

# 基于亲和传递聚类的多类物体识别方法

代松, 李伟生

(重庆邮电大学计算机科学与技术研究所, 重庆 400065)

**摘要:** 多类物体识别在提取特征之后, 样本的数量会呈指数倍增加, 为减少计算量同时, 不降低识别率, 采用亲和传递算法对样本数据进行聚类形成视觉字典, 帮助并提升物体识别效率。在 Sowerby 图像数据库上进行实验证明, 该方法与使用  $k$  均值聚类建立视觉字典方法相比, 在同等条件下具有更高的识别率。

**关键词:** 亲和传递聚类; 多类物体识别; 视觉字典

## Multi-class Object Recognition Method Based on Affinity Propagation Clustering

DAI Song, LI Wei-sheng

(Institute of Computer Science and Technology, Chongqing University of Posts and Telecommunications, Chongqing 400065)

**【Abstract】** The multi-class object recognition method usually suffers from the problem of curse dimensionality after extracting feature. This paper proposes a new approach for learning a discriminative model of object classes combined affinity propagation clustering. By using Affinity Propagation(AP), a representative visual vocabulary can be obtained. Experiment in Sowerby databases shows that it is superior to the recognition rate of k-means algorithms.

**【Key words】** Affinity Propagation(AP) clustering; multi-class object recognition; visual vocabulary

### 1 概述

近几年来基于场景的多类物体识别<sup>[1-2]</sup>已经成为国内外计算机视觉与模式识别领域研究的热点, 它可以广泛地应用于智能视频监控、车辆辅助驾驶、为盲人提供视觉伺服以及各类身份识别和认证系统等。在多类物体识别中, 由类内物体的差异性而带来的样本个数大量增加的问题已经成为多类物体识别中的难点之一。许多学者用各类滤波器响应、SIFT描述等方法<sup>[1-2]</sup>提取图像的原始信息之后, 通过  $k$  均值聚类方法形成视觉字典, 以方便以后基于小样本的学习和检测。但是  $k$  均值聚类函数是高度非线性和多峰值函数, 在初始化较好时, 算法才会收敛到全局最优解。对于样本空间是球形或超球形的数据具有较好的性能, 而在空间分布复杂的情况下聚类效果会很差。因此, 采用  $k$  均值聚类会使数据字典倾向于数据高密度区域, 样本数量少但识别特征显著的区域有可能被合并成一个区域, 使数据字典可区分性降低, 从而导致物体的识别率不高。

本文采用亲和传递聚类的算法, 该算法同属于  $k$  均值聚类方法, 但是它克服了  $k$  均值聚类方法的一些缺点, 在其迭代过程中不断地搜索适合的聚类中心, 使得聚类目标函数最大化。经过聚类后的数据变成了小样本, 减少了计算量, 保证了系统实验具有可行性, 这类似于人类感知机理, 人们能从少量几个到近百个样本中学习而推广到一般。仿真结果证明了该方法能够在减少计算时间的同时, 有效提高识别率。

### 2 亲和传递聚类算法

#### 2.1 算法思想

聚类是将待处理的数据分配到相应的类, 使得属于同一类别的个体之间的相似度尽可能大, 而不同类别的个体之间的相似度尽可能小。聚类分析预先不知道任何关于类如何分

类的先验知识, 无教师指导, 仅靠事物间的相似性作为类属性划分的准则, 属于无监督机器学习的范畴。

亲和传递聚类(Affinity Propagation, AP)是由 Frey<sup>[3-4]</sup>等人于 2007 年在 science 上提出的一种新的无监督聚类算法。亲和传递聚类算法的快速、有效性体现在处理大数据集的聚类问题上, 例如对数千个手写邮政编码的图片, 该算法只花费了 5 min 就可以找出能准确解释各种笔迹类型的少量图片, 而  $k$  均值算法到同样的时间内的达到的精度却很低。

