

基于二次星座聚类提取信号特征参数的方法

单剑锋, 岳震林

(南京邮电大学 电子科学与工程学院, 江苏 南京 210003)

摘要:针对传统单一的聚类算法在低信噪比时对调制信号识别率低,以及进行特征提取的算法复杂并且难以实现的问题,为了提高星座调制信号在低信噪比的识别率,在对比了不同的 MPSK 和 MQAM 的星座图的差异性的基础上,提出了一种二次星座聚类提取信号特征值的方法。其利用改进核函数的 DENCLUE(基于密度的聚类)提取信号星座图中密度最大点作为 k 均值聚类的初始聚类中心,并通过 k 均值聚类实现了一种新型的提取特征值方法,提取出一组特征值,采用支持向量机构造信号识别模型,并进行分类识别。仿真结果表明,所提出的方法相比传统单一的聚类算法,尤其是在低信噪比下对星座图调制信号的识别率和健壮性较好,并且基于该算法的实际系统简单可靠,具有广阔的发展空间。

关键词:星座调制;聚类算法;调制识别;特征提取

中图分类号:TP911

文献标识码:A

文章编号:1673-629X(2017)06-0091-04

doi:10.3969/j.issn.1673-629X.2017.06.019

A Method of Extracting Feature Parameters of Signals Based on Two Constellation Clustering

SHAN Jian-feng, YUE Zhen-lin

(School of Electronic Science and Engineering, Nanjing University of Posts and Telecommunications,
Nanjing 210003, China)

Abstract: Aimed at the problems that traditional single clustering algorithm has the low recognition rate of the modulation signal in low SNR (Signal-Noise Ratio) and feature extraction algorithm is complex and difficult to achieve, in order to improve the recognition rate of the constellation modulation signal in low SNR, an innovative method for twice constellations clustering has been proposed based on the comparison of the differences of MPSK and MQAM constellation diagram, where improved kernel function DENCLUE (density-based clustering) is used to extract the maximum density of the signal constellation as the initial clustering center of K -means clustering. The new feature value extraction method has been implemented for acquisition of characteristic values. The support vector machine has been used to construct the signal recognition model and its classification and identification has been conducted. Simulation results show that the method is better than traditional single clustering algorithm, which is good at the recognition rate and robustness of constellation modulation signal in low SNR especially and that the practical system with this proposed algorithm is simple and reliable with broad development space in future.

Key words: constellation modulation; clustering algorithm; modulation recognition; feature extraction

0 引言

作为现代移动通信技术中的关键点—调制识别一直是通信领域的研究热点,特别进入 21 世纪以来,随着 3G、4G 通信的兴起,世界各个国家和地区的科学家、学者提出了很多调制识别通信信号方法^[1]。近年来, MPSK 和 MQAM 的研究发展日趋成熟,作为新兴的数字通信技术,其拥有频带利用率高、调制方式灵

活、抗噪性能优越等众多优点^[2]。未来将会在商用民用 LTE 通信、认知无线电、军用通信系统等领域内广泛使用^[3]。然而,对 MPSK 和 MQAM 信号识别率的高低依然是目前能否进行实际使用的一个关键所在。因此,学者们对 QAM 调制信号的识别也进行了广泛而深入的研究^[4-8]。例如,文献[9]使用减法聚类和粒子群结合提取星座图分类特征的方法,并使用函数进

收稿日期:2016-06-19

修回日期:2016-09-22

网络出版时间:2017-03-13

基金项目:国家自然科学基金面上项目(GZ212015)

作者简介:单剑锋(1967-),男,副教授,研究方向为电路与系统、智能信息处理、电路故障诊断;岳震林(1992-),男,硕士研究生,研究方向为智能信息处理。

网络出版地址: <http://cnki.net/kcms/detail/61.1450.tp.20170313.1546.058.html>

行判决分类。文献[10]基于遗传算法的自适应 C 均值聚类算法对 MQAM 进行星座识别。文献[11]使用模糊 C 均值聚类提取特征值,并使用神经网络进行机器学习和分类识别。然而上述文献只涉及了一种聚类识别方法,均值聚类法容易受初始聚类中心随机选取的影响,而陷入局部最优;DENCLUE(基于密度的聚类)必须选用样本中的点代替重构的星座图点,会导致聚类的结果不是最优;神经网络中也存在学习过程复杂缓慢等问题。

