

## 基于支持向量聚类的多分量线性调频信号检测

王令欢<sup>①</sup> 马红光<sup>①</sup> 张欣豫<sup>①</sup> 张葛祥<sup>②</sup>

<sup>①</sup>(第二炮兵工程学院 西安 710025)

<sup>②</sup>(西南交通大学电气工程学院 成都 610031)

**摘要:** 为了精确获取多分量线性调频(Linear FM, LFM)信号中分量的数量, 该文引入支持向量聚类(Support Vector Clustering, SVC)算法对 LFM 信号的 Radon-时频分析结果进行聚类分析, 完成多个分量的检测; 并通过减少 SVC 算法中输入集样本数量和改进聚类标识方法为“直接聚类标识”法, 提高了 SVC 算法的计算效率。仿真结果表明: 在较低信噪比条件下, Radon-时频分析和 SVC 结合的方法可有效地检测多分量 LFM 信号中分量数和进行参数估计。

**关键词:** 支持向量聚类; 聚类标识; 信号检测; 参数估计

中图分类号: TN911.23

文献标识码: A

文章编号: 1009-5896(2007)11-2661-04

## Multi-component Linear FM Signal Detection Based on Support Vector Clustering

Wang Ling-Huan<sup>①</sup> Ma Hong-Guang<sup>①</sup> Zhang Xin-Yu<sup>①</sup> Zhang Ge-Xiang<sup>②</sup>

<sup>①</sup>(The Second Artillery Engineering College, Xi'an 710025, China)

<sup>②</sup>(School of Electrical Engineering, Southwest Jiaotong University, Chengdu 610031, China)

**Abstract:** The Support Vector Clustering (SVC) algorithm is introduced to get the number of the pinnacles in the result of the time-frequency analysis and Radon transform of the multi-component Linear FM (LFM) signal, and to finish the detection of the components of the LFM signal. Meanwhile, the preprocessing to reduce the points' number of the input data-set for SVC is proposed to improve the computation efficiency. And a novel cluster labeling method is developed to improve the SVC algorithm. The simulation results depict that the SVC-Radon-time-frequency approach is efficient for the detection and parameter estimation of the multi-components LFM signal with low SNR.

**Key words:** Support Vector Clustering(SVC); Cluster labeling; Signal detection; Parameter estimation

### 1 引言

线性调频 (Linear FM, LFM) 信号多分量检测通常使用 Radon-时频分析<sup>[1-4]</sup>并结合 CLEAN<sup>[2-4]</sup>方法, 这个过程中存在如下问题: Radon 变换搜索的结果存在误差, 由此引起解线调的不精确; 窄带滤波器带宽需要人为设定, 且信号频率存在交叉时, 滤波将对剩余分量的参数估计进一步带来误差。本文针对多分量 LFM 信号的 Radon-时频分析输出结果中具有多个尖峰的特点, 引入 SVC 算法<sup>[5-8]</sup>对多个尖峰进行聚类, 完成 LFM 信号中多个分量的检测; 同时改进原始聚类标识方法<sup>[5]</sup>为“直接聚类标识”法来减少聚类标识过程中的计算量。低 SNR 下的仿真试验表明了本文方法的有效性。

### 2 Radon-时频分析与支持向量聚类算法理论

设 LFM 信号形式为

$$s(t) = \sum_i A_i \cos\left(2\pi f_i t + \frac{1}{2} k_i t^2\right) + n(t) \quad (1)$$

式中  $f_i$  是 LFM 信号的初始频率,  $k_i$  是 LFM 信号的调频斜率,  $A_i$  是信号的幅度,  $n(t)$  是噪声。

本文采用短时 Fourier 变换(STFT)作为时频分析方法<sup>[9]</sup>。先进行 STFT, 然后进行 Radon<sup>[9]</sup>变换, 设其结果为  $\text{Radon}_W(u, \alpha)$ , 其中存在多个尖峰, 每个尖峰对应为一个 LFM 信号。在尖峰中获得峰值对应坐标  $(u_i, \alpha_i)$ , 便可获得 LFM 信号的初始频率  $f_i$  和斜率  $k_i$ :

$$f_i = u_i \cos(\alpha_i), \quad k_i = -\cot(\alpha_i) \quad (2)$$

支持向量聚类算法<sup>[5,6]</sup>主要理论如下: 将样本集非线性映射到高维特征空间并在其中寻找包含其所有点且半径  $R$  最小的球; 运用优化、对偶等相关理论, 可以获得特征空间球体半径最小值; 由此可以定义支持向量(SVs)和边界支持向量(BSVs), 最后通过聚类标识获得信号聚类。