亲和传递聚类算法把所有的数据点都作为潜在的聚类中心点, 假设把每个数据点都看作是有向图中的一个节点, 任意节点之间传递责任度和适合度 2 种消息, 在该图的有向边上不断递归地传递这 2 种消息和修改它们的值, 直到一个很好的中心和相应的聚类出现, 如图 1 所示。

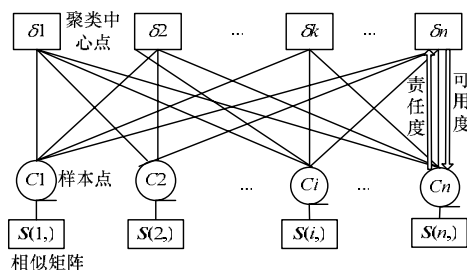


图 1 AP 算法主要思想

**基金项目:** 重庆市自然科学基金资助项目(CSTC.2006BB2365); 重庆市教委科学技术基金资助项目(KJ060504)

**作者简介:** 代松(1983-), 女, 硕士研究生, 主研方向: 模式识别, 人工智能; 李伟生, 教授

**收稿日期:** 2009-01-20 **E-mail:** daisong\_1@163.com

## 2.2 亲和传递聚类算法的过程

### 2.2.1 相似度矩阵的建立

亲和传递聚类算法首先建立一个  $N \times N$  相似度矩阵  $S$  作为工作基础,此相似度矩阵是由  $N$  个数据点之间的相似度组成的。数据之间的相似性  $s(i,k)$  表示下标为  $k$  的数据(下文简称为节点  $k$ )作为节点  $i$  的中心的适合程度。值越大,节点  $k$  越适合作为节点  $i$  的中心。其值可以由用户指定,也可以用负的误差平方来指定。对于每个节点  $k$  初始时定义一个优先值  $s(k,k)$ , 值越大,表明节点  $k$  越易于被选作聚类中心点。为了使所有数据点平等地参与聚类,让所有节点的优先值都相同,一般选择相似度矩阵  $S$  的均值为优先值。

### 2.2.2 节点责任度的计算

在构造好相似度矩阵和优先值之后,需要计算 2 类在节点间传递的消息。节点间的责任度  $r(i,k)$  是从数据点到候选中心点的一条边,表示含义为:考虑到其他潜在的中心点,节点  $k$  适合作为数据点  $i$  的类代表的程度。它的值计算如下:

$$r(i,k) = s(i,k) - \max_{j \neq k} \{a(j,i) + s(i,j)\} \quad (1)$$

其中,责任度  $r(k,k)$  表示节点  $k$  是否适合作为一个中心,如果  $r(k,k)$  小于 0,说明节点  $k$  更适合作为一般的数据点而不是中心点。

### 2.2.3 计算节点的可用度

节点间的可用度  $a(i,k)$  是从候选中心点到数据点的一条边,它表示含义是在考虑其他节点对候选中心  $k$  的支持度的同时,数据点  $i$  选择数据点  $k$  作为其类代表的适合程度。计算公式如下:

$$a(i,k) = \min \left( 0, r(i,i) + \sum_{j \in \{k, \cdot\}} \max \{0, r(j,i)\} \right) \quad (2)$$

其中,自可用度  $a(k,k)$  反映了节点  $k$  是否适合作为一个中心。定义如下:

$$a(k,k) = \max_{j \neq k} \{0, r(j,k)\} \quad (3)$$

### 2.2.4 算法过程表述

亲和传递聚类算法(对  $n$  个样本进行聚类):

**Step1** 计算出  $n$  个样本点任意 2 点间相似度矩阵  $s(i,j) = -\|x_i - x_j\|^2$ ;

**Step2** 对  $\forall i, k$ , 初始化可用度  $r(i,k)$  和适用度  $a(k,i)$  分别为 0;

**Step3** 根据式(1)更新责任度  $r(i,k)$ ;

**Step4** 根据式(3)更新可用度  $a(k,i)$ ;

**Step5** 给定一个节点  $i$ , 存在节点  $k$ , 使得  $a(i,k) + r(i,k)$

最大,那么可以确定节点  $k$  是  $i$  的中心;

