

图强化典型相关分析及在图像识别中的应用

苏树智^{*①②} 谢 军^① 平昕瑞^③ 高鹏连^①

^①(安徽理工大学计算机科学与工程学院 淮南 232001)

^②(合肥综合性国家科学中心能源研究院(安徽省能源实验室) 合肥 230031)

^③(安徽理工大学数学与大数据学院 淮南 232001)

摘 要: 典型相关分析(CCA)作为一种传统特征提取算法已经成功应用于模式识别领域, 其旨在找到使两组模态数据间相关性最大的投影方向, 但其本身为一种无监督的线性方法, 无法利用数据内在的几何结构和监督信息, 难以处理高维非线性数据。为此该文提出一种新的非线性特征提取算法, 即图强化典型相关分析(GECCA)。该算法利用数据中的不同成分构建多个成分图, 有效保留了数据间的复杂流形结构, 采用概率评估的方法使用类标签信息, 并通过图强化的方式将几何流形和监督信息融合嵌入到典型相关分析框架。为了对该算法进行评估, 分别在人脸和手写体数字数据集上设计了针对性实验, 良好的实验结果显示该算法在图像识别中的优势。

关键词: 模式识别; 典型相关分析; 成分图; 概率评估

中图分类号: TN911.73; TP391.4

文献标识码: A

文章编号: 1009-5896(2021)11-3342-08

DOI: 10.11999/JEIT210154

Graph Enhanced Canonical Correlation Analysis and Its Application to Image Recognition

SU Shuzhi^{①②} XIE Jun^① PING Xinrui^③ GAO Penglian^①

^①(School of Computer Science and Engineering, Anhui University of Science & Technology, Huainan 232001, China)

^②(Institute of Energy, Hefei Comprehensive National Science Center, Hefei 230031, China)

^③(School of Mathematics and Big Data, Anhui University of Science & Technology, Huainan 232001, China)

Abstract: As a traditional feature extraction algorithm, Canonical Correlation Analysis (CCA) has been excellently used for the field of pattern recognition. It aims to find the projection direction that makes the maximum of the correlation between two groups of modal data. However, since the algorithm is an unsupervised linear method, it can not use intrinsic geometry structures and supervised information hidden in data, which will cause difficulty in dealing with high-dimensional nonlinear data. Therefore, this paper proposes a new nonlinear feature extraction algorithm, namely Graph Enhanced Canonical Correlation Analysis (GECCA). The algorithm uses different components of the data to construct multiple component graphs, which retains effectively the complex manifold structures between the data. The algorithm utilizes the probability evaluation method to use class label information, and the graph enhancement method is utilized to integrate the geometry manifolds and the supervised information into the typical correlation analysis framework. Targeted experiments are designed on the face and handwritten digital image datasets to evaluate the algorithm. Good experimental results show the advantages of GECCA in image recognition.

Key words: Pattern recognition; Canonical Correlation Analysis (CCA); Component graph; Probability evaluation

收稿日期: 2021-02-18; 改回日期: 2021-05-20; 网络出版: 2021-06-04

*通信作者: 苏树智 sushuzhi@foxmail.com

基金项目: 国家自然科学基金(61806006), 中国博士后科学基金(2019M660149), 合肥综合性国家科学中心能源研究院项目(19KZS203)
Foundation Items: The National Natural Science Foundation of China (61806006), The China Postdoctoral Science Foundation (2019M660149), The Institute of Energy, Hefei Comprehensive National Science Center (19KZS203)

1 引言

特征提取^[1,2]是解决分类和识别任务的一类主流方法, 从模态数据种类角度划分, 可分为单模态方法和多模态方法。在单模态方法中主成分分析(Principal Component Analysis, PCA)^[3,4]使用最为广泛, 其通过正交变换从原始单模态数据^[5,6]的高维空间中抽取数据的低维表示形式, 同时最大化低维数据的差异。但是PCA是一种线性特征提取方

法, 没有考虑数据的内在结构信息, 因此在很多实际应用中难以抽取强鉴别力的低维特征。而多模态方法则是处理两种或两种以上模态数据的方法, 典型相关分析(Canonical Correlation Analysis, CCA)^[7]是多模态特征提取方法中的代表算法, 能够实现高维多模态数据的特征提取和融合。

CCA旨在学习两组模态数据的相关投影方向, 使投影后的两组模态间相关性最大, 目前已应用于图像处理^[8]、特征融合^[9]等领域。但CCA本身和PCA均为线性方法, 针对其难以提取高维非线性数据有效特征的问题, 文献^[10,11]中提出局部保持CCA(Locality Preserving Canonical Correlation Analysis, LPCCA), 该方法考虑局部邻域间的关系, 利用局部的几何结构解决一些非线性问题。虽然LPCCA比CCA更好地揭示了内在数据结构, 但是保留的局部信息多包含噪声, 揭示的近邻结构存在失真现象, 因此仅在姿态估计任务中获得了良好的实验结果, 而分类能力较差。为了解决LPCCA的弱分类性能问题, 文献^[12]提出一种新的可替代局部保持CCA(Alternative Locality-Preserving Canonical Correlation Analysis, ALPCCA), 该算法对LPCCA整体近邻几何结构的嵌入方式加以改进, 进而提升了识别性能。

此外使用核的思想也可以解决非线性问题, 文献^[13,14]提出的核CCA(Kernel Canonical Correlation Analysis, KCCA)利用核函数在更高维空间中表示原高维数据, 令原高维数据在更高维空间中线性可分, 尽管KCCA在分类性能上有所提升, 但将高维数据映射到更高维大大增加了运算过程的计算量。

除了考虑数据结构信息外还可以利用标签信息和图的思想提升算法的性能。标签信息即一种有效的监督信息, 鉴别CCA(Discriminative Canonical Correlation Analysis, DCCA)在文献^[15,16]中被提出, 将监督信息融合到特征提取的框架中, 在子空间中类内样本分布更加紧密, 类间样本离散程度更大, 从而使其在分类任务中拥有较好的类分离性。

而图多视角CCA(Graph Multiview Canonical Correlation Analysis, GMCCA)^[17,18]则通过使用图诱导的方式直接分析数据间的内在关系, 减少规范变量与常见的低维表示形式之间的距离。

