

基于稀疏逼近的图像编码算法分析与改进

吕艳琳,陶玉婷,张 燕

(金陵科技学院软件工程学院,江苏 南京 211169)

摘 要:在图像表示领域,矩阵分解法近年来得到了越来越多的关注。然而,该类算法大多属于无监督学习方法,故难以有效捕捉数据中固有的几何结构。为此,提出了一种综合应用矩阵分解和稀疏表示的稀疏逼近编码新方法,通过综合应用矩阵分解与稀疏约束构建矩阵优化模型,并通过实验进行验证。实验结果表明:该方法通过稀疏表示,可有效捕捉图像空间的内在几何结构,在识别语义结构方面具有一定的优越性,能更有效地进行图像表示和编码。

关键词:图像表示;图像编码;稀疏逼近编码;矩阵分解

中图分类号:TP3

文献标识码:A

文章编号:1672-755X(2020)04-0018-04

Analysis and Improvement of Image Coding Algorithm Based on Sparse Approximation

LYU Yan-lin, TAO Yu-ting, ZHANG Yan

(Jinling Institute of Technology, Nanjing 211169, China)

Abstract: In the field of image representation, matrix decomposition has received more and more attention in recent years. However, as an unsupervised learning method, it cannot effectively capture the inherent geometric structure of the data. Therefore, this paper proposes a new method of sparse approximation coding that comprehensively applies matrix decomposition and sparse representation. A matrix optimization model is constructed by applying matrix decomposition combined with sparse constraints, and then the corresponding code is written to verify the proposed model. The experimental results of image clustering show that the new method can effectively capture the inherent geometric structure of the image space through sparse representation. It has clear advantages in recognizing semantic structure, and it is more effective for image representation and coding.

Key words: image representation; image coding; sparse approximation coding; matrix decomposition

图像表示是图像处理中的基本问题。长期以来,研究人员一直在寻找有效的图像表示方法。对于一个给定的图像数据库,可能有数千个不同的特性,然而,每个图像的自由度可能要小得多。因此,人们希望找到一个隐藏的语义“概念”空间(而不是原始的特征空间)来对图像进行特征表示^[1]。这个“概念”空间的维度将比特征空间小得多。为了实现这一目标,基于矩阵分解的方法在过去几十年中引起了相当多的关

收稿日期:2020-10-12

基金项目:教育部产学合作协同育人项目

作者简介:吕艳琳(1977—),女,江苏徐州人,讲师,硕士,主要从事算法设计与应用——机器学习、人工智能等方面的研究。

注^[2]。其中一个著名的矩阵分解方法是潜在的语义分析(latent semantic analysis, LSA)^[3],它在根本上是基于奇异值分解的方法。LSA在重构误差的意义上是最优的,因此它所获得的系数对于欧几里得空间是一种最优表示系数。另一个流行的矩阵分解方法是非负矩阵分解(non-negative matrix factorization, NMF)^[4],这要求分解矩阵(\mathbf{U} 和 \mathbf{A})是非负的。非负约束只允许不同基向量之间的加性组合,并且认为NMF可以学习基于部分数据的表示系数^[5]。NMF的有效性已在许多图像分析任务中得到证实^[6]。受生物视觉系统的启发,人们一直在争论数据点的稀疏特征在数据学习过程中所起到的作用。稀疏编码(sparse coding, SC)^[7]是最近流行的矩阵分解方法,要求矩阵 \mathbf{A} 稀疏。 \mathbf{A} 的稀疏表示每个图像只与几个概念(与非零系数对应的基向量)相关^[8]。然而,这些流行的矩阵分解方法只考虑图像空间的欧几里得结构。

以上讨论的三个矩阵分解方法只考虑图像空间的欧几里得结构。然而,最近的研究表明,人类生成的图像数据很可能是从环境欧几里得空间的一个分支中取样的^[9]。事实上,人类生成的图像数据不可能均匀地“填充”高维欧几里得空间。因此,在学习基础的同时,需要考虑数据固有的流形结构^[10]。

