对称稳定分布的相关熵及其在时间延迟估计上的应用

宋爱民[①②]　　邱天爽*[①]　　佟祉谏[①]
[①](大连理工大学电子信息与电气工程学部　大连　116024)
[②](大连交通大学理学院　大连　116028)

摘　要：相关熵是一个表示随机变量局部相似性的统计量。该文首先研究对称 α-稳定(SαS)分布的相关熵的参数表示，利用该参数表示证明了对于位置参数为零的SαS分布，最大相关熵准则与最小分散系数准则是等价的。最后将研究结果应用于稳定分布噪声环境下自适应时间延迟估计。仿真实验表明，该文算法性能优于最小均方误差时间延迟估计与最小平均 P-范数时间延迟估计。

关键词：信号处理；相关熵；对称稳定分布；最大相关熵准则；最小分散系数准则

中图分类号：TN911.7　　**文献标识码**：A　　**文章编号**：1009-5896(2011)02-0494-05

DOI: 10.3724/SP.J.1146.2010.00309

Correntropy of the Symmetric Stable Distribution and Its Application to the Time Delay Estimation

Song Ai-min[①②]　　Qiu Tian-shuang[①]　　Tong Zhi-jian[①]
[①](*Faculty of Electronic Information and Electrical Engineering*, *Dalian University of Technology*, *Dalian* 116024, *China*)
[②](*School of Science*, *Dalian Jiaotong University*, *Dalian* 116028, *China*)

Abstract: Correntropy is a localized similarity measure between two scalar random variables. This paper presents the parametric representation of the symmetric α-stable (SαS) distribution's correntropy. The equivalency of the maximum correntropy criterion and the minimum dispersion criterion is derived from the parametric representation for zero location SαS distributions. This result is used to propose the adaptive time delay estimation in SαS noise. Simulations show that the algorithm based on correntropy works better than the least mean square and the least mean p-norm approaches.

Key words: Signal processing; Correntropy; Symmetric stable distributions; Maximum correntropy criterion; Minimum dispersion criterion

1　引言

近年来的研究发现，雷达杂波[1−3]，水声信号[1]，移动无线电信道中的电磁噪声[4]等，在时域表现出短时大幅度脉冲的特性，他们的概率密度函数的尾部比高斯概率密度函数的尾部衰减得更慢，利用对称 α-稳定(SαS)分布建模更加准确[1−6]。

一个随机变量 S 称为SαS分布，如果它的特征函数具有形式 $\varphi(w)=\exp(jaw-\gamma|w|^{\alpha})$，其中 $0<\alpha\le 2$ 是特征指数，a 是位置参数，$\gamma>0$ 为分散系数。当 $\alpha=2$ 时，S 是Gauss随机变量；$\alpha=1$ 时，S 是Cauchy随机变量。

最小分散系数准则是稳定分布噪声环境下信号处理的常用准则[1,6,7]，SαS 分布的分散系数与方差的作用类似，它决定分布偏离中心的离散程度，通常用来刻画SαS 噪声的功率。由于稳定分布的分散系数与其分数低阶矩成正比，最小分散系数准则常常利用最小化分数低阶矩来实现[1,3,4,8−10]。

但是分数低阶矩不是解决稳定分布信号处理的唯一统计量，只要是适用于稳定分布，并且能够反映其分散系数信息的统计量都是值得研究的。相关熵作为随机变量间局部相似性的度量，近些年来受到广泛关注[11−17]。相关熵最初是针对随机过程定义的[11]，后来这个概念被进一步推广，对两个随机变量 X 与 Y，定义其相关熵为[12]

$$V_{\sigma}(X,Y)=E(k_{\sigma}(X-Y)) \tag{1}$$

其中 $k_{\sigma}(\cdot)=\dfrac{1}{\sqrt{2\pi}\sigma}\exp\left(-\dfrac{(\cdot)^2}{2\sigma^2}\right)$，$\sigma>0$ 是核长参数。

文献[12]证明了相关熵可以诱导一个距离并据此给出最大相关熵准则(maximum correntropy criterion)。相关熵既可以看成是基于 Parzen 核估计

