

# 基于 Overcomplete ICA 的声音压缩模型<sup>\*</sup>

江宇闻<sup>1</sup> 黄榕波<sup>2</sup> 朱思铭<sup>1</sup>

(中山大学数学系 广州510275)<sup>1</sup> (广东药学院数学教研室 广州510224)<sup>2</sup>

**摘要** 独立成分分析(ICA)方法是近几年发展起来的一种新统计方法,旨在将所观测到的多维随机向量转换成统计上尽可能独立的成分。本文基于 Overcomplete(过完备)ICA 算法(SCO),提出了一种新的声音压缩模型。我们的实验实现了 SCO 的混合压缩与分离解压功能。

**关键词** Overcomplete ICA, 声音压缩

## Sound Compression Model Based on Overcomplete ICA

JIANG Yu-Wen<sup>1</sup> HUANG Rong-Bo<sup>2</sup> ZHU Si-Ming<sup>1</sup>

(Department of Mathematics, Sun Yat-Sen University, Guangzhou 510275)<sup>1</sup>

(Department of Mathematics, Pharmaceutical College, Guangzhou510224)<sup>2</sup>

**Abstract** ICA is a statistical method for transforming an observed multidimensional random vector into components that are mutually independent as possible. In this paper we introduce Algebra ICA algorithm (AICA) which is the efficient approach to solve Overcomplete ICA algorithm problem. Then we put forward a new sound compression structure based on the Overcomplete ICA algorithm(SCO). Finally we present an experiment on sound signal compression and decompression.

**Keywords** Overcomplete ICA, Sound compression

## 1 引言

独立成分分析(ICA)方法是近几年发展起来的一种新方法,它的目的是:为非高斯数据找到一种线性变换,使得成分与成分之间统计独立或者尽可能独立。由于其高阶统计特性,被广泛应用于信号处理、生物医学、分子学、电子通信、声音处理、航天物理、纹理结构等<sup>[8~10]</sup>。

根据混合向量  $X$  的维数  $m$  和源信号向量  $S$  的维数  $n$  的不同情形,ICA 分为如下三类问题:

(1)  $m=n$ ,称为二次 ICA 问题(quadratic ICA)或者标准 ICA 问题(simply ICA)。

(2)  $m>n$ ,由于其更高的混合投影维数空间,因此该类问题被称为不完备 ICA 问题(undercomplete ICA)。

(3)  $m<n$ ,由于其在混合空间上的维数不足,因此该类问题被称为过完备 ICA 问题(overcomplete ICA)。

Overcomplete ICA 由于混合信号向量维数  $m$  小于源信号向量维数  $n$ ,矩阵  $A$  不是一个方阵,其分离算法与传统的 ICA 算法不同。Lewicki<sup>[1]</sup>提出了采用线性规划(linear program)和自然梯度(Nature Gradient)方法的一步算法。随后 Fabian<sup>[4~6]</sup>提出了一种基于几何方法的两步学习算法,其分离效果不错。随后 Khurram<sup>[8]</sup>提出了一种与几何方法相似的代数学习算法,其学习本质与 Fabian 的算法相类似,但是在高维向量空间上表现出比 Fabian 的算法更好的扩展性。

本文提出一种基于 Overcomplete ICA 的声音压缩(Sound Compression Using Overcomplete, SCO)模型。传统的声音压缩算法都是基于对信号本身重新编码来实现压缩目

的,通过解码来恢复原信号;而 SCO 通过把  $n$  道声音源信号混合成  $m$  道声音混合信号进行压缩,注意  $n<m$ ,通过分离技术从混合信号中恢复源信号。本文第2节提出基于 Overcomplete ICA 的声音压缩模型;第3节介绍 Overcomplete ICA 的模型和算法;第4节给出实验结果;最后进行总结。