上述方法是基于原始高维数据确定近邻关系或构建图, 但因噪声和冗余信息, 这种近邻关系或图揭示的几何结构会存在失真现象, 降低特征的鉴别性。本文提出一种新的图强化典型相关分析(Graph Enhanced Canonical Correlation Analysis, GECCA)算法。该算法使用谱聚类方法对原始高维数据进行分割, 获取多种数据成分, 并针对每种数据成分构建对应成分图, 可从不同成分的角度出发更好地揭示高维数据间复杂几何流形; 为了获得有益于分类和识别任务的信息, 采用相似准则构建成分图的权重矩阵, 借助概率评估方法保留成分图对应的类监督信息, 以此构建成分图的类系数矩阵; 通过图强化方式将成分图的权重矩阵和类系数矩阵进行融合得到强化矩阵, 并将其嵌入到典型相关分析的框架中。GECCA更好地揭示和保留了隐藏在原始高维数据中的几何流形和鉴别信息, 从而能够更好地指导和辅助特征的提取。GECCA有以下特点: (1)能更好地揭示数据的本质几何流形, 解决非线性特征提取问题; (2)有效地利用了监督信息, 使其在图像识别任务中拥有良好的类分离性。

2 算法回顾

本节将简要回顾CCA算法的主要内容。CCA为两种模态数据学习一个相关子空间, 使得在子空间中两种模态数据的相关性最大。两种模态数据集可分别用 $\mathbf{X}$ 和 $\mathbf{Y}$ 表示, 且所有样本均一一对应, 其中 $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n] \in \mathbf{R}^{p \times n}$, $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n] \in \mathbf{R}^{q \times n}$ 。每种模态数据集都有 n 个样本, $\mathbf{x}_i, \mathbf{y}_i$ ($i = 1, 2, \dots, n$)表示数据集 $\mathbf{X}, \mathbf{Y}$ 中第 i 个样本, 且样本均为均值化后的样本数据, $\mathbf{X}$ 数据集中数据维度是 p , $\mathbf{Y}$ 数据集中数据维度是 q 。CCA可简化为使模态间相关性最大的投影方向 (α, β) 学习问题

$$\begin{aligned}
 (\alpha, \beta) &= \max_{\alpha, \beta} \frac{\alpha^T \mathbf{E}[\mathbf{x}\mathbf{y}] \beta}{\sqrt{\alpha^T \mathbf{E}[\mathbf{x}^2] \alpha} \cdot \sqrt{\beta^T \mathbf{E}[\mathbf{y}^2] \beta}} = \max_{\alpha, \beta} \frac{\sum_{i=1}^n \alpha^T (\mathbf{x}_i - \bar{\mathbf{x}}) (\mathbf{y}_i - \bar{\mathbf{y}}) \beta}{\sqrt{\sum_{i=1}^n \alpha^T (\mathbf{x}_i - \bar{\mathbf{x}}) (\mathbf{x}_i - \bar{\mathbf{x}})^T \alpha} \cdot \sqrt{\sum_{i=1}^n \beta^T (\mathbf{y}_i - \bar{\mathbf{y}}) (\mathbf{y}_i - \bar{\mathbf{y}})^T \beta}} \\
 &= \max_{\alpha, \beta} \frac{\alpha^T \mathbf{X} \mathbf{Y}^T \beta}{\sqrt{\alpha^T \mathbf{X} \mathbf{X}^T \alpha} \cdot \sqrt{\beta^T \mathbf{Y} \mathbf{Y}^T \beta}}
 \end{aligned} \quad (1)$$

由于投影方向的尺度不变性, 上述投影方向的学习问题能够进一步简化如式(2)的优化模型

$$\begin{aligned}
 &\max_{\alpha, \beta} \alpha^T \mathbf{X} \mathbf{Y}^T \beta \\
 \text{s.t. } &\alpha^T \mathbf{X} \mathbf{X}^T \alpha = 1, \beta^T \mathbf{Y} \mathbf{Y}^T \beta = 1
 \end{aligned} \quad (2)$$

3 图强化典型相关分析(GECCA)

3.1 CCA的等价描述

CCA的优化函数可等价于两模态间样本距离最小, 因此可将其等价描述为

$$\begin{aligned} \min_{\alpha, \beta} & \sum_{i=1}^n \|\alpha^T(x_i - \bar{x}) - \beta^T(y_i - \bar{y})\|^2, \\ \text{s.t.} & \sum_{i=1}^n \|\alpha^T(x_i - \bar{x})\|^2 = 1, \sum_{i=1}^n \|\beta^T(y_i - \bar{y})\|^2 = 1 \end{aligned} \quad (3)$$

其中, 目标min函数可以进一步等价转化为

$$\begin{aligned} & \sum_{i=1}^n \|\alpha^T(x_i - \bar{x}) - \beta^T(y_i - \bar{y})\|^2 \\ &= \frac{1}{2n} \sum_{i=1}^n \sum_{j=1}^n \left[\alpha^T(x_i - x_j)(x_i - x_j)^T \alpha \right. \\ & \quad \left. + \beta^T(y_i - y_j)(y_i - y_j)^T \beta \right] \\ & \quad - \frac{1}{2n} \sum_{i=1}^n \sum_{j=1}^n 2\alpha^T(x_i - x_j)(y_i - y_j)^T \beta \end{aligned} \quad (4)$$

由于式(4)中 $1/2n$ 对所求投影 (α, β) 没有影响, 因此可以被忽略, 式(3)能够转化为

$$\begin{aligned} \max_{\alpha, \beta} & \sum_{i=1}^n \sum_{j=1}^n \alpha^T(x_i - x_j)(y_i - y_j)^T \beta, \\ \text{s.t.} & \sum_{i=1}^n \sum_{j=1}^n \alpha^T(x_i - x_j)(x_i - x_j)^T \alpha = 1, \\ & \sum_{i=1}^n \sum_{j=1}^n \beta^T(y_i - y_j)(y_i - y_j)^T \beta = 1 \end{aligned} \quad (5)$$

3.2 成分图权重矩阵的构建

对于给定两组对应模态数据集 $\mathbf{X}$ 和 $\mathbf{Y}$, 数据样本间通常包含不同的几何结构信息, 为了有效利用这些信息本文提出图强化的方法。

