

DOI: 10.13382/j.jemi.B2407476

任意三角形结构 2DPCA 在水下光学图像识别中的应用*

陈璇 毕鹏飞 胡志远

(南京信息工程大学人工智能学院 南京 210044)

摘要:在实际应用中,受观测条件和采集场景等诸多因素的综合作用,水下光学图像通常呈现出高维小样本的特性,与此同时,这类图像还极易伴随着各类噪声信息的干扰。导致许多降维方法在其识别过程中的鲁棒表现力不足。为解决上述问题,提出一种创新的任意三角形结构二维主成分分析方法(ATS-2DPCA)应用于水下光视觉图像识别。该方法在构建过程中,充分考虑了投影数据的重构误差和方差两者之间的关系,在此基础上成功匹配到了灵活的鲁棒距离度量机制。通过这种方式,能够切实有效地提升在面临噪声干扰时水下光学图像数据的识别精度,并且实现对于数据几何结构的合理保护。从理论层面证明了该方法的可用性和收敛性。同时,选取了3个水下光学图像数据库进行了实验验证,得出的最优识别精度分别为:89.07%、88.52%、86.00%。一系列实验结果有力地表明,ATS-2DPCA在同类方法中展现出了更为卓越的性能表现。

关键词: 二维主成分分析;任意三角形结构;鲁棒距离度量;水下光学图像识别;降维

中图分类号: TP391.41;TN06

文献标识码: A

国家标准学科分类代码: 520.2040

Arbitrary triangle structure 2DPCA and its application to underwater optical image recognition

Chen Xuan Bi Pengfei Hu Zhiyuan

(School of Artificial Intelligence, Nanjing University of Information Science and Technology, Nanjing 210044, China)

Abstract: Influenced by factors such as observation conditions and acquisition scenarios, underwater optical image data usually presents the characteristics of high-dimensional small samples and is easily accompanied with noise interference. Resulting in insufficient robust performance of many dimensionality reduction methods in their recognition process. To solve this problem, we propose a novel 2DPCA method for underwater optical image recognition, called arbitrary triangle structure 2DPCA (ATS-2DPCA). On the basis of considering the relationship between reconstruction error and variance of projection data, ATS-2DPCA can successfully match the flexible robust distance metric mechanism, which effectively improves the accuracy of underwater optical image recognition under noise interference environment and achieves reasonable protection of the geometric structure of the data. In this paper, we theoretically prove the availability and convergence of the proposed method and use three underwater optical image databases for experimental verification. The optimal recognition accuracy is 89.07%, 88.52%, and 86.00%, respectively. The extensive experimental results show that ATS-2DPCA has more outstanding performance than other 2DPCA-based methods.

Keywords: two-dimensional principal component analysis (2DPCA); arbitrary triangle structure; robust distance metric; underwater optical image recognition; dimensionality reduction

0 引言

当前,随着光学技术和采集设备的快速发展,面向计算机视觉和模式识别领域的应用范围已逐步从大气环境

向水下光视觉进行拓展^[1]。然而,由于水体作业环境的复杂多变,使得获取到的图像数据一方面呈现出高维形态(属于图像类数据的共性),另一方面具有小样本特性且易伴随噪声信息的干扰(属于水下光学图像数据的固有属性)。因此,如何找到一种有效的方法以提高水下光

学图像识别精度已成为一个重要的研究课题。近二三十年来,子空间学习技术因其模型结构简单、可解释性强及泛化能力好等特点,已在小样本图像数据的特征描述中获得成功应用,并逐步成为了图像识别的主流方法之一。在基于子空间学习技术的众多方法中,主成分分析^[2](principal component analysis, PCA),线性判别分析^[3](linear discriminant analysis, LDA),局部保持投影^[4](locality preserving projection, LPP)是 3 种最为经典的方法。PCA 用于提取图像数据中最具表现力的特征信息,LDA 寻找嵌入在图像数据中最具辨别力的特征信息,LPP 试图挖掘图像数据之间的局部几何结构信息。

在应用上述方法对图像进行描述之前,需要预先将二维(2D)图像矩阵转化为一维(1D)向量数据,然而,此操作不仅会增加投影向量的计算成本,还可能导致像素间隐含的空间结构信息丢失。为了克服这个困难,刘玉珍等^[5]提出了典范的二维主成分分析方法(2DPCA),受到 2DPCA 在图像特征提取方面获得成功的启发,许多基于矩阵的子空间学习方法相继提出,包括二维线性判别分析^[6](2DLDA),二维局部保持投影^[7](2DLPP)等。尽管上述方法利用二维矩阵信息能够直接提取图像的主要特征,但其均使用欧式距离的平方作为数据之间相似性的度量。众所周知,大距离度量对噪声干扰不具有鲁棒性,容易导致实际投影方向与期望解之间出现显著偏差。因此,面向二维子空间学习方法的鲁棒性展开研究变得十分重要。

目前,采取鲁棒范数度量策略被认为是有效提升二维子空间学习方法性能的重要技术手段,且大多考虑在经典的 2DPCA 中率先进行应用。例如 Wang 等^[8]利用 l_1 范数提出基于非贪婪算法的 2DPCA 方法(2DPCA based on l_1 -norm, 2DPCA-L1)。Wang 等^[9]在 2DPCA-L1 鲁棒性的基础上,引入稀疏诱导回归策略,进而提出具有稀疏约束的 2DPCA 方法(2DPCA-L1 with sparsity, 2DPCA-L1-S)。Wang 等^[10]考虑将 l_1 范数的使用广度进行延拓,提出使用 l_p 范数度量的广义 2DPCA 方法(generalized 2-D principal component analysis, G2DPCA)。然而,在上述方法中,其距离度量机制选用的是非欧式空间范数度量(l_1 范数),由于其不具有旋转不变性,导致获取到的解与保护数据几何结构的协方差矩阵无关^[11-13]。为了解决该问题,许多选用欧式空间的鲁棒范数度量方法得以相继发展,比较经典的有 Li 等^[14]在目标函数中使用 F 范数替代 F 范数的平方策略提出方法(F -norm 2DPCA),Gao 等^[15]提出 R_1 -2DPCA 用于提取图像关键特征,考虑通过 R_1 范数最大化投影数据的方差。Zhou 等^[16]将中心化权重矩阵作用到投影数据中并提出广义中心 2DPCA(generalized centered 2DPCA, GC-2DPCA),有效保护了数据的结构信息。尽管这些方法在度量机制属性比 l_1 范

数更具优势,但其在结构上无法同时满足最小化重构误差和最大化方差的双重诉求。因此,结构体系合理化建设将进一步提升方法的鲁棒性。自然地,学者们会将重构误差与方差之间可能存在的关联纳入到考虑范围之内,得到更具鲁棒性的体系结构,如 Gao 等^[17]提出的角度 2DPCA(Angle-2DPCA),Wang 等^[18]在此基础之上通过使用可调参数提出的余弦 2DPCA(Cos-2DPCA),其本质均是基于直角三角形体系结构,求解到可同时满足最小化重构误差和最大化方差的最优解。然而,该方法在取得自身结构优势的前提下,间接制约了鲁棒度量模式的可选择性,在一定程度上削弱了噪声环境中对于主要特征准确描述的能力。因此,现存具有鲁棒属性的 2DPCA 方法未能在真正意义上有效地抑制噪声干扰。鉴于上述改进方法在鲁棒表现力方面的不足,决定提出 ATS-2DPCA 方法,并将其在易受噪声干扰的水下光学图像识别中进行应用。该方法主要贡献包括如下几个方面:

1) ATS-2DPCA 的模型结构具有出色的鲁棒性,其在确保输入数据,重构误差以及方差能够形成三角形的前提下,不仅合理构建了重构误差与方差之间的关系,还成功突破了距离度量机制选择的局限性。

2) ATS-2DPCA 考虑灵活的 $l_{2,p}$ 范数作为衡量数据之间相似性的策略,进一步提升对于噪声干扰鲁棒性的同时,兼顾了欧式范数在距离度量方面的优势。

