

基于稀疏重构的超图谱聚类方法

王灿田¹ 孙玉宝^{1,2} 刘青山¹

(南京信息工程大学 南京 210044)¹

(南京理工大学江苏省光谱成像与智能感知重点实验室 南京 210094)²

摘要 超图谱聚类方法由于能很好地描述数据点间的高阶信息,近年来受到了广泛的关注。不同于传统图结构,超图结构中的超边不是两两数据点间的连接,而是一组具有某种相同特性的数据子集。在实际应用中,常用K-近邻来构建超图中的超边,因此,并没有考虑到数据内在的关联性。提出一种新的基于稀疏重构的超图构建方法。对每一样本,用稀疏表示来找到与其最有关联的近邻样本,以此形成基于稀疏重构的超图模型,使得每个超边内的样本都具有很强的关联性。最后通过对超图拉普拉斯矩阵进行谱分解得到聚类结果。在人脸数据库、手写体数据库上的实验结果验证了算法的有效性。

关键词 超图,稀疏表示,谱聚类

中图法分类号 TP391.4 文献标识码 A

Hypergraph Spectral Clustering with Sparse Representation

WANG Can-tian¹ SUN Yu-bao^{1,2} LIU Qing-shan¹

(Nanjing University of Information Science and Technology, Nanjing 210044, China)¹

(Jiangsu Key Laboratory of Spectral Imaging & Intelligent Sense, Nanjing University of Science and Technology, Nanjing 210094, China)²

Abstract Hypergraph spectral clustering method attracts much attention, because it can effectively describe high-order information among the data. Different from traditional graph model, hyperedge in hypergraph is not a pair-wise link between two data points, while it is a subset of data points sharing with some attribute. In practices, hyperedge is usually built by simple K-NN clustering, so it does not consider inherent relationship among the data. We proposed a new hypergraph spectral clustering algorithm with sparse representation. For each data point, sparse representation was used to seek its related neighbors to form a hyperedge, so the data points in a hyperedge have strong dependency. Finally, the spectral decomposition was performed on the Laplace matrix of the hypergraph to obtain the clustering result. Extensive experiments on face database and handwriting database demonstrate the effectiveness of the proposed method.

Keywords Hypergraph, Sparse representation, Spectral clustering

1 引言

近年来,谱聚类成为机器学习和计算机视觉领域非常活跃的一个研究方向。谱聚类算法建立在图模型理论的基础上,根据给定的样本数据集,通过相似度计算,在数据点与点之间建立连接关系(边),从而形成直观的图模型。图谱聚类就是将聚类问题转化为图的最优划分问题,在数学上是通过对图拉普拉斯矩阵进行谱分解来实现的。在机器学习和视觉领域已经提出了很多经典的谱分析方法,如等尺度映射(ISO-MAP)^[17]、局部线性嵌套(LLE)^[2]、拉普拉斯特征映射(Laplacian eigenmap)^[1]、规范化分割(Normalized Cut)^[4,5]等。Normalized Cut方法不仅能满足样本内的相似度最大,

而且让样本间的相异度最大,因此受到了广泛的关注。考虑到基于高斯距离的相似性度量对噪声特别敏感,文献[10]中提出一种基于稀疏编码的图模型(L1-graph),它主要通过系数分解中的稀疏系数矩阵构建图的权重矩阵进行相似性度量,达到从数据自身信息构建图结构的目的。但是,传统图模型仅仅考虑数据间的成对关系,即二元关系。在实际应用中,数据通常是非常复杂的,因此,基于点对的图模型不能有效描述数据间的复杂关系,从而影响了最终的聚类分析性能。

为了有效克服传统图模型在多元关系知识表示上的缺陷,超图模型应运而生,拓展了经典图模型中两两顶点组建边的构图方式,以具有某种相似属性的数据子集构建超边,从而可以有效刻画数据间的高阶关系。超图学习算法最初应用于