针对上述缺陷,提出了一种基于核函数估计的 DENCLUE 和 k 均值聚类的二次星座聚类算法,以提取 QAM、PSK 信号星座图中的特征值,并采用改进的支持向量机进行分类识别,提高了 MPSK、MQAM 等信号的调制识别率。

1 基于二次星座聚类提取调制信号特征参数的原理

对于某些基于星座图调制方式 MPSK/MQAM 的信号,采用基于二次星座聚类与支持向量机的新方法,实现对信号的调制识别,其识别系统模型如图 1 所示。



图 1 识别系统模型

1.1 基于二次星座聚类的特征提取

DENCLUE 是一种基于密度的快速独立的近似聚类方法^[4]。 k 均值聚类是一种传统的基于距离的聚类算法^[7]。针对 MPSK/MQAM 信号的调制识别领域,上述两种算法有明显的不足之处: DENCLUE 计算后得到的聚类中心个数与初始给定半径 r_a 影响较大,半径稍微增加就会导致密度的显著改变,不能准确还原星座图的真实图像(星座图可能存在变形、残缺),因此采用核函数估计方法,把每个测量对象都当作一个对周围区域的高密度指示器。 k 均值聚类中,聚类中心 k 的数目是事先给定的,但是在 MQAM、MPSK 信号的识别过程中,由于事先并不知道 M 是多少,无法给定聚类中心 k 的值,因此利用 DENCLUE 聚类算法给定一个初始的聚类中心,以提高 k 均值聚类的精度并且减少搜索聚类中心次数。在分析上述两种聚类算法的优缺点的基础上,将其结合并提出了一种二次星座聚类提取信号特征参数的方法:

(1) 利用 DENCLUE 选择密度最大的 k 个初始聚类中心。

step1: 对 MPSK/MQAM 信号进行信号预处理,得到信号具有 n 个数据的星座图。定义星座图中样本点 x_i

的密度指标为:

$$D_i = \sum_{j=1}^n \exp \left[- \frac{\|x_i - x_k\|^2}{(0.5 r_a)^2} \right] \quad (1)$$

step2: 利用式(1)计算每个点的密度指标,选择密度指标最高的样本点 x_{c_1} 作为第一个聚类中心, D_{c_1} 为其密度指标。其中, r_a 是一个正数,定义了该点的邻域半径,半径以外的数据点对该点的密度指标贡献较小,这里取:

$$r_a = \frac{1}{2} \min_k \{ \max_i \|x_i - x_k\| \} \quad (2)$$

令 x_{c_i} 为第 i 次选取的聚类中心, D_{c_i} 为其密度指标,利用式(3)对其他数据点的密度指标进行修正。

$$D_i = D_i - D_{c_i} \left[- \frac{\|x_i - x_{c_i}\|^2}{(0.5 r_b)^2} \right] \quad (i \neq c_i) \quad (3)$$

其中,常数 r_b 定义了一个密度指标显著减小的领域。通常设定 r_b 大于 r_a ,以避免出现太近的聚类中心对后续算法的干扰,这里取 $r_b = 1.4 r_a$ 。

step3: 在修正每个数据点的密度指标后,选定下一个聚类中心并不断重复上述过程,直至确定 k 个聚类中心。

(2) 利用 k 均值聚类进行特征提取。

step1: 对于上述给定的星座图,令 $J = 1$, 并选择上面确定的 k 个初始聚类中心 $Z_j(J)$, $j = 1, 2, \dots, k$; 并计算星座图中每个点与聚类中心的欧氏距离 $D(x_i, Z_j(J))$, $i = 1, 2, \dots, n, j = 1, 2, \dots, k$ 。

$$D(x_i, z_j(J)) = \sqrt{(Z_x - x_x)^2 + (Z_y - x_y)^2} \quad (4)$$

其中, Z_x, Z_y, x_x, x_y 分别为聚类中心和星座图中其他点的纵横坐标。

如果欧氏距离满足式(5),则 $x_i \in \omega_j$ 。

$$D(x_i, z_j(J)) = \min \{ D(x_i, z_j(J)) \}, j = 1, 2, \dots, k \quad (5)$$

step2: 令 $J = J + 1$, 计算 k 个新的聚类中心和误差平方和准则 J_c 的值:

$$z_k(J + 1) = \frac{1}{n} \sum_{i=1}^{n_j} x_i^{(j)}, j = 1, 2, \dots, k \quad (6)$$

