过完备 ICA 算法在语音信号提取中的应用①

郭水旺，李白燕

(黄淮学院 信息工程系，驻马店 463000)

摘　要：在语音信号处理中常用麦克风采集语音，然后用算法进行提取和分离，目前常用的有独立分量分析(Independent component Analysis，ICA)算法。但是当麦克风个数少于说话人的个数时，即欠定情形，此时语音信号的提取需采用过完备 ICA 算法。提出了一种基于过完备 ICA 算法的两步算法：估计混合矩阵的几何算法和估计源矩阵的最短路径法。该算法能在欠定情形下对语音信号的提取有很好的作用，仿真实验验证了这一结果。

关键词：独立分量分析；过完备；欠定；语音信号提取；两步算法

Application of Overcomplete ICA to Speech Signal Extraction

GUO Shui-Wang, LI Bai-Yan

(Department of Information Engineering, Huanghuai University, Zhumadian 463000, China)

Abstract: In speech signals processing, speech signals are collected by microphones, and then use algorithms to extract and separate, ICA (Independent component Analysis, ICA) algorithm is in common use.When the number of microphone is less than speakers number, that is called underdetermined, then speech signals’ extraction needs overcomplete ICA algorithm . A two-step approach is introduced:geometric algorithm to estimate a mixing matrix and shortest path to estimate sources. The algorithm of speech signal extraction is very good, the simulation results verify the results.

Key words: independent component analysis(ICA); overcomplete ICA; underdetermined; speech signal extraction; two-step approach

独立成分分析(ICA)方法为非高斯数据找到一种线性变换，使成分之间统计独立或者尽可能独立。由于其高阶统计特性，生物医学信号处理、混合语音信号分离、无线通信、图像处理、地震、声纳等领域已得到良好的应用。根据观测信号和源信号的个数，ICA 分为标准 ICA 问题(simply ICA)或者正定、不完备 ICA 问题（undercomplete ICA）或者过定和过完备 ICA 问题(Overcomplete ICA)或者欠定（underdetermined）三类问题。如图 1 所示的情况（忽略噪声），在语音信号提取中，麦克风的个数少于说话人的个数，这就是一种过完备状态，需采用过完备 ICA 算法。当然对于 ICA 分离输出的各种语音信号和噪声信号，还需要进行特征检测后处理，才能从多个 ICA 输出信号中提取出所需的语音信号。由于篇幅的关系，本文不作讨论。

1　Overcomplete ICA算法介绍

Ovecomplete ICA 的数学模型可以表述为:

$$x = As + \varepsilon \tag{1}$$

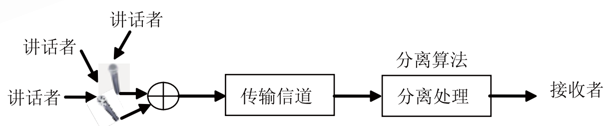


图 1　过完备语音信号提取示意图

其中，A 是一个 $m \times n$ 矩阵，且 $m < n$，A 称为混合矩阵。ε 是噪声信号，s 是 N 维未知独立信号矢量，x 是混合后的观测矢量。Lewicki 等学者提出了几种解决 Overcomplete ICA 问题的方法[1]，此类方法均可归纳为两步的 Overcomplete 算法[2]：首先在计算的过程中估

① 基金项目:河南省科技攻关计划(102102210411)
收稿时间:2010-10-28;收到修改稿时间:2010-11-15

计混合矩阵 A,即矩阵恢复(Matrix Recover)；然后通过 A 来还原出源信号向量 s，即源信号恢复(Source Recover)。本文将介绍用于欠定情况下对语音信号提取的一类两步算法。

1.1 矩阵恢复（Matrix Recover)

估计混合矩阵通常采用聚类算法[3]，该类方法要求源信号满足一定的稀疏性。TheisF J 提出的几何算法从信号的几何特性出发，对信号的稀疏特性没有特别要求，因此得到了更广泛的应用，对语音信号的提取本文将采用几何 ICA 算法[4]。

在单位球上 $S^{M-1}\subset R^M$ 选择 2N 个单位根 $\omega_1,\omega_1',\cdots,\omega_N$,，使得 $\omega_i=-\omega_i'(i=1,2,\cdots,N)$，这些根称为神经元，且每对根彼此线性独立。取固定的学习率 $\eta:N\to R$，满足 $\eta(n)>0,\sum_{n\in N}\eta(n)=\infty$ 和 $\sum_{n\in N}\eta(n)^2<\infty$ 。