到稿日期:2013-05-20 返修日期:2013-08-16 本文受国家自然科学基金项目(21004401),模式识别国家重点实验室开放课题基金(201204234),江苏省杰出青年基金(SBK201210296),中国博士后基金(20110491429),江苏省光谱成像与智能感知重点实验室(南京理工大学)基金(30920130122003)以及江苏省优势学科建设工程资助。

王灿田(1987-),男,硕士,主要研究方向为基于图/超图模型的图像分析;孙玉宝(1983-),男,博士,讲师,主要研究方向为矩阵低秩稀疏分解模型及其应用、视频处理与分析等,E-mail:sunyb@nuist.edu.cn(通信作者);刘青山(1975-),男,教授,博士生导师,主要研究方向为图像分析、基于图/超图模型的视觉理解、医学图像分析、基于事件的视频分析等。

超大规模集成电路的设计和合成^[25]以及网络节点问题当中。S. Agarwal 等人^[6,7]第一次将超图的想法应用到计算机视觉领域,通过“clique average”的方式将超图转化为简单图,再利用归一化割的思想来解决聚类问题。在文献[15]中,D. Zhou 等人第一次将谱聚类思想和超图结合,并且从随机游走的思想出发对超图分离算法做了详细的证明,为后来超图学习算法的应用奠定了很好的基础。Y. Huang, Q. Liu^[12,16,21]将图像聚类问题看成超图分离问题,将每一幅图像和它的 K 近邻基于形状和外观描述子形成两种不同的超边。在文献[8]中, Yu 提出了一种基于自适应超图学习的方法并将其应用于图像分类上,它主要是通过改变近邻的尺寸生成一组超边,通过半监督学习的方式自动调节不同超边对图像分类结果产生的影响。对超图学习来说,最关键的问题就是如何构建超边。不同的超边构建方式就会得到不同的超图结构模型。目前,最常用的超边构建方法就是使用 K 近邻(K-nearest-neighbor/ KNN)^[8,12,16,21]。以一个数据点为中心,把它和它的 K 个近邻放在一起形成一条超边。虽然,基于 KNN 的超边构建很简单直接,但是它并没有考虑到数据点的分布特性和数据点间的内在关联性,因此,可能会导致如下 3 个问题:

1) 噪声的鲁棒性。噪声对于视觉数据而言是不可避免的。基于 K 近邻构建的超图是通过计算特征空间的欧氏距离来进行相似性度量,它对噪声数据十分敏感,因此鲁棒性较差。

2) 数据近邻的自适应。由于不同特征空间的数据分布概率相差很大,导致每个数据都有自己的近邻结构。由于 K 近邻超图的 K 值固定,因此它无法自适应于数据的局部结构,对于分布不均匀的数据,用简单的 K 近邻来构建超边是不合理的。

3) 权重矩阵的稀疏性。权重矩阵的稀疏性对于超图谱分解的计算复杂度影响很大,K 近邻超图权重矩阵稠密,谱分解复杂度高,不利于数据的处理。

为了解决这些问题,本文提出一种新的基于稀疏重构的超图谱聚类方法。通过稀疏学习,得到每个数据的稀疏重构数据点,我们把每个数据点和它的稀疏重构点形成一条超边,因此,超边内的数据点间就存在着很强的关联性。而且,基于稀疏重构的超图谱聚类方法还具有如下两点良好特性:1) 通过对数据样本求解 L1 范数优化问题得到数据的重建系数,利用非零系数的分布结构代替两点之间的欧氏距离进行相似性度量,有效提高了对噪声的鲁棒性;2) 利用稀疏分解中少量的非零系数构建超边,形成稀疏的超图,在达到表示高阶耦合关系的同时实现了数据近邻自适应的效果。

2 超图的基本概念

一个超图 $G=(V, E, W)$ 是由顶点集合 V 、超边集合 E , 以及超边权重矩阵 W 构成。每一个超边 e_i 都有一个权重 w (e_i)。如图 1(b) 所示,超边 e_1 由顶点 v_1, v_2 组成, e_2 由 v_5, v_6, v_7 3 个点组成, e_3 由 v_2, v_3, v_4, v_6 4 个点组成。超图 G 的连接关系表示为 $|V| \times |E|$ 的关联矩阵 H , 定义如下(见图 1(c)):

$$H(v, e) = \begin{cases} 1, & \text{if } v \in e, \\ 0, & \text{if } v \notin e \end{cases} \quad (1)$$