## 2 Overcomplete ICA 声音压缩模型

基于 Overcomplete ICA 的声音压缩模型 SCO 如图1所示。

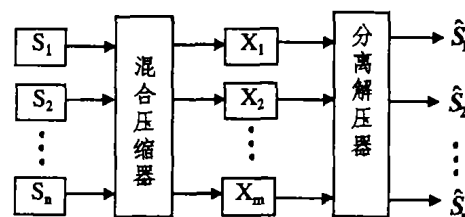


图1 SCO 模型

混合压缩器把  $n$  个源声音信号  $S_1, S_2, \dots, S_n$  混合压缩成一组  $m$  维的混合信号  $X_1, X_2, \dots, X_m$ 。这是 SCO 的压缩过程。解压分离器从混合信号中分离得到源信号,这是 SCO 的解压分离过程。在 SCO 的混合压缩过程中,  $n$  维源信号向量  $S$  通过左乘一个  $m$  行  $n$  列混合矩阵  $A$  混合成  $m$  维的混合向量  $X$ , 由于  $m<n$ , 因此实现了压缩目标;SCO 的分离解压过程就是 Overcomplete ICA 的实现过程,目前已出现了许多有效的 Overcomplete ICA 算法,本文采用 Khurram 提出的代数学习算法。

<sup>\*</sup> 本文得到国家自然科学基金资助(10371135)。江宇闻 博士研究生,主要研究方向:计算机网络,人工智能,动力系统。黄榕波 博士研究生,主要研究方向:人工智能,动力系统。朱思铭 博士生导师,主要研究方向:计算机网络与人工智能,动力系统,混沌理论。

### 3 Overcomplete ICA 算法

#### 3.1 模型及记号

任取  $m, n \in N$  令  $Mat(m \times n)$  是  $R$  上的  $m \times n$  维矩阵实向量空间,  $Gl(n) := \{W \in Mat(n \times n) | \det(W) \neq 0\}$  上的一般线性群。

在盲源信号分离(BSS)问题中,由  $n$  个相互独立的信号源产生的随机向量  $S: \Omega \rightarrow R^n$ , 通过个混合矩阵  $A = (a_1, \dots, a_n)^T \in Mat(m \times n)$  生成了混合向量  $X: \Omega \rightarrow R^m$ :

$$X = AS$$

其中  $\Omega$  表示一个特定的概率空间。所谓的盲源信号分离是指,当仅仅知道混合向量  $X$  的时候,要去通过  $X$  来得到源信号向量  $S$  和混合矩阵  $A$ 。令  $a_i = Ae_i$  表示矩阵  $A$  的列向量,其中  $i$  表示单位向量。设矩阵  $A$  是满秩的,任取两个列向量  $i, j$  都是线性独立的,  $i \neq j$ 。

#### 3.2 Overcomplete ICA 算法

在 Overcomplete 的自主学习过程中要同时求得  $A$  和  $s$  是相当困难的,因此 Fabian<sup>[5]</sup> 提出两步的 Overcomplete 算法,称为 Geo-ICA。该算法先求出  $A$  在单位球上的投影  $A'$ , 再通过它来近似地求出  $s$ 。Geo-ICA 分为两步:

- 1) 矩阵恢复(Matrix Recover Step)。
- 2) 源信号恢复(Source Recover Step)。

代数 Overcomplete ICA 算法(AICA)和 Geo-ICA 在思想上是基本一致的,都是通过先学习基向量矩阵  $A$  在单位圆上的投影再来恢复源向量。AICA 算法依据点积运算来测算向量间的距离,向量  $\alpha, \beta$  间的点积可以定义为:

$$\alpha \cdot \beta \triangleq \alpha^T \beta = |\alpha| |\beta| \cos \theta_{\alpha, \beta} \quad (1)$$

其中  $\theta_{\alpha, \beta}$  表示向量  $\alpha$  和  $\beta$  间的方向角。当  $\alpha$  和  $\beta$  表示两个不同的向量在单位圆上的投影的时候,则它们之间的距离可以完全通过方向角  $\theta_{\alpha, \beta}$  来刻画。使用点积的另一个好处是点积运算满足交换率,即:  $\alpha \cdot \beta = \beta \cdot \alpha$ 。