3) 设计一种有效的非贪婪迭代算法以获取 ATS-2DPCA 的最优解,使其能够快速收敛至局部最优解。

在此基础上,3 个水下光学图像数据库中的大量实验结果也有力佐证了所提出方法的优异性能。因此,提出的方法具有重要的理论研究价值和广泛的实际应用前景。

1 相关工作

通过构建投影数据的方差及其重构误差之间的关系,提出了 ATS-2DPCA,该方法的结构建立具有良好的鲁棒性和普适性,能够匹配到更为灵活的距离度量模式,行之有效地突破了基于 Angle-2DPCA 模型体系自身的局限性。与此同时,针对模型的优化求解,还设计了一种具有非贪婪属性的迭代算法。

假设有 N 幅 $m \times n$ 维训练样本图像数据集 $\mathbf{A}_i \in \mathbf{R}^{m \times n}$ ($i = 1, 2, \dots, N$), \mathbf{A}_i 表示第 i 个训练样本。通常来说,需要预先完成对于输入数据的中心化处理任务,即 $\sum_{i=1}^N \mathbf{x}_i - \mathbf{m} = \mathbf{0}$, 其中 $\mathbf{m} = 1/N \sum_{i=1}^N \mathbf{x}_i$ 是所有训练样本的均值数据。基于 l_1 范数的 2DPCA 思想本质上是通过求解以下优化模型来找到最优投影矩阵:

$$\operatorname{argmax}_{\mathbf{V}^T \mathbf{V} = \mathbf{I}_k} \sum_{i=1}^N \|\mathbf{A}_i \mathbf{V}\|_{l_1} \quad (1)$$

其中, $\|\cdot\|_{l_1}$ 和 $\mathbf{I}_k \in \mathbf{R}^{k \times k}$ 分别表示矩阵的 l_1 范数运算和 k 维单位矩阵。

相比于经典的 2DPCA 方法, 2DPCA-L1 在距离度量方案的选择中表现出更好的鲁棒性, 但其同样也存在一些不足。首先, l_1 范数的使用不具备旋转不变性这一良好属性, 并且无法保证所获取到的解与保护数据几何结构的协方差矩阵有关。其次, 2DPCA-L1 在执行优化任务过程中的计算成本相对较高。最重要的是, 优化模型 (1) 没有充分考虑投影数据的方差和相应的重建误差之间的关系, 这直接削弱了对噪声干扰的有效抑制。

为了克服 2DPCA-L1 存在的不足, 并进一步提高基于 2DPCA 模型的鲁棒性, Gao 等^[17] 提出了 Angle-2DPCA, 相应的优化模型可表示为:

$$\operatorname{argmin}_{\mathbf{V}^T \mathbf{V} = \mathbf{I}_k} \sum_{i=1}^N \frac{\|\mathbf{A}_i - \mathbf{A}_i \mathbf{V} \mathbf{V}^T\|_F}{\|\mathbf{A}_i \mathbf{V}\|_F} \quad (2)$$

其中, $\|\cdot\|_F$ 指出一个矩阵的 F 范数运算。

在整体结构建立以及距离度量方式的使用方面, Angle-2DPCA 表现出比 2DPCA-L1 更好的鲁棒性。然而, 其在一定程度上仍具有明显的局限性。首先, Angle-2DPCA 所建立的直角三角形结构 (如图 1 所示) 束缚了鲁棒度量机制的选择广度。其次, 选用 F 范数作为距离度量准则缺少灵活性, 没能在真正意义上实现最小化重构误差的需求, 严重影响了模型的性能。最后, Angle-2DPCA 在求解过程中涉及更多的中间变量和除法运算, 从而增加了存储空间的需求。

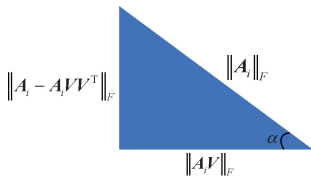


图 1 直角三角形结构

Fig. 1 The right triangle structure

鉴于传统鲁棒 2DPCA 方法在表现力方面存在的不足, 提出了创新的任意三角形结构 2DPCA 方法 (ATS-2DPCA), 其目标函数可表示为:

$$\operatorname{argmax}_{\mathbf{V}^T \mathbf{V} = \mathbf{I}_k} \sum_{i=1}^N \sum_{j=1}^m \|\mathbf{A}_i(j, :) \mathbf{V}\|_2^p - \|\mathbf{A}_i(j, :) - \mathbf{A}_i(j, :) \mathbf{V} \mathbf{V}^T\|_2^p \quad (3)$$

其中, $\mathbf{A}_i(j, :)$ 指出矩阵 \mathbf{A}_i 的第 j 行元素。

目标函数式 (3) 的构建是考虑在非单一鲁棒范数作为距离度量条件下 (如 l_1 范数和 F 范数等), 投影数据的方差, 重构误差以及输入数据三者均能够形成任意三角

型结构体系 (如图 2 所示)。基于此想法, 选用灵活的 $l_{2,p}$ 范数作为距离度量, 科学构建了投影数据方差与其重构误差之间的关系, 同时切实提升了模型在噪声环境中整体结构的鲁棒性。此外, 针对提出模型的求解问题, 还设计了一种快速的非贪婪迭代算法, 使其在每次迭代过程中均能获得封闭形式的解。最后, 与同类算法相比, 该算法在显著提高识别精度的基础上, 并未增加计算复杂度。

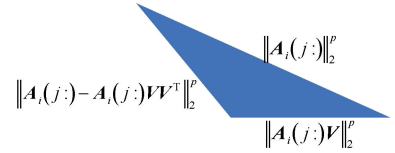


图 2 任意三角形结构

Fig. 2 The arbitrary triangle structure

2 算法求解

本节给出 ATS-2DPCA 获取最优投影矩阵 \mathbf{V} 的求解过程。基于线性代数相关知识, 可得:

$$\begin{aligned} & \sum_{i=1}^N \sum_{j=1}^m \frac{\|\mathbf{A}_i(j, :) \mathbf{V}\|_2^{2-p} - \|\mathbf{A}_i(j, :) - \mathbf{A}_i(j, :) \mathbf{V} \mathbf{V}^T\|_2^{2-p}}{\|\mathbf{A}_i(j, :) \mathbf{V}\|_2^{2-p} - \|\mathbf{A}_i(j, :) - \mathbf{A}_i(j, :) \mathbf{V} \mathbf{V}^T\|_2^{2-p}} = \\ & \sum_{i=1}^N \sum_{j=1}^m \operatorname{tr}((\mathbf{A}_i(j, :) \mathbf{V})^T (\mathbf{A}_i(j, :) \mathbf{V})) \mathbf{g}_{ij} - \\ & \operatorname{tr}((\mathbf{A}_i(j, :) - \mathbf{A}_i(j, :) \mathbf{V} \mathbf{V}^T)^T (\mathbf{A}_i(j, :) - \mathbf{A}_i(j, :) \mathbf{V} \mathbf{V}^T)) \mathbf{s}_{ij} \end{aligned} \quad (4)$$

其中, $\operatorname{tr}(\cdot)$ 指出矩阵的迹运算, $\mathbf{g}_{ij} = 1 / \|\mathbf{A}_i(j, :) \mathbf{V}\|_2^{2-p}$ 和 $\mathbf{s}_{ij} = 1 / \|\mathbf{A}_i(j, :) - \mathbf{A}_i(j, :) \mathbf{V} \mathbf{V}^T\|_2^{2-p}$ 。在实际应用中, 输入数据 $\mathbf{A}_i(j, :)$ 存在为 0 的可能。因此, 为了防止分母为 0 的现象出现, 需要对 \mathbf{g}_{ij} 和 \mathbf{s}_{ij} 的表达式进行重新定义为:

$$\mathbf{g}_{ij} = \begin{cases} \frac{1}{\|\mathbf{A}_i(j, :) \mathbf{V}\|_2^{2-p}}, & \|\mathbf{A}_i(j, :) \mathbf{V}\|_2^{2-p} \neq 0 \\ 0, & \|\mathbf{A}_i(j, :) \mathbf{V}\|_2^{2-p} = 0 \end{cases} \quad (5)$$

$$\mathbf{s}_{ij} = \begin{cases} \frac{1}{\|\mathbf{A}_i(j, :) - \mathbf{A}_i(j, :) \mathbf{V} \mathbf{V}^T\|_2^{2-p}}, & \|\mathbf{A}_i(j, :) - \mathbf{A}_i(j, :) \mathbf{V} \mathbf{V}^T\|_2^{2-p} \neq 0 \\ 0, & \|\mathbf{A}_i(j, :) - \mathbf{A}_i(j, :) \mathbf{V} \mathbf{V}^T\|_2^{2-p} = 0 \end{cases} \quad (6)$$