根据关联矩阵 H , 每一个顶点 $v \in V$ 的度表示为:

$$d(v) = \sum_{e \in E} w(e) H(v, e) \quad (2)$$

每个超边 $e \in E$ 的度表示为:

$$\delta(v) = \sum_{v \in V} H(v, e) \quad (3)$$

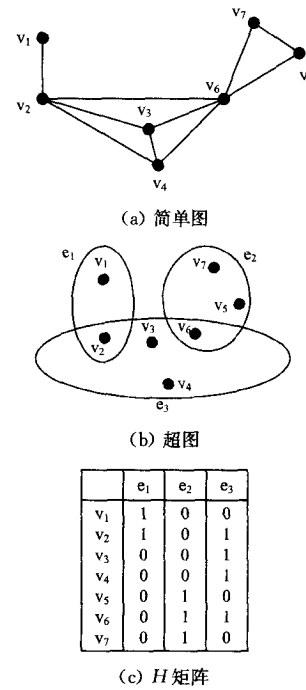


图 1 简单图与超图的边连接关系比较

图 1(a) 表示的是传统的简单图结构。相对于传统的图模型来说,超边的构建与超边权重计算是超图建模中的两个关键问题,当前广泛使用的 K 近邻(K-nearest-neighbor)超图通过搜索样本的 K 个近邻构建超边,超边的权重计算通常是基于超边内顶点间相似度的和。但是,邻域 K 值的选取并没有考虑数据的实际分布结构,并且超边的数量也会随着顶点数的增加而增加,从而大大增加了计算复杂度。

3 本文工作

在实际问题中,数据往往都存在一定的结构关联性。稀疏性是结构化数据(如人脸、手写体数字等)的一种重要的先验属性,即样本数据可以通过样本空间中的少数样本进行稀疏表示。基于这一先验,本文提出一种新的基于稀疏重构的超图谱学习方法。

3.1 基于稀疏重构的超边构建

不同于基于 K 近邻的超边构建方法,基于稀疏重构的超边构建思想如下:

1) 输入: 给定一个已知的样本数据集 $X = [x_1, x_2, \dots, x_n] \in \mathbb{R}^{d \times n}$ 。

2) 鲁棒稀疏表示: 对于样本数据中的每一列 x_i , 通过求解 L1-范数优化问题来进行鲁棒稀疏重构:

$$\begin{aligned} \min_{Z, E} \|Z\| + \lambda \|E\|_{2,1} \\ \text{s. t. } X = AZ + E \end{aligned} \quad (4)$$

其中, $Z = [z_1, z_2, \dots, z_n]$ 是重建系数矩阵, A 是一个完备的基矩阵(也可以看作字典), $\|E\|_{2,1} = \sum_{j=1}^n \sqrt{\sum_{i=1}^m ([E]_{i,j})^2}$ 是 $l_{2,1}$ -范数, $\lambda > 0$ 为正则化参数。E 为稀疏的误差分量,用于分离样本集中的少数异常样本点与显著误差,进而形成鲁棒的表示。在计算机视觉中,结构化的样本数据,如人脸、手写体等,通常呈现出子空间聚类分布结构,相同类型的样本数据近

似位于同一低维子空间,每一样本点可以通过同类型的样本进行有效表示。因此本文选用数据样本集自身作为字典进行稀疏分解,所以式(4)的鲁棒分解模型可改写为:

$$\min_{Z,E} \|Z\| + \lambda \|E\|_{2,1} \quad (5)$$

$$\text{s. t. } X=Z+E, \text{diag}(Z)=0$$

其中, Z 的对角线元素为零,要求在稀疏表示过程中每个数据点 x_i 只能使用除自身以外的数据样本 $x_j(1 \leq j \leq n, j \neq i)$ 进行稀疏表示,通过求解上述优化问题,得到系数矩阵 Z 。