结合稀疏编码方法的思想,本文认为一个好的矩阵分解应该考虑以下两个方面:模型所学习的基向量应该捕捉数据的内在几何结构,模型所学习的基向量应该是稀疏的。基于此,本文提出了一种新的矩阵分解方法,称为稀疏逼近编码(sparse approximation coding, SAC),用于图像表示。SAC是一种两步法,包括基底学习和稀疏表示学习。在第一步中,SAC利用数据的内在几何结构来学习基础。通过对最近邻图的谱分析,SAC对基向量的语义结构进行编码。在第二步中,SAC使用LASSO^[11]求解器来对每个图形的基向量进行学习,从而获得一个稀疏向量值。因此,SAC比传统的矩阵分解方法具有更强的识别能力,而传统的矩阵分解方法只考虑了数据的欧几里得结构。

1 稀疏逼近编码

SAC使用LASSO,这是一种L1正则化的回归模型,用于学习每个图像的稀疏表示。

SAC的目标是解决如下稀疏回归的优化问题:

$$\min_{\mathbf{U}, \mathbf{A}} \|\mathbf{X} - \mathbf{U}\mathbf{A}\|^2 + \|\mathbf{A}\|_1 \quad (1)$$

其中,数据矩阵 \mathbf{X} 的每一列为单个数据样本,在此基础上,矩阵 \mathbf{U} 通过分解得到原始数据矩阵 \mathbf{X} 的固有几何结构,而矩阵 \mathbf{A} 的每一列向量被限定为稀疏^[12]。

式(1)中的基底矩阵 \mathbf{U} 的求解可以看作经典的最小二乘回归问题,通过令导数为0可以获得 \mathbf{U} 的解析解如下:

$$\mathbf{U} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{X}^T \mathbf{A} \quad (2)$$

在得到了基底矩阵 \mathbf{U} 之后,可以通过下面的最小化问题,独立地通过列来计算表示 \mathbf{A} 。

$$\min_{\mathbf{a}_i} \|\mathbf{x}_i - \mathbf{U}\mathbf{a}_i\|^2 + \beta \|\mathbf{a}_i\|_1 \quad (3)$$

其中 \mathbf{x}_i 和 \mathbf{a}_i 分别是 \mathbf{X} 和 \mathbf{A} 的第 i 列。 $\|\mathbf{a}_i\|_1$ 表示 \mathbf{a}_i 的L1范数,而L1-norm正则化项用以保证 \mathbf{a}_i 的稀疏性。上述优化问题有如下的等价优化形式:

$$\min_{\mathbf{a}_i} \|\mathbf{x}_i - \mathbf{U}\mathbf{a}_i\|^2 \quad \text{s. t. } \|\mathbf{a}_i\|_1 \leq \gamma \quad (4)$$

LARs(least angle regression)算法^[13]可以用于解决这个优化问题。LARs在 $O(k^3 + mk^2)$ 的时间复杂度内,可以计算出所有解,即包含所有可能的 \mathbf{a}_i 基数的解。求解基底矩阵 \mathbf{U} 与求解稀疏矩阵 \mathbf{A} 的总时间复杂度为:

$$O(n^2 s + n^2 p + k^3 + mk^2)$$

最后,将迭代求解基底矩阵 \mathbf{U} 和矩阵 \mathbf{A} 的过程总结为图1。

与传统的稀疏编码方法相比,SAC有以下优点:

1)传统的稀疏编码方法考虑了欧几里得空间的分解,而SAC通过利用数据的内在几何特性来学习基础向量。SAC学习的基础可以具有更强的识别能力。

2)传统的稀疏编码方法多次反复迭代才能获得两个分解后的数值矩阵,这是非常耗时的。而本文提

出的 SAC 方法在单次迭代中具有闭式解,只需要解决一个稀疏的特征问题和两个回归问题,这在计算上是非常高效的。

2 结果与分析

为了验证本文方案的有效性和高效性,对本实验的方法进行编码验证。本次实验的平台为 MATLAB R2014b 软件平台,微型计算机处理器为 Intel(R) Core (TM) i7-4770k CPU@3.50 GHz,内存为 16.0 GB。实验数据为 COIL20 数据库,包含 1 440 张 32×32 像素的灰度图像,分别为 20 个物体在 5 个角度拍摄的照片,每个物体包含 72 张。由于基于编码的无监督矩阵分解算法不需要预训练过程,因此将数据集的所有图像直接用于测试,以聚类准确度作为实验评价指标。

