

一类新的独立性度量及其在盲信号分离中的应用

陈 阳¹ 何振亚¹

(1 东南大学无线工程系, 南京市四牌楼2号, 210096)

摘要: 提出了一类无穷多种称为准熵的新的独立性度量, 它们用严格凸函数对原变量经分布函数变换再量化后得到的变量的联合概率的均匀性进行度量, 并提出了基于准熵的盲分离算法, 可分离任意连续分布的信号, 包括峭度为零的信号。通过与前人算法的对比试验, 证实了基于准熵的算法的优越性。

关键词: 盲分离; 独立性; 独立元分析

中图法分类号: TN911.7

盲信号分离(blind signal separation, BSS)在通信和信号处理中已有的以及潜在的广泛应用使其在近十几年来得到了关注。目前, 已提出的盲分离准则和算法绝大多数都对源信号的概率分布有限制。例如, 最大似然算法^[1]要求源信号近似符合某个分布模型。而较流行的基于峭度(4阶累积量)的算法^[2]不能应付有两个或更多源信号的峭度同时为零的情况, 这是因为零峭度源信号混合得到的信号峭度依然为零, 单从峭度上无法区别混合前后的信号。设计能分离零峭度源信号的盲分离算法是重要的, 因为会遇到一些信号, 如图像, 其峭度恰好为零或非常接近零。

1 一类新的独立性度量

笔者用 $q_r(\cdot)$ 表示变量 r 的(累积)分布函数, $q_r(u) = \text{prob}(r \leq u) = \int_{-\infty}^u p_r(v)dv$, 其中 $\text{prob}(A)$ 表示事件 A 发生的概率。

考虑两个连续变量 r_1, r_2 , 将它们分别用各自的分布函数变换得到两个变量:

$$z_1 \equiv q_{r_1}(r_1), \quad z_2 \equiv q_{r_2}(r_2) \quad (1)$$

则有^[3] 引理 1。

引理 1 r_1, r_2 相互独立的充要条件是 $p_{z_1 z_2}(u, v) = 1, \forall (u, v) \in [0, 1] \times [0, 1]$ 。

设 l 为一个大于 1 的正整数, 在 $[0, 1]$ 上定义 l 级均匀量化算子 $D_l(\cdot)$ 如下:

$$D_l(u) = \begin{cases} \lceil ul \rceil, & 0 < u \leq 1 \\ 1, & u = 0 \end{cases} \quad (2)$$

式中, $\lceil v \rceil$ 为不小于 v 的最小整数。

定义离散变量:

$$k_1 \equiv D_l(z_1), \quad k_2 \equiv D_l(z_2) \quad (3)$$

设 k_1, k_2 的联合概率为 $p_{k_1 k_2}(i, j) = \text{prob}(k_1 = i \text{ 且 } k_2 = j)$, 其中 $(i, j) \in \{1, \dots, l\} \times \{1, \dots, l\}$ 。则准熵:

$$\beta(r_1, r_2) = \sum_{i=1}^l \sum_{j=1}^l f(p_{k_1 k_2}(i, j)) \quad (4)$$

式中, $f(\cdot)$ 为 $[0, 1]$ 上的严格凸函数。将 $\beta(r_1, r_2)$ 命名为“准熵”, 是由于其与信息论中熵的一些相似之处^[3]。

由 Jensen 不等式, 则有:

$$\beta(r_1, r_2) \geq l^2 f(1/l^2) \quad (5)$$

且其中等号成立的充要条件是 $p_{k_1 k_2}(i, j) = 1/l^2, \forall (i, j) \in \{1, \dots, l\} \times \{1, \dots, l\}$, 即 (k_1, k_2) 在 $\{1, \dots, l\} \times \{1, \dots, l\}$ 上均匀分布。

若 r_1, r_2 相互独立, 则由引理 1, (z_1, z_2) 在 $[0, 1] \times [0, 1]$ 上均匀分布, (k_1, k_2) 在 $\{1, \dots, l\} \times \{1, \dots, l\}$ 上均匀分布。