3) 超边构建:通过鲁棒稀疏表示,得到重建系数矩阵 Z ,每一列系数 Z_i 揭示了样本 x_i 同其余样本间的稀疏重构关系,少数具有非零大系数的样本同 x_i 的关联最为密切,为 x_i 的近邻样本,可与其组合为一条超边。因此,针对每个数据样本,以其为中心与每一列中的非零大系数对应的样本数据组合为一条超边,进而构建一个加权的无向超图 $G=(V,E,W)$ 。其中 $V=\{v_i\}_{i=1}^n$ 是一个点的集合,每一个顶点 v_i 对应于样本数据集中的 x_i 。 E 是超边集合,每一个超边 e_i 对应一个超边权重 $w(e_i)$ 。 W 是超图权重矩阵,它是由 $w(e_i)$ 构成的对角矩阵。

3.2 超边权重计算

实际中,样本数据 x_i 中的噪声通过稀疏编码之后常常表现为系数向量 z_i 中一些较小的值。因此,需要定义一个准则来消除噪声的干扰,从而构建关联矩阵 H 。这个准则定义如下:

$$z_{ij} = \begin{cases} z_{ij}, & \text{if } z_{ij} > \theta \\ 0, & \text{otherwise} \end{cases} \quad (6)$$

其中, θ 是一个经验阈值。经过这样处理之后得到一个稀疏的系数矩阵 Z^* 并通过其得到 $|V| \times |V|$ 的相似矩阵 $A=[Z^*+(Z^*)^T]/2$ 。

图2比较了YALE人脸数据集中 K 近邻超图和稀疏超图的相似矩阵,可以发现相比于 K 近邻超图,由于稀疏超图的超边是由少量非零样本构成,其相似矩阵是稀疏的,因此它对噪声有更好的鲁棒性。同时相似矩阵的稀疏性,对于超图谱分解的计算复杂度影响很大, K 近邻超图权重矩阵稠密,谱分解复杂度高,因此稀疏超图的方法更利于数据的处理。



(a) K 近邻超图的相似矩阵

(b) 稀疏超图的相似矩阵

图2 稀疏超图与KNN超图的相似矩阵比较

针对每个数据样本,以其为中心构建一条超边,超边中的顶点由本样本及稀疏分解中系数超过阈值的样本组成。因此稀疏超图关联矩阵定义如下:

$$h(v_i, e_j) = \begin{cases} A(j, i), & \text{if } v_i \in e_j \\ 0, & \text{otherwise} \end{cases} \quad (7)$$

由关联矩阵 H 得到了超图结构之后,下面由相似矩阵 A

计算超边权重 W :

$$w(e_i) = \sum_{v_i, v_j \in e_i, i > j} A(i, j) \quad (8)$$

其中,权重的计算为超边中两两顶点相似性的累加和。

3.3 超图谱聚类算法

给定超图模型,目前常用的超图谱聚类算法有3种:Boila's Laplacian^[26]是一种无权重超图的拉普拉斯矩阵,通过对拉普拉斯矩阵进行特征值分解,利用特征向量与超图在欧氏空间的一一对应关系得到聚类结果。文献[9]中,Rodriguez将无权重超图转化为加权的图模型,利用简单图的谱聚类算法得到聚类结果。相比于前两种基于超图最小割的谱聚类算法,Zhou^[15]吸收归一化割的思想,定义一个超图归一化拉普拉斯矩阵,对其进行特征分解,利用最小的 K 个特征向量构成新的矩阵,最后进行K-means聚类。由于文献[15]方法简单有效,本文采用文献[15]的超图谱聚类算法,具体步骤如下:

定义归一化的超图拉普拉斯矩阵 $L=I-D_v^{-1/2}/HWD_e^{-1}H^TD_v^{-1/2}$ ^[15], D_v 表示顶点的度构成的对角矩阵, D_e 表示超边的度构成的对角矩阵, W 表示超边权重构成的对角矩阵, I 为单位矩阵。