每个模态数据集在原始高维数据中通常包含多种独立数据成分, 例如一张照片中人物和背景即为两种不同数据成分, 利用原始数据直接构图的方法无法反映不同数据成分间的关系。GECCA先将上述数据集 $\mathbf{X}$ 和 $\mathbf{Y}$ 采用谱聚类方法进行分割, 将不同数据中同一成分的数据分割在一块, 分割后得到原始高维数据的多重数据成分集, 并为每种数据成分构建对应成分图, 以此保留数据间的内在联系, 使用相似准则构建成分图权重矩阵, 揭示原始数据中的复杂内在流形结构, 流程如图1所示。

对数据集 $\mathbf{X}$ 使用谱聚类方法, 可将 p 行元素分割成 H 个独立子空间, 此时数据集 $\mathbf{X} \in R^{p \times n}$ 可用成分集 $\mathbf{X} = \{\mathbf{X}^1, \mathbf{X}^2, \dots, \mathbf{X}^H\}$ ($h = 1, 2, \dots, H$)形式表示; $\mathbf{X}^h$ 用于表示数据集 $\mathbf{X}$ 的第 h 个成分的成分集, 记作 $\mathbf{X}^h = \{\mathbf{x}_1^h, \mathbf{x}_2^h, \dots, \mathbf{x}_n^h\} \in R^{j \times n}$, ($\sum_{h=1}^H j = p$); $\mathbf{x}_i^h$ ($i = 1, 2, \dots, n$)用于表示第 h 个成分集 $\mathbf{X}^h$ 中的第 i 个成分数据。同理数据集 $\mathbf{Y} \in R^{q \times n}$ 可用

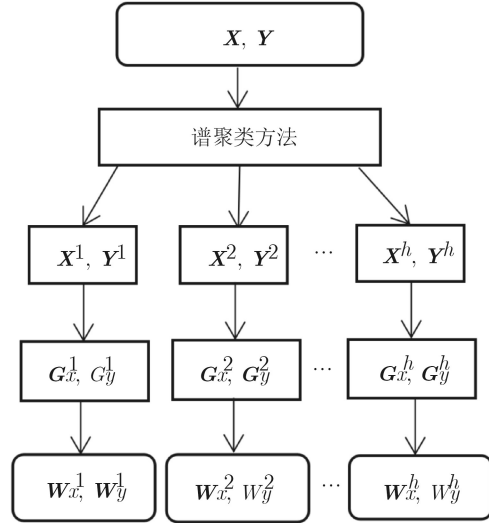


图1 成分图权重矩阵流程图

$\mathbf{Y} = \{\mathbf{Y}^1, \mathbf{Y}^2, \dots, \mathbf{Y}^H\}$ ($h = 1, 2, \dots, H$)来表示; $\mathbf{Y}^h = \{\mathbf{y}_1^h, \mathbf{y}_2^h, \dots, \mathbf{y}_n^h\} \in R^{k \times n}$, ($\sum_{h=1}^H k = q$)表示数据集 $\mathbf{Y}$ 中第 h 个成分对应的成分集; $\mathbf{y}_i^h$ ($i = 1, 2, \dots, n$)表示成分集 $\mathbf{Y}^h$ 中的第 i 个成分数据。

由数据集 $\mathbf{X}$ 和 $\mathbf{Y}$ 的成分集构建成分图, 以此保留数据间内在联系, 揭示成分集数据中的几何流形结构。以成分集 $\mathbf{X}^h$ ($h = 1, 2, \dots, H$)为例, 构建的成分图可记作 $\mathbf{G}_x^h = \{\mathbf{X}^h, \mathbf{W}_x^h\}$ ($h = 1, 2, \dots, H$), 每个成分图均为无向加权图, $\mathbf{X}^h$ 表示成分图的各个顶点, $\mathbf{W}_x^h$ 表示成分图边上的权重系数, 即成分图的权重矩阵。本文采用余弦相似准则来揭示各个成对顶点之间的相似性, 计算方法为

$$(\mathbf{W}_x^h)_{uv} = \frac{\mathbf{x}_u^h \mathbf{x}_v^h}{\|\mathbf{x}_u^h\| \cdot \|\mathbf{x}_v^h\|} \quad (6)$$

成分集 $\mathbf{Y}^h$ 的成分图构建方法与之相同, $\mathbf{Y}^h$ 的成分图记作 $\mathbf{G}_y^h = \{\mathbf{Y}^h, \mathbf{W}_y^h\}$ ($h = 1, 2, \dots, H$), 其权重矩阵计算方法为

$$(\mathbf{W}_y^h)_{uv} = \frac{\mathbf{y}_u^h \mathbf{y}_v^h}{\|\mathbf{y}_u^h\| \cdot \|\mathbf{y}_v^h\|} \quad (7)$$

式(6)和式(7)中 $\mathbf{W}_x^h, \mathbf{W}_y^h$ 表示成分集 $\mathbf{X}^h$ 和 $\mathbf{Y}^h$ 的成分图权重矩阵, $(\mathbf{W}_x^h)_{uv}$ 和 $(\mathbf{W}_y^h)_{uv}$ 表示成分图中第 (u, v) 个顶点间边的权重, $\|\theta\|$ 表示 θ 的L2范数。由式(6)和式(7)可以分别得到数据集 $\mathbf{X}$ 的 H 个成分图权重矩阵 $\mathbf{W}_x = \{\mathbf{W}_x^1, \mathbf{W}_x^2, \dots, \mathbf{W}_x^H\}$ ($h = 1, 2, \dots, H$)和数据集 $\mathbf{Y}$ 的 H 个成分图权重矩阵 $\mathbf{W}_y = \{\mathbf{W}_y^1, \mathbf{W}_y^2, \dots, \mathbf{W}_y^H\}$ ($h = 1, 2, \dots, H$)。 $\mathbf{W}_x$ 和 $\mathbf{W}_y$ 中任意成分图权重矩阵表示两两成分数据之间的对应关系, 因此具有对称性, 构成的矩阵为对称矩阵。若直接采用图强化(即将成分图直接融合为增强图),