将式 (4) 代入到目标函数式 (3) 中, 进行合理的整合, 可得:

$$\begin{aligned} & \operatorname{argmax}_{\mathbf{V}^T \mathbf{V} = \mathbf{I}_k} \sum_{i=1}^N \sum_{j=1}^m \operatorname{tr}((\mathbf{A}_i(j, :) \mathbf{V})^T (\mathbf{A}_i(j, :) \mathbf{V})) \mathbf{g}_{ij} - \\ & \operatorname{tr}((\mathbf{A}_i(j, :) - \mathbf{A}_i(j, :) \mathbf{V} \mathbf{V}^T)^T (\mathbf{A}_i(j, :) - \mathbf{A}_i(j, :) \mathbf{V} \mathbf{V}^T)) \mathbf{s}_{ij} = \\ & \sum_{i=1}^N \sum_{j=1}^m \operatorname{tr}(\mathbf{V}^T \mathbf{A}_i(j, :)^T \mathbf{A}_i(j, :) \mathbf{V}) \mathbf{g}_{ij} + \\ & \operatorname{tr}(\mathbf{V}^T \mathbf{A}_i(j, :)^T \mathbf{A}_i(j, :) \mathbf{V}) \mathbf{s}_{ij} - \operatorname{tr}(\mathbf{A}_i(j, :)^T \mathbf{A}_i(j, :)) \mathbf{s}_{ij} \end{aligned} \quad (7)$$

通过观察目标函数式(7),不难发现其内部共包含 3 个未知变量 \mathbf{V} , \mathbf{g}_{ij} 和 \mathbf{s}_{ij} ,且 \mathbf{g}_{ij} 和 \mathbf{s}_{ij} 的值均与投影矩阵 \mathbf{V} 存在着紧密的关联。因此,想要直接求解目标函数式(7)将会变得十分困难。针对这个问题,设计了一个有效的迭代更新策略去获取 \mathbf{V} (固定 \mathbf{g}_{ij} 和 \mathbf{s}_{ij}), \mathbf{g}_{ij} (固定 \mathbf{V} 和 \mathbf{s}_{ij}) 以及 \mathbf{s}_{ij} (固定 \mathbf{g}_{ij} 和 \mathbf{V})。具体来说,在第 $t+1$ 次迭代过程中,当 $\mathbf{g}_{ij}^{(t)}$ 和 $\mathbf{s}_{ij}^{(t)}$ 是已知时,执行最大化目标函数式(7)完成对于 \mathbf{V} 的更新操作。在这种情况下,目标函数式(7)的最后一项可视为固定值。因此,目标函数式(7)可进一步改写为

$$\mathbf{V} = \underset{\mathbf{V}^T \mathbf{V} = \mathbf{I}_k}{\operatorname{argmaxtr}} (\mathbf{V}^T \mathbf{A}^T (\mathbf{G} + \mathbf{S}) \mathbf{A} \mathbf{V}) = \underset{\mathbf{V}^T \mathbf{V} = \mathbf{I}_k}{\operatorname{argmaxtr}} (\mathbf{V}^T \mathbf{A}^T \mathbf{U} \mathbf{A} \mathbf{V}) \quad (8)$$

其中, $\mathbf{A} = [\mathbf{A}_1(1, :)^T, \mathbf{A}_1(2, :)^T, \dots, \mathbf{A}_N(1, :)^T, \mathbf{A}_N(m, :)^T]^T \in \mathbf{R}^{m \times N \times n}$ 是由输入数据 $\mathbf{A}_i(j, :)$ 所组成的样本增广矩阵。 \mathbf{G} , \mathbf{S} 和 $\mathbf{U} = \mathbf{G} + \mathbf{S}$ 分别是由对角元素 \mathbf{g}_{ij} , \mathbf{s}_{ij} 和 \mathbf{u}_{ij} 构成的对角矩阵。根据拉格朗日乘子法,目标函数式(8)的解 \mathbf{V} 是由矩阵 $\mathbf{A}^T \mathbf{U} \mathbf{A}$ 前 k 个最大特征值所对应特征向量所组成。在此基础上,利用已更新的 \mathbf{V} 完成对 \mathbf{g}_{ij} 和 \mathbf{s}_{ij} 的计算。最后,反复执行上述迭代过程直至发现满足收敛条件的最优投影矩阵 \mathbf{V} ,具体求解过程详见算法 1。

表 1 ATS-2DPCA 的算法步骤

Table 1 Algorithm Steps of ATS-2DPCA

算法 1: ATS-2DPCA 求解最优投影矩阵 \mathbf{V} 算法

输入: 样本增广矩阵 $\mathbf{A} \in \mathbf{R}^{m \times N \times n}$, 特征维度 k 和 $p \in (0, 2)$, 其中样本数据 \mathbf{A}_i 已完成数据中心化处理。

初始化: 满足 $\mathbf{V} \mathbf{V}^T = \mathbf{I}_k$ 的 $\mathbf{V}^{(t-1)} \in \mathbf{R}^{n \times k}$, $\mathbf{g}_{ij}^{(t-1)}$, $\mathbf{s}_{ij}^{(t-1)}$, $t = 1$, $\delta = 0.01$ 。

while not converged **do**

1) 分别利用式(5)和(6)计算对角矩阵 $\mathbf{G}^{(t-1)}$ 和 $\mathbf{S}^{(t-1)}$ 的对角元素 $\mathbf{g}_{ij}^{(t-1)}$ 和 $\mathbf{s}_{ij}^{(t-1)}$ 。

2) 计算对角矩阵 $\mathbf{U}^{(t-1)}$ 的对角元素 $\mathbf{u}_{ij}^{(t-1)}$, $\mathbf{U}^{(t-1)} = \mathbf{G}^{(t-1)} + \mathbf{S}^{(t-1)}$ 以及 $\mathbf{u}_{ij}^{(t-1)} = \mathbf{g}_{ij}^{(t-1)} + \mathbf{s}_{ij}^{(t-1)}$

3) 计算加权协方差矩阵 $\mathbf{A}^T \mathbf{U}^{(t-1)} \mathbf{A}$

4) 求解目标函数式(8)的最优投影矩阵 $\mathbf{V}^{(t)}$, $\mathbf{V}^{(t)}$ 是由 $\mathbf{A}^T \mathbf{U}^{(t-1)} \mathbf{A}$ 的前 k 个最大特征值所对应特征向量组成。

5) 检验收敛条件 $J(\mathbf{V}^{(t)}) - J(\mathbf{V}^{(t-1)}) \leq \delta$ 是否满足; 如果满足, 执行步骤 6); 否则执行步骤 7)。

6) 使用 Armijo 线性搜索策略^[19]去探寻满足 $J(\mathbf{V}^{(t)}) \leq J(\mathbf{V}^{(t-1)})$ 的投影矩阵 $\mathbf{V}^{(t)}$, 如果有解存在, 则执行步骤 7); 否则结束循环。

7) 根据获取到的 $\mathbf{V}^{(t)}$ 完成对角矩阵 $\mathbf{G}^{(t)}$, $\mathbf{S}^{(t)}$ 和 $\mathbf{U}^{(t)}$ 中的每个对角元素 $\mathbf{g}_{ij}^{(t)}$, $\mathbf{s}_{ij}^{(t)}$ 和 $\mathbf{u}_{ij}^{(t)}$ 的计算。