1) 求出超图拉普拉斯矩阵 L 中 K 个最小的非零特征值所对应的特征向量,并将它们构成矩阵 $C=[c_1, c_2, \dots, c_k]$ 。

2) 矩阵 C 中的每一行看作 K 维空间上的一个数据点,利用K-means算法将它们聚成 K 个类。

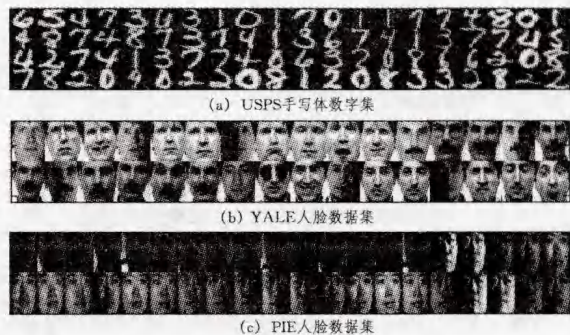
3) 最后,如果矩阵 C 中第 i 行属于第 j 类,那么 x_i 就属于第 j 类。

4 实验设计及结果分析

为了验证本文的方法,我们在3种不同的数据集上对其进行测试,并把本文的稀疏超图谱聚类方法(SHG)和基于主成分分析(PCA)+K-means的方法^[23]、基于高斯核的图谱聚类方法(G-graph)^[5]、拉普拉斯特征映射的谱聚类方法(LE-graph)^[1]、稀疏编码的图谱聚类方法(L1-graph)^[10],以及基于KNN的超图谱聚类方法(KNN-HG)^[8,12,16,21]进行实验比较。

4.1 数据集

实验中使用了图3中的3个不同数据集。



(a) USPS 手写体数字集

(b) YALE 人脸数据集

(c) PIE 人脸数据集

图3 实验中的3个不同数据集

1) USPS 手写体数据集。USPS 美国邮政服务手写数据集,共有9298个手写数字图像,数据集中均为 16×16 像素的灰度图像的值,且灰度已被归一化。

2) YALE 的人脸数据集。它由耶鲁大学计算视觉与控制中心创建,包含15位志愿者的165张图片,图片包含光照、表情和姿态的变化。为了验证实验结果,将所有的人脸图片裁剪为 32×32 像素和 64×64 像素两种大小的图片,构建两个

数据集。

3)PIE 的人脸数据集。数据集包含 41368 张人脸图片,每个人在 43 种不同光照下,有 13 种不同的姿势和 4 种不同的表情。本文选取的数据集仅仅包含 5 种近额姿势(C05, C07, C09, C27, C29),每个人有 170 幅图片,每一幅图片都有不同的光照条件和不同的表情。本文选择 C27 进行实验。

4.2 实验结果及其分析

本文采用文献[11]的方法来评价聚类结果,即使用准确性和标准化互信息来评价聚类结果。准确性(Accuracy/AC)主要用一个最优映射函数来使得 X 匹配 Y ,其中 X 是聚类结果的指示向量, Y 是标定结果的指示向量。标准化互信息(Normalized mutual information/NMI)用来度量聚类结果的鲁棒性。与 AC 相比,NMI 对类的选取没有偏置,不需要预先匹配 X 和 Y ,当两个类标签一一对应时,NMI 值取最大值 1;当两个类别结果相差很多时,NMI 的值取最小值 0。

表 1 和表 2 分别给出了 6 种算法在两个数据集上获得的 AC 和 NMI 值,每组实验获得的最高 AC 和 NMI 值用粗体字标出。比较表 1 和表 2 中的结果,我们可以明显看出:

1)基于稀疏超图的谱聚类算法得到的聚类结果所获得的 AC 和 NMI 基本高于其他算法,这说明了稀疏超图学习算法在数据聚类上的有效性。

2)比较 G-graph、LE-graph 和 L1-graph 3 个基于简单图划分的谱聚类结果的 AC 和 NMI 值,L1-graph 得到较好的聚类结果,这是因为前两者通过两两数据点之间欧氏距离进行相似性度量,对噪声很敏感,而后者通过稀疏编码进行相似性度量更能反映数据的结构信息,对噪声更具鲁棒性。