从优化角度来看没有区分各成分的重要性，所以本文采用线性融合策略进行图强化，根据标签信息获取类系数矩阵以此作为图强化的优化系数。

3.3 类系数矩阵的构建

标签信息作为有效的监督信息可以大幅提升算法的分类效果，本文采用概率评估的方法来解析成分图中数据的分类情况。先计算出成分集 $\mathbf{X}^h$ 中 $\mathbf{x}_v^h$ 属于 l_u^h 的概率用 $P(l_u^h, \mathbf{x}_v^h)$ 表示， l_u^h 表示 $\mathbf{x}_u^h$ 的类标签，计算方法为

$$P(l_u^h, \mathbf{x}_v^h) = \frac{m_x^h}{m} \quad (8)$$

同理计算成分集 $\mathbf{Y}^h$ 中 $\mathbf{y}_v^h$ 属于 l_u^h 的概率，计算方法与式(8)类似

$$P(l_u^h, \mathbf{y}_v^h) = \frac{m_y^h}{m} \quad (9)$$

在式(8)和式(9)中 m_x^h 和 m_y^h 分别表示 $\mathbf{x}_v^h$ 和 $\mathbf{y}_v^h$ 的前 m 个权重近邻样本中属于 l_u^h 类的个数， m 为一个经验参数，通常设为 l_u^h 类中样本的数， $\mathbf{x}_v^h$ 和 $\mathbf{y}_v^h$ 取自 $\mathbf{X}^h$ 和 $\mathbf{Y}^h$ 的第 h 个成分集。权重近邻样本的确定是基于权重近邻方案，即两个样本的相似权重越大，两个样本越相似，属于一类的概率越大，采用概率评估的方法为每个成分集构建类系数矩阵，用于揭示数据中类间关系，以成分集 $\mathbf{X}^h$ 为例其类系数矩阵计算方法为

$$(O_x^h)_{uv} = \begin{cases} \gamma \left(\frac{P(l_u^h, \mathbf{x}_v^h) + P(l_v^h, \mathbf{x}_u^h)}{2} \right)^2, & l_u^h = l_v^h \\ \gamma \left(1 - \frac{P(l_u^h, \mathbf{x}_v^h) + P(l_v^h, \mathbf{x}_u^h)}{2} \right)^2, & l_u^h \neq l_v^h \end{cases} \quad (10)$$

同样成分集 $\mathbf{Y}^h$ 的类系数矩阵的计算方法为

$$(O_y^h)_{uv} = \begin{cases} \gamma \left(\frac{P(l_u^h, \mathbf{y}_v^h) + P(l_v^h, \mathbf{y}_u^h)}{2} \right)^2, & l_u^h = l_v^h \\ \gamma \left(1 - \frac{P(l_u^h, \mathbf{y}_v^h) + P(l_v^h, \mathbf{y}_u^h)}{2} \right)^2, & l_u^h \neq l_v^h \end{cases} \quad (11)$$

其中， O_x^h 和 O_y^h ($h = 1, 2, \dots, H$)分别是 $\mathbf{X}^h$ 和 $\mathbf{Y}^h$ 对应的类系数矩阵，也是图强化时与 $\mathbf{W}_x^h$ ， $\mathbf{W}_y^h$ 进行融合的优化系数。式(10)中当 l_u^h 和 l_v^h 为同一类时 $(O_x^h)_{uv}$ 为 $P(l_u^h, \mathbf{x}_v^h)$ 与 $P(l_v^h, \mathbf{x}_u^h)$ 和的均值的平方，当 l_u^h 和 l_v^h 为不同类时 $(O_x^h)_{uv}$ 为1减去 $P(l_u^h, \mathbf{x}_v^h)$ 与 $P(l_v^h, \mathbf{x}_u^h)$ 和的均值的平方，其中 γ 是 $\sum_{h=1}^H (O^h)_{uv} = 1$ 的权重平衡因子，用于确保类系数矩阵中任意行列的和为1，和成分图权重矩阵一样，类系数矩阵也为对称方阵。式(11)中符号含义与式(10)中含义相同，由此获得数据集 $\mathbf{X}$ ， $\mathbf{Y}$ 的 H 个成分图对应类系数矩阵。

3.4 GECCA的优化模型

由式(10)和式(11)得到数据集 $\mathbf{X}$ 和 $\mathbf{Y}$ 的成分图类系数矩阵，结合式(6)、式(7)所得的成分图权重矩阵可进行图强化，得到强化图和对应该强化矩阵 $\mathbf{M}$ ，图强化具体为在类系数矩阵的指导下将成分图权重矩阵进行线性融合得到强化矩阵，并以此作为强化图各边的权重，具体计算方法为

$$\mathbf{M}_x = \mathbf{O}_x^1 \otimes \mathbf{W}_x^1 + \mathbf{O}_x^2 \otimes \mathbf{W}_x^2 + \dots + \mathbf{O}_x^H \otimes \mathbf{W}_x^H \quad (12)$$

$$\mathbf{M}_y = \mathbf{O}_y^1 \otimes \mathbf{W}_y^1 + \mathbf{O}_y^2 \otimes \mathbf{W}_y^2 + \dots + \mathbf{O}_y^H \otimes \mathbf{W}_y^H \quad (13)$$

符号 $\otimes$ 表示两矩阵中对应下标的元素相乘，例如 $(\mathbf{A} \otimes \mathbf{B}) = \mathbf{A}_{ij} \mathbf{B}_{ij}$ ，矩阵 $\mathbf{A}$ 和 $\mathbf{B}$ 为同型矩阵；其中 $h = 1, 2, \dots, H$ 。从式(12)和式(13)中可知强化矩阵 $\mathbf{M}$ 仍为对称矩阵，此时将包含成分图全部几何流形和监督信息的GECCA模型和约束条件写作

$$\begin{aligned} \max_{\alpha, \beta} \quad & \sum_{i=1}^n \sum_{j=1}^n \alpha^T (\mathbf{M}_x)_{ij} (\mathbf{x}_i - \mathbf{x}_j) (\mathbf{M}_y)_{ij} (\mathbf{y}_i - \mathbf{y}_j)^T \beta, \\ \text{s.t.} \quad & \sum_{i=1}^n \sum_{j=1}^n \alpha^T (\mathbf{M}_x)_{ij} (\mathbf{x}_i - \mathbf{x}_j) (\mathbf{M}_x)_{ij} (\mathbf{x}_i - \mathbf{x}_j)^T \alpha = 1, \\ & \sum_{i=1}^n \sum_{j=1}^n \beta^T (\mathbf{M}_y)_{ij} (\mathbf{y}_i - \mathbf{y}_j) (\mathbf{M}_y)_{ij} (\mathbf{y}_i - \mathbf{y}_j)^T \beta = 1 \end{aligned} \quad (14)$$