由于篇幅所限, 本文对 STFT, Radon 变换和 SVC 及其聚类标识方法不再叙述, 请见参考文献[5-9]。

### 3 多分量信号 LFM 检测方法

为了进行多分量信号的检测和参数估计本文引入 SVC

算法对  $\text{Radon}_W(u, \alpha)$  多个尖峰进行聚类。在此过程中, 在时频平面上进行了去噪处理; 并对  $\text{Radon}_W(u, \alpha)$  进行处理, 使用信号尖峰作为 SVC 输入, 以减少 SVC 输入样本数; 采用“直接聚类标识”的方法改进 SVC 中的聚类标识, 提高聚类运算效率。

在以下的分析中对信号及信号分析作出如下的假设: 信号定义如式(1)所示, 信号长度为  $10 \mu\text{s}$ , 信号最大带宽为  $B_{\max} = 100 \text{ MHz}$ , 采样频率为  $f_s = 1 \text{ GHz}$ ; 不同分量信号幅度之比最大值为  $A_{r_{\max}} = 10$ ; 并设置 STFT 的参数为: FFT 长度为  $L = 4096$ , 窗函数长度为  $1024$ , 调整迭代步长, 使得时间轴长度为  $300$  点。

### 3.1 时频平面去噪

假定信号中包含白噪声, 对信号进行时频分析, LFM 信号的聚集性能体现在信号能量分布在一定的带宽的一条直线上, 而白噪声信号的能量均匀分布在全时频平面上, 信号在时频平面的强度将远大于噪声的强度。为了在  $Z$  的结果中减弱噪声对信号所形成的尖峰的影响, 在时频平面滤除噪声。具体做法是直接时将时频分析的结果中小于某个阈值  $T_{if}$  的值置 0,  $T_{if}$  的值必须大于噪声在时频平面的强度, 但也不宜过大, 否则可能将有效的弱信号分量滤除。如下式所示:

$$A_{ij}(i, j) = (a_{ij}) = \begin{cases} \lg(w_{ij} - T_{if} + 1), & w_{ij} \in W_{if}(i, j), w_{ij} > T_{if} \\ 0, & w_{ij} \in W_{if}(i, j), w_{ij} \leq T_{if} \end{cases} \quad (3)$$

式中  $W_{if}$  为 STFT 的结果。  $T_{if}$  以  $\text{STFT}_s(t, \omega)$  的均值  $M_{\text{stft}}$  为参考, 乘以系数  $p$ 。通过仿真统计, 在假定的信号参数条件下, 通常取  $p = 1.5$  能保证有效滤除噪声而保留弱信号。

### 3.2 聚类输入集预处理

如果采用 Radon-STFT 输出的全部数据点作为 SVC 的输入, 那将是较大的输入数据集, 如果只输入包含尖峰最大值的少数的信号尖峰部分, 输入数据集将会减小。直观的想法就是取一个阈值来获取信号尖峰部分, 如图 1 所示。显然, 应当选择合适的阈值保证保留  $Z$  中所有信号尖峰, 而又不至于使 SVC 的输入数据集过大。影响阈值选择的主要因素是幅度最小的尖峰的最大值。以下进行定性分析。

首先分析信号的时频谱。时频谱由不同时间段的频谱构成, 其基本要素是窗口 Fourier 变换的频谱。 LFM 信号的频谱近似为矩形, 离散后的频谱<sup>[10]</sup>幅度为  $U(f) = AN_w / 2\sqrt{D}$ ,  $f \in [f_0, f_0 + B]$ 。其中,  $N_w$  为被截信号的有效长度(点数),  $D$  为脉冲压缩比, 且上式须满足  $D \gg 1$ 。而当  $k = 0$  时,  $U_{\max} = N_w / 2$ , 在假定的信号参数和 STFT 条件下可得  $D_{\max} = 10$ , 那么信号在时频平面上直线的最大强度之间的差异最大值为  $(A_{f1} - A_{f2})_{\max} = 10 \lg(U_{\max} / U_{\min}) = 10 \lg(A_{r_{\max}}) + 10 \lg(\sqrt{D_{\max}}) = 15(\text{dB})$ 。计算频谱幅度的最小值为  $(A_{f1})_{\min} = 10 \lg(N_w / 2\sqrt{D_{\max}})$ 。设  $r_{if}$  为时频面上两直线