选择随机变量 X 中的一个样本 $x(t)\in R$，且 $x(t)\neq 0$，将 $x(t)$ 投影到单位球，得到 $y(t)=x(t)/|x(t)|$。假设 ω_i 或 $\omega_i^{'}$ 是所有根中根据欧氏距离算出的离 $y(t)$ 最近的点，则按如下（2）式进行迭代，直到遇到收敛条件。

$$\omega_i(t+1)=\pi\{\omega_i(t)+\eta(t)\operatorname{sgn}[y(t)-\omega_i(t)]\} \tag{2}$$

式中：$\pi:R^M\setminus\{0\}\to S^{M-1}$ 表示在 R^M 中的 $M-1$ 单位球上 S^{M-1} 的投影。另外有：

$$\omega_i'(t+1)=-\omega_i(t+1) \tag{3}$$

在本次迭代中其它的神经元则不进行更新，而其它的所有神经元均按照这种迭代方式进行更新。迭代的最终结果将获得一组 $\omega_1,\omega_1',\cdots,\omega_N,\omega_N'$，将 $\omega_1,\omega_2,\cdots,\omega_N$ 组成向量矩阵 A，即为待恢复的矩阵。当观测信号维数的增长时，样本和收敛次数呈指数增长，因此几何 ICA 不适合高维情况。

1.2 源信号恢复(Source Recover)

源信号的恢复可以表示成这样一个过程：给定一个随机矢量 $x:\Omega\to R^m$ 和一个混合函数 f，找到一个独立的矢量 $s:\Omega\to R^n$ 满足另外的假设条件如 $x=f(s)$。选用模型 $x=As$，这里忽略了噪音，此时 x 完全由 A 和 s 决定。因此给定 A 和 s，观测 x 的概率可以表示为 $P(x|s,A)$。利用贝叶斯理论可以写出 s 的后验概率：

$$P(s|x,A)=\frac{P(x|s,A)P(s)}{P(x)} \tag{4}$$

已知 x 和 A 后，估计出 s 的概率。给定一些 x 的采样数据，重建 s 的一个优秀的方法是最大似然算法，利用源信号的后验概率 $P(s|x,A)$，通过解下式可以得到这个未知的源信号的估计。

$$\begin{aligned}s&=\arg\max_{x=As}P(s|x,A)\\&=\arg\max_{x=As}P(x|s,A)P(s)\end{aligned} \tag{5}$$

因为 x 完全由 s 和 A 决定，$P(x|s,A)$ 是不太重要的，从而可以得出

$$s=\arg\max_{x=As}P(s) \tag{6}$$

如果 $P(s)$ 假定是拉普拉斯算子，即 $P(s_i)(t)=a\exp(-|t|)$，那么可以得出：

$$\begin{aligned}s&=\arg\max_{x=As}\exp(-|s_1|-\ldots-|s_n|)\\&=\arg\max_{x=As}|s_1|+\ldots+|s_n|\\&=\arg\max_{x=As}|s|_1\end{aligned} \tag{7}$$

这里 $|s|_1:=\sum_i|s_i|$ 是 1-范数。经证明 s 在 1-范数是唯一的，而在其他的范数情况下不可能是唯一的[2,5]。源信号恢复的一般算法是在约束条件 $x=As$ 下对 $P(s)$ 的取最大化。这是个线性规划问题，可以使用各种优化算法处理[6]，但是进行线性规划计算，效率会比较低。上文已经假设 s 是拉普拉斯先验分布的，而且具有观测传感信号的稀疏编码的特性，所以对所有的采样值 x_λ 在 $As_\lambda=x_\lambda$ 约束条件下，可以把问题看作是最小化这 1-范数 $|s_\lambda|1$。因为一个矢量的 1-范数可以表示成和坐标轴平行的路径的长度，Bofill 和 Zibulevsky 把它叫做寻找最短路径的分解——实际上，表示 s_λ 是在 R^m 内 x_λ 最短的路径，这条路径是沿着 A 矩阵的列 $a_i=Ae_i$ 所给定的线而产生的，如图 2 所示 x_λ 的最短路径分解仅用 a_1 和 a_2 就能得到。

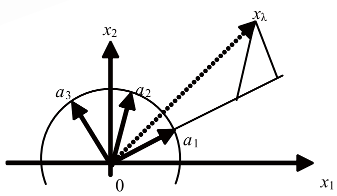


图 2　最短路径算法

至此，源矩阵的恢复已经转化为最小化 1-范数，也就是找到最短路径。本文仅处理 2 维的混合矩阵，即 $m=2$。目的是找到一个 $\arg\min_{x_\lambda=As}|s|_1$，如同上文所说的，设 $A=(a_1|\ldots|a_n)$ 表示规范化的 A 的列向量。

最短路径算法（Shortest-Path Algorithm）描述如下：设 $a_1,a_2\ldots,a_n\in R^2,n>1$，是两两线性独立的。设 $x\in R^2$ 和 $\hat{s}\in R^n$，这样，$x_\lambda=\sum_{i=1}^{n}\hat{s}_ia_i=As$。设 $j,k\in\{1,\ldots,n\}$，