为优化模型，可将式(14)在代数运算后重写为(忽略一般性常数)

$$\begin{aligned} \max_{\alpha, \beta} \quad & \alpha^T \mathbf{X} \mathbf{L}_{xy} \mathbf{Y}^T \beta, \\ \text{s.t.} \quad & \alpha^T \mathbf{X} \mathbf{L}_{xx} \mathbf{X}^T \alpha = 1, \\ & \beta^T \mathbf{Y} \mathbf{L}_{yy} \mathbf{Y}^T \beta = 1 \end{aligned} \quad (15)$$

其中， $\mathbf{X}$ 和 $\mathbf{Y}$ 表示均值化后的样本数据， (α, β) 表示需要求得的相关投影方向，其中 $\mathbf{L}$ 表示类拉普拉斯矩阵，计算方法为

$$\left. \begin{aligned} \mathbf{L}_{xx} &= \mathbf{D}_{xx} - \mathbf{M}_x \otimes \mathbf{M}_x \\ \mathbf{L}_{yy} &= \mathbf{D}_{yy} - \mathbf{M}_y \otimes \mathbf{M}_y \\ \mathbf{L}_{xy} &= \mathbf{D}_{xy} - \mathbf{M}_x \otimes \mathbf{M}_y \end{aligned} \right\} \quad (16)$$

其中， $\mathbf{D}_{xx}$ ($\mathbf{D}_{yy}$, $\mathbf{D}_{xy}$)表示 $n \times n$ 的对角矩阵，其对角线上第 i 位的值为 $\mathbf{M}_x \otimes \mathbf{M}_x$ ($\mathbf{M}_y \otimes \mathbf{M}_y$, $\mathbf{M}_x \otimes \mathbf{M}_y$)第 i 行 (或第 i 列，由于对称) 值的和。若假设 $\mathbf{S} = \mathbf{M}_x \otimes \mathbf{M}_x$, $\mathbf{S} \in R^{n \times n}$, $\mathbf{D}_{xx} = \text{diag}(\sum_{i=1}^n \mathbf{S}_{i1}, \dots, \sum_{i=1}^n \mathbf{S}_{in}) = \text{diag}(\sum_{j=1}^n \mathbf{S}_{1j}, \dots, \sum_{j=1}^n \mathbf{S}_{nj})$ ， $\mathbf{D}_{yy}$ 和 $\mathbf{D}_{xy}$ 的计算方法与之类似。由于矩阵 $\mathbf{M}_x$ 和 $\mathbf{M}_y$ ， $\mathbf{L}_{xx}$ ， $\mathbf{L}_{yy}$ ， $\mathbf{L}_{xy}$ 均为对称矩阵，因此可以得到 $\mathbf{L}_{yx} = \mathbf{L}_{xy}$ 且 $\mathbf{L}_{yx}$ 也为对称矩阵。

3.5 GECCA的求解

对式(15)中模型进行求解，首先为该模型构建拉格朗日乘子函数

$$F(\alpha, \beta) = \alpha^T X L_{xy} Y^T \beta - \frac{\lambda_1}{2} (\alpha^T X L_{xx} X^T \alpha - 1) - \frac{\lambda_2}{2} (\beta^T Y L_{yy} Y^T \beta - 1) \quad (17)$$

其中, λ_1 和 λ_2 为拉格朗日函数的乘子, 对上述函数 $F(\alpha, \beta)$ 的 α 和 β 分别求偏导, 令结果为0, 可得到

$$\left. \begin{aligned} \frac{\partial F(\alpha, \beta)}{\partial \alpha} &= X L_{xy} Y^T \beta - \lambda_1 X L_{xx} X^T \alpha = 0 \\ \frac{\partial F(\alpha, \beta)}{\partial \beta} &= Y L_{yx} X^T \alpha - \lambda_2 Y L_{yy} Y^T \beta = 0 \end{aligned} \right\} \quad (18)$$

由式(15)中约束条件, 可得

$$\left. \begin{aligned} \alpha^T X L_{xy} Y^T \beta &= (\alpha^T X L_{xy} Y^T \beta)^T = \beta^T Y L_{yx} X^T \alpha \\ \lambda_1 \alpha^T X L_{xx} X^T \alpha &= \lambda_2 \beta^T Y L_{yy} Y^T \beta \end{aligned} \right\} \quad (19)$$

由(19)得到 $\lambda_1 = \lambda_2$, 可令 $\lambda_1 = \lambda_2 = \lambda$, 此时可将优化模型的求解转化为如下广义特征值分解问题

$$\begin{bmatrix} X L_{xy} Y^T \beta \\ Y L_{yx} X^T \alpha \end{bmatrix} = \lambda \begin{bmatrix} X L_{xx} X^T \alpha \\ Y L_{yy} Y^T \beta \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \quad (20)$$

对式(20)特征值分解问题进行求解, 得到特征值 λ 和其对应特征向量 $[\alpha^T, \beta^T]^T$, 对求得特征值 λ 进行降序, 取前 d 个最大特征值 λ_i ($i = 1, 2, \dots, d$)对应的特征向量 $\{[\alpha_i^T, \beta_i^T]^T \in \mathbf{R}^{1 \times (p+q)}\}_{i=1}^d$ 作为投影方向, 由此得到数据集 X 和 Y 的相关投影空间 $A = \{\alpha_1, \alpha_2, \dots, \alpha_d\} \in \mathbf{R}^{p \times d}$ 和 $B = \{\beta_1, \beta_2, \dots, \beta_d\} \in \mathbf{R}^{q \times d}$ 。GECCA的计算步骤如表1所示。

3.6 特征融合方法

利用求得的相关投影矩阵 A 和 B 可得到数据集 X 和 Y 对应的低维相关特征训练集 $A^T X \in \mathbf{R}^{d \times n}$ 和 $B^T Y \in \mathbf{R}^{d \times n}$ 。本文采用并行特征融合策略完成特征融合

$$Z = \begin{bmatrix} A \\ B \end{bmatrix}^T \begin{bmatrix} X \\ Y \end{bmatrix} \quad (21)$$

融合后获得的训练集在子空间低维融合特征集可记作 $Z = \{Z_1, Z_2, \dots, Z_n\} \in \mathbf{R}^{d \times n}$ 。采用同样方法可得到测试集低维融合特征集。