**Step 6** 如果满足小于固定迭代次数,数据低于某个阈值,得到目标聚类个数 3 个条件之一,则算法停止,否则,返回 Step 3。

## 3 实验分析

### 3.1 实验数据及设备

为了验证本文算法的有效性,本实验使用了英国航空航天索尔研究中心的 Sowerby 图像数据库(该数据库可以免费使用。如果想获取,请联系 Dr. Andy Wright 或者 Dr. Gareth Rees, email: gareth-s.rees@baesystem.com)(简称 SID)。它取材于室外近郊公路附近两旁的典型场景,共有 104 张彩色图片,并用不同的颜色手工标定了这些图片中类的区域,标定物体

的类别为 7 类:天空,草地,树木,道路,建筑物,汽车,未标注的区域。实验图片的分类采用标准的分类方法,随机抽取 45% 的图片作为训练集,45% 的图片作为测试集,10% 的图片作为验证集。每幅图像的分辨率都是 96 像素  $\times$  64 像素。图 2 是 Sowerby 图片库中一副图像,图 3 是 Sowerby 图片库中人工标注图像。实验使用 CPU 为 2.4 GHz、内存为 1 GB 的计算机,算法使用 C#和 Matlab 实现。



图 2 Sowerby 图片库图片



图 3 手工标定图片

### 3.2 实验过程

本文为了验证亲和传递聚类在识别率上的效果,采用了文献[1]中使用的方法,在相同的实验方法下,做了对比实验。训练时首先通过滤波器组获得滤波响应,共产生 17 维的特征向量,见表 1。

表 1 17 维滤波器的构成

滤波器名称	滤波器的尺度
高斯滤波器	$\kappa$ 和 $2\kappa$
$x$ 和 $y$ 高斯导数滤波器	$2\kappa, 4\kappa$
拉普拉斯高斯滤波器	$\kappa, 2\kappa, 4\kappa, 8\kappa$

接着用亲和传递聚类算法、原始的  $k$  均值、模糊  $c$  均值算法分别形成视觉字典。随后结合图片的纹理、外观、颜色、位置、边缘轮廓信息使用条件随机域建立物体计算模型,可以得到任意一副图片中  $x$  像素点标记为  $c$  的条件概率为

$$\log P(c|x, \theta) = \sum_i \overbrace{\varphi_i(c_i, x; \theta_\varphi)}^{\text{纹理外观}} + \overbrace{\pi(c_i, x_i; \theta_\pi)}^{\text{颜色}} + \overbrace{\lambda(c_i, i; \theta_\lambda)}^{\text{位置}} + \sum_{(i,j) \in \mathcal{E}} \overbrace{\phi(c_i, c_j, g_{ij}(x); \theta_\phi)}^{\text{边缘}} - \log Z(\theta, x) \quad (4)$$

其中,  $Z(\theta, x)$  是分配函数,它保证模型有适合的归一化概率;  $\theta$  是模型参数;  $\varphi$  是特征函数;  $\theta_\varphi$  和  $\theta_\phi$  是需要学习的参数。

CRF 模型通过下式标记的最大似然函数来估计:

$$L(\theta) = \sum_n \log P(c_n | x_n, \theta) + \log P(\theta) \quad (5)$$

把式(4)带入式(5)可以估计出参数  $\theta_\varphi$  和  $\theta_\phi$ 。

测试时使用由 Antonio Torralba<sup>[5]</sup>提出的 JointBoosting 学习物体识别算法,该算法能在不同的类之间实现特征共享。识别率的计算方法是所有测试图片标记为正确的像素的百分率的平均值。训练过程与测试过程如图 4、图 5 所示。

