

# 通信中数字调制信号自动识别算法<sup>①</sup>

江 艳, 刘宏立, 张俊臣, 郭湘勇

(湖南大学 电气与信息工程学院, 长沙 410082)

**摘要:** 将信号的高阶累积量作为信号的特征参数具有很好的抗噪声性能。然而, 不同调制信号提取的高阶累积量参数可能相同, 造成不能完全识别。针对这种情况, 引进了分形理论。提出了将信号的高阶累积量和分形维数结合共同作为识别的特征参数。这种方法不仅很好的避免了高斯噪声对调制识别产生的影响, 并且解决了单一的将高阶累积量作为识别的特征参数中存在的不能完全识别所有调制类型的问题。仿真结果证明了算法的有效性。

**关键词:** 调制识别; 分形维数; 高阶累积量; 抗噪声性能; 特征参数

## Automatic Recognition Algorithm for Digital Modulation Signal in Communication

JIANG Yan, LIU Hong-Li, ZHANG Jun-Chen, GUO Xiang-Yong

(Department of Electrical and Information Engineering, Hunan University, Changsha 410082, China)

**Abstract:** It has a good noise immunity to take the high-order cummulants of the signal as it has characteristic parameters. However, different modulation type may have the same characteristic parameters, which result in partial identified. so, it is necessary to bring in fractal theory. The paper takes the combination of the high-order cummulants and the fractal dimension as the characteristic parameters of the signal. it not only avoid the Gaussian noise impact on the modulation recognition, but also solved the partial identified problem exist in the situation that take the order cummulants as the single characteristic parameters. Simulation results show the effectiveness of the algorithm.

**Keywords:** modulation recognition; fractal dimension; high-order cummulants; noise immunity; characteristic parameter

## 1 引言

对于通信系统来说, 调制方式是一个重要特征, 采用不同调制方式传输的信号往往体现出不同的信号特性。接收方如果要获取通信信号的信息内容, 就必须知道信号的调制方式和调制参数。通信信号调制的自动识别是指在未知调制信息内容以及调制参数的前提下, 判断出信号所采用的调制方式, 并估计出某些调制参数, 为解调器正确选择解调算法提供参数依据, 从而最终获得有用的信息内容。

近几十年来, 许多学者在通信信号的识别方面做了大量有益的探索, 提出了很多新思路和新方法<sup>[1-3]</sup>。但是, 由于非协作通信的复杂传输环境、新的调制方式不断涌现, 以及调制识别本身内在特性, 调制识别领

域远未发展成熟。表现为尚缺乏一个完整的研究体系, 如现有方法缺乏普适性、对假设条件依赖较多、识别性能没有一个统一的评价标准等。由于将信号的高阶累积量(大于二阶)作为信号的特征参数具有很好的抗噪声性能, 因而, 近些年来对基于高阶累积的调制识别算法的研究受到了广泛重视<sup>[4]</sup>, 然而不同调制信号提取的高阶累积量的参数可能相同, 造成不能完全识别的情况。因此, 针对这种情况, 本文将信号的高阶累积量和分形维数结合共同作为识别的特征参数, 解决了单一的将高阶累积量作为识别的特征参数存在的不能完全识别所有调制类型的问题。本文分析了算法的有效性, 并利用计算机仿真评估了算法的识别性能。

① 基金项目: 国家高技术研究发展计划(863)(2007AA1A121)

收稿时间: 2010-09-25; 收到修改稿时间: 2010-10-25

## 2 高阶累积量和分形盒维数基础知识

### 2.1 高阶累积量

受噪声污染的数字调制信号表达式为:

$$r(n) = s(n) + n(n) \quad (1)$$

式中,  $s(n)$  为发送序列,  $n(n)$  为零均值的复高斯白噪声, 且  $s(n)$  与  $n(n)$  相互独立。由文献[5], 累积量具有半不变性, 即:

$$cum(r(n)) = cum(s(n)) + cum(n(n)) \quad (2)$$

由于零均值高斯白噪声二阶以上的累积量均为零<sup>[6]</sup>, 因此, (2)式可以写为:

$$cum(r(n)) = cum(s(n)) \quad (3)$$

也即, 接收信号的高阶累积量等于发送信号的高阶累积量, 而不受高斯白噪声的影响。信号的调制方式决定其高阶累积量的值。故对被零均值高斯白噪声污染的信号建模, 从接收信号中提取高阶累积量作为特征参数可以有效地识别信号的调制方式。