在本次实验中,首先和 5 种常见方法进行比较,分别为:原始空间的 K-means 聚类(Baseline),基于奇异值分解的潜在语义分析(SVD)^[14],非负矩阵分解(NMF)^[15],稀疏表示方法(SC),拉普拉斯稀疏表示(LapSC)^[16]。实验中的 K-means 聚类的个数设置为 5,岭回归的正则化参数为 0.1。为了避免实验的偶然性,在 K-means 聚类过程中随机挑选 10 个点作为起始点,并将实验重复 10 次,获取平均值作为实验的结果。首先,探究了第一步中获取的向量个数对实验结果的影响,结果如图 2 所示。

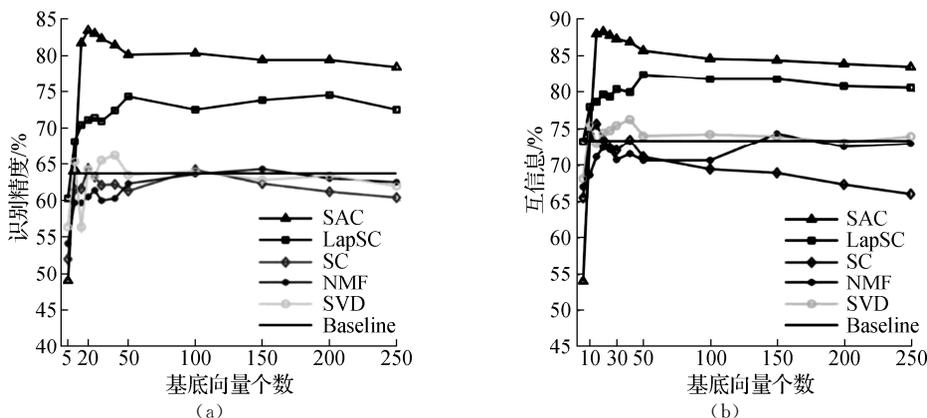


图 2 识别精度与互信息随着基底向量个数增加的变化情况

其次,由于绝大多数的编码算法最终的聚类效果会受到 K-means 算法中设定的聚类个数的影响,因此本文还探究了 K-means 算法聚类个数对所提算法的影响,结果如图 3 所示。从图 2 和图 3 可以得出以下结论:

- 1) 实验提出的方法可以取得最好的聚类精度,比排名第二的基于拉普拉斯图的稀疏表示方法精度提高了 8~17 个百分点,比其他 3 种方法平均高出 14.5 个百分点,验证了本实验方法的有效性和高效性;
- 2) 在本组实验中,当测试阶段所采用 K-means 算法的聚类个数 $k=14$ 的时候,可以取得最高的精度,精度为 92.4%;
- 3) 通过对基于拉普拉斯图的稀疏方法和本实验的方法与其他非图模型的方法进行比较,发现基于图的方法在性能上均有较大的提升,说明通过图的方法能很好地抓住数据的内在集合信息。