的幅度之比值, 则有:

$$(r_{if})_{\max} = \left( \frac{A_{f1}}{A_{f2}} \right)_{\max} \leq \frac{(A_{f1} - A_{f2})_{\max} + (A_{f1})_{\min}}{(A_{f1})_{\min}} = 1.6719$$

其最大值不超过 2。显然,  $r_{if}$  的最大值随着分量信号  $B_{\max}$  (或  $D_{\max}$ ) 增大而增大。

再分析 Radon 变换。影响结果的主要因素是 Radon 变换过程中直线的强度和累加的点数。分析 Radon 变换原理可知, 累加的点数由直线相对时间轴的倾角决定: 倾角越大, 累加点数越多, 累加结果增大; 反之亦然。设两直线长度分别为  $l_1, l_2$ , 倾角分别为  $\alpha_1, \alpha_2$ , 直线强度分别为  $A_{f1}, A_{f2}$ , 则有  $l_1 \cos \alpha_1 = l_2 \cos \alpha_2$ 。变换结果比值为  $r = 10 \lg(l_1 A_{f1} / l_2 A_{f2}) \leq 10 \lg((r_{if})_{\max} \cos \alpha_2 / \cos \alpha_1)$ 。设最小倾角为 0, 其最大倾角为  $\alpha_{\max} = \arctg(B_{\max} / (f_s / L) / 300) = 0.2988\pi$ 。

设  $(r_{if})_{\max} = 2$ , 则:  $r_{\max} = 10 \lg[(r_{if})_{\max} / \cos \alpha_{\max}] = 5.285 \text{ dB}$ 。

在  $Z = \text{Radon}_W(u, \alpha)$  中找出最大值为  $z_{\max}$ , 取阈值为  $h = z_{\max} - r_{\max} + 0.5$ , 更新  $Z$  为

$$z_{ij} = \begin{cases} z_{ij} - h, & z_{ij} > h \\ 0, & z_{ij} \leq h \end{cases} \quad (4)$$

以上讨论了信号在 STFT 后在时频平面所形成的直线的强度变化情况, 通过对信号的部分参数进行假设, 计算了 LFM 信号在时频平面所形成直线经过滤波和取对数处理后, 其强度变化最大不超过 2 倍; 随后讨论了不同倾角的直线进行 Radon 变换后最大之的变化情况; 最后的结论为: 在 LFM 信号调频斜率不超过  $10 \text{ MHz} / \mu\text{s}$ , 信号带宽不超过  $100 \text{ MHz}$ , 经过设定参数的 STFT 和 Radon 变换, 强信号和弱信号的变换结果不超过  $5.285 \text{ dB}$ 。图 1 是预处理结果。

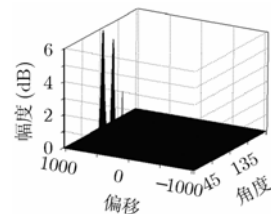


图 1 预处理结果

### 3.3 直接聚类标识方法

基于文献[5]的 CG 聚类标识方法中对每两点之间检查相邻关系, 然后进行聚类赋值和聚类标识, 这是原始算法中主了部分改进。

新的标识方法基于这样的—个事实: 对数据进行聚类后, 每个数据点只能属于聚类后的某一个类中, 那么它映射到特征空间的像对应地只能属于一个球中。这样, 可以得出结论: 和某个数据点  $x$  相邻的所有数据点必然和点  $x$  同属于—类, 而不属于其它类。当获得一个点的所有相邻点时, 就可以将这些点判为—类, 并做出标识; 这一类点不再参与是

否和其它点相邻的判断。设输入集样本数为  $N$ ，则有：

**算法 1**

步骤 1 设  $0 < n \leq N$ ，且  $c_n = 0$ ，令  $k = 1$ ，设  $E = \{x_l | c_l = 0, 0 < l \leq N\}$ ；

步骤 2 任取一点  $x_i \neq 0$  且  $x_i \in E$ ，令  $c_i = k$ ，更新  $x_i$ ：  
 $x_i = 0$ ；

步骤 3 取遍  $E$  中所有不为 0 的点  $x_j$ ，计算  $a_{ij}^{[5]}$ ，更新  $c_j$ ：  
 $c_j = \begin{cases} k, & a_{ij} = 1 \\ 0, & a_{ij} = 0 \end{cases}$  和  $x_j$ ：  
 $x_j = \begin{cases} 0, & a_{ij} = 1 \\ x_j, & a_{ij} = 0 \end{cases}$ ；