4 实验部分

为评估GECCA的图像识别性能, 分别在人脸数据集XM2VTS和手写体数字数据集Semeion上设计了针对性实验。利用模态策略^[19]获得上述数据集的模态数据, 具体为采用Coiflets和Daubechies两种小波变换得到每个图像的低频子图, 为了弱化小样本问题, 可用PCA将低频子图约简到100维, 以此作为每幅图形的两种模态数据。在实验中将GECCA算法与DCCA^[16], CCA^[7], LPCCA^[10], ALPCCA^[12]算法作对比分析, 并且所有算法的最终识别率均利用基于欧氏距离的最近邻分类器来获得。

4.1 XM2VTS人脸图像数据集上的实验

XM2VTS人脸数据集包含来自295名志愿者每人4个会话中的8幅图像, 图2为XM2VTS数据集中部分人脸图片。本实验从每类样本中进行10次随机抽取, 每次抽取 t ($t=3, 4, 5, 6$)个样本作为训练样本, 其余样本作为本次实验的测试样本。图3展示了在人脸数据集上识别率随维度变化情况。

CCA没有考虑数据间内在几何关系和类标签信息, 只是将投影后两模态间的相关性最大化, 无法掌握数据中非线性几何结构, 学习的低维特征在图像识别中鉴别力不足。LPCCA和ALPCCA都在一定程度上考虑了高维数据中的内在联系, 在相关子空间中尽可能保留数据的几何结构; 但LPCCA直接利用原始的高维数据揭示的局部近邻结构, 会受到高维数据中包含的噪声和冗余信息的影响, 保留的局部近邻结构容易失真, 所以识别效果较差; ALPCCA也使用原始高维数据, 与LPCCA不同, 采用的是两模态整体近邻结构的等权线性表示, 因此识别率要高于LPCCA。DCCA在相关分析框架中加入了监督信息, 对高维数据进行约束, 具有较高识别率。

表1 GECCA的算法步骤

算法: 图强化典型相关分析
(1) 输入: $\{X, Y\}, C_x$ 和 C_y 。
(2) 通过式(12)和式(13)构建图强化矩阵 M_x 和 M_y 。利用式(16)构建类拉普拉斯矩阵 L_x, L_y, L_{xy} 和 L_{yx} 。
(3) 构建拉格朗日乘子函数: $F(\alpha, \beta) = \alpha^T X L_{xy} Y^T \beta - \frac{\lambda_1}{2} (\alpha^T X L_{xx} X^T \alpha - 1) - \frac{\lambda_2}{2} (\beta^T Y L_{yy} Y^T \beta - 1)$ 。
(4) 按照式(20)求出特征值 λ 和对应特征向量 (α, β) 。
(5) 取前 d 个特征值对应特征向量 $\{[\alpha_i^T, \beta_i^T]^T \in \mathbf{R}^{1 \times (p+q)}\}_{i=1}^d$ 。
(6) 输出: $A = \{\alpha_1, \alpha_2, \dots, \alpha_d\} \in \mathbf{R}^{p \times d}$ 和 $B = \{\beta_1, \beta_2, \dots, \beta_d\} \in \mathbf{R}^{q \times d}$ 。



图2 XM2VTS部分人脸图像

GECCA利用成分图的方式揭示数据中各成分之间的联系,用成分图权重矩阵保留数据的几何流形,并借助类系数矩阵优化成分图权重矩阵,使得低维相关特征更具鉴别力。从图3中不难看出,在不同个数训练样本下GECCA均展示出最佳识别率,且识别率在较低维度时增幅最快说明GECCA提取的特征鉴别能力更强。

4.2 在Semeion手写体数字数据集上的实验

Semeion手写体数字数据集包含大约80个人的1593个手写体数字,每人写出由0到9的所有数字。该数据集中每类有大量数据样本,因此本实验从每类中选择 $t(t=40, 60, 80, 100)$ 个样本用于训练,剩余样本用作实验测试,并进行10次随机性测试。图4为每次随机实验各算法的最佳识别率立体图,表2

记录各算法平均最佳识别率及标准差。

CCA是一种线性方法,无法发现隐藏在高维数据中的非线性几何结构,因此会影响识别性能。LPCCA是将全局非线性转化为局部线性,解决姿态估计中的非线性问题,但鉴别能力受原始高维数据中的噪声和冗余信息的影响,表现出了较差的识别性能。ALPCCA对LPCCA邻域信息的利用加以改进,使之鉴别能力得以提升。类标签作为一种重要的监督信息,可提升算法的识别和分类性能,但CCA, LPCCA和ALPCCA均未用到标签信息,因此识别率均低于包含类标签约束的DCCA。

GECCA通过图强化方法将成分图权重矩阵和类系数矩阵线性融合,并将其嵌入到相关特征学习的理论框架中,使其在保留成分数据间关系的同时具有鉴别效果;此外,GECCA对模态内样本的散布进行约束,使得同类样本间的距离更近,分类性更强。如图4所示,不难发现在训练样本情况不同时,10次随机实验中GECCA均表现出最佳鉴别能力,这表明GECCA的识别效果优于其他算法。在表2中,GECCA具有最高的平均识别率和更小的标准差,说明GECCA在相同训练样本下更加稳定,有更小的误差,实验对随机样本有更好的鲁棒性。