引理 2 若 r_1, r_2 相互独立, 则 $\beta(r_1, r_2)$ 取最小值 $l^2 f(1/l^2)$ (式(5)右边)。

引理 3 若 r_1, r_2 不独立, 则存在 l_0 , 使得对任意 $l > l_0$, $\beta(r_1, r_2)$ 都不能取最小值。

图 1 是准熵的示意图。图 1(a)和图 1(b)分别为两个独立源信号(s_1, s_2)和两个不独立混合

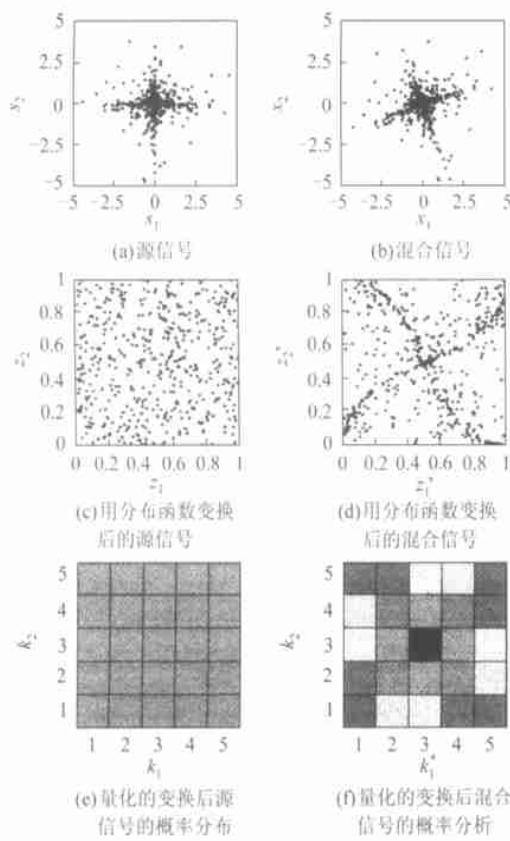


图1 准熵的示意图

Fig. 1 Quasi-entropy

信号 (x_1, x_2) 的观测点;图1(c)和图1(d)分别为 (z_1, z_2) 和 (z_1^*, z_2^*) 的观测点,其中 $z_1 \equiv q_{s_1}(s_1)$, $z_2 \equiv q_{s_2}(s_2)$, $z_1^* \equiv q_{x_1}(x_1)$, $z_2^* \equiv q_{x_2}(x_2)$ 。图1(c)中的点分布均匀,而图1(d)中的点分布不均匀,符合引理1。笔者将图1(c)和1(d)中的 $[0, 1] \times [0, 1]$ 区域分别划分为 $l \times l$ 个(如 5×5 个)相同大小的方格,则图1(e)、1(f)中的点落入每个方格的概率相等(不等),如图1(e)、1(f)所示,图中灰度表示概率的大小,灰度越深,概率越大。根据图1(c)、1(e)、1(d)、1(f), $k_1 \equiv D_5(z_1)$, $k_2 \equiv D_5(z_2)$, $k_1^* \equiv D_5(z_1^*)$, $k_2^* \equiv D_5(z_2^*)$,图1(e)、1(f)中用灰度表示的概率即为 $p_{k_1 k_2}^{k_1^* k_2^*}$,它在 $\{1, \dots, 5\} \times \{1, \dots, 5\}$ 上均匀(不均匀),因此, $\beta(s_1, s_2) = l^2 f(1/l^2) = 25 f(1/25) < \beta(x_1, x_2)$,独立性的大小在 β 值上体现出来。

因为 $[0, 1]$ 上的严格凸函数有无穷多种,因此准熵实际上是一类无穷多种独立性度量的总称。

在实际情况下,分布函数的解析形式一般是不知道的,不过这并不妨碍准熵的计算。根据分布函数的定义易得,若将变量 r 的 N 个观测值 $r(1), \dots, r(N)$ 按升序排列且其中 $r(t)$ 排在第 i 个位置,则 i/N 是 $q_r(r(t))$ 的无偏估计。相应地, $\lceil i/N \rceil$ 就

是 $D_l[q_r(r(t))]$ 的估计。这样就可以将 (r_1, r_2) 的观测值直接映射到 (k_1, k_2) 的观测值(的估计),且在此过程中,不牵涉任何浮点运算,详见文献[4]。那么,计算准熵所需的 $p_{k_1 k_2}(i, j)$ 即可用 (k_1, k_2) 的观测值中 (i, j) 出现的比率估计。注意到凸函数 $f(\cdot)$ 的定义域限制在 $[0, 1]$ 上,因此,这可以通过查表来实现。