2010-03-29 收到，2010-10-15 改回
国家自然科学基金(60872122，60940023)资助课题
*通信作者：邱天爽　qiutsh@dlut.edu.cn

的 Renyi 二次熵的一种退化表示，又能够反映两个随机变量的相似度，因此在信道的盲均衡[11]，风力预报[13]，模式识别[14]，噪声抵消[15]，非线性检测[16]等领域得到广泛的应用，特别是 Pokharel 利用相关熵设计了脉冲噪声下的鲁棒检测器[17]。

由式(1)知相关熵是误差 $e = X - Y$ 的函数，本文将讨论 SαS 分布相关熵的存在性及其参数表示，解释利用相关熵作为稳定分布自适应信号处理的代价函数的合理性，并利用最大相关熵准则解决稳定分布噪声下的自适应时间延迟估计，仿真试验表明，本文算法在处理强脉冲时表现出更好的性能。

2　SαS 分布的相关熵

定理 1　设 e 是特征指数 α，位置参数 a，分散系数 γ 的 SαS 分布，则 e 的相关熵 $V_\sigma(e)$ 存在，并可表示为

$$V_\sigma(e) = \frac{1}{\pi}\int_0^{+\infty} \exp\left(-\frac{1}{2}\omega^2\sigma^2 - \gamma w^\alpha\right)(\cos aw)\mathrm{d}w \tag{2}$$

证明　由式(1)知 e 的相关熵为

$$V_\sigma(e) = \int_{-\infty}^{+\infty} k_\sigma(e) f(e)\mathrm{d}e \tag{3}$$

设 $f(e)$ 与 $\varphi(w)$ 分别是 e 的概率密度函数与特征函数，则有

$$f(e) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \varphi(w)\exp(-jwe)\mathrm{d}w \tag{4}$$

将式(4)代入式(3)有

$$\begin{aligned} V_\sigma(e) &= \frac{1}{2\pi}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{e^2}{2\sigma^2} - jwe\right)\varphi(w)\mathrm{d}w\mathrm{d}e \\ &= \frac{1}{2\pi}\int_{-\infty}^{+\infty}\exp\left(-\frac{1}{2}w^2\sigma^2\right)\varphi(w)\mathrm{d}w \\ &= \frac{1}{2\pi}\int_{-\infty}^{+\infty}\exp\left(-\frac{1}{2}w^2\sigma^2\right)\exp(-jaw - \gamma|w|^\alpha)\mathrm{d}w \\ &= \frac{1}{\pi}\int_0^{+\infty}\exp\left(-\frac{1}{2}w^2\sigma^2 - \gamma w^\alpha\right)(\cos aw)\mathrm{d}w \end{aligned}$$

由于 $\frac{1}{\pi}\int_0^{+\infty}\exp\left(-\frac{1}{2}w^2\sigma^2 - \gamma w^\alpha\right)(\cos aw)\mathrm{d}w \le \sqrt{\frac{\pi}{2\sigma^2}}$，故 $V_\sigma(e)$ 存在。　证毕

推论1　设 e_G 表示Gauss随机变量，则有

$$V_\sigma(e_G) = \frac{1}{\sqrt{2\pi(\sigma^2 + \sigma_{e_G}^2)}}\exp\left(-\frac{a^2}{2(\sigma^2 + \sigma_{e_G}^2)}\right) \tag{5}$$

证明　将 $\alpha = 2$，$\gamma = (1/2)\sigma_{e_G}^2$ 代入整理可得式(5)。

由于分散系数 γ 通常用来刻画 SαS 噪声的功率，以下将讨论 $V_\sigma(e)$ 与 γ 的关系。

定理 2　对于 $a = 0$ 的 SαS 随机变量 e，给定核长参数 σ，那么最大化 e 的相关熵等价于最小化 e 的分散系数。

证明　将 $a = 0$ 代入式(2)，有 $V_\sigma(e) = 1/\pi \cdot \int_0^{+\infty}\exp\left(-\frac{1}{2}\omega^2\sigma^2 - \gamma w^\alpha\right)\mathrm{d}w$ 成立。定义

$$U(\gamma) \triangleq \frac{1}{\pi}\int_0^{+\infty}\exp\left(-\frac{1}{2}\omega^2\sigma^2 - \gamma w^\alpha\right)\mathrm{d}w \tag{6}$$