5 结束语

CCA作为一种线性多模态特征学习方法,无

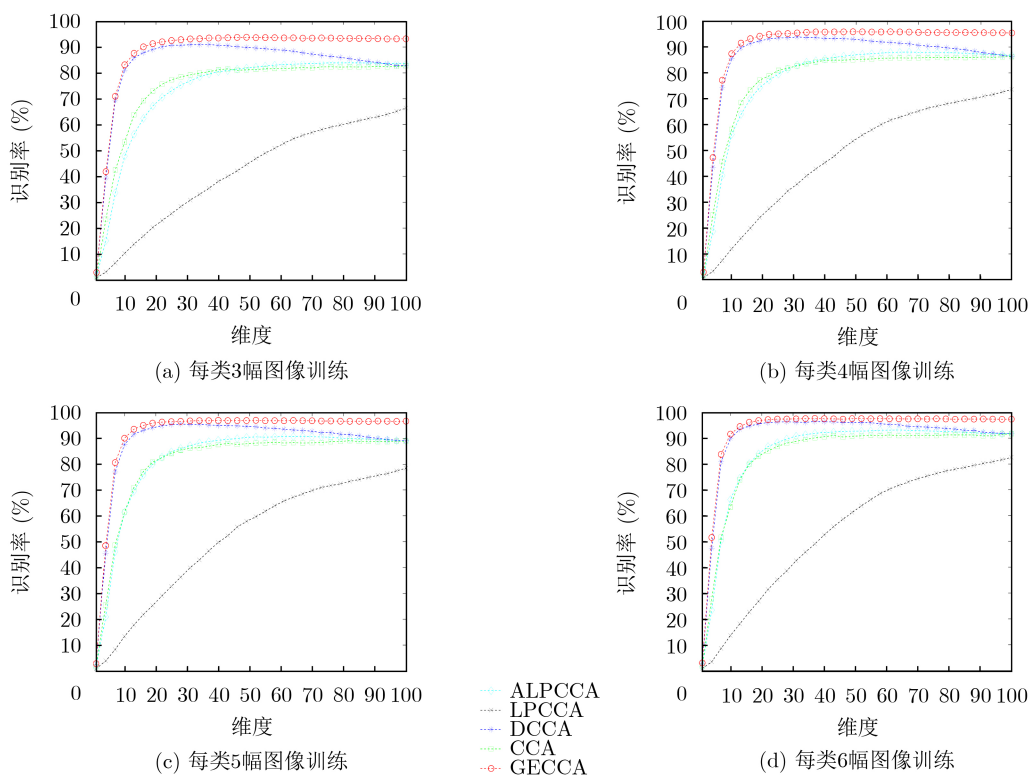


图3 在XM2VTS人脸数据集上识别率随维度变化情况

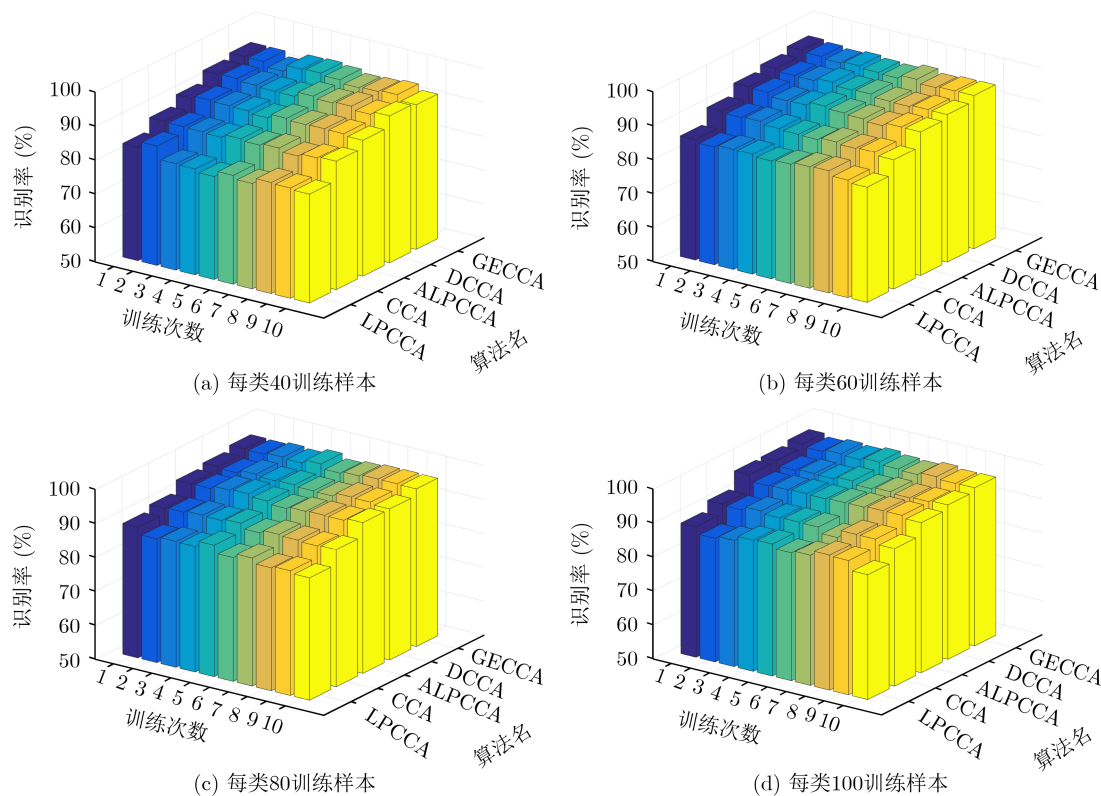


图4 在Semeion数据集中每次随机实验的最佳识别率

表2 在Semeion手写体数字数据集上的识别率及标准差

	40训练样本	60训练样本	80训练样本	100训练样本
GECCA	86.93±0.92	88.89±0.98	90.69±0.77	91.38±0.70
DCCA	85.72±1.02	87.57±1.04	89.14±0.82	80.25±0.91
ALPCCA	80.82±0.75	84.76±1.19	87.92±0.77	89.38±1.03
LPCCA	69.84±2.24	74.83±1.51	77.33±1.47	80.56±2.10
CCA	74.20±0.84	77.93±1.19	80.92±0.85	82.34±1.44

$A \pm B$: A 表示平局识别率(%), B 表示标准差

法解决高维数据的非线性问题。而基于局部领域的特征提取方法的几何结构往往会因噪声和冗余信息存在失真现象,从而影响低维特征的分类分离性。在原始高维数据中通常包含多种独立的数据成分,仅使用一个图难以反映不同数据成分之间的差异。为此,本文提出了一种新的GECCA算法,该方法考虑了原始数据中的几何信息和监督信息。采用成分图权重矩阵揭示数据内在本质几何流形;以类系数矩阵方式增加数据的可分类性;以此为基础,通过图强化将成分图权重矩阵和类系数矩阵进行融合,并嵌入到相关特征提取框架中,减小直接使用原始数据而保留噪声的影响,同时增强算法的识别和分类性能。在XM2VTS人脸数据集和Semeion手写体数字数据集上进行实验,良好的实验结果显示GECCA是一种有效的图像识别算法。

本文为未来的研究开辟了几个有趣的方向。将

两模态数据融合扩展至3个或3个以上模态值得进一步探究。将全监督改为半监督也具有十分重要的现实意义。

参考文献

- [1] LIU Zhonghua, LAI Zhihui, OU Weihua, *et al.* Structured optimal graph based sparse feature extraction for semi-supervised learning[J]. *Signal Processing*, 2020, 170: 107456. doi: [10.1016/j.sigpro.2020.107456](https://doi.org/10.1016/j.sigpro.2020.107456).
- [2] 王肖锋, 孙明月, 葛为民. 基于图像协方差无关的增量特征提取方法研究[J]. *电子与信息学报*, 2019, 41(11): 2768–2776. doi: [10.11999/JEIT181138](https://doi.org/10.11999/JEIT181138).
WANG Xiaofeng, SUN Mingyue, and GE Weimin. An incremental feature extraction method without estimating image covariance matrix[J]. *Journal of Electronics & Information Technology*, 2019, 41(11): 2768–2776. doi: [10.11999/JEIT181138](https://doi.org/10.11999/JEIT181138).