并作以下判断:若 $Z_j(J + 1)$ 不等于 $Z_j(J)$, 则 $J = J + 1$, 并且重复进行该步骤。否则算法结束,得到 k 个聚类中心和每个聚类中心的类内点的集合。

step3: 当聚类中心数目为 k 时,计算每个类内的点到该类中心点的欧氏距离之和 s_1, s_2, \dots, s_k , 并求出距离之和的平均值 S_k :

$$S_k = (s_1 + s_2 + \dots + s_k) / k \quad (7)$$

step4: 对 S_k 进行归一化,定义单位欧氏距离的 S_k 为特征值 T_c 。

由于研究的是 MPSK/MQAM 信号的调制识别,所以聚类中心数 k 分别取 2, 4, 8, 16, 32, 64。即对接收到

的信号点分别在以上6种不同的信号点进行上述二次星座聚类运算,并分别计算出不同 k 值的特征值 T_k 。在不同的信噪比下,针对不同的聚类中心 k ,分别计算出6种调制方式的特征参数 $T_2、T_4、T_8、T_{16}、T_{32}、T_{64}$,如表1所示。

表1 特征参数

特征参数	调制方式	-2	0	2	5	10
T_2	BPSK	3 803	2 531	2 047	1 652	1 526
	8PSK	634.5	410.8	331.1	260.6	212.2
	4QAM	1 911	1 232	985.9	744.1	564.6
	16QAM	1 219	783.8	618.7	498.2	402.2
	32QAM	833.8	560.2	453.9	361.5	298.8
	64QAM	808.28	544.4	448.6	346.6	283.2
T_4	BPSK	999.4	652.2	519.4	407.6	280.8
	8PSK	163.1	104.5	80.34	57.95	40.34
	4QAM	489.2	298.9	217.3	131.8	47.22
	16QAM	316.12	198.2	149.2	112.49	80.65
	32QAM	216.6	143.8	112.4	85.11	65.02
	64QAM	211.78	137.8	111.5	79.86	59.26
T_8	BPSK	279.8	177.7	135.4	85.26	35.35
	8PSK	44.71	29.22	21.97	16.15	9.711
	4QAM	137.42	82.88	63.33	41.14	17.62
	16QAM	85.15	55.21	42.11	31.39	22.13
	32QAM	60.52	39.11	30.83	23.76	17.65
	64QAM	59.86	37.55	30.76	22.02	16.18
T_{16}	BPSK	75.57	48.67	37.26	24.55	11.22
	8PSK	12.03	7.778	5.875	4.342	2.806
	4QAM	37.06	22.71	17.32	10.89	4.878
	16QAM	23.61	14.69	11.23	8.471	5.068
	32QAM	16.25	10.41	8.288	6.212	4.627
	64QAM	15.24	10.11	7.886	5.791	4.219
T_{32}	BPSK	19.87	12.54	9.625	6.644	2.954
	8PSK	3.186	2.006	1.507	1.092	0.7332
	4QAM	9.535	5.915	4.478	2.876	1.228
	16QAM	6.176	3.777	2.883	2.167	1.427
	32QAM	4.188	2.721	2.132	1.616	1.149
	64QAM	4.113	2.623	2.117	1.491	1.088
T_{64}	BPSK	5.103	3.258	2.514	1.708	0.768 9
	8PSK	0.801 9	0.521 9	0.389 7	0.287 6	0.189 9
	4QAM	2.477	1.501	1.16 1	0.739 1	0.307 8
	16QAM	1.603	0.965 6	0.731 7	0.556 5	0.372 2
	32QAM	1.057	0.706 3	0.557 7	0.415 5	0.292 3
	64QAM	1.042	0.663 5	0.519 6	0.387 2	0.271 7