由于式(6)右端的积分是绝对可积的，对式(6)两端关于 γ 求导，得

$$\frac{\mathrm{d}U(\gamma)}{\mathrm{d}\gamma} = -\frac{1}{\pi}\int_0^{+\infty}\exp\left(-\frac{1}{2}w^2\sigma^2 - \gamma w^\alpha\right)w^\alpha\mathrm{d}w \tag{7}$$

由式(7)知，$\forall\gamma \in (0, +\infty)$，$\frac{\mathrm{d}U(\gamma)}{\mathrm{d}\gamma} < 0$ 成立，故 $U(\gamma)$ 在 $(0, +\infty)$ 上关于 γ 单调递减，因此对于 $a = 0$ 的 SαS 随机变量 e，在核长参数 σ 固定的情况下，最大化 e 的相关熵等价于最小化 e 的分散系数。　证毕

由定理 2 知，在自适应信号处理中，如果误差是服从 $a = 0$ 的 SαS 分布，那么最大化误差的相关熵能够起到最小化误差功率的目的，并且定理 2 的证明过程表明特征指数 α 不会改变相关熵在 $(0, +\infty)$ 关于 γ 的单调递减性这一事实。

3　应用

作为相关熵在稳定分布噪声下信号处理中的一个简单应用，我们将讨论稳定分布下基于最大相关熵准则的时间延迟估计问题。

本文的时间延迟估计模型为：$x_1(n) = s(n) + v_1(n)$，$x_2(n) = s(n - D) + v_2(n)$。其中，$n$ 为离散时间变量，D 是延迟时间，$s(n)$ 是感兴趣信号，$v_1(n)$ 与 $v_2(n)$ 是独立的 SαS 噪声，$x_1(n)$ 与 $x_2(n)$ 是观测信号。

由于延迟信号 $s(n - D)$ 可以视为原信号 $s(n)$ 与一个 sinc 函数作为权系数的横向滤波器卷积得来，因此在自适应时间延迟估计中，常常将时间延迟的估计转化为对有限脉冲响应(FIR)滤波器系数的估计，进而搜索估计得到的权系数的峰值获取时间延迟信息[8,18]。

设延迟时间的范围是 $[-M, M]$，自适应 FIR 滤波器的权系数是 $\boldsymbol{H} \triangleq \{h(-M), \cdots, h(0), \cdots, h(M)\}^{\mathrm{T}}$，则滤波器的误差[18] $e(n) = x_2(n) - \sum_{i=-M}^{M} h(i)x_1(n+i)$。设在某一准则下，经过自适应训练得到的滤波器权系数为 $\boldsymbol{H}^* \triangleq \{h^*(-M), \cdots, h^*(0), \cdots, h^*(M)\}^{\mathrm{T}}$，那么时间延迟估计为 $\widehat{D} = -\arg\max_i(h^*(i))$。

由于非高斯的稳定分布不存在二阶统计量，传统的基于最小均方误差时延估计 LMSTDE(Least

Mean Square Time Delay Estimation)算法[18]性能将出现退化，因此文献[8]对 LMSTDE 算法进行改进，得到 LMPTDE(Least Mean P-norm Time Delay Estimation)算法。

3.1 基于最大相关熵准则的时间延迟估计

根据定理2，如果误差是 $a=0$ 的 SαS 随机变量，最大化误差的相关熵能够起到最小化误差功率的目的，因此本文将采用误差的相关熵

$$J_{\mathrm{MCC}}(\boldsymbol{H})=E\left[k_\sigma\left(x_2(n)-\sum_{i=-M}^{M}h(i)x_1(n+i)\right)\right] \tag{8}$$

作为代价函数，通过 $\max\limits_{\boldsymbol{H}} J_{\mathrm{MCC}}(\boldsymbol{H})$ 求得 FIR 滤波器的权系数 $\boldsymbol{H}$。$J_{\mathrm{MCC}}(\boldsymbol{H})$ 的估计为

$$\hat{J}_{\mathrm{MCC}}(\boldsymbol{H})=\frac{1}{N-2M}\sum_{n=M+1}^{N-M}k_\sigma\left(x_2(n)-\sum_{i=-M}^{M}h(i)x_1(n+i)\right) \tag{9}$$