- [3] YI Shuangyan, LAI Zhihui, HE Zhenyu, *et al.* Joint sparse principal component analysis[J]. *Pattern Recognition*, 2017, 61: 524–536. doi: [10.1016/j.patcog.2016.08.025](https://doi.org/10.1016/j.patcog.2016.08.025).
- [4] BRO R and SMILDE A K. Principal component analysis[J]. *Analytical Methods*, 2014, 6(9): 2812–2831. doi: [10.1039/C3AY41907J](https://doi.org/10.1039/C3AY41907J).
- [5] 刘政怡, 段群涛, 石松, 等. 基于多模态特征融合监督的RGB-D图像显著性检测[J]. 电子与信息学报, 2020, 42(4): 997–1004. doi: [10.1199/JEIT190297](https://doi.org/10.1199/JEIT190297).
LIU Zhengyi, DUAN Quntao, SHI Song, *et al.* RGB-D image saliency detection based on multi-modal feature-fused supervision[J]. *Journal of Electronics & Information Technology*, 2020, 42(4): 997–1004. doi: [10.1199/JEIT190297](https://doi.org/10.1199/JEIT190297).
- [6] 姜庆超, 颜学峰. 基于局部-整体相关特征的多单元化工过程分层监测[J]. 自动化学报, 2020, 46(9): 1770–1782.
JIANG Qingchao and YAN Xuefeng. Hierarchical monitoring for multi-unit chemical processes based on local-global correlation features[J]. *Acta Automatica Sinica*, 2020, 46(9): 1770–1782.
- [7] YANG Xinghao, LIU Weifeng, LIU Wei, *et al.* A survey on canonical correlation analysis[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2021, 33(6): 2349–2368. doi: [10.1109/TKDE.2019.2958342](https://doi.org/10.1109/TKDE.2019.2958342).
- [8] YANG Xinghao, LIU Weifeng, TAO Dapeng, *et al.* Canonical correlation analysis networks for two-view image recognition[J]. *Information Sciences*, 2017, 385/386: 338–352. doi: [10.1016/j.ins.2017.01.011](https://doi.org/10.1016/j.ins.2017.01.011).
- [9] SHEN Xiaobo, SUN Quansen, and YUAN Yunhao. Orthogonal canonical correlation analysis and its application in feature fusion[C]. The 16th International Conference on Information Fusion, Istanbul, Turkey, 2013: 151–157.
- [10] EL MADANY N E D, HE Yifeng, and GUAN Ling. Human action recognition by fusing deep features with globality locality preserving canonical correlation analysis[C]. 2017 IEEE International Conference on Image Processing, Beijing, China, 2017: 2871–2875.
- [11] ELMADANY N E D, HE Yifeng, and GUAN Ling. Information fusion for human action recognition via biset/multiset globality locality preserving canonical correlation analysis[J]. *IEEE Transactions on Image Processing*, 2018, 27(11): 5275–5287. doi: [10.1109/TIP.2018.2855438](https://doi.org/10.1109/TIP.2018.2855438).
- [12] WANG Fengshan and ZHANG Daoqiang. A new locality-preserving canonical correlation analysis algorithm for multi-view dimensionality reduction[J]. *Neural Processing Letters*, 2013, 37(2): 135–146. doi: [10.1007/s11063-012-9238-9](https://doi.org/10.1007/s11063-012-9238-9).
- [13] GAO Xizhan, NIU Sujie, and SUN Quansen. Two-directional two-dimensional kernel canonical correlation analysis[J]. *IEEE Signal Processing Letters*, 2019, 26(11): 1578–1582. doi: [10.1109/LSP.2019.2939986](https://doi.org/10.1109/LSP.2019.2939986).
- [14] MELZER T, REITER M, and BISCHOF H. Appearance models based on kernel canonical correlation analysis[J]. *Pattern Recognition*, 2003, 36(9): 1961–1971. doi: [10.1016/S0031-3203\(03\)00058-X](https://doi.org/10.1016/S0031-3203(03)00058-X).
- [15] SUN Tingkai, CHEN Songcan, YANG Jingyu, *et al.* A supervised combined feature extraction method for recognition[C]. The IEEE International Conference on Data Mining, Pisa, Italy, 2008: 1043–1048.
- [16] GAO Lei, QI Lin, CHEN Enqing, *et al.* Discriminative multiple canonical correlation analysis for information fusion[J]. *IEEE Transactions on Image Processing*, 2018, 27(4): 1951–1965. doi: [10.1109/TIP.2017.2765820](https://doi.org/10.1109/TIP.2017.2765820).
- [17] CHEN Jia, WANG Gang, and GIANNAKIS G B. Graph multiview canonical correlation analysis[J]. *IEEE Transactions on Signal Processing*, 2019, 67(11): 2826–2838. doi: [10.1109/TSP.2019.2910475](https://doi.org/10.1109/TSP.2019.2910475).
- [18] CHEN Jia, WANG Gang, SHEN Yanning, *et al.* Canonical correlation analysis of datasets with a common source graph[J]. *IEEE Transactions on Signal Processing*, 2018, 66(16): 4398–4408. doi: [10.1109/TSP.2018.2853130](https://doi.org/10.1109/TSP.2018.2853130).
- [19] SU Shuzhi, FANG Xianjin, YANG Gaoming, *et al.* Self-balanced multi-view orthogonality correlation analysis for image feature learning[J]. *Infrared Physics & Technology*, 2019, 100: 44–51.

苏树智: 男, 1987年生, 副教授, 研究方向为多模态模式识别。

谢 军: 男, 1996年生, 硕士生, 研究方向为模式识别。

平昕瑞: 男, 1994年生, 硕士生, 研究方向为多模态信息融合、数据处理。

高鹏连: 男, 1996年生, 硕士生, 研究方向为模式识别。

责任编辑: 余 蓉