对于一个零均值的复平稳随机过程  $X(t)$ , 定义其  $p$  阶混合矩为<sup>[7]</sup>:

$$M_{pq} = E\{[X(t)]^{p-q}[X^*(t)]^q\} \quad (4)$$

其累积量为:

$$C_{40} = cum(x, x, x, x) = M_{40} - 3M_{20}^2 \quad (5)$$

$$C_{42} = cum(x, x, x, x^*) = M_{42} - |M_{20}|^2 - 2M_{21}^2 \quad (6)$$

$$C_{63} = cum(x, x, x, x^*, x^*, x^*) = M_{63} - 6M_{20}M_{41} - 9M_{42}M_{21} + 18(M_{20})^2M_{21} + 12(M_{21})^3 \quad (7)$$

### 2.2 分形盒维数

分形维数是分形理论中的重要参数, 它有多种定义和算法, 它定量描述了分形集的复杂性, 能度量信号的不规则度, 常用的是 Hausdorff 维数、盒维数、信息维数, 相似维数、关联维数和广义维数等, 其中, Hausdorff 维数是分形理论中一种最基本的分形维数, 但在实际应用中, Hausdorff 维数的计算比较困难, 所以采用盒维数描述分形信号的几何尺度信息。

设  $(F, d)$  是一个度量空间,  $R$  是  $F$  的非空紧集族,  $\epsilon$  是一个非负实数, 令  $B(f, \epsilon)$  表示一个中心在  $f$ 、半径是  $\epsilon$  的闭球, 设  $A$  是  $F$  中的一个非空子集, 对于每个正数  $\epsilon$ , 令  $N(A, \epsilon)$  表示覆盖  $A$  的最小闭球的数目, 即:

$$N(A, \epsilon) = \left\{ M : A \subset \bigcup_{i=1}^M B(f_i, \epsilon) \right\} \quad (8)$$

其中,  $f_1, f_2, \dots, f_M$  是  $F$  中的不同点。

分形维数定义如下:

$$D(g) = \lim_{\epsilon \rightarrow 0} \left\{ \sup \frac{\log N(G, \tilde{\epsilon})}{-\log \tilde{\epsilon}} : \tilde{\epsilon} \in (0, \epsilon) \right\} \quad (9)$$

其中,  $G = \{(x, y) : x \in T \subset R, y = g(x) \subset R\} \subset R^2$ ,  $g$  为定义在  $R$  的闭集  $T$  上的连续函数,  $G$  为  $R^2$  上的集合。

根据盒维数<sup>[8,9]</sup>的定义, 对于数字化离散空间信号点集的分维值, 其简化式如下。设信号的采样序列为  $f(t_1), f(t_2), \dots, f(t_N), f(t_{N+1})$ ,  $N$  为偶数。令

$$d(\Delta) = \sum_{i=1}^N |f(t_i) - f(t_{i+1})| \quad (10)$$

$$d(2\Delta) = \sum_{i=1}^{N/2} (\max\{f(t_{2i-1}), f(t_{2i}), f(t_{2i+1})\} - \min\{f(t_{2i-1}), f(t_{2i}), f(t_{2i+1})\}) \quad (11)$$

则盒维数为:

$$D_B(f) = 1 + \log_2(d(\Delta)/d(2\Delta)) \quad (12)$$

计算的点数对盒维数有影响, 点数越多, 盒维数越稳定, 但太多的点数计算量太大。同时, 分形盒维数对噪声不敏感。

## 3 特征提取

### 3.1 信号模型

假定接收信号已经过预处理, 在已达到载波频率、相位、定时同步的前提下, 则经下变频在接收机匹配器输出端得到的基带信号可表示为:

MPSK 信号:

$$S_{MPSK} = \sqrt{E} e^{j\theta} a_k + n_0, \quad a_k \in \{e^{j2\pi(m-1)/M}, m=1, 2, \dots, M\} \quad (13)$$

MFSK 信号:

$$S_{MFSK} = \sqrt{E} e^{j\theta} e^{jw_k t} + n_0, \quad w_k \in \{(2m-1-M)\Delta w, m=1, 2, \dots, M\} \quad (14)$$

其中,  $\Delta w$  为 MFSK 信号的频偏。

16QAM 信号(矩形星座图):

$$S_{16QAM} = \sqrt{E}(a_k + jb_k)e^{j\theta} + n_0, \quad a_k, b_k \in \{-3, -1, 1, 3\} \quad (15)$$

其中,  $K=1, 2, \dots, N$ ,  $N$  为发送码元序列的长度,  $E$  是信号的平均功率,  $a_k$  为接收信号中信号有用信号平均功率归一化的码元序列,  $\theta$  是载波相位偏差。