利用随机梯度下降法[15] 得到 $\boldsymbol{H}$ 的更新关系式为

$$\begin{aligned}\boldsymbol{H}_{n+1}&=\boldsymbol{H}_n+\mu\left(\frac{\partial \hat{J}_{\mathrm{MCC}}(\boldsymbol{H})}{\partial \boldsymbol{H}}\right)_n\\&=\boldsymbol{H}_n+\mu\frac{1}{\sqrt{2\pi}\sigma^3}\exp\left(-\frac{e(n)^2}{2\sigma^2}\right)\cdot e(n)\cdot\boldsymbol{x}_1\end{aligned} \tag{10}$$

其中 μ 是收敛因子，$\boldsymbol{x}_1=[x_1(n-M),\cdots,x_1(n),\cdots,x_1(n+M)]^{\mathrm{T}}$。为方便讨论，本文基于相关熵准则的时间延迟估计算法简记为 MCCTDE(Maximum Correntropy Criterion Time Delay Estimation)。

3.2 MCCTDE 算法的性能

本文的仿真条件：感兴趣的信号 $s(n)$ 是零均值高斯白噪声。观测的数据长度是 1064 点，真实的时延为 $D=20$ 个单位采样点。FIR 滤波器长度是125，初始权系数 $\boldsymbol{H}_0$ 取零向量。由于非高斯稳定分布不存在二阶矩，所以通常使用广义信噪比[8]来刻画信号与噪声的功率比，它的定义是 GSNR $\triangleq 10\lg(\sigma_s^2/\gamma_w)$ dB，其中 σ_s^2 是 $s(n)$ 的方差，γ_w 是 $v_1(n)$ (或 $v_2(n)$)的分散系数。

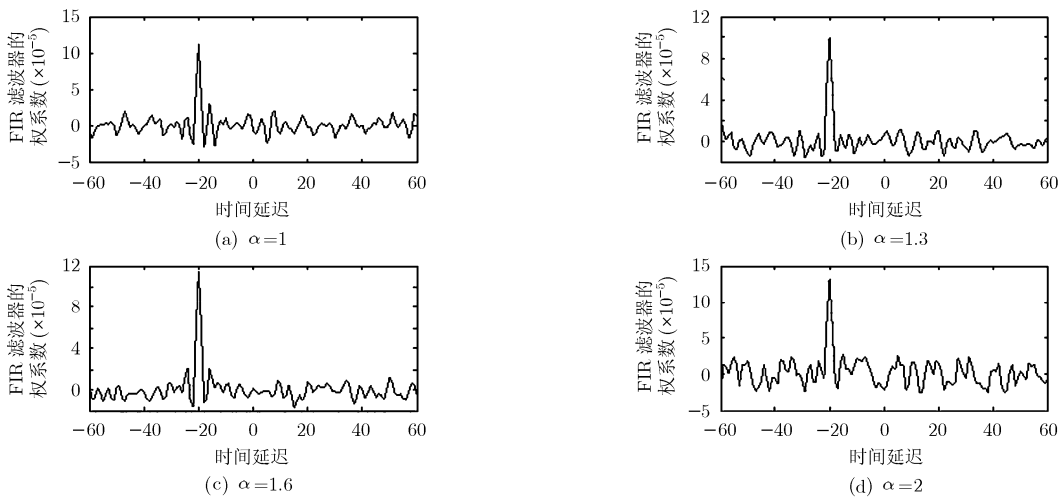
图 1 是在 GSNR=1.5 dB 下对 FIR 滤波器权系数的估计，其中收敛因子是 $\mu=10^{-5}$，核长参数 $\sigma=1$。由图 1 知权系数的最大值出现在-20，因此延迟时间的估计是 20。

为进一步测试 MCCTDE 的性能，本文将通过时延估计偏差 $\phi=\dfrac{1}{W}\sum_{i=1}^{W}(\hat{D}_i-D)$ 以及归一化估计误差功率 $\zeta=\begin{cases}(1/W)\sum_{i=1}^{W}\left(\hat{D}_i/D-1\right)^2, & D\neq 0\\(1/W)\sum_{i=1}^{W}\hat{D}_i^{\,2}, & D=0\end{cases}$ 来衡量算法的性能，其中 W 是仿真试验的次数，$\hat{D}_i\ (i=1,\cdots,W)$ 是第 i 次估计的结果，D 是时延的真值。