2 盲分离算法

先用一些标准分布来验证一下前面的结果。在盲分离中,若混合信号已被白化^[2],则混合阵和分离阵都简化为旋转阵。将两个等方差相互独立的源信号 s_1, s_2 用一个 2×2 旋转阵混合,得到信号 r_1, r_2 :

$$\begin{bmatrix} r_1 \\ r_2 \end{bmatrix} = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} \quad (6)$$

令

$$\beta(\theta) = \beta(r_1, r_2) \quad (7)$$

并且作出准熵随旋转角 θ 变化的图像。

采用的源信号有均匀分布、Laplace 分布、正弦波和 Rayleigh 分布。此外,还这样产生了一个零峭度源:设有一个正峭度源 s^+ ,其峭度为 k^+ ;还有一个与之独立的负峭度源 s^- ,其峭度为 k^- 。令 $s = as^+ + bs^-$,则当两个系数满足

$$a^4/b^4 = |k^-|/|k^+| \quad (8)$$

时, s 的峭度就等于0(当然也可用3个或更多独立源信号的组合来产生零峭度源)。若选取的正、负峭度源分别为 Laplace 和均匀分布,产生的零峭度分布如图2中实线所示,图中还用虚线绘出了与其方差相等的高斯分布,虽然高斯分布也具有零峭度,但从图2中可见,该分布与高斯分布是不同的。总之,上述这些源信号包括了正、负、零峭度,对称和不对称分布。

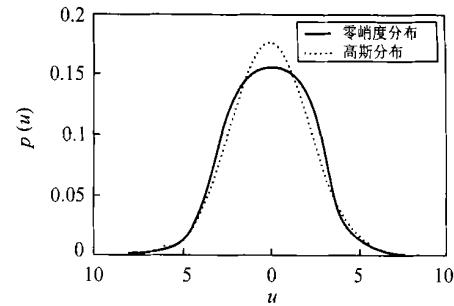


图2 1种零峭度概率密度函数

Fig. 2 Zero Kurtosis

图3为采用上述源信号的不同组合,以及利用几种不同凸函数处理得到的 $\beta(\theta)$ 图像。各函数图线都在旋转角 $\theta=0$,也就是 r_1, r_2 相互独立时达到

最小值,实际上此时的 r_1, r_2 即分别为 s_1, s_2 。此外,还计算出了图 3 中准熵的理论最小值,列于表 1。对

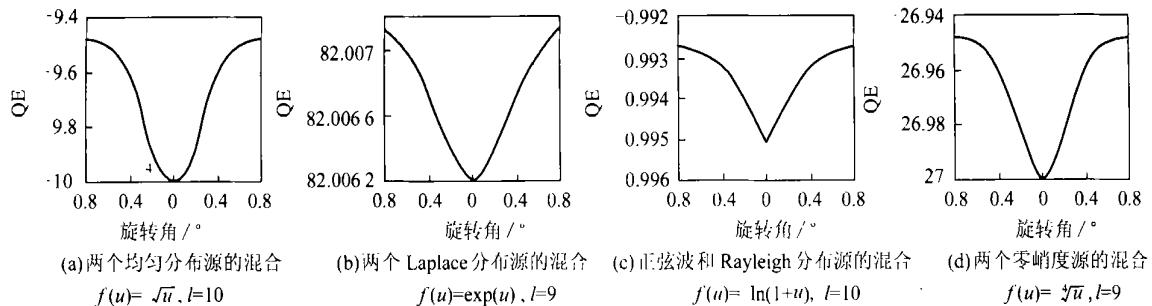


图 3 两个等方差独立源经旋转阵得到的信号的准熵(QE)随旋转角 θ 变化的图像

Fig. 3 Quasi-entropies Corresponding to the Changes of Rotation Angle θ

表 1 图 3 中准熵的理论最小值 $l^2/f(1/l^2)$

Tab. 1 Theoretical Minimum Value of Quasi-entropy

图 3 中的子图	图 3(a)	图 3(b)	图 3(c)	图 3(d)
凸函数 $f(u)$	$-\sqrt{u}$	$\exp(u)$	$-\ln(1+u)$	$-\sqrt[3]{u}$
量化级数 l	10	9	10	9
准熵的理论最小值	-10	82.0062	-0.9950	-27

图 3 表明,对于两个源信号的混合,将它们白化后乘以一个 2×2 旋转阵,找出使输出信号的准熵最小的旋转角就能达到分离。那么,对于 n 个源信号的混合,采用一种逐对信号迭代的方式^[3] 最小化准熵,就得到了如下的盲分离算法。