8) $t \leftarrow t + 1$

end while

输出: $\mathbf{V}^{(t)} \in \mathbf{R}^{n \times k}$

3 实验结果与分析

选用 NF、WildFish 数据库和 EPIDHEU 3 组水下光学图像数据库作为验证平台, 将所提方法与现存主要具有低维鲁棒特性的 2DPCA 模型 2DPCA-L1、F-2DPCA、R1-2DPCA、Angle-2DPCA、GC-2DPCA 和 Cos-2DPCA 进行了对比。此外, 针对 3 组数据库中的实验, 考虑通过改变特征维数(变化范围设定在 0~200)的策略以实现对方方法识别精度以及重构性能的验证, 其中识别精度可利用最近邻(1-nearest neighbor, 1NN)分类器的结果进行判断, 而重构误差(reconstruction error, RE)可如式(9)所示。

$$\text{RE} = \frac{1}{N} \sum_{i=1}^N \|\mathbf{A}_i^{\text{clean}} - \mathbf{A}_i^{\text{clean}} \mathbf{V} \mathbf{V}^T\|_2 \quad (9)$$

式中: N 为无噪声样本的数量, $\mathbf{A}_i^{\text{clean}}$ 为第 i 个无噪声样本数据。最后, 为了最大限度地避免实验结果的偶然性, 决定对每个数据库中的实验过程进行 10 次重复。

3.1 实验数据

NF 数据库^[20]包含超过 50 000 个与海洋和水下生物相关的视频数据, 其中可用的水下生物视频数据不低于 8 000 个。基于这些视频, 挑选出具有不同类别和物理变化的 45 种鱼类视频作为研究基础, 并进一步截取每个视频中具有不同姿态的 40 帧图像数据作为本章的实验数据。在实验中, 将每个图像样本统一裁剪为 30×50 像素, 并从每个水下生物所包含的 40 帧图像中任意选取 12 帧添加椒盐噪声, 噪声强度设定在 0.005~0.050 之间且随机分布。接下来, 随机选取每个水下生物的 24 帧图像用于训练, 其中包括 16 帧干净图像和 8 帧含噪图像, 其余图像全部用于测试, 图 3 展示了 NF 数据库中的一些样本图像示例。

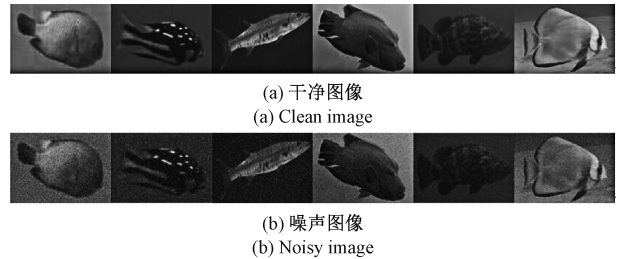


图 3 NF 数据库中一些图像样本示例, 包含干净图像(a)以及相对应的噪声图像(b)

Fig. 3 Examples of some image samples on the NF database, including clean image (a) and corresponding noisy image (b)

WildFish 数据库^[21]; 此数据库是目前最大的野生鱼类数据库, 共包含 1 000 种鱼类的 54 459 幅无约束样本

图像,所谓无约束图像就是不限制拍摄条件,是指这些图像不受拍摄条件的限制,展示了鱼类样本在多种姿态、不同照明强度等条件下的图形特性。此外,部分图像还包含水印信息和较为复杂的背景环境。为了能够让数据库中的样本图像更好的服务于实验需求,随机选取 28 种无水印信息并且背景相对简单的鱼类数据用于实验分析,其中每种鱼类包含 15 幅样本图像,同时将其尺寸裁剪成 30×50 像素的大小并做灰度化处理。在实验中,随机从每种鱼类的 15 幅样本图像中挑选出 6 幅并加入尺寸为 8×13 像素的水下物体(珊瑚)进行遮挡处理,这些遮挡随机分布于样本图像有效区域的任意位置。图 4 展示了数据库中的部分干净图像以及与之对应的遮挡图像示例。在实验中,考虑从每类样本中任意选取 9 张样本图像(包含 3 幅遮挡图像和 6 幅干净图像)用于训练,剩余的样本图像用于测试。

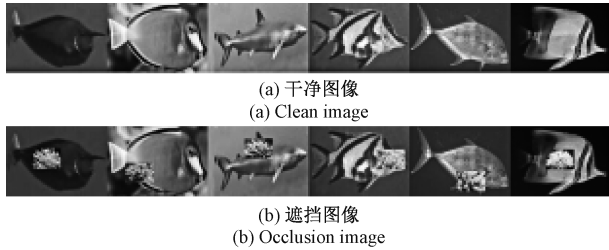
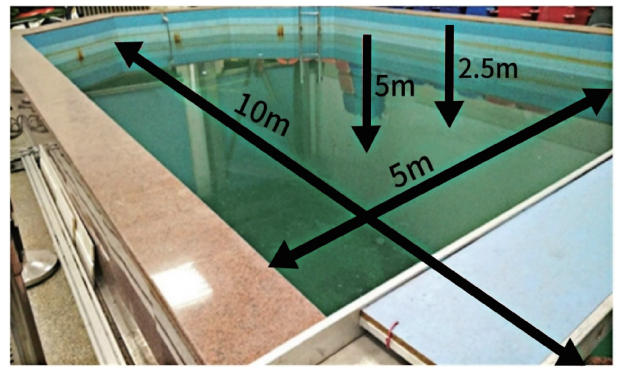


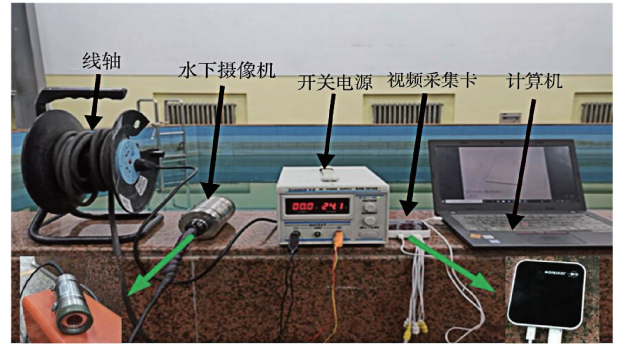
图 4 WildFish 数据库中一些图像样本示例,包含干净图像(a)以及相对应的遮挡图像(b)

Fig. 4 Examples of some image samples on the WildFish database, including clean image (a) and corresponding occlusion image (b)

EPIDHEU 数据库由包含 30 类水下物体的 600 张光学图像所组成,每类物体拥有 20 幅具有不同状态的图像,且存在微小的光照变化,其实际采集环境以及采集过程中所使用的主要仪器设备参数分别如图 5 和表 2 所示。在实验过程中,每幅图像的尺寸自动从 355×288 像素标准化到 44×36 像素。鉴于水下图像不可避免地会面临遮挡(如水下生物及岩石等)和高斯噪声(如水下光学传输特性以及水体环境影响等)同时出现的情况,本研究从每类物体的样本数据中随机选取 7 幅图像并引入混合噪声(遮挡与高斯噪声叠加)。其中遮挡选择为黑色的矩形块,占据图像目标区域的 $5\% \sim 25\%$,而椒盐噪声的波动强度则设定为在 $0.005 \sim 0.035$ 之间任意变化且呈随机分布,图 6 展示了 EPIDHEU 数据库中部分物体的样本图像示例。实验从每个物体的 20 幅样本图像中随机选取 12 幅图像用于训练(包括 4 幅噪声图像),其余 8 张图像用于测试。



(a) 实验环境
(a) Experimental environment



(b) 实验设备
(b) Experimental equipment

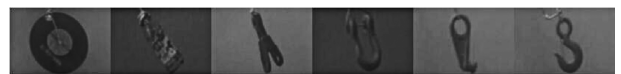
图 5 EPIDHEU 数据库的实际建立过程

Fig. 5 Real establishment process of EPIDHEU database

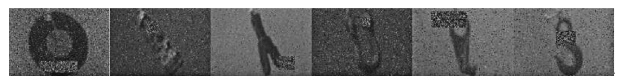
表 2 视频采集卡和水下摄像机的主要参数

Table 2 Main performance parameters for the video capture card and the underwater camera

名称	水下摄像机	名称	视频采集卡
型号	Tornado	型号	JVS-C301
尺寸	直径:78 mm 长度:178 mm	尺寸	长度×宽度: 96 mm×90 mm
重量	6 000 g	主机接口	USB
有效像素	768×576	最大视频分辨率	PAL:960×576 NTSC:960×487
分辨率	>570 lines	播放帧率	PAL:25 fps NTSC:30 fps
最小照度	0.02 Lux	视频压缩	标准 H. 264
电压	24VDC	采样能力	10 bit
光圈	自动光圈	板卡功耗	<3 W