矩阵恢复过程采用了 win-takes-all 规则,可以归纳为下面的步骤:

1. 在  $m$  维单位圆上选择  $n$  个向量组成一个权重矩阵  $W_0 = [w_1, w_2, \dots, w_n]$ 。在该矩阵初始化的时候,初始的点可以在单位球上随机地选择也可以按照某种分布选择。

2. 在观测向量组中选择一个  $m$  维的样本  $x_i$ , 使得  $x_i \neq 0, i = 1, 2, \dots, N$ 。

3. 将选出的样本  $x_i$  投影到单位球上, 即:  $y_i = \frac{x_i}{\|x_i\|}, i = 1, 2, \dots, N$ 。

4. 计算  $y_i$  与  $W_0$  的列向量间的点积。

5. 通过下面的公式选出与  $y_i$  距离最近的列向量:

$$w_j^* = \arg \max_{w_j} |y_i^T \cdot w_j| \quad (2)$$

$$j = 1, 2, \dots, n; i = 1, 2, \dots, N$$

6. 根据距离最近的原则进行迭代更新:

$$w_j(k+1)^* = \begin{cases} \psi(w_j(k)^* + \eta(t) \operatorname{sgn}(y_i - w_j(k)^*), y_i^T \cdot w_j^* > 0; \\ \psi(w_j(k)^* + \eta(t) \operatorname{sgn}(-y_i - w_j(k)^*), y_i^T \cdot w_j^* < 0 \end{cases} \quad (3)$$

其中  $\eta(t)$  是一个依赖于迭代次数的学习率。 $\Psi(\cdot)$  是将权重矩阵中的列向量投影到单位圆上的非线性方程。

7. 更新权重矩阵:

$$\Delta W_i = W_i - W_{i-1}; i = 1, 2, \dots, N \quad (4)$$

8. 若  $|\Delta W_i| < \epsilon$ , 其中  $\epsilon > 0$ , 则该算法收敛到它的结束状态,最后得到的  $W_i$  即为所求的基向量矩阵。

9. 若算法未达到收敛条件,则重复进行2~7步运算直到达到算法收敛。

源向量还原过程使用了最短路径算法(shortest-path algorithm)。为简单起见,不妨假设  $m=2$ , 该算法的目标是使得  $\arg \min_{x_1, \dots, x_n} |S|_1$ 。其中混合矩阵  $A$  可以通过矩阵恢复过程求得。令:  $A = (a_1 | \dots | a_n)$  表示矩阵  $A$  的列向量。Fabian<sup>[7]</sup> 证明了只要通过最靠近  $x_1$  的两个列向量  $a_i$  和  $a_k$  就可以还原出源向量  $s_1$ 。算法的运算公式可以表述为:

$$(s_1)_i = \begin{cases} ((a_j | a_k)^{-1} x_1)_i, i = j \\ ((a_j | a_k)^{-1} x_1)_k, i = k \\ 0, \text{otherwise} \end{cases} \quad (5)$$

我们不难验证通过上述运算所得到的  $A$  和  $s_1$  有如下关系  $As_1 = x_1$ 。

### 4 实验结果

我们采用 AICA 算法进行声音压缩-解压的实验。所有的实验我们都采用 Matlab6.5 进行编码,并在 PIV1.7G, 256M 的 PC 上运行。

在实验中,仅考虑  $n=3$  和  $m=2$  的情形。实验中3组声音信号均采自同一个人的声音,所有声音样本均以8kHz 和8位的方式进行采样。我们将3组声音样本信号混合成为2组混合信号,其混合矩阵我们选择:

$$A = \begin{pmatrix} 0.7071 & -0.4472 & -0.9487 \\ 0.7071 & 0.8944 & 0.3162 \end{pmatrix}$$

引入评价混合矩阵(mixing matrix)的还原效果的指标:代数矩阵距离指标(AMDI)