1) 白化。通过如 PCA(主元分析),可确定源信号的个数 n ,并且找到一个 $n \times d$ 矩阵 W ,使得白化后的混合信号

$$y(t) = Wx(t), \quad t = 1, \dots, N \quad (9)$$

的样本协方差阵等于单位阵。

2) 迭代。重复直到收敛。

角扫描: 对所有 $1 \leq i \leq n-1$ 及 $i+1 \leq j \leq n$,令

$$\begin{bmatrix} \hat{y}_i(t) \\ \hat{y}_j(t) \end{bmatrix} = \begin{bmatrix} \cos \theta_{ij} & \sin \theta_{ij} \\ -\sin \theta_{ij} & \cos \theta_{ij} \end{bmatrix} \begin{bmatrix} y_i(t) \\ y_j(t) \end{bmatrix} \quad (10)$$

式中, $t=1, \dots, N$ 。令

$$\beta(\theta_{ij}) = \beta(\hat{y}_i, \hat{y}_j) \quad (11)$$

在 $(-\pi/4, \pi/4)$ 上通过黄金分割法找出:

$$\theta_{ij}^{pt} = \arg \min_{\theta_{ij}} \beta(\theta_{ij}) \quad (12)$$

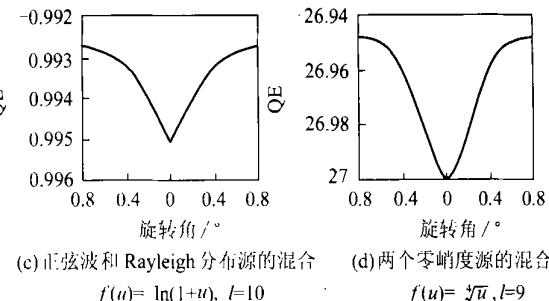
将与 θ_{ij}^{pt} 对应的 $\hat{y}_i(t)$ 和 $\hat{y}_j(t)$ 分别赋值给 $y_i(t)$ 和 $y_j(t)$,其中 $t=1, \dots, N$ 。

{ i, j 循环结束}
{整个循环结束}

3 试验结果

图 4 为分别具有负的、正的以及两个非常接近

比图 3 和表 1,可见试验值与理论值符合得很好。



零的峭度的 4 幅 256 像素 \times 256 像素的标准图像的幅度分布直方图及其峭度。试验结果如图 5 所示,图 5(a)为源图像。将它们用一个矩阵混合,得到图 5(b)的混合图像。图 5(c)是用本文算法,凸函数取 $f(u) = -\sqrt{u}, l=100$ 分离出的源信号。还用两种著名的算法——fastICA^[2] 和扩展 infomax^[1] 进行了对比试验,它们的输出图像分别如图 5(d)、5(e)所示。显然,基于峭度的 fastICA 算法不能分离最后两个峭度几乎为零的图像。扩展 infomax 算法是在两个分布模型之间切换的最大似然算法能分离某些亚高斯(峭度小于 0)和超高斯(峭度大于 0)源信号的混合信号,但在这里,它对后 3 幅图像的分离效果都不好。表 2 总结对比了采用几种不同的凸函数的准熵算法,以及上述两种算法分离图像的信噪比(SNR)和峰值信噪比(peak SNR, PSNR)。基于准熵的算法的复杂度约为 $n^2 N \lg N$ 阶,其中, n 为源信号个数, N 为样本点个数。在 P4 2.0G-M 的电脑上,用 Matlab 编程,得表 2 中的每一组结果($l=100$, 4 轮角扫描),耗时约 140 s。

已对基于准熵的算法的鲁棒性进行了充分的测试。例如,曾用类似于文献 [1] 中的如下 20 路信号——Pearlmutter 和 Parra 的 10 路音乐^[4]、6 路语音、3 路均匀分布噪声以及 1 路高斯噪声,进行 Monte Carlo 试验。每一次运行采用一个随机产生的不同的混合阵。应用上述任意一个凸函数和仅仅 $l=10$ 个量化级,通常只需 2 轮角扫描就能很好地恢复出全部源信号。限于篇幅,详细结果从略。