(a) 干净图像
(a) Clean image



(b) 噪声图像
(b) Noisy image

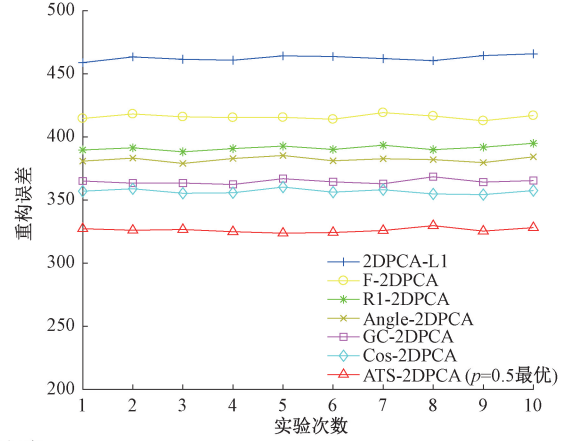
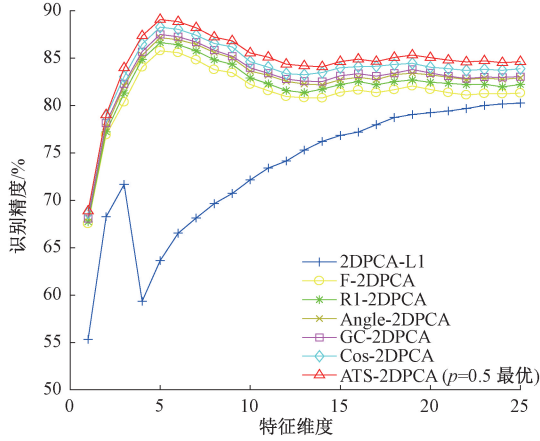
图 6 EPIDHEU 数据库中一些图像样本示例,包含干净图像(a)以及相对应的噪声图像(b)

Fig. 6 Examples of some image samples on the EPIDHEU database, including clean image (a) and corresponding noisy image (b)

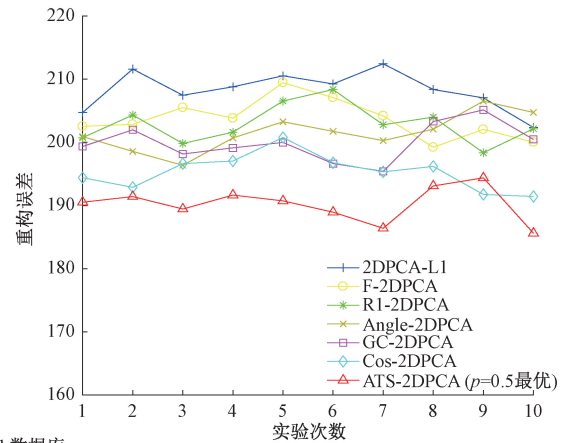
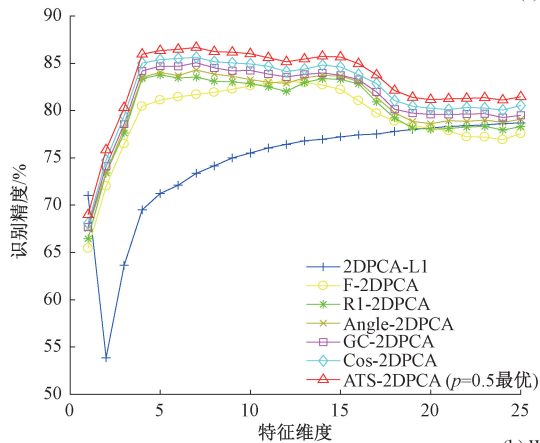
3.2 图像的识别与重构实验

为了直观且充分说明 ATS-2DPCA 的性能,本小节考虑将其与 6 种同类型对比方法 2DPCA-L1、F-2DPCA、R1-2DPCA、Angle-2DPCA、GC-2DPCA 和 Cos-2DPCA 在识别精度和重构结果等方面进行对比和验证。表 3 展示了

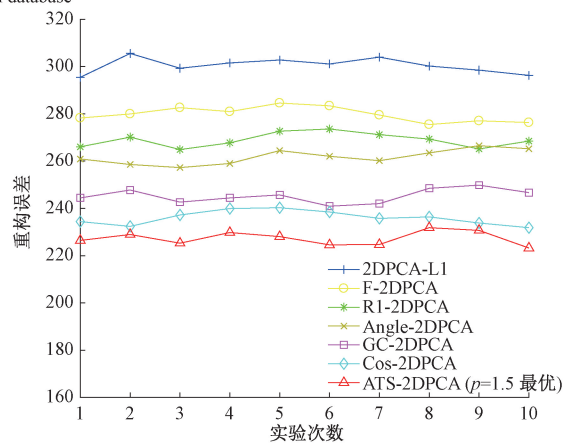
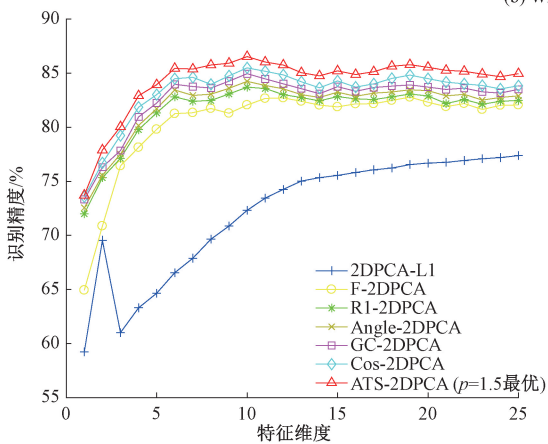
NF 数据库中每种方法的平均最优识别精度 (%) 和平均最小重构误差及其所对应的标准差。表 4 展示了 WildFish 数据库中每种方法的平均最优识别精度 (%) 和平均最小重构误差及其所对应的标准差。表 5 展示了 EPIDHEU 数据库中每种方法的平均最优识别精度 (%)



(a) NF数据库
(a) NF database



(b) WildFish数据库
(b) WildFish database



(c) EPIDHEU数据库
(c) EPIDHEU database

图 7 3 个水下图像数据库中所有对比方法在不同特征维度下的识别精度以及 10 组实验下的最小重构误差
Fig. 7 The average recognition accuracy of all the contrast methods under different feature dimensions and the minimum reconstruction error under ten experiments on three underwater optical image databases

表 3 NF 数据库中每种方法的平均最优识别精度 (%) 和平均最小重构误差及其所对应的标准差

Table 3 The average optimal recognition accuracy (%) and average minimum reconstruction error with the corresponding standard deviation (Std) of each method on NF database

方法	2DPCA-L1	F-2DPCA	R1-2DPCA	Angle-2DPCA	GC-2DPCA	Cos-2DPCA	ATS-2DPCA		
							$p=0.5$	$p=1$	$p=1.5$
识别精度	80.25±0.76	85.77±0.69	86.63±0.73	87.16±0.82	87.50±0.85	88.27±0.66	89.07±0.63	87.78±0.71	88.53±0.67
重构误差	462.47±2.14	415.86±1.92	402.60±2.08	391.25±2.01	382.04±1.97	364.59±1.88	343.91±1.85	377.42±1.93	370.73±1.90

表 4 WildFish 数据库中每种方法的平均最优识别精度 (%) 和平均最小重构误差及其所对应的标准差

Table 4 The average optimal recognition accuracy (%) and average minimum reconstruction error with the corresponding standard deviation (Std) of each method on WildFish database

方法	2DPCA-L1	F-2DPCA	R1-2DPCA	Angle-2DPCA	GC-2DPCA	Cos-2DPCA	ATS-2DPCA		
							$p=0.5$	$p=1$	$p=1.5$
识别精度	80.48±0.78	85.61±0.76	85.95±0.70	86.31±0.79	86.85±0.71	87.68±0.69	88.52±0.75	87.26±0.70	88.04±0.71
重构误差	208.26±3.07	203.68±3.10	202.87±3.04	201.52±2.92	199.96±2.95	195.32±2.83	190.22±2.74	197.48±2.86	193.14±2.78