### 3.2 高阶累积量特征

由 1.1 节的 (5) (6) (7) 式可得到各种调制信号的理论值, 则 2.1 中不同信号的值见表 1:

表 1 各调制信号的累积量值

信号类型	$ C_{40} $	$ C_{42} $	$ C_{63} $
BPSK	$2E^2$	$2E^2$	$16E^3$
QPSK	$E^2$	$E^2$	$4E^3$
2FSK	0	$E^2$	$4E^3$
4FSK	0	$E^2$	$4E^3$
8FSK	0	$E^2$	$4E^3$
16QAM	$0.68E^2$	$0.68E^2$	$2.08E^3$

为去除信号平均功率 E 的影响，目前已有的算法大多用  $C_{21}$  对各累积量进行功率归一法，而高斯白噪声的二阶累积量并不为零，所以用  $C_{21}$  对各累积量进行功率归一法引入了噪声的影响。从而违背了用高阶累积量作为特征参数以抑制噪声的初衷。因此本文选取的联合特征参数为：

$$M_1 = |C_{40}| / |C_{42}| \quad (16)$$

$$M_2 = |C_{63}|^2 / |C_{42}|^3 \quad (17)$$

由式(14)(15)得到各调制信号的联合特征参数见表 2。不同信噪比下的联合特征参数 M1、M2 仿真结果分别如图 1、图 2：

表 2 各调制信号的特征参数值

信号类型	$M_1$	$M_2$
BPSK	1	32
QPSK	1	16
2FSK	0	16
4FSK	0	16
8FSK	0	16
16QAM	1	13.7594

不同信噪比下的联合特征参数 M1、M2 仿真图如下：

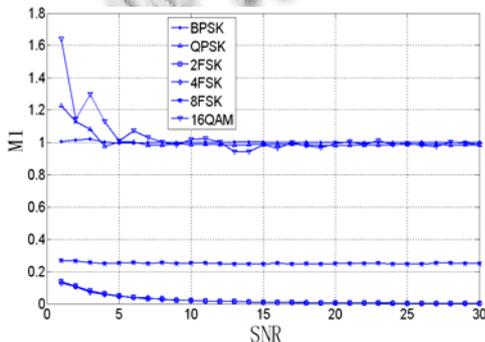


图 1 不同信噪比下 M1 的值

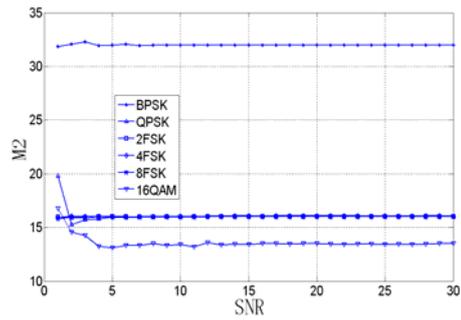


图 2 不同信噪比下 M2 的值

由图可以看出，仿真结果与理论值基本一致，由图 1 可看出，{BPSK, QPSK, 16QAM} 和 {MFSK} 可以明显区分开来，在图 2 中，{BPSK}、{QPSK}、{16QAM} 的值各不相同，这三种信号可以区分开来，而 MFSK 的各阶累积量相同，特征参数值 M1、M2 也相同，因而，单纯的利用这些高阶累无法进行 MFSK 的类内识别，为进一步进行 MFSK 的类内识别，可以计算它们的分形盒维数。

### 2.3 分形盒维数

MFSK 中用 M 个载频来表征 M 进制信号，它们的瞬时频率的变化数量不同，也可理解为其复杂度不同。因此，可以用表征复杂度的分形盒维数来表征这三种信号的频率上的差异。先提取信号的瞬时频率，再根据 1.2 节中的(10)式计算瞬时频率的分形盒维数。仿真结果如图 3，由仿真图可以看出，将 MFSK 信号的瞬时频率的分形盒维数作为识别的特征参量，在一定的信噪比范围内，可以较好的区分。