图 2 与图 3 比较 MCCTDE,LMSTDE 以及 LMPTDE 等算法在相同的仿真条件下独立运行 $W=5000$ 次的时延估计偏差以及归一化估计误差功率，其中 GSNR=1.5 dB，收敛因子均取 $\mu=10^{-5}$，LMPTDE 的参数 $p=\alpha-0.1$。

由图 2 与图 3 可知，MCCTDE 算法在零均值高斯($\alpha=2$)噪声下可以有效工作。其原因在于：将 $a=0$，$\sigma_{e_G}^2=E(e_G^2)$ 代入式(5)，得到 $V_\sigma(e_G)=1\big/\sqrt{2\pi(\sigma^2+E(e_G^2))}$，故此时最大相关熵准则等价于最小均方误差准则。LMPTDE 算法的参数 p 通常被限制在 $0\le p<\alpha$ 的范围内，因此 p 值的确定依赖于对稳定分布的特征指数 α 的先验知识的获取，本文算法则不需要。

同时，图 2 与图 3 也反映了核长参数对 MCCTDE 算法性能的影响：即使对于不同的稳定分布噪声(特征指数 α 变动)，核长参数在区间[1,3]上任取，均可得到较好的估计结果。并且噪声的特征指数越大，本文算法对核长的选取依赖越小。



(a) $\alpha=1$ (b) $\alpha=1.3$ (c) $\alpha=1.6$ (d) $\alpha=2$

图 1 $\sigma=1$ 时基于最大相关熵准则的 FIR 滤波器权系数估计结果(GSNR=1.5 dB)

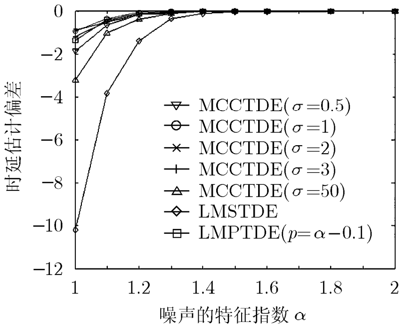


图 2 在 GSNR=1.5 dB 下各种算法对于不同的稳定分布噪声的时延估计偏差

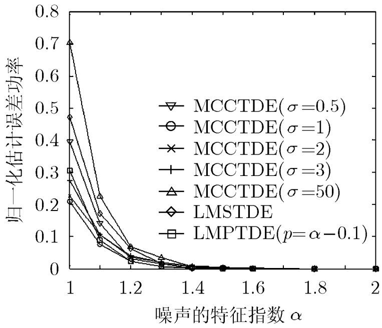


图 3 在 GSNR=1.5 dB 下各种算法对于不同的稳定分布噪声的归一化估计误差功率

进一步地，我们测试在噪声的特征指数为 $\alpha = 1.4$ 时，各种算法在不同 GSNR 下的性能。其中 MCCTDE 算法中的核长参数 $\sigma = 1$，LMPTDE 算法中的参数 $p = 1.3$。

表 1 说明 MCCTDE 算法在不同的 GSNR 下，性能依然不逊于其他两种算法。

4 结束语

本文讨论 SαS 分布的相关熵的存在性及其参数表达，分析了在自适应信号处理中如果误差是 $a = 0$ 的 SαS 随机变量，那么最大化误差的相关熵能够起到最小化误差功率的目的，并利用这一原理设计了一种稳定分布噪声下的自适应时间延迟估计算法。今后的工作将进一步讨论利用相关熵处理稳定分布噪声下的时延算法问题。

表 1 在不同的 GSNR 下各种算法性能比较

GSNR (dB)	时延估计偏差			归一化估计误差功率		
	MCCTDE	LMSTDE	LMPTDE	MCCTDE	LMSTDE	LMPTDE
1.5	−0.0058	−0.1178	−0.0182	0.0014	0.0056	0.0042
1.2	−0.0602	−0.4784	−0.0948	0.0125	0.0213	0.0164
1.0	−0.1230	−0.5538	−0.2044	0.0288	0.1154	0.0374
0.8	−0.2056	−1.6378	−0.4156	0.0440	0.3303	0.0861
0.6	−0.3074	−4.8244	−0.8918	0.0674	1.0204	0.1925
0.4	−1.5054	−10.0462	−1.7316	0.3253	2.1648	0.3797

参 考 文 献

[1] Nikias C L and Shao M. Signal Processing with Alpha-Stable Distributions and Applications. New York: Wiley, 1995: 98–128.