3)比较超图和简单图的谱聚类的 AC 和 NMI 值,基本可以看出基于超图的算法更有效,说明超图反映的数据的高阶信息对聚类有一定的影响。但是 L1-graph 的结果却比 KNN-HG 好,也说明稀疏表示能够有效反映数据间的结构信息。

4)总体来看,AC 和 NMI 值之间有很强的联系,即高的 NMI 一般能够得到高的 AC 值,且一般 NMI 要高于 AC。但是也有例外情况,例如 PCA+K-means 和 G-graph 在数据集 YALE 上 NMI 值高于 AC 值,但在数据集 USPS 上的 NMI 值却略低于 AC 值。这正说明了聚类问题的困难性,即直观上可行的聚类准则不一定得到最优解,而一个好的聚类准则函数的获得需要对不同准则函数的本质和使用这些函数生成的类簇有深刻的理解。

表 1 不同算法的 AC 值

数据集	PCA+ K-means	G- graph	LE- graph	L1- graph	KNN- HG	S- HG
USPS	0.6754	0.3340	0.6684	0.8262	0.7894	0.8736
YALE(32×32)	0.3818	0.2789	0.5091	0.5758	0.5152	0.5455
YALE(64×64)	0.5152	0.3515	0.4909	0.6181	0.5697	0.6224
PIE	0.3305	0.1736	0.9020	0.8887	0.7536	0.9195

表 2 不同算法的 NMI 值

数据集	PCA+ K-means	G- graph	LE- graph	L1- graph	KNN- HG	S- HG
USPS	0.6138	0.2350	0.8150	0.8303	0.8221	0.8982
YALE(32×32)	0.4480	0.3215	0.5397	0.6114	0.5597	0.6776
YALE(64×64)	0.5453	0.4115	0.5292	0.6285	0.5590	0.7158
PIE	0.6595	0.4869	0.9793	0.9710	0.9287	0.9831

结束语 本文提出一种新的基于稀疏重构的超图构建方法。对每一样本,用稀疏表示来找到与其最有关联的紧邻样

本,与其构建超边,形成基于稀疏重构的超图模型,再通过超图拉普拉斯矩阵进行谱分解得到聚类结果。通过在人脸数据库、手写体数据库上的实验结果表明,与其他基于图划分的算法相比,基于稀疏重构的超图谱聚类算法显示了更好的性能。

值得指出的是,本文提出的算法对于结构化数据显示了很好的性能。本文的算法同样可以用于解决其他应用领域的聚类问题,前提条件是对于不同领域的不同数据,如果能具有一定结构关联性,那么应用本文的方法将会得到很好的结果。