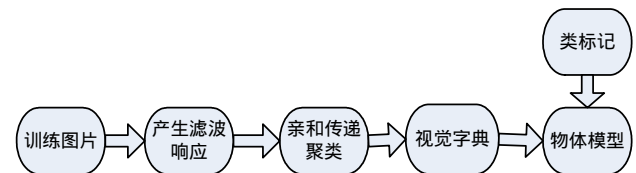


图 4 训练过程

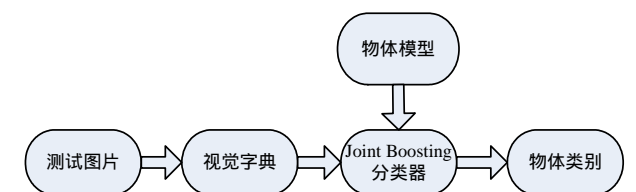


图 5 测试过程

### 3.3 实验结果及分析

本文将亲和传递聚类算法分别与原始的  $k$  均值、模糊  $c$  均值进行了比较。其中,亲和传递聚类算法的参数设置如下:亲和传递聚类中亲和成熟条件迭代次数为 500 次。FCM 的参数为:隶属度矩阵  $U$  的指数 2.0,最大迭代次数 100,隶属度最小变化量 0.000 01。分别用不同的聚类方法与亲和传递聚类算法进行了比较,如表 2 所示。

表 2 实验结果

方法	样本个数	聚类中心个数	识别率/(%)
$k$ 均值	3 290	116	74.81
FCM	3 290	116	75.90
AP	3 290	116	76.70
$k$ 均值	4 136	138	73.61
FCM	4 136	138	73.62
AP	4 136	116	75.58

由上面的实验可以得出下面的结论:

(1)在相同的样本个数的条件下, $k$  均值聚类效果最差,模糊  $c$  均值居于中间,亲和传递聚类算法的识别率效果最好。这是因为  $k$  均值聚类的初始化中心样本是随机选取的,在此基础上进行聚类,不一定会产生很好的效果。FCM 聚类虽然是一种无监督聚类,但是 FCM 算法却要求聚类原型参数的先验知识,在不知道数据样本的先验知识的情况下只能假设一些参数,而这种假设未必与样本数据相一致。亲和传递聚类算法克服了  $k$  均值和 FCM 的一些缺点,不仅选择实际的点作为聚类中心点,而且初始时不需要指定聚类的个数,完全是样本根据自己特点自发的聚类,使得物体识别效率得到了提高。

(2)在不同样本个数的条件下,识别率并不随样本增加。原因是由于在特征选择时,仅通过纹理滤波器的滤波,不一定能够很好地获得到代表类别之间差异性最好的信息。因此,

以后可以在选择图像的原始信息上加以改进。

### 4 结束语

本文采用了一种新的 AP 聚类算法应用于多类物体识别算法中,该算法在无监督的条件下快速有效地对大量的样本进行聚类,形成了较好的代表样本信息的视觉字典,帮助并提高了物体识别率。下一步可以在如何获得更多的结合图片的纹理、空间外观的轮廓、上下文信息的特征,如何建立好的模型来有效表达大量类内结构变化大的物体等问题上做进一步深入的研究。

### 参考文献

- [1] Shotton J, Winn J, Rother C, et al. TextonBoost: Joint Appearance, Shape and Context Modeling for Multi-class Object Recognition and Segmentation[C]//Proc. of European Conf. on Computer Vision. [S. l.]: Springer, 2006: 1-15.
- [2] Win J, Criminisi A, Minka T. Object Categorization by Learned Universal Visual Dictionary[C]//Proc. of IEEE Int'l. Conf. on Computer Vision. Beijing, China: [s. n.], 2005: 1800-1807.
- [3] Frey B J, Dueck D. Clustering by Passing Messages Between Data Points[EB/OL]. (2007-02-20). [http://www.psi.toronto.edu/affinity\\_propagation/FmyDueckScience07.pdf](http://www.psi.toronto.edu/affinity_propagation/FmyDueckScience07.pdf).
- [4] 王开军,李健,张军英,等.半监督的仿射传播聚类[J].计算机工程,2007,33(23):197-199.
- [5] Torralba A, Murphy K P, Freeman W T. Sharing Visual Features for Multiclass and Multiview Object Detection[J]. IEEE Transactions on Pattern and Pattern Analysis and Machine Intelligence, 2007, 19(5): 854-869.