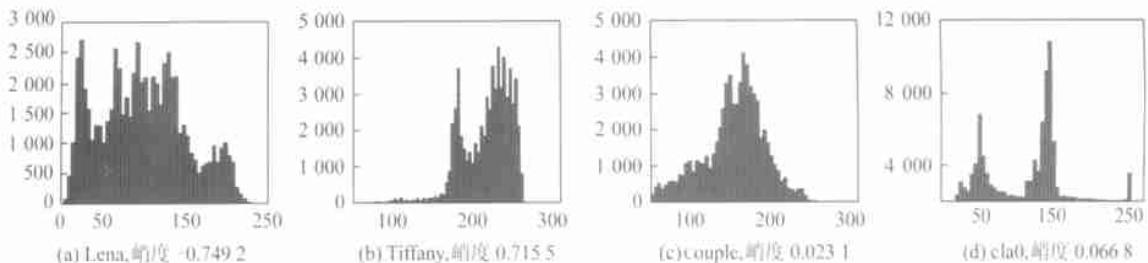


图4 图像信号的幅度分布统计直方图和峭度

Fig. 4 Distribution Statistics of Signals and Their Kurtoses



图5 图像信号盲分离结果

Fig. 5 Result of BSS

表2 基于准熵的算法、fastICA、扩展 infomax 分离图像的 SNR 和 PSNR /dB

Tab. 2 Result of SNR and PSNR Based on Quasi-entropy

凸函数 $f(u)$	Lena		Tiffany		Couple		cla0		平均	
	SNR	PSNR	SNR	PSNR	SNR	PSNR	SNR	PSNR	SNR	PSNR
$-\sqrt{u}$	24.9	36.7	19.9	31.0	15.0	28.8	30.7	41.7	22.6	34.6
$\exp(u)$	26.7	38.2	20.8	32.2	14.3	28.0	52.6	60.8	28.6	39.8
$-\ln(1+u)$	27.0	38.1	20.7	31.9	13.9	27.8	45.6	53.9	26.8	37.9
$-\sqrt[4]{u}$	26.4	37.7	19.1	30.4	16.0	29.7	28.4	39.3	22.5	34.3
fastICA	14.1	26.2	15.6	26.6	3.5	18.6	1.9	16.4	8.8	21.9
扩展 infomax	21.8	30.7	2.0	14.5	6.0	20.5	3.0	16.4	8.2	20.5

由于准熵对连续变量的分布函数的形状或统计特性没有作先验假设, 而是直接利用从样本中得出的分布函数值的无偏估计, 因此基于准熵的盲分离算法可以分离任意连续分布的信号, 包括峭度为零的信号。文献[4]提出了基于占格率的盲分离算法, 而占格率的期望也是一种准熵。

参 考 文 献

- Lee T W, Girolami M, Sejnowski T J. Independent Component Analysis Using an Extended Infomax Algorithm for Mixed Subgaussian and Supergaussian Sources. *Neural Computation*, 1999(11): 417~441
- Hyvärinen A, Oja E. A Fast Fixed-Point Algorithm for

Independent Component Analysis. *Neural Computation*, 1997(9): 1 483~1 492

- 陈 阳. 新的独立性度量及其在混沌信号分析中的应用. *东南大学学报*, 2003, 33(1): 13~18
- Chen Y, He Z Y. A New Method for Blind Source Separation. *ICONIP 2001*, Shanghai, 2001
- Comon P. Independent Component Analysis, A New Concept? *Signal Processing*, 1994, 36: 287~314
- Pearlmutter B A, Parra L C. A Context-sensitive Generalization of ICA. *ICONIP'96*, Hong Kong, 1996

第一作者简介: 陈阳, 博士后。现主要从事盲信号处理、混沌、神经网络等方面的研究。

E-mail: cheny@seu.edu.cn

A Class of New Independence Measures and Their Application to Blind Signal Separation

CHEN Yang¹ HE Zhenya¹

(1 Department of Radio Engineering, Southeast University, 2 Sipailou, Nanjing 210096, China)

Abstract: A class of new infinitely many independence measures named quasi-entropy (QE) is proposed, in which strictly convex functions are used to evaluate the uniformity of the joint probability of the variables. In QE, none of priori assumptions is made on the shape or statistical features of the distribution functions of continuous variables but unbiased estimates of the values of distribution functions are obtained from the samples. Therefore, blind separation algorithms based on QE can separate signals with arbitrary continuous distributions, including those with zero kurtoses. The superior performance of the QE-based algorithms is verified by comparison experiments with previous algorithms.

Key words: blind separation; independence; independent component analysis

About the first author: CHEN Yang, post-doctoral fellow, concentrated on the research in blind signal processing, chaos and neural networks, etc.

E-mail: cheny@seu.edu.cn

(责任编辑: 晓晨)