MDI 是用来测量矩阵之间的“距离”的指标,有如下定义:

$$D(W, H) = \frac{n - \sum_{\text{rows}} \max(|W' \cdot H|)}{n} + \frac{n - \sum_{\text{cols}} \max(|W' \cdot H|)}{n}$$

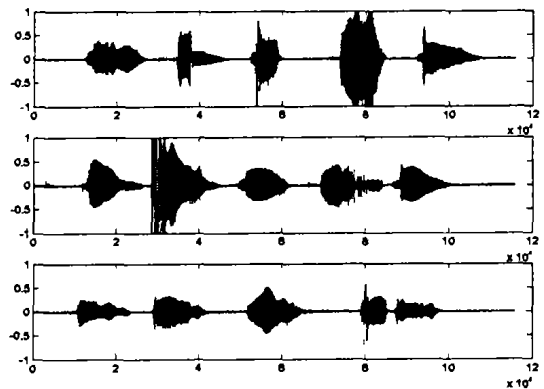


图2 混合前的声音信号

其中,  $W$  和  $H$  两个具有相同的维数且列向量归一化矩阵。MDI 有如下的性质:

$0 \leq D(W, H) \leq 2; D(W, H) = 2$  当且仅当  $W$  和  $H$  都是  $n$  维空间上的正交矩阵。

$$D(W, H) = D(H, W)$$

通过学习后得到的矩阵为:

$$W_{AICA} = \begin{pmatrix} 0.9475 & 0.9505 & 0.5274 \\ 0.3199 & -0.3108 & -0.8496 \end{pmatrix}$$

$$D(A, W) = 0.0613$$

(下转第128页)

误的碰撞检测。图4c中,我们选取最小阈值后,使其随距离作适应性变化,两个周期内仅产生13个状态更新消息,并且在近me点有较好地估算轨迹。

试验表明,我们的最小阈值适应性预测算法较好地解决了位置预测精确性与网络代价之间的矛盾,并且满足 AIMNET 通讯结构中消息的匿名性要求。

**结束语** 为降低分布式虚拟环境中大量用户间的通讯量,位置预测技术已成为分布式虚拟环境的关键技术之一。本文提出的用于 AIMNET 系统中的最小阈值适应性位置预测算法,充分利用阈值与实体间距离的关系,通过阈值的适应性变化,部分解决了预测精度与网络代价之间的矛盾;同时通过最小阈值的选取,满足 AIMNET 系统消息匿名性的要求。算法经实验验证,取得了较好的效果。

随着分布式虚拟环境在不同领域的广泛应用,为不同的应用设计特定的预测算法将成为今后的需求。总结预测算法的共性与个性,寻找统一的算法框架,为预测算法的设计提供模板,以简化算法的设计与实现,是我们以后的工作方向。

(上接第98页)

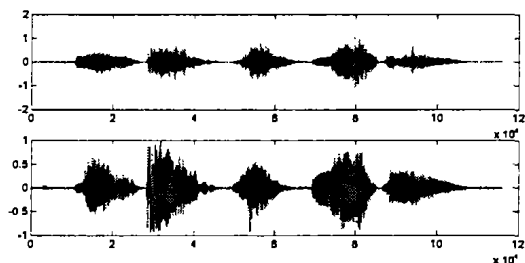


图3 混合了的声音信号

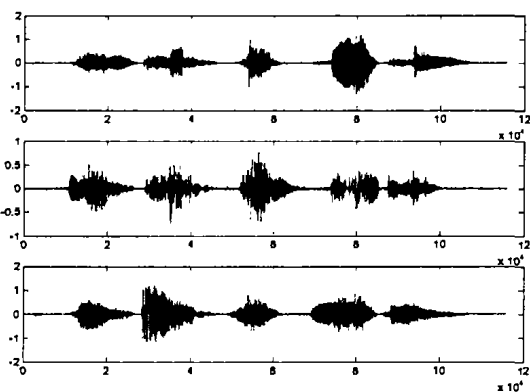


图4 使用 AICA 分离后的声音信号

可以看到图3和图4中还原得到的声音信号和原信号不是完全相同,而是混入了一定的噪音。这些噪音是由混合过程中参与混合的其他声音信号引起的,目前使用的 Overcomplete ICA 算法还不能完全消除引入的噪音。