编辑 索书志

(上接第 172 页)

三元鉴别体制在 WLAN 中的访问控制通过对通信三方的环环相扣实现身份鉴别,每个实体都有 2 次签名验证过程,有效防止了伪造实体的侵入。

### 5 安全性分析

本文基于 ECDSA 密码体制的三元鉴别方法能够满足无线环境中的安全需要,抵挡常见的攻击手段。

(1)中间人攻击

窃听器伪装成 AP 试图将 AS $\rightarrow$ AP 分组中的 AP 签名字段  $S_{AP}(R1||R2||Cert_{STA})$  和  $Pub_{AP}$  替换为  $S_{AP'}(R1' ||R2' ||Cert_{STA})$  和  $Pub_{AP'}$ ,以修改来自 AS 的鉴别结果,但 STA 收到分组后将采用 ECDSA 的验证步骤检验随机数  $R1, R2$  的值,因此,窃听器无法通过冒充 AP 影响验证结果。窃听器如果同时伪装成 AP' 和 AS',进而引诱 STA 进入一个预先设置的无线网络陷阱中,STA 将通过保存的 AS 公钥验证 4.2 节步骤(5)中的  $S_{AS}(R2 || Res_{AP})$  判断 AS 的身份。因此,本方案可以很好地抵抗中间人攻击,抑制特殊情况下对 WLAN 的破坏行为。

(2)针对 ECDSA 的攻击

对 ECDSA 的主要攻击行为是自主选择消息攻击,攻击者访问签名机,随机地选择消息进行签名而获得合法签名/消息对,但是 ECDSA 通过选择大 Mersenne 素数,能够抵抗攻击者的各种模仿攻击。因此,在 WLAN 环境限制下,针对 ECDSA 的攻击显得很实际。

### 6 结束语

本文从 ECC 数学基础、密码机制和工程应用 3 个方面对 ECDSA 在无线网络中的应用做了研究,从 WLAN 工程实现的角度引入了有限域上的 ECC,描述了理想的椭圆曲线生成方法,并就略加改进的一种普通椭圆曲线数字签名算法的实现过程进行了介绍。以此为基础,介绍了在无线网络环境下的三元实体鉴别方法,通过该方法,提高了对 AP 和 AS 的身份认证,防范了中间人攻击等攻击手段。目前,正在对 ECDSA 的抗攻击型和其在 WLAN 环境下对数据带宽的影响进行定量的实验分析,在加强安全性的同时尽量避免数据在空中接口处造成的速率瓶颈。

### 参考文献

- [1] 赵美玲,张少武.基于 ECC 的组合公钥技术的安全性分析[J].计算机工程,2008,34(1):156-157.
- [2] 肖攸安.网络信息安全中的椭圆曲线公钥密码体系的研究[D].武汉:武汉理工大学,2003.
- [3] Silverman J H. The Arithmetic of Elliptic Curves[M]. New York: Springer-Verlag, 1986.
- [4] 国家质量监督检验检疫总局. GB 15629.11-2003 无线局域网媒体访问控制和物理层规范[S]. 2003.
- [5] Boussakta S, Holt A G J. New Transform Using the Mersenne Numbers[J]. IEE Proceedings: Vision, Image and Signal Processing, 1995, 142(6): 381-388.

编辑 张帆