步骤 4 判断  $E$  是否为全 0 集合？是：结束；否：  
 $k = k + 1$ ，转到步骤 2。

以下以“计算相邻<sup>[5]</sup>”为基本计算单位估计计算的复杂度，设聚类结果包含  $n$  类，每个类包含数据点数为  $m_i, 0 < i \leq n, m_1 + m_2 + \dots + m_n = N$ 。计算的总次数  $N_C$ ：

$$N_C = n(N - 1) - [(n - 1)m_1 + (n - 2)m_2 + \dots + m_{n-1}] \quad (5)$$

分析式 (5) 可知： $N_C$  随  $n$  增加而增大，反之亦然。 $N$  不变，当  $n = 1$  时， $N_C$  取得最小值  $N - 1$ ； $n = N$  时， $N_C$  取得全局最大值  $N(N - 1)/2$ ，这时就变成了 CG 方法。一般情况下有  $N_C < n(N - 1)$ ，与 CG 算法计算量<sup>[5]</sup>的比值小于  $2n/N$ 。

在本文应用中，经过近似零均值处理后，聚类结果中  $n$  随信号分量数而增加，但远小于  $N$ ，因此，聚类标识的计算量远远小于原始聚类标识算法的计算量。可见在本文的具体应用中，直接聚类标识的方法可以明显减少聚类时间，为多分量信号检测提高了计算效率。

**3.4 多分量信号检测方法**

综合 Radon-STFT 和 SVC 算法，多分量 LFM 检测和参数估计如下：

**算法 2**

步骤 1 计算短时 Fourier 变换： $W(t, f) = \text{STFT}(s(t))^{[9]}$ ；

步骤 2 取  $p = 1.5$ ，计算  $T_{lf} = p \cdot \text{mean}(W(t, f))$ ，按照式(3)计算  $A_{lf}$  并更新  $W(t, f)$ ；

步骤 3 计算  $Z = \text{Radon}(W(t, f))^{[9]}$ ，取  $h = 6$ ，并按照式(4)更新  $Z$ ；

步骤 4 将  $Z$  中大于 0 的点设置 SVC 算法的输入集：  
 $X = \{x | x(z, u, \alpha), z = z_{u, \alpha} > 0\}$ ；

步骤 5 选择核函数  $K(x_i, x_j)^{[5]}$ ： $K(x_i, x_j) = e^{-q \|x_i - x_j\|^2}$ ，其中  $q$  为参数；通过训练选取核函数参数；

步骤 6 对  $X$  进行优化，计算支持向量和最小球半径；

步骤 7 应用算法 1 对  $X$  进行聚类标识，获得聚类结果；

步骤 8 在每类中求  $z$  的最大值所对应的坐标  $(u_i, \alpha_i)$ ，按式(2)估计信号参数。

**3.5 算法复杂度**

从本文算法整体的处理过程来看，其运算包含一次 STFT，Radon 线性变换运算和 SVC。STFT 和 Radon 线性

变换的运算量可参考文献[9]；SVC 算法运算量主要集中在聚类标识部分。设 SVC 输入集包含  $n$  类，支持向量的数量为  $n_v$ ，本文标识方法以计算两点相邻为单位计算量小于  $O(nN)$ 。由文献[5]可知而计算两点是否相邻计算量为  $O(n_v)$ 。

**4 实验与结果**

试验的过程包括通过训练确定参数  $q$ ， $C^{[5]}$  和对多分量信号检测方法的仿真试验。

**4.1 参数确定**

文献[5,6]对参数  $q$ ， $C$  选取与聚类结果之间的关系已做出明确地分析： $q$  增加，SVs 增加，聚类结果中类的数量增加；控制  $C$  的取值(小于 1)，将可能把部分 SVs 转化为 BSVs，类的数量将有所减少，而 BSVs 对应的输入数据点可以认为是由噪声引起的。选择  $q$ ， $C$  的较好的原则是：产生较少的 SVs 保证聚类边界的光滑性；同时，在两个参数的一定的范围内，使得聚类分配保持一定的稳定性。

通过仿真，得到  $q$  的范围为  $80 \leq q \leq 110$ 。

**4.2 仿真与结果**

根据以上的分析，本文用多分量信号检测方法进行了仿真试验。实验中采用 3 分量 LFM 信号，其参数如表 1 所示，在其中加入高斯白噪声，最弱分量 LFM 信号的 SNR 为 -10dB。

