.820 1         | <b>0.884 0</b>         | 0.753 5         | 0.795 5         | 0.871 5         | *     |
| USPS      |         | <b>0.953 7</b>  | 0.949 2         | 0.951 2                | 0.948 2         | 0.928 7         | 0.940 2         | *     |

Notes: The best performance in each line has been bolded; “#  $N_i$ ” means number of samples for training in each class; “\*” means out of memory (similarly hereinafter).

数据集获得最优,后者在 MNIST350 与 MNIST2k2k 上获得最优识别率和最高的稳定性;2) LIMED-ISR-S 稳定地优于 ISR 方法,略低于 SSSL.

3.3 实验结果分析

上述一系列实验结果表明:1)无论 GIMED-ISR-S, BLF-ISR-S 还是 LIMED-ISR-S 在 7 个数据集上的分类识别率大部分优于 ISR-s, ISR-b, SSSL, 2DMLV. 其中, GIMED-ISR-S 与 BLF-ISR-S 在所有数据集中均显著优于其他方法,这表明本文的框架在隐性利用局部空间信息的同时对图像显性全局光滑能够更加充分且有效地利用图像的空间结构信息,提升分类性能. 2) LIMED-ISR-S 在 7 个数据集 中的 4 个数据集上性能优于对比方法,在有遮挡人脸数据集 AR 上的分类性能显著最优,故在 Step1 采用局部光滑的策略相比全局光滑策略对一般图像分类性能提升较弱,但对于有遮挡的图像分类提升效果显著.

4 结 论

本文兼顾考虑图像空间结构信息的充分利用和计算代价的降低,提出了 1 种结合显性与隐性空间光滑的高效二维图像判别特征抽取框架(2D-CISSE). 在 2D-CISSE 中,仅需简单地在 ISR 策略上融合原图像的显性光滑,图像的局部光滑性得到了充分保持和学习利用,结果仅需结合单边 2DLDA 即可获得比 ISR 结合双边 2DLDA 显著更优的性能. 由 3 种图像光滑策略嵌入该框架得到的 3 个算法在标准数据集的实验结果验证了本文框架的有效性和性能的优越性,同时也体现了其一般适应性. 相比其他方法,2D-CISSE 具有 3 点优势:1)与 SSSL

相比,本文的框架是基于图像的 2D 方法,性能优于 SSSL 的同时降低了计算复杂度;2)与 2DMLV 相比,2D-CISSE 不局限于图像欧氏距离单一光滑,其双重空间光滑使得空间信息得以更充分地利用,性能显著更优,其次,2DMLV 的计算空间复杂度依赖于样本数量,2D-CISSE 则避免了其训练样本数量较大导致的超出内存问题;3)与非重叠 ISR 相比,2D-CISSE 使单边 2DLDA 即可获得比 ISR+双边 2DLDA 更好的分类性能,计算复杂度得到显著降低. 尽管 2D-CISSE 需要在 Step1 中付出图像空间光滑的计算代价,但这种代价相对于整体计算复杂度的降低很小. 在对光滑图像划分重组后,并非每个空间窗口都能对分类起判别作用. 因此,下一步工作一是重组矩阵的稀疏化,二是设计判别准则筛选空间窗口.

参 考 文 献

[1] Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks [C] // Proc of the Advances in Neural Information Processing Systems 25. Cambridge, MA: MIT Press, 2012: 1097-1105

[2] Sun Yi, Wang Xiaogang, Tang Xiaoou. Deeply learned face representations are sparse, selective, and robust [C] //Proc of the 28th IEEE Conf on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2015: 2892-2900

[3] Dahl G E, Yu Dong, Deng Li, et al. Context-dependent pre-trained deep neural networks for large-vocabulary speech recognition [J]. IEEE Trans on Audio, Speech, and Language Processing, 2012, 20(1): 30-42

[4] Collobert R, Weston J. A unified architecture for natural language processing: Deep neural networks with multitask learning [C] //Proc of the 25th Int Conf on Machine Learning. New York: ACM, 2008: 160-167