由表1可以看出,在一定的信噪比下,不同信号的六个特征参数有明显不同。因此,可以通过特征参数

$T_2、T_4、T_8、T_{16}、T_{32}、T_{64}$ 来实现对BPSK,8PSK,4QAM,16QAM,32QAM,64QAM等6种信号的调制识别。

1.2 支持向量机的训练与调制识别

(1)支持向量机的训练。

由于提取的信号特征值较多,因此利用改进的一对一支持向量机方法来解决多种信号的分类问题,并使用多个二分类器共同组成MQAM/MPSK信号的分类器^[12]。首先将上面经过二次星座聚类算法所提取的6个特征参数 $T_2、T_4、T_8、T_{16}、T_{32}、T_{64}$ 送入支持向量机,对支持向量机进行机器学习,将出现最多的那个类别作为该样本最终的分类结果^[13]。

(2)支持向量机的调制识别。

支持向量机训练完成后,将上述6种中的一种未知信号经过预处理后,利用上面的二次星座聚类进行特征提取,然后将6个特征参数 $T_2、T_4、T_8、T_{16}、T_{32}、T_{64}$ 作为输入,送入训练好的支持向量机,输出得到测试信号的调制类型。

2 仿真结果与分析

使用 Matlab 2014, Intel Core i5 四核处理器进行仿真实验,选取的待检测信号分别是 BPSK, 8PSK, 4QAM, 16QAM, 32QAM, 64QAM, 载波信号为正弦波, 载波频率为 2 000 Hz, 采样频率为 12 000 Hz, 码元速率为 500 bps, 在信噪比为 -2 dB, 0 dB, 2 dB, 5 dB, 10 dB 时, 对模糊 C 均值聚类算法和所提出的基于二次星座聚类进行信号特征提取的算法进行对比, 均使用基于支持向量机的多分类器进行分类识别, 每种情况各进行 500 次, 计算出在不同信噪比下该方法对各种 MQAM 调制的平均识别率, 结果如图 2 和图 3 所示。

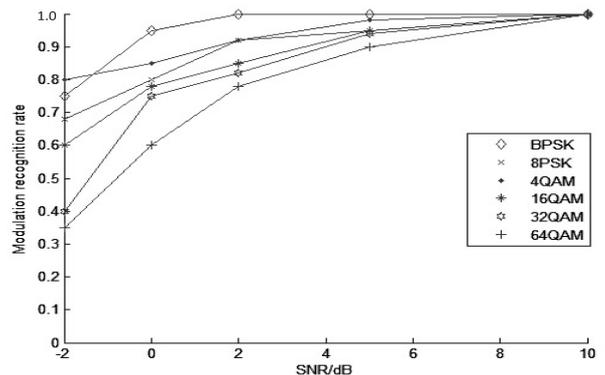


图2 采用模糊C均值聚类算法进行特征提取的调制识别率仿真图

图2、图3表明,当信噪比为2 dB时,二次星座聚类算法对BPSK,8PSK,4QAM,16QAM的识别率基本接近100%,有了大幅提高,对32QAM的识别率达80%以上,对64QAM的识别率将近80%,高于基于模糊C均值聚类单独的聚类算法;在信噪比为-2 dB时,对

BPSK, 8PSK, 4QAM, 16QAM 的识别率高于 90%, 32QAM, 64QAM 也有 70% 左右, 高于其他单一的聚类算法。充分体现了二次星座聚类算法在低信噪比下的健壮性, 并且保持了较高的信号识别率。