我们采用信噪比(signal-to-noise, SNR)来衡量 Overcomplete ICA 声音压缩方面的性能:

$$SNR(s_j, \hat{s}_j) = -10 \log_{10} MSE(s_j, \hat{s}_j)$$

其中,  $s_j$  表示源信号,  $\hat{s}_j$  表示恢复后的信号,  $MSE(s_j, \hat{s}_j)$  是  $s_j$  和  $\hat{s}_j$  之间的均方误差。在上述实验中我们得到的 SNR 分别

## 参考文献

- 1 潘志庚,姜晓红,张明敏,石教英. 分布式虚拟环境综述[J]. 软件学报, 2000, 11(4): 461~467
- 2 Pope A R. The SIMNET Network and Protocols: [Report No. 7262, BBN Systems and Technologies]. Cambridge MA, June 1991
- 3 Cai Wentong, Lee F B S, Chen L. Adaptive Dead Reckoning algorithms for Distributed Interactive Simulation. In: 13<sup>th</sup> Workshop on Parallel and Distributed Simulation, Atlanta, Georgia, May 1999. 82~89
- 4 Lin K-C. Dead Reckoning and Distributed Interactive Simulation. In: Proc. of SPIE Conf. (AeroSense'95), Orlando Florida, April 1995
- 5 Krumm-Heller A. Using Determinism to Improve the Accuracy of Dead-Reckoning Algorithms. In: SimTecT 2000 Conference
- 6 赵沁平. DVENET: 分布式虚拟环境. 北京: 科学出版社, 2002. 14~18

为: 16. 138dB, 16. 115dB, 21. 755dB.

**总结** Overcomplete ICA 将高维的信息通过矩阵混合为低维的信息, 本文提出了一种基于 Overcomplete ICA 的声音压缩模型 SCO, 实现了声音的压缩。我们的实验实现了 SCO 的混合压缩和分离解压功能。SCO 的压缩和解压效果与混合矩阵  $A$  及 Overcomplete ICA 的算法密切相关, 一般地, 混合矩阵  $A$  的列数  $n$  与行数  $m$  的差越大, 则压缩比越大, 但源信号的失真越大; 因此如何选择一个合适的混合矩阵  $A$  是一个值得进一步研究的问题。

## 参考文献

- 1 Lewicki M S, Sejnowski T J. Learning Overcomplete Representations. Neural Computation, 1998
- 2 Lee T W, Lewicki M S, Giorlami M. Blind Source Separation of More Sources Than Mixture Using Overcomplete Representations. IEEE Signal Processing Letters, 1999, 6(4): 87~90
- 3 Lewicki M S, Sejnowski T J. Learning nonlinear overcomplete representations for efficient coding
- 4 Theis F J, Lang E W. Formalization of the Two-Step Approach to Overcomplete BSS
- 5 Theis F J. A Geometric Algorithm For Overcomplete Linear ICA
- 6 Theis F J, Jung A, Puntonent C G, Land E W. Linear Geometric ICA: Fundamentals and Algorithms. Neural Computation 2003, 15: 419~439
- 7 Theis F J. A Geometric Algorithm for Overcomplete ICA
- 8 Waheed K, Salem F M. Algebraic Independent Component Analysis: An Approach for Separation of Overcomplete Speech Mixtures. IEEE 2003. 775~780
- 9 Zhang L Q, Cichocki A, Amari S. Nature Gradient Algorithm for Blind Separation of Overdetemind Mixture with Additive Noise
- 10 Scott, Chen ShaoBing, Donoho D L, Saunders M A. Atomic Dcomposition by Basis Pursuit
- 11 Sparse Coding with an Overcomplete Basis Set: A Strategy Employed by V1?, Elsevier Science, 1997, 37(23): 3311~3325