表 5 EPIDHEU 数据库中每种方法的平均最优识别精度 (%) 和平均最小重构误差及其所对应的标准差

Table 5 The average optimal recognition accuracy (%) and average minimum reconstruction error with the corresponding standard deviation (Std) of each method on EPIDHEU database

方法	2DPCA-L1	F-2DPCA	R1-2DPCA	Angle-2DPCA	GC-2DPCA	Cos-2DPCA	ATS-2DPCA		
							$p=0.5$	$p=1$	$p=1.5$
识别精度	77.38±1.57	82.79±1.81	83.46±1.67	83.71±1.73	84.25±1.59	84.96±1.55	85.63±1.53	84.58±1.51	86.00±1.60
重构误差	300.47±3.21	279.80±3.08	272.33±3.14	268.92±3.03	261.76±3.10	245.28±2.93	240.36±2.98	259.05±2.90	236.54±2.95

和平均最小重构误差及其所对应的标准差。图 7 分别展示 NF、WildFish 和 EPIDHEU 3 个数据库中所有对比方法在不同特征维度下的识别精度以及 10 组实验下的最小重构误差。图 8 展示了 EPIDHEU 数据库中所有对比方法在最优特征维度下的重构可视化结果。图 9 展示了 3 个水下光学图像数据库中所提方法初始投影矩阵选取与识别精度的关系。针对上述实验结果进行观测和分析,可以总结得出如下有价值的结论:

1) 根据表 3~5, 图 7 和 8 中所呈现出的内容, 能够确切无疑地得出结论: ATS-2DPCA 在所有图像数据库上的识别精度和重构误差都明显优于其他 6 种对比方法, 这主要归因于 ATS-2DPCA 采用了任意三角形结构, 此结构不仅能够同时兼顾投影数据方差与重构误差之间的关系, 而且能够匹配到更具鲁棒性的小距离度量模式。因此, 该方法能有效地抑制噪声信息所带来的负面影响。另一个重要原因是 ATS-2DPCA 选用了灵活的 $l_{2,p}$ 范数作为距离度量, 此度量策略在继承了欧式距离度量的旋转不变性基础上, 可通过合理选取 p 值, 实现基于 2DPCA 方法最小化重构误差的目标, 进一步提升模型在噪声环境中的特征提取能力。

2) 进一步深入观察表 3~5, 图 7 和 8 所反映出的实验结果可知, 尽管 2DPCA-L1 在距离度量方式上选用了具有鲁棒特性的 l_1 范数作为标准, 但其整体性能均逊色于其他基于矩阵的 2DPCA 方法, 出现此结果可能主要包

含两点原因, 其一是 2DPCA-L1 在优化模型中均采用最大化方差的策略, 而最大化方差并不是基于 2DPCA 方法的本质诉求。其二是 l_1 范数的度量模式是在丢失了诸如旋转不变性等一些宝贵性质的前提下实现了鲁棒性的提升, 此种改变并没有达到预期目标。此外, 在使用欧式距离度量策略的方法中, F-2DPCA 在识别和重构能力方面要逊色于其他对比方法, 原因是 F-2DPCA 仅考虑了模型度量准则的鲁棒性, 而忽略了对于结构建立与识别机制的探索, 如 R1-2DPCA 在识别阶段使用了削弱背景干扰的识别方案, GC-2DPCA 减少了噪声信息对于均值数据的影响, Angle-2DPCA 和 Cos-2DPCA 在结构上同时兼顾了最小化重构误差与最大化方差的双重诉求。

3) 图 7 中的实验结果表明, 当特征维数处于相对较低的某个阶段时, 2DPCA-L1 的识别率曲线出现了明显下降趋势, 可能是因为该方法前 k 个特征维度主要描述了噪声数据特征信息, 形成类内样本之间差异大于类间样本之间差异的现象发生, 进而导致识别精度变差。然而, 与 2DPCA-L1 表现结果不同, F-2DPCA、R1-2DPCA、Angle-2DPCA、GC-2DPCA 以及 Cos-2DPCA 均采用非贪婪算法来获取最优解, 该最优解与保护数据几何结构的协方差矩阵密切相关, 因此能够在很大程度上降低噪声数据带来的负面影响。

3.3 算法性能相关性实验

设计非贪婪迭代算法 1 的核心目标是为了快速准确

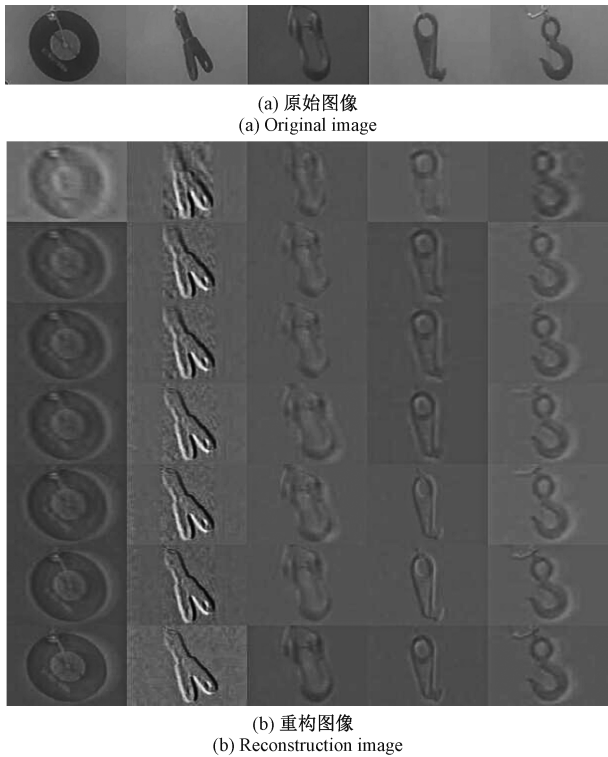


图 8 EPIDHEU 数据库中所有对比方法在最优识别精度下的重构可视化结果. 第 1 行是干净的原始图像, 从第 2 行到第 8 行分别为 2DPCA-L1, F-2DPCA, R1-2DPCA, Angle-2DPCA, GC-2DPCA, Cos-2DPCA 以及 ATS-2DPCA 所对应的重构图像

Fig. 8 The reconstruction visualization results of all the contrast methods under optimal recognition accuracy on EPIDHEU database. The first row is the clean original image, from the second row to the eighth row are the reconstructed images corresponding to 2DPCA-L1, F-2DPCA, R1-2DPCA, Angle-2DPCA, GC-2DPCA, Cos-2DPCA and ATS-2DPCA, respectively

获取到 ATS-2DPCA 的最优解, 因此从检验算法 1 相关性角度出发, 本小节实验从这两个方面进行验证, 其一是考察算法 1 中初始投影矩阵选取对于识别结果的影响, 其二是考察算法 1 的收敛速率, 具体验证过程与分析如下所述。

1) ATS-2DPCA 的最优解是利用算法 1 数次迭代更新所得, 而在执行首次迭代任务的过程中需要预先给定一个具有正交特性的投影矩阵 V 。为了验证初始投影矩阵 V 对于 ATS-2DPCA 识别精度的影响尺度, 实验考虑针对 3 个水下光学图像数据库 NF, WildFish 以及 EPIDHEU 中的数据随机给出 10 组不同的初始投影矩阵 V , 并依次选择在 $p=0.5$, $p=1$ 和 $p=1.5$ 时完成关于 ATS-2DPCA 的识别精度获取, 具体实验结果如图 9 所示。图 9 中每个数据库包含的 10 个绿色圆点代表 10 个不同初始矩阵

作用下的识别精度, 红色基准线代表平均识别精度。从图 9 的实验结果不难发现, NF 数据库中 10 组识别精度的波动范围为 87.92%~90.42% ($p=0.5$), 86.53%~88.61% ($p=1$) 以及 87.75%~89.69% ($p=1.5$); WildFish 数据库中 10 组识别精度的波动范围为 87.80%~90.12% ($p=0.5$), 86.13%~87.92% ($p=1$) 和 87.20%~89.23% ($p=1.5$); EPIDHEU 数据库中 10 组识别精度的波动范围为 84.58%~86.46% ($p=0.5$), 83.77%~85.73% ($p=1$) 和 84.90%~87.19% ($p=1.5$)。上述实验结果充分表明, 投影矩阵不同初值的选取不会导致识别精度出现显著波动, 因此初始值的随机给定对算法整体性能影响不大, 且在可接受的范围以内。