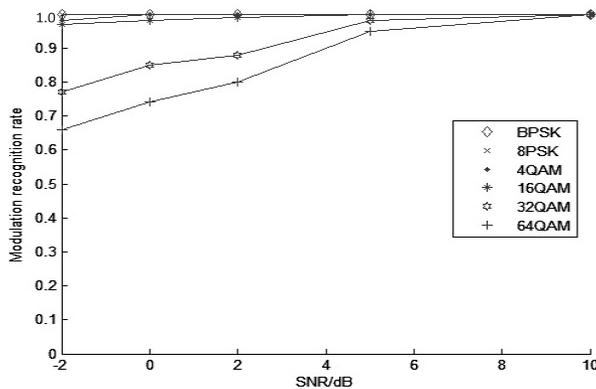


图 3 采用二次星座聚类算法进行特征提取的调制识别的仿真图

由于不同 PSK/QAM 信号的星座图的中心位置不同, 导致当聚类中心为 k 时, 其聚类中心位置不同, 因此不同星座图各点到该点的聚类中心的距离不同, 取平均并将其作为特征值, 据此来判断不同的信号。当信号的星座图的中心位置差异越大时, 该算法的正确率越高, 并且在低信噪比下, 通过两次星座聚类, 可以有效去除噪声的影响, 从而更加精确地获得星座图上信号的点, 因此该算法在低信噪比下依然保持了良好的识别率。并且首先使用了 DENCLUE 算法找出了密度极大值点, 在一定程度上给予了 k 均值聚类正确的初始聚类中心, 避免了 k 均值聚类随机选取聚类中心不正确而导致多次计算选取, 因此可以有效减少算法的迭代次数以及运算时间。

3 结束语

为了提高星座调制信号在低信噪比下的识别率, 提出了基于二次星座聚类提取信号特征值的算法。选取了通信中几种常用的基于星座图的调制方式, 利用二次星座聚类算法提取了具有明显差别的 6 种特征参数作为支持向量机的输入, 并采用多个二分类器结合的支持向量机对其进行训练, 以实现 MQAM、MPSK 调制信号的调制识别。

基于 Matlab 的多次仿真结果表明, 与传统单一的

聚类算法相比, 提出的识别方法在低信噪比下的抗噪声能力强, 信号识别率较高, 在实际的信号识别系统中具有一定的应用价值。

参考文献:

- [1] 杨杰, 刘珩, 卜祥元, 等. 通信信号调制识别—原理与算法[M]. 北京: 人民邮电出版社, 2014.
- [2] 党月芳, 徐启建, 张杰. 数字调制识别的关键技术与发展趋势[J]. 传感器与微系统, 2012, 31(10): 1-5.
- [3] 李苹苹, 孙钢灿, 申金媛, 等. 基于半监督聚类理论的 MQAM 信号的首识别[J]. 青岛科技大学学报: 自然科学版, 2014, 35(4): 405-409.
- [4] 杨发权, 李赞, 罗中良. 混合调制信号调制识别方法[J]. 中山大学学报: 自然科学版, 2014, 53(1): 42-46.
- [5] 盛宇波, 周澜, 程维文, 等. Efficient entanglement purification in quantum repeaters[J]. Chinese Physics B, 2012, 21(3): 67-73.
- [6] 王婷婷, 龚晓峰. 基于星座图的 PSK、QAM 信号联合识别算法应用[J]. 计算机应用研究, 2015, 32(7): 2116-2118.
- [7] Ericson K, Pallickara S. On the performance of high dimensional data clustering and classification algorithms[J]. IEEE Communications Letters, 2010, 12(5): 801-813.
- [8] Ahmadi N, Berangi R. Modulation classification of QAM and PSK from their constellation using genetic algorithm and hierarchical clustering[C]//ICITA. Cairns: IEEE, 2008: 98-103.
- [9] Li Yanling, Li Bingbing, Yin Changyi. Modulation classification of MQAM signals using particle swarm optimization and subtractive clustering[C]//IEEE 10th international conference on signal processing. [s. l.]: IEEE, 2010.
- [10] 吴月娟, 葛临东, 许志勇, 等. 基于遗传算法的自适应聚类与 MQAM 星座识别[J]. 计算机工程, 2007, 33(22): 39-41.
- [11] Yang Faquan, Li Zan. Method of neural network modulation recognition based on clustering and Polak-Ribiere algorithm[J]. Journal of Systems Engineering and Electronics, 2014, 25(5): 742-747.
- [12] 程佳. 支持向量机与 K-均值聚类融合算法研究[D]. 大连: 辽宁师范大学, 2008.
- [13] Zhu Zhechen, Aslam M W, Nandi A K. Support vector machine assisted genetic programming for MQAM classification[C]//10th international symposium on signals, circuits and systems. [s. l.]: [s. n.], 2011.