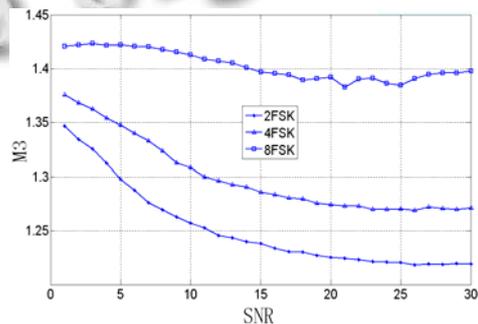


图 3 不同信噪比下 MFSK 的盒维数值

## 4 算法及仿真结果

### 4.1 算法描述

由前面的分析可知，要进行这六种调制信号的识别，首先要计算各信号的累积量，并利用它们的累积

量构造所最能表征信号最大差异的特征参数 [M1, M2], 利用 M1, M2 可以明显区分出 {BPSK}, {QPSK}, {16QAM} 和 {MFSK}, 单一的利用高阶累积量无法识别 MFSK, 对于区分不出来的 MFSK 可以用表征复杂度的分形盒维数 M3 来识别, 这样就实现了各信号的调制识别。分类识别算法如下:

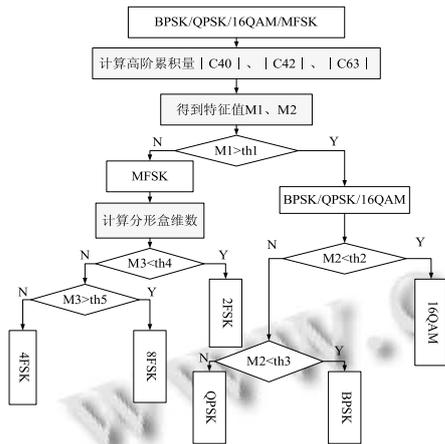


图 4 分类识别算法流程图

阈值的选取:

$$th1 = \frac{mean(M1_{BPSK}) - mean(M1_{8FSK}) + mean(M1_{8FSK})}{2} \quad (18)$$

$$th2 = \frac{mean(M2_{QPSK}) - mean(M2_{16QAM}) + mean(M2_{16QAM})}{2} \quad (19)$$

$$th3 = \frac{mean(M2_{BPSK}) - mean(M2_{QPSK}) + mean(M2_{QPSK})}{2} \quad (20)$$

$$th4 = \frac{mean(M3_{4FSK}) - mean(M3_{2FSK}) + mean(M3_{2FSK})}{2} \quad (21)$$

$$th5 = \frac{mean(M3_{8FSK}) - mean(M3_{4FSK}) + mean(M3_{4FSK})}{2} \quad (22)$$

其中, mean()表示取不同信噪比下信号的累积量的均值, 用于将六类信号区分为 {MFSK} 和 {PSK/QPSK/16QAM} 两大类, 和用于进一步识别 {BPSK/QPSK/16QAM} 中的三类信号, 同时, 对于 MFSK, 用高阶累积量是无法直接识别的, 因此, 计算它们的分形盒维数, 并设置阈值和识别 {MFSK} 中的三类信号, 这样就成功的对六类信号进行了识别。

#### 4.2 实验仿真

根据上述分析, 采用 matlab 对算法进行仿真验证。设信号的载波频率为 2kHz, 采样频率为 10kHz, 码元

速率为 500bps., 码元个数取 250 个。对六种信号随机抽取进行 1000 次仿真, 取正确识别的次数和抽取的次数之比作为正确的识别率, 且其中提取的用来识别的特征参数是经过 100 次计算后取平均值得到的。各信号在不同的信噪比下的识别率如图 5。

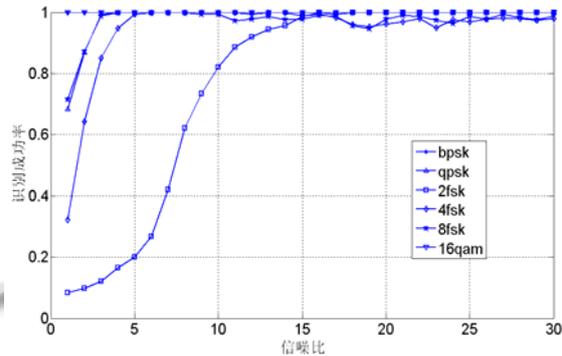


图 5 不同信噪比下各信号的识别率

由图 5 可以看出, 在整个信噪比范围内, 16QAM 和 BPSK 的识别率几乎达到了百分之百, QPSK、4FSK、8FSK 在 5dB 以上也接近完全识别, 识别效果相当好, 2FSK 在 10dB 以上识别效果较好。相比于单一的使用高阶累量来识别调制类型类算法, 本文方法能识别更多类型。而相比于单一使用分形维数识别调制类型类算法、本文方法识别性能更好、算法更简单。