使得 a_j 或 $-a_j$ 从下面最靠近 x_λ， a_k 或 $-a_k$ 从上面最靠近 x_λ，如果 $x_\lambda=0$ 则它们到固定的轴的角度是任意的。当且仅当 $\hat{s}_i=0(i\neq j,k)$ 时， $\hat{s}=\arg\min_{x_\lambda=As}|s|_1$ 而且这个 $\hat{s}$ 是唯一的。

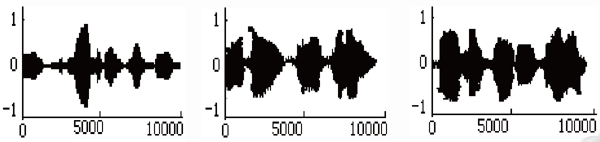
证明从略[5]。这个算法可以进行归一化，对于一个给定的采样值 x_λ 选择 A 的列 a_j 和 a_k，它们最靠近 x_λ，那么 $s_\lambda\in R^n$ 可以定义为：

$$m=2(s_\lambda)_i=\left\{\begin{matrix}((a_j\,|\,a_k)^{-1}x_\lambda)_j,i=j\\((a_j\,|\,a_k)^{-1}x_\lambda)_k,i=k\\0\qquad\qquad ,otherwise\end{matrix}\right\}\tag{8}$$

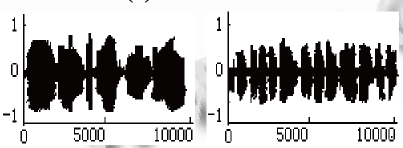
容易验证 $As_\lambda=x_\lambda$，所以利用最短路径算法，s_λ 是源矩阵的恢复。

2 语音信号提取实验

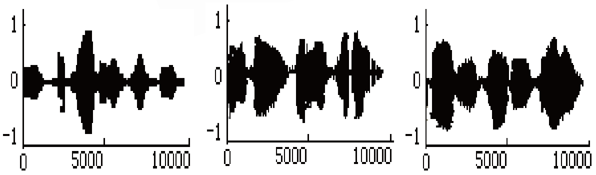
实验采用上文介绍的 Overcomplete ICA 算法进行声音混合一分离的实验，采用 Matlab 7.0 进行编码仿真，并在 PC 上运行。在安静的环境中用 PC 自带的录音机采集，得到 3 组真实的语音信号，如图 3(a)所示。然后将 3 组语音信号进行混合，从而得到两组混合信号，如图 3(b)所示,这时 $n=3$ 和 $m=2$ 的情形。采用本文描述的算法，得到如图 3(c)所示语音信号。从仿真结果可以看出，利用几何算法和最小路径法在语音信号的分离中能起到很好的作用。



(a) 3 路语音信号



(b) 2 路混合信号



(c) 对源信号的估计

图 3 利用几何方法和最短路径法的仿真结果

3 结语

在过完备状态下对语音信号实行盲源分离采用本文所采用的两步分离方法。首先，利用几何方法估计出原始的混合矩阵，然后，通过利用拉普拉斯的先验算子通过最大似然的方法恢复源矩阵，本文重点介绍了最短路径法。最后，通过仿真实验验证了两步算法对欠定情况下的语音信号的提取作用。

本文所介绍的算法中，几何 ICA 算法当观测信号维数的增长时，样本和收敛次数呈指数增长，不适合高维情况，另外对源矩阵的恢复也作了一定的假设。所以在以后的研究工作中，应该找到新的更多的混合矩阵的恢复算法，而且源矩阵恢复这一步也有待进一步论证对数据的恢复状况。

参考文献

1 Lewicki MS, Sejnowski TJ. Learning overcomplete representtations. Neural Computation, 2000,12(2):337–365.

2 Theis F. E. Lang. Formalization of the two-step approach to overcomplete BSS. Proc. of SIP, Hawaii Citeseer, 2002. 207–212.

3 Li YQ, Cichocki A, Amari SI. Sparse Component Analysis for Blind Source Separation with Less Sensors than Sources. ICA2003, Nara Japan, Citeseer 2003.89–94.

4 Fabian J, Andreas J, Carlos G, et al. Linear Geomatric ICA: Fundamentals and Algorithms. Neural Computation, 2003, 15(12):4l 9–439

5 李拥军,江宇闻.基于最短路径和自然梯度的过完备I CA算法.计算机工程,2006,32(15):16–19.

6 Chen SS, Donoho DL. Saunders MA, Atomic decomposition by basis pursuit. SIAM review, 2001,43(1):129–159.