2) 衡量算法 1 性能的另一个重要指标是收敛速度的快慢, 其直接影响到了 ATS-2DPCA 在执行特征提取任务过程中的效率。尽管算法 1 存在良好的收敛特性, 但本着将其迭代变化结果进行直观展示的目的, 最终通过图 10 明确刻画了 ATS-2DPCA 在经过算法 1 每次迭代时所获取到的目标函数值。根据图 10 的实验结果可知, 随着迭代次数的逐渐增加, ATS-2DPCA 在 3 个数据库中的目标函数值也都同时呈现出单调递减的趋势, 且均在 10 次以内趋于平稳, 即能够快速寻找到 ATS-2DPCA 的最优解。这说明了算法 1 拥有较快的收敛速度, 从另一个角度展现了其在性能上的杰出表现力。

3.4 计算效率实验

本小节主要考虑从两个方面检验所提模型 ATS-2DPCA 的计算效率, 其一是考察 ATS-2DPCA 自身的计算复杂度, 其二是检验所有对比方法在 3 个水下光学图像数据库中的运行时间, 相关具体描述如下所示:

1) 在算法 1 执行每一次迭代更新任务过程中, ATS-2DPCA 的计算复杂度主要体现在对具有低维属性投影矩阵 $V \in R^{n \times k}$ 的求解, 而 V 的获取方式为计算协方差矩阵 $A^T U A$ 的特征值分解操作, 所以其计算复杂度不会超过 $O(n^3)$, 这与经典的 2DPCA 方法保持一致。综上分析可知, ATS-2DPCA 经过算法 1 每一次迭代后的整体时间复杂度最大限度为 $O(n^3)$, 如果共需 t 次迭代才能使之完成收敛, 那么所提方法的最终时间复杂度便为 $O(tn^3)$ 。

2) 呈现在表 6 中的实验结果表明, 在所有具有迭代特性的算法中, ATS-2DPCA 的运行时间接近于 F-2DPCA, R1-2DPCA 和 Cos-2DPCA, 因为针对其模型结构所设计的求解算法求解更为便捷; 比 Angle-2DPCA 和 GC-2DPCA 略快, 原因是 Angle-2DPCA 在同时兼顾投影数据方差和重构误差的前提下, 每次迭代过程中需使用奇异值分解操作, GC-2DPCA 要引入较多的中间变量去获取保护模型几何结构的权重矩阵; 明显优于 2DPCA-L1, 出现此结果是因为 2DPCA-L1 中基于 l_1 范数的非贪

婪求解算法较为复杂,从而在计算过程中增大了存储空间需求和时间损耗。因此,综合 ATS-2DPCA 在噪声环境

中的识别结果,可知其在对比方法中拥有最为出色的整体性能。

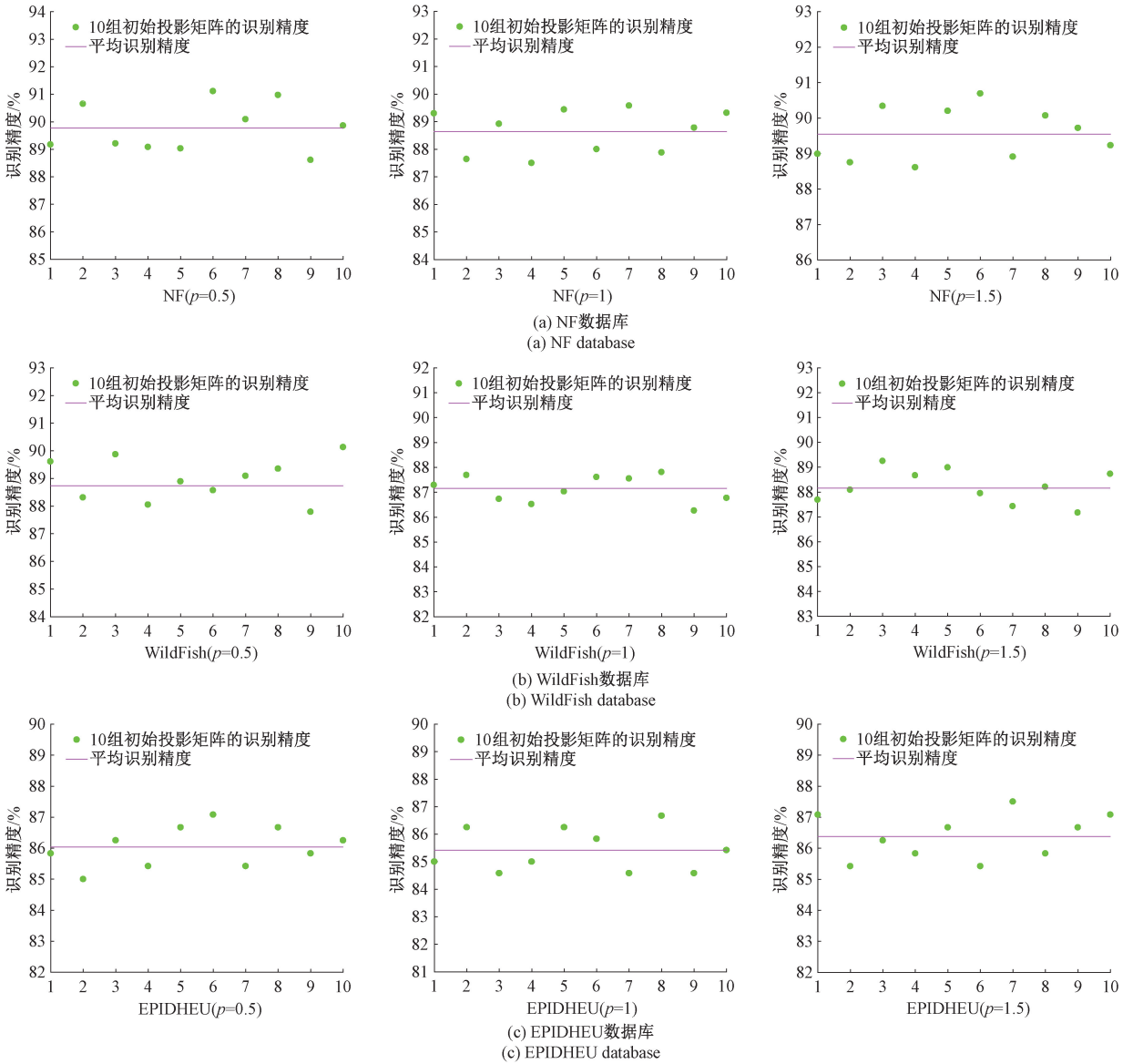


图 9 3 个水下图像数据库中所提方法初始投影矩阵选取与识别精度的关系

Fig. 9 The relationship between the initial projection matrix selection and recognition accuracy of the proposed method on three underwater image databases

表 6 3 个水体环境数据库中每种方法的平均运行时间与对应的标准差
Table 6 The average running time and the corresponding standard deviation on the three water environment databases

方法	2DPCA-L1	F-2DPCA	R1-2DPCA	Angle-2DPCA	GC-2DPCA	Cos-2DPCA	ATS-2DPCA (p 取最优值)
NF	11.26±0.58	2.13±0.01	2.86±0.18	5.60±0.44	4.79±0.37	3.08±0.30	3.19±0.25
WildFish	4.72±0.54	0.43±0.01	0.48±0.01	0.63±0.05	0.59±0.04	0.50±0.03	0.51±0.02
EPIDHEU	6.24±0.77	1.52±0.02	1.75±0.28	2.64±0.61	2.21±0.43	1.80±0.39	1.87±0.33