#### 4 结论

本文针对单一的使用高阶累积量作为信号的识别参数并不能够完全识别信号的缺点, 将具有很好抗噪声性能的高阶累积量和对噪声不敏感的分形盒维数共同作为信号的特征参数, 利用提取的特征参数用决策树分类器对六种信号进行识别。理论分析和仿真实验证实了该方法的有效性。算法中特征参数较少, 算法较简单, 在一定的信噪比的条件下识别性能较好。

#### 参考文献

- 1 Swami A, Sadler BM. Hierarchical digital modulation classification using cumulants. IEEE Trans. on Communications, 2000,48(3):416-429.
- 2 Wong MLD, Nandi AK. Automatic digital modulation recognition using spectral and statistical features with multi-layer perceptrons. Signal Processing and Its Applications,

(下转第 160 页)

取不同值时的分类错误率如表4所示:

表4 COST(A)取不同值时模型预测准确率

COST(A)	1.0	1.2	1.4	1.6	1.8
总错误率(%)	14.71	15.69	17.64	20.58	21.56
A类错误率(%)	7.54	5.66	3.77	3.77	1.89
B类错误率(%)	22.44	26.53	32.65	38.78	42.86

从表4可以看出,随着COST(A)的不断增大,模型的预测总错误率是不断上升的,而A类错误率则呈下降趋势。接下来我们通过分析比较,找出最佳的COST(A):

1) 当COST(A) > 1.4时,总错误率超过了20%,这样的错误率太高了,显然不符合我们的选择标准。

2) 通过比较COST(A)分别为1.2和1.4的模型,发现当COST(A)等于1.4时,不仅总错误率相对较小,而且A类错误率也是最小的,因此,选择COST(A)为1.4是相对比较合适的。

参数调整后决策树的预测结果如表5所示:

表5 改进后模型预测结果

		预测类别		合计	错误率(%)	正确率(%)
		正常	异常			
实际类别	正常	33	16	49	32.65	67.35
	异常	2	51	53	3.77	96.23

改进后的模型与初始模型预测结果错误率对比表如表6所示:

表6 改进模型与初始模型预测错误率对比

	初始模型	改进模型	差额
总错误率(%)	14.71	17.64	-2.93
A类错误率(%)	7.55	3.77	3.78
B类错误率(%)	22.45	32.65	-10.2

(上接第64页)

2001,8(5):390-393.

- 陈卫东,杨绍全.利用累量不变量对MPSK信号分类.西安电子科技大学学报,2002,29(2):229-232.
- 一种基于高阶累积量的数字调相信号识别方法.系统工程与电子技术,2008,30(9):1611-1615.
- 张贤达.现代信号处理.第2版.北京:清华大学出版社,2002.
- 包锡锐,吴瑛,周欣.基于高阶累积量的数字调制信号识别算

表6可以看出,模型改进后,在总错误率稍有增加的情况下降低了A类错误率,是异常预测准确率达到96.23%,这种结果与我们的建模目标是吻合的。

## 5 结语

输电线路巡检工作需要有一个科学可靠的智能决策系统来辅助巡检工作的实施,本文尝试将粗糙集理论与C5.0决策树算法应用在输电线路工作状态预测中,初步取得了较为满意的效果,有较大的研究价值和实用价值,对于输电线路巡检计划的制定有较强的指导意义。本文采用决策树算法形成匹配规则,相对于参考文献一中的知识匹配方法,匹配效率有了较大的提升。

## 参考文献

- 徐家明,张树友,万昌江,傅顺强.基于知识的线路巡检计划层次推理技术研究.计算机工程,2003,29(19):45-47.
- 吴成东,许可.基于粗糙集与决策树的数据挖掘方法.东北大学学报,2006,(5):481-484.
- Pawlak Z. Rough sets. International Journal of Computer and Information Science, 1982,11(5):341-356. Ernesto Vazquez.
- Quinlan JR. C4.5: Programs for Machine Learning. San Mateo, CA: Morgan-Kaufmann Publishers, 1993:34-106.
- An online expert system for fault section diagnosis in power systems. IEEE Trans. on Power Systems, 1997,12(1): 357-362.
- Yao YY. On generalizing pawlak approximation operators. Lecture Notes in Artificial Intelligence 1998,1424:1998, 298-307.

法.信息工程大学学报,2007,8(4):463-467.

- 陈筱倩,王宏远.基于联合特征向量的自动数字调制识别算法.计算机应用研究,2009,26(7):2478-2480.
- 郭双冰.基于小波和分形理论的调制信号特征提取方法研究.信号处理,2005,21(3):316-318.
- 吕铁军,郭双冰,肖先赐.基于复杂度特征的调制信号识别.通信学报,2002,23(1):111-115.