- [5] Belhumeur P N, Hespanha J P, Kriegman D J. Eigenfaces vs fisherfaces: Recognition using class specific linear projection [J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 1997, 19(7): 711-720
- [6] Turk M, Pentland A. Eigenfaces for recognition [J]. Journal of Cognitive Neuroscience, 1991, 3(1): 71-86
- [7] Murase H, Nayar S K. Visual learning and recognition of 3-D objects from appearance [J]. International Journal of Computer Vision, 1995, 14(1): 5-24
- [8] Tian Qing, Chen Songcan, Tan Xiaoyang. Comparative study among three strategies of incorporating spatial structures to ordinal image regression [J]. Neurocomputing, 2014, 136: 152-161
- [9] Wang Liwei, Zhang Yan, Feng Jufu. On the Euclidean distance of images [J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2005, 27(8): 1334-1339
- [10] Li Jing, Lu Baoliang. A framework for multi-view gender classification [C] //Proc of ICONIP'07. Berlin: Springer, 2008: 973-982
- [11] Li Jing, Lu Baoliang. An adaptive image Euclidean distance [J]. Pattern Recognition, 2009, 42(3): 349-357
- [12] Chen Songcan, Wang Zhe, Tian Yongjun. Matrix-pattern-oriented Ho-Kashyap classifier with regularization learning [J]. Pattern Recognition, 2007, 40(5): 1533-1543
- [13] Wang Zhe, Chen Songcan. New least squares support vector machines based on matrix patterns [J]. Neural Processing Letters, 2007, 26(1): 41-56
- [14] Tian Yongjun, Chen Songcan. Matrix-pattern-oriented Ho-Kashyap classifier with regularization learning [J]. Journal of Computer Research and Development, 2005, 42(9): 1628-1632 (in Chinese)  
(田永军, 陈松灿. 面向矩阵模式的正则化 Ho-Kashyap 算法 [J]. 计算机研究与发展, 2005, 42(9): 1628-1632)
- [15] Wang Zhe, Zhu Changming, Gao Daqi, et al. Three-fold structured classifier design based on matrix pattern [J]. Pattern Recognition, 2013, 46(6): 1532-1555
- [16] Wang Zhe, Chen Songcan, Liu Jun, et al. Pattern representation in feature extraction and classifier design: Matrix versus vector [J]. IEEE Trans on Neural Networks, 2008, 19(5): 758-769
- [17] Yang Jian, Zhang D, Frangi AF, et al. Two-dimensional PCA: A new approach to appearance-based face representation and recognition [J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2004, 26(1): 131-137
- [18] Ye Jieping. Generalized low rank approximations of matrices [J]. Machine Learning, 2005, 61(1/2/3): 167-191
- [19] Li Ming, Yuan Baozong. 2D-LDA: A statistical linear discriminant analysis for image matrix [J]. Pattern Recognition Letters, 2005, 26(5): 527-532
- [20] Ye Jieping, Janardan R, Li Qi. Two-dimensional linear discriminant analysis [C] //Proc of the Advances in Neural Information Processing Systems 17. Cambridge, MA: MIT Press, 2004: 1569-1576
- [21] Luo Luo, Xie Yubo, Zhang Zhihua, et al. Support matrix machines [C] //Proc of the 32nd Int Conf on Machine Learning. New York: ACM, 2015: 938-947
- [22] Cai Deng, He Xiaofei, Hu Yuxiao, et al. Learning a spatially smooth subspace for face recognition [C] //Proc of the 20th IEEE Conf on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2007: 1-7
- [23] Zuo Wangmeng, Liu Lei, Wang Kuanquan, et al. Spatially Smooth Subspace Face Recognition Using LOG and DOG Penalties [M]. Berlin: Springer, 2009: 439-448
- [24] Lotte F, Guan C. Spatially regularized common spatial patterns for EEG classification [C] //Proc of the 20th Int Conf on Pattern Recognition. 2010. Piscataway, NJ: IEEE, 2010: 3712-3715
- [25] Gao Quanyue, Gao Feifei, Zhang Hailin, et al. Two-dimensional maximum local variation based on image Euclidean distance for face recognition [J]. IEEE Trans on Image Processing, 2013, 22(10): 3807-3817
- [26] Zhu Yulian, Chen Songcan, Tian Qing. Spatial regularization in subspace learning for face recognition: Implicit vs. explicit [J]. Neurocomputing, 2016, 173: 1554-1564
- [27] Yan Shuicheng, Xu Dong, Zhang Benyu, et al. Graph embedding and extensions: A general framework for dimensionality reduction [J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2007, 29(1): 40-51
- [28] Franklin J N. Matrix Theory [M]. New York: Dover, 2000
- [29] Tomasi C, Manduchi R. Bilateral filtering for gray and color images [C] //Proc of the 6th Int Conf on Computer Vision. Piscataway, NJ: IEEE, 1998: 839-846
- [30] Chen Songcan, Zhu Yulian, Zhang Daoqiang, et al. Feature extraction approaches based on matrix pattern: MatPCA and MatFLDA [J]. Pattern Recognition Letters, 2005, 26(8): 1157-1167



**Zhu Kuaikuai**, born in 1992. Master candidate. His main research interests include machine learning and pattern recognition.



**Tian Qing**, born in 1984. PhD candidate. Student member of CCF. His main research interests include machine learning and pattern recognition.



**Chen Songcan**, born in 1962. Professor and PhD supervisor. Senior member of CCF. His main research interests include pattern recognition, machine learning and neural networks.