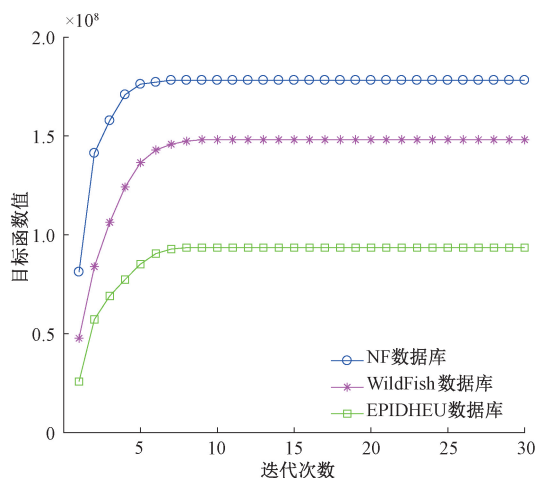


图 10 3 个水下图像数据库中 ATS-2DPCA 的收敛曲线

Fig. 10 The convergence curve of ATS-2DPCA on three underwater image databases

4 结 论

本文提出了一种具有低维表达属性的任意三角形结构 2DPCA (ATS-2DPCA) 方法并将其应用于水下光视觉观测环境下的图像识别研究,该方法在考虑建立最大化方差与最小化重构误差双重诉求的基础上,成功匹配到了更具灵活性和鲁棒性的 $l_{2,p}$ 范数距离度量模式,极大抑制了水下环境噪声干扰所带来的不利影响。与此同时,ATS-2DPCA 的解与加权协方差矩阵有着紧密关联,实现了对于数据几何结构的合理保护。最为重要的是,ATS-2DPCA 中采用的架构存在良好的普适性和延拓性,可为数据降维方法提供新的探索方向。在 NF、WildFish 和 EPIDHEU 3 个水下光学图像数据库上的实验结果表明,ATS-2DPCA 不仅具有良好的收敛速度,而且在同类方法中展现了出色的识别性能。因此,所提出的方法具有显著的创新性与实用价值。

参考文献

[1] 庄苏锋,屠大维,张旭,等. 水下双目立体视觉对应点匹配与三维重建方法研究[J]. 仪器仪表学报, 2022, 43(5): 147-154.
ZHUANG S F, TU D W, ZHANG X, et al. Research on corresponding point matching and 3D reconstruction of underwater binocular stereo vision[J]. Chinese Journal of Scientific Instrument, 2022, 43 (5): 147-154.

[2] REHMAN A, KHAN A, ALI M A, et al. Performance analysis of pca, sparse pca, kernel pca and incremental pca algorithms for heart failure prediction [C]. 2020 International Conference on Electrical, Communication, and Computer Engineering (ICECCE). IEEE, 2020;

1-5.

[3] 胡帅,程迎新,顾艳. PCA 和 Elman 网络在移动学习策略分类中的应用[J]. 电子测量技术, 2016, 39(5): 182-186.
HU SH, CHENG Y X, GU Y. Application of PCA and Elman network in mobile learning strategy classification[J]. Electronic Measurement Technology, 2016, 39(5): 182-186.

[4] DENG X G, TIAN X M. Sparse kernel locality preserving projection and its application in nonlinear process fault detection[J]. Chinese Journal of Chemical Engineering, 2013(2): 163-170.

[5] 刘玉珍,王鑫磊,林森. 加权自适应 MULBP 与 2DPCA 结合的掌纹识别方法[J]. 电子测量与仪器学报, 2021, 35(1): 142-150.
LIU Y ZH, WANG X L, LIN S. Palmprint recognition method by combining weighted adaptive MULBP and 2DPCA [J]. Journal of Electronic Measurement and Instrumentation, 2021, 35(1): 142-150.

[6] LI M, YUAN B. 2D-LDA: A statistical linear discriminant analysis for image matrix [J]. Pattern Recognition Letters, 2005, 26(5): 527-532.

[7] 高全学,高菲菲,郝秀娟,等. 基于图像欧氏距离的二维局部多样性保持投影[J]. 自动化学报, 2013, 39(7): 1062-1070.
GAO Q X, GAO F F, HAO X J, et al. 2D local diversity preservation projection based on image euclidean distance[J]. Journal of Automation, 2013, 39 (7): 1062-1070.

[8] WANG R, NIE F P, YANG X J, et al. Robust 2DPCA with non-greedy L1-norm maximization for image analysis [J]. IEEE Transactions on Cybernetics, 2014, 45 (5): 1108-1112.

[9] WANG H X, WANG J. 2DPCA with L1-norm for simultaneously robust and sparse modelling[J]. Neural Networks, 2013, 46: 190-198.

[10] WANG J. Generalized 2-D principal component analysis by Lp-norm for image analysis[J]. IEEE Transactions on Cybernetics, 2015, 46(3): 792-803.

[11] 王肖锋,陆程昊,酆金祥,等. 广义余弦二维主成分分析[J]. 自动化学报, 2022, 48(11): 2836-2851.
WANG X F, LU CH H, LI J X, et al. Generalized cosine two-dimensional principal component analysis [J] Journal of Automation, 2022, 48(11): 2836-2851.

[12] WANG Q, GAO Q, GAO X, et al. Optimal mean two-dimensional principal component analysis with F-norm minimization [J]. Pattern Recognition, 2017, 68: 286-294.

- [13] WANG Q Q, GAO Q X, GAO X B, et al. Robust 2DPCA and its application[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2016: 79-85.
- [14] LI T, LI M Y, GAO Q X, et al. F-norm distance metric based robust 2DPCA and face recognition[J]. Neural Networks, 2017, 94: 204-211.
- [15] GAO Q, XU S, CHEN F, et al. R1-2DPCA and face recognition[J]. IEEE Transactions on Cybernetics, 2018, 49(4): 1212-1223.
- [16] ZHOU G, XU G, HAO J, et al. Generalized centered 2-D principal component analysis[J]. IEEE Transactions on Cybernetics, 2019, 51(3): 1666-1677.
- [17] GAO Q X, MA L, LIU Y, et al. Angle 2DPCA: A new formulation for 2DPCA [J]. IEEE Transactions on Cybernetics, 2017, 48(5): 1672-1678.
- [18] WANG X, SHI L, LIU J, et al. Cosine 2DPCA with weighted projection maximization [J]. IEEE Transactions on Neural Networks and Learning Systems, 2022, 34(12): 9643-9656.
- [19] ZHANG B J, GONG X L, WANG J, et al. Nonstationary fuzzy neural network based on FCMnet clustering and a modified CG method with Armijo-type rule [J]. Information Sciences, 2022, 608: 313-338.
- [20] BILEN Ç, ZEPEDA J, PÉREZ P. The CNN news footage datasets: Enabling supervision in image retrieval[C]. 2016 24th European Signal Processing Conference (EUSIPCO). IEEE, 2016: 1901-1905.
- [21] ZHUANG P, WANG Y, QIAO Y. Wildfish: A large benchmark for fish recognition in the wild [C]. Proceedings of the 26th ACM International Conference on Multimedia, 2018: 1301-1309.

作者简介



陈璇, 现为南京信息工程大学人工智能学院本科生, 主要研究方向为机器学习、深度学习、水下目标识别。

E-mail: Xuan565888@outlook.com

Chen Xuan is currently a B. Sc. candidate at the School of Artificial Intelligence, Nanjing University of Information Science and Technology. Her main research interests include machine learning, deep learning, and underwater target recognition.



毕鹏飞 (通信作者), 2021 年于哈尔滨工程大学获得博士学位, 现为南京信息工程大学讲师, 主要研究方向为机器学习、深度学习、模式识别、水下目标感知。

E-mail: pfcx@nuist.edu.cn

Bi Pengfei (Corresponding author) received his Ph. D. degree from Harbin Engineering University in 2021 and is currently a lecturer at Nanjing University of Information Science and Technology. His main research interests include machine learning, deep learning, pattern recognition, and underwater target perception.



胡志远, 2023 年于泰州学院获得学士学位, 现为南京信息工程大学在读研究生, 主要研究方向为水下图像增强、水下目标识别。

E-mail: 2360739256@qq.com

Hu Zhiyuan received his Ph. D. degree from Taizhou University in 2023 and is currently a M. Sc. candidate at Nanjing University of Information Science and Technology. His main research interests include underwater image enhancement and underwater target recognition.