参考文献

- [1] Belkin M, Niyog P. Laplacian eigen maps for dimensionality reduction and data representation [J]. *Neural Comput*, 2002, 6(15): 1373-1396
- [2] Roweis S, Saul L. Nonlinear dimensionality reduction by locally linear embedding [J]. *Science*, 2000, 290(5500): 2323-2326
- [3] Tenenbaum J, Silva V, Langford J. A global geometric framework for nonlinear dimensionality reduction [J]. *Science*, 2000, 290(5500): 2319-2323
- [4] Shi J B, Malik J. Motion segmentation and tracking using normalized cuts [C]// *IEEE International Conference on Computer Vision (ICCV)*. 1998: 1154-1160
- [5] Shi J B, Malik J. Normalized cuts and image segmentation [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2000, 22(8): 888-905
- [6] Agarwal S, Branson K, Belongie K. Higher order learning with graphs [C]// *Int. Conf. Mach. Learn. Pittsburgh, PA*, 2006: 17-24
- [7] Agarwal S, Lim J, Zelnik M L, et al. Beyond pairwise clustering [C]// *Int. Conf. Comput. Vis. Pattern Recog. San Diego, CA*, 2005: 838-845
- [8] Yu J, Tao D C, Wang M. Adaptive Hypergraph Learning and its Application in Image Classification [J]. *IEEE Transactions on image processing*, 2012, 21(7): 3262-3272
- [9] Rodr'iguez J. On the laplacian spectrum and walk-regular hypergraphs [J]. *Linear and Multilinear Algebra*, 2003, 15(3): 285-297
- [10] Cheng B, Yang J, Yan S, et al. Learning with L1-graph for image analysis [J]. *IEEE Trans. Image Process*, 2010, 19(4): 858
- [11] Cai D, He X F, Han J. Active subspace learning [C]// *ICCV'09: IEEE International Conference on Computer Vision*. 2009: 911-916
- [12] Huang Y, Liu Q S, Metaxas D. Video object segmentation by hypergraph cut [C]// *Int. Conf. Comput. Vis. Pattern Recog. Miami, FL*, 2009: 1738-1745
- [13] Sun L, Ji S, Ye J. Hypergraph spectral learning for multilabel classification [C]// *Proc. Int. Conf. Know. Discov. Data Mining. Las Vegas, NV*, 2008: 668-676
- [14] Guan N, Tao D, Luo Z, et al. Manifold regularized discriminative non-negative matrix factorization with fast gradient descent [J]. *IEEE Trans. Image Process.*, 2011, 20(7): 2030-2048
- [15] Zhou D, Huang J, Schölkopf B. Learning with hypergraphs: Clustering, classification, and embedding [C]// *Neural Inf. Process. Syst. Vancouver, BC, Canada*, 2006: 1601-1608
- [16] Huang Y, Liu Q S, Lv F, et al. Unsupervised image category zation by hypergraph partition [J]. *IEEE Trans. Pattern Anal. Mach. Intell*, 2011, 33(6): 1266-1273

(下转第 156 页)

习速率 η 对剪枝算法以及测试精度的影响。实验结果显示, 本文提出的算法是行之有效的, 用该算法对网络进行剪枝不需要重复训练神经网络, 得到的网络结构紧凑, 具有较高的泛化能力。