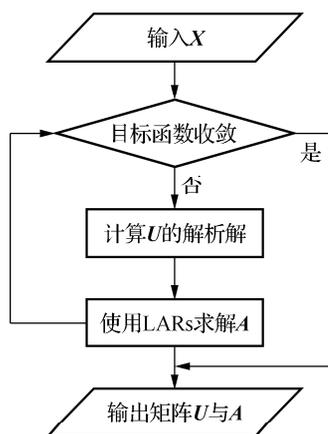


图 1 迭代求解基底矩阵 U 与表示系数 A 的流程

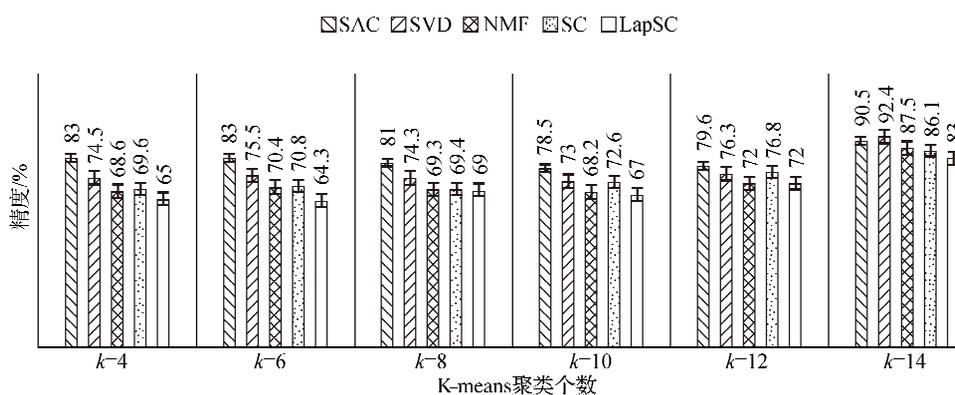


图3 不同聚类个数设置对实验精度的影响

3 结 语

图像分析与表示算法在过去几十年里受到了广泛关注,并在模式识别等相关任务中得到了卓有成效的应用。奇异值分解、非负矩阵分解等算法是经典的图像表示算法,但在高维欧氏特征空间内对图像的代表能力欠佳。因此,本文提出了一种新型的稀疏逼近编码(SAC)模型,利用矩阵分解与稀疏约束构建了一个矩阵优化模型,并使用交替迭代法对优化模型进行求解。编写了相应的 MATLAB 代码对所提出的模型进行了验证,在 COIL20 数据集上与非负矩阵分解、稀疏编码等经典图像编码算法进行了对比,充分验证了算法的有效性。实验结果证明,本文提出的模型在不同聚类个数设定环境下,均具有准确而稳健的识别效果。本文方法对数据具有一定的稀疏假设,因此在不满足稀疏假设的情形下不一定能保证识别性能良好。

参考文献:

- [1] 王珏. 机器学习及其应用[M]. 北京:清华大学出版社,2006
- [2] 杨淑莹. 图像模式识别:VC++技术实现[M]. 北京:北京交通大学出版社,2005
- [3] 肖延辉. 基于矩阵分解的图像表示理论及其应用研究[D]. 北京:北京交通大学,2014
- [4] 姜小燕. 基于非负矩阵分解的图像分类算法研究[D]. 锦州:辽宁工业大学,2016
- [5] 王科俊,左春婷. 非负矩阵分解特征提取技术的研究进展[J]. 计算机应用研究,2014(4):970-975
- [6] 舒振球,赵春霞. 基于图正则化的受限非负矩阵分解算法及在图像表示中的应用[J]. 模式识别与人工智能,2013(3):300-306
- [7] 赵康. 利用稀疏编码结合深度学习的人体姿态估计[J]. 信息技术,2020(9):61-65
- [8] 练秋生,张伟. 基于图像块分类稀疏表示的超分辨率重构算法[J]. 电子学报,2012(5):920-925
- [9] GAO S, SANG I T, CHIA L, et al. Local features are not lonely-laplacian sparse coding for image classification[J]. IEEE Conference on Computer Vision and Pattern Recognition(CVPR'10), 2010(4):3555-3561
- [10] 李映,张艳宁,许星. 基于信号稀疏表示的形态成分分析:进展和展望[J]. 电子学报,2009(1):146-152
- [11] 朱杰,杨万扣,唐振民. 基于字典学习的核稀疏表示人脸识别方法[J]. 模式识别与人工智能,2012(5):859-864
- [12] 浦剑,张军平. 基于词典学习和稀疏表示的超分辨率方法[J]. 模式识别与人工智能,2010,23(3):335-340
- [13] 胡正平,李静. 基于低秩子空间恢复的联合稀疏表示人脸识别算法[J]. 电子学报,2013,41(5):987-991
- [14] EFRON B, HASTIE T, JOHNSTONE I, et al. Least angle regression[J]. Annals of Statistics, 2004, 32(2):407-499
- [15] CAI D, HE X, HAN J, et al. Document clustering using locality preserving indexing[J]. IEEE Transactions on Knowledge and Data Engineering, 2005, 17(12):1624-1637
- [16] GOLUB G H, LOAN C F V. Matrix computations[M]. 3rd ed., Maryland:Johns Hopkins University Press, 1996

(责任编辑:湛 江)