[2] Liao Ming-sheng, Wang Chang-cheng, and Wang Yong, *et al*.. Using SAR images to detect ships from sea clutter. *Geoscience and Remote Sensing Letters*, 2008, 5(2): 194–198.

[3] Georgiou P G, Tsakalides P, and Kyriakakis C. Alpha-stable modeling of noise and robust time-delay estimation in the presence of impulsive noise. *IEEE Transactions on Multimedia*, 1999, 1(3): 291–301.

[4] Rupi M, Tsakalides P, and Del Re E, *et al*.. Constant modulus blind equalization based on fractional lower-order statistics. *Signal Processing*, 2004, 84(5): 881–894.

[5] Tang Xiao-tong, Ma Meng, and Ostry D I, *et al*.. Characterizing impulsive network traffic using truncated α -stable processes. *IEEE Communications Letters*, 2009, 13(12): 980–982.

[6] Samorodnitsky G and Taqqu M S. Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance. New York, London: Chapman & Hall, 1994: 1–48.

[7] Stuck B W. Minimum error dispersion linear filtering of scalar symmetric stable processes. *IEEE Transactions on Automatic Control*, 1978, 23(3): 507–509.

[8] Ma X Y and Nikias C L. Joint estimation of time delay and frequency delay in impulsive noise using fractional lower order statistics. *IEEE Transactions on Signal Processing*, 1996, 44(11): 2669–2687.

[9] Li S, Qiu T S, and Zhang S F. Space-time blind equalisation in impulsive noise. *IET Signal Processing*, 2009, 3(6): 445–458.

[10] Zha D. Robust multiuser detection method based on least

p-norm state space criterion. *Wireless Personal Communications*, 2007, 40(2): 191–204.

[11] Santamaria I, Pokharel P P, and Principe J C. Generalized correlation function: definition, properties, and application to blind equalization. *IEEE Transactions on Signal Processing*, 2006, 54(6): 2187–2197.

[12] Liu W F, Pokharel P P, and Principe J C. Correntropy: properties and applications in non-Gaussian signal processing. *IEEE Transactions on Signal Processing*, 2007, 55(11): 5286–5298.

[13] Bessa R J, Miranda V, and Gama J. Entropy and correntropy against minimum square error in offline and online three-day ahead wind power forecasting. *IEEE Transactions on Power Systems*, 2009, 24(4): 1657–1666.

[14] Jeong K H, Liu W F, and Han S, *et al*.. The correntropy MACE filter. *Pattern Recognit*, 2009, 42(5): 871–885.

[15] Singh A and Principe J C. Using correntropy as a cost function in linear adaptive filters. Proceedings of the 2009 international joint conference on Neural Networks, Atlanta, USA, Jun.14–19, 2009: 2950–2955.

[16] Gunduz A and Principe J C. Correntropy as a novel measure for nonlinearity tests. *Signal Processing*, 2009, 89(1): 14–23.

[17] Pokharel P P, Liu W F, and Principe J C. A low complexity robust detector in impulsive noise. *Signal Processing*, 2009, 89(10): 1902–1909.

[18] Youn D H, Ahmed N, and Carter G C. On using the LMS algorithm for time delay estimation. *IEEE Transactions on Acoustics, Speech and Signal Process*, 1982, 30(5): 798–801.

宋爱民： 男，1978 年生，博士生，研究方向为非高斯信号处理.

邱天爽： 男，1954 年生，教授，博士生导师，研究方向为数字信号处理理论、生物医学信号处理、非平稳与非高斯信号处理等.

佟祉谏： 男，1983 年生，硕士生，研究方向为时间延迟估计.