参考文献

- [1] Kumar S. Neural networkw [M]. Beijing: Tsinghua University Press, 2006
- [2] Bishop C M. Neural networks for pattern recognition [M]. Oxford: Clarendon Press, 1996
- [3] Zhang G P. An investigation of neural networks for linear time-series forecasting [J]. Computers & Operations Research, 2001, 28(12): 1183-1202
- [4] Zanchettin C, Ludermir T B, Almeida L M. Hybrid training method for MLP: optimization of architecture and training [J]. IEEE Transactions on Systems, Man, and Cybernetics-Part B: Cybernetics, 2011, 41(4): 1097-1109
- [5] Yang S H, Chen Y P. An evolutionary constructive and pruning algorithm for artificial neural networks and its prediction applications [J]. Neurocomputing, 2012, 86: 140-149
- [6] Kwok T Y, Yeung D Y. Constructive algorithms for structure learning in feedforward neural networks for regression problems [J]. IEEE Transactions on Neural Networks, 1997, 8(3): 630-645
- [7] Reed R. Pruning algorithms-a survey [J]. IEEE Transactions on Neural Networks, 1993, 4(5): 740-747
- [8] Redding N J, Kowalczyk A, Downs T. Constructive higher-order network that is polynomial time [J]. Neural Networks, 1993, 6(7): 997-1010
- [9] Tsoi A C, Tan S. Recurrent neural networks; A constructive algorithm, and its properties [J]. Neurocomputing, 1997, 15(3/4): 309-326
- [10] Liu D R, Chang T S, Zhang Y G. A constructive algorithm for feedforward neural networks with incremental training [J]. IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications, 2002, 49(12): 1876-1879
- [11] Subirats J L, Franco L, Jerez J M. C-Mantec: A novel constructive neural network algorithm incorporating competition between neurons [J]. Neural Networks, 2012, 26: 130-140
- [12] Zhang R, Lan Y, Huang G B, et al. Universal approximation of extreme learning machine with adaptive growth of hidden nodes [J]. IEEE Transactions on Neural Networks and Learning Systems, 2012, 23(2): 365-371
- [13] Karnin E D. A simple procedure for pruning back-propagation trained neural networks [J]. IEEE Transactions on Neural Networks, 1990, 1(2): 239-242
- [14] Hagiwara M. A simple and effective method for removal of hidden units and weights [J]. Neurocomputing, 1994, 6: 207-218
- [15] Castellano G, Fanelli A M, Pelillo M. An Iterative Pruning Algorithm for Feedforward Neural Networks [J]. IEEE Transactions on Neural Networks, 1997, 8(3): 519-531
- [16] Zeng X Q, Yeung D S. Hidden neuron pruning of multilayer perceptrons using a quantified sensitivity measure [J]. Neurocomputing, 2006, 69(7-9): 825-837
- [17] Egrioglu E, Aladag C H, Gunay S. A new model selection strategy in artificial neural networks [J]. Applied Mathematics and Computation, 2008, 195(2): 591-597
- [18] Xing H J, Hu B G. Two-Phase Construction of Multilayer Perceptrons Using Information Theory [J]. IEEE Transactions on Neural Networks, 2009, 20(4): 715-721
- [19] Aladag C H. A new architecture selection method based on tabu search for artificial neural networks [J]. Expert Systems with Applications, 2011, 38(4): 3287-3293
- [20] Jeong C, Min J H, Kim M S. A tuning method for the architecture of neural network models incorporating GAM and GA as applied to bankruptcy prediction [J]. Expert Systems with Applications, 2012, 39(3): 3650-3658
- [21] 郑绪枝, 雷靖, 夏薇. 基于快速确定隐层神经元数的 BP 神经网络算法[J]. 计算机科学, 2012, 39(26): 432-436
- [22] 姚立忠, 李太福, 易军, 等. 神经网络模型的透明化及输入变量约简[J]. 计算机科学, 2012, 39(9): 247-251
- [23] Hornik K, Stinchcombe M, White H. Multilayer feedforward networks are universal approximators [J]. Neural Networks, 1989, 2: 359-366
- [17] Wright J, Yang A, Ganesh A, et al. Robust face recognition via sparse representation [J]. IEEE Trans. on PAMI, 2008, 31(2): 210-227
- [18] Zheng X, Cai D, He X F, et al. Locality preserving clustering for image database[C]//ACM Int. Conf. Multimedia. 2004: 885-891
- [19] Lin Z, Liu R, Su Z. Linearized alternating direction method with adaptive penalty for low rank representation [C]//NIPS. 2011: 612-620
- [20] 陈丽敏, 杨静, 张健沛. 一种基于加速迭代的大数据集谱聚类方法 [J]. 计算机科学, 2012, 39(5): 172-175
- [21] Huang Y, Liu Q S, Zhang S, et al. Image retrieval via probabilistic hypergraph ranking [C]// Int. Conf. Comput. Vis. Pattern Recog. San Francisco, CA, 2010: 3376-3383
- [22] Zheng M, Bu J, Chen C, et al. Graph Regularized Sparse Coding for Image Representation [J]. IEEE Trans. Image Process, 2011, 20(5): 1057-1149
- [23] Jolliffe I. Principal Component Analysis [M]. New York: Springer-Verlag, 1986: 1580-1584
- [24] Chen S, Donoho D, Saunders D. Atomic decomposition by basis pursuit [J]. Soc. Ind. Appl. Math. Rev. , 2011, 43(1): 129-159
- [25] Alpert C J, Kahng A B. Recent directions in netlist partitioning: A survey [J]. Integration: The VLSI Journal, 1995, 19(1/2): 1-81
- [26] Bolla M. Spectra euclidean representations and clustering of hypergraphs [J]. Discrete Mathematics, 1993, 117(1-3): 19-39
- [27] Zhuang L, Gao H, Lin Z, et al. Non-Negative Low Rank and Sparse Graph for Semi-Supervised Learning[C]//Proceedings of IEEE Conference on CVPR. Providence, RI, 2012: 2328-2335

(上接第 148 页)