STFT 参数设置如上节所示；Radon 变换的角度旋转范围为  $0^\circ \leq \alpha \leq 179^\circ$ ，旋转间隔为  $1^\circ$ 。Radon 变换预处理结果如图 1 所示。选择聚类参数分别为  $q = 100$ ， $C = 1$  进行聚类训练，运用直接聚类标识方法进行聚类标识，聚类的结果如图 2 所示。图中 3 个类分别对应为 3 个信号尖峰；从每个类中获取幅度最大值，并求得其坐标  $(u, \alpha)$  以及对应的 LFM 信号的参数。

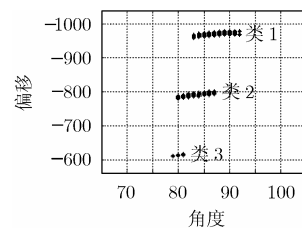


图 2 聚类结果

仿真过程中与基于 CLEAN 的方法<sup>[3,4]</sup>作了对比，所求得结果如表 1 所示。由表 1 可知，当 SNR 低于 -7dB 时，CLEAN 方法所获得信号参数误差较大，而采用聚类方法在 SNR 不低于 -10dB 时，仍能准确地进行多分量 LFM 信号检测和参数估计，通过聚类的方法可以提高弱信号的检测检测能力。本文提出的基于 SVC 的多分量 LFM 信号检测和参数估计方法具有优良的性能。

**5 结束语**

本文结合 Radon-时频分析和 SVC 完成了多分量 LFM

**表 1 Radon-STFT-SVC 检测多分量 LFM 信号**  
(其中,  $f_0$  的单位是 MHz,  $k$  的单位是 MHz/ $\mu$ s)

信号幅度	真实值	检测结果	
		聚类方法	CLEAN 方
1	$f_0$	10	10.3025
	$k$	0.5	0.49647
0.2	$f_0$	50	49.9172
	$k$	1.0	1.0042
0.1	$f_0$	90	90.0342
	$k$	1.5	1.4959

信号检测。其中, 经过预处理减少了 SVC 输入数据集点数和通过改进聚类标识, 提高 SVC 算法的计算速度; 在一定的 SNR(大于-10dB)条件下, 通过聚类的方法, 可以准确地检测到 Radon 变换后信号中所含分量个数, 为进行准确的参数估计提供基础。分析和试验的结果表明, 这是一种检测多分量线性调频信号和参数估计的有效方法。

### 参 考 文 献

- [1] Wood J C and Barry D T. Linear signal synthesis using the Radon-Wigner transform. *IEEE Trans. on Signal Processing*, 1994, 42(8): 2105-2111.
- [2] Li Y X and Yi M. Recursive filtering Radon-Ambiguity transform algorithm for multi-LFM signals detection. Proc. Of IEEE 2002 International Conference on Communications, Circuits and Systems and West Sino Expositions, Chengdu, China, 2002, (2): 1050-1053.
- [3] 郭汉伟, 王岩. 基于小波 Radon 变换检测线性调频信号. 国防科技大学学报, 2005, 25(1): 91-94.
- [4] Guo H W and Wang Yan. Linear chirp signals detection by wavelet Radon transform. *Journal of National University of Defense Technology*, 2005, 25(1): 91-94.
- [5] 章步云, 刘爱芳. 基于 Radon-STFT 的多分量线性调频信号检测与参数估计. 探测与控制学报, 2003, 25(3): 30-33.
- [6] Zhang B Y and Liu A F. Multicomponent LFM signal detection and parameter estimation based on Radon-STFT. *Journal of Detection & Control*, 2003, 25(3): 30-33.
- [7] Ben-Hur A and Horn D. Support vector clustering. *Journal of Machine Learning Research*, 2001, (2): 125-137.
- [8] Ben-Hur A and Horn D. Support vector method for hierarchical clustering. *Advances in Neural Information Processing Systems*, 2001, (13): 367-373.
- [9] Tax D M J and Duin R P W. Support vector domain description. *Pattern Recognition Letters*, 1999, 20(11 ~ 13): 1191-1199.
- [10] Yang J, Estivill-Castro V, and Chalup S K. Support vector clustering through proximity graph modelling. Proc. Ninth Int'l Conf. Neural Information Processing, 2002: 898-903. Singapore: IEEE Press, 2002: 898-903.
- [11] 张贤达, 保铮. 非平稳信号分析与处理[M]. 北京: 国防工业出版社, 1998: 1-180.
- [12] 向敬成, 张明友. 雷达系统. 北京: 电子工业出版社, 2001. 5: 112-130.

王令欢: 男, 1973 年生, 博士生, 研究方向为雷达信号处理.

马红光: 男, 1959 年生, 教授, 博士生导师, 主要研究方向为信号处理、人工智能等.