

## 摘要

盲源分离是在源信号与传输信道未知的情况下仅利用接收天线的观测数据提取或恢复统计独立的源信号。盲源分离因为在语音处理、医学信号处理, 图像增强与无线通信等诸多领域具有广泛的应用前景, 从而引起了信号处理学界和神经网络学界的共同兴趣。本文围绕这一热点课题展开, 并把研究重点放在非平稳环境中的盲源分离算法研究, 本文的主要工作包括以下几个方面:

1. 提出一种变步长、变动量项因子的自适应改进的自然梯度算法实现源信号瞬时混合的盲分离。在后向传播算法的启发下, 在自然梯度学习过程中结合动量项以加快收敛速度, 同时改善自然梯度算法的稳定性。然后在分离模型中提出一个合适的测度函数自适应控制步长和动量项因子, 由此得到的变步长和变动量项因子的改进的自然梯度算法适合解决时变环境下的盲源分离问题。实验表明与经典的自然梯度算法和其它改进的自然梯度算法相比, 即使在信源个数很多的情况下, 本文所提自适应改进的自然梯度算法有更快的收敛速度和更好的稳态精确性, 当混合矩阵突变或信号功率突变时自适应改进的自然梯度算法依然有较好的跟踪能力。此外, 结合自然梯度和非线性主分量分析提出了一种块递归的盲源分离方法, 构造出按块递归更新的矩阵方程, 然后用 QR 分解和回代法求解该矩阵方程得到最优分离矩阵。所提方法与已有递归型盲源分离算法相比适于实时处理且遗忘因子的选择相对简单, 与其它块处理算法相比有较快的初始收敛速度。

2. 针对混合矩阵发生突变的情况, 提出一种基于时变遗忘因子递归广义特征分解的非白源盲分离算法。首先给出一种新的协方差矩阵的逆矩阵和时延相关矩阵乘积的递归更新方程, 广义特征向量的估计经由近似幂迭代法和压缩处理在线实现, 得到的在线算法可以避免同时估计协方差矩阵及其逆矩阵。为跟踪混合矩阵的突然变化, 提出一种新颖的基于广义特征向量关于协方差矩阵正交特性的在线决策规则, 可以判别混合矩阵是否发生突变, 从而利用时变遗忘因子的递归广义特征分解算法在混合矩阵发生突变后能获得较好的跟踪能力及精确的稳态性能。

3. 充分利用人类发音的特点和语音信号的非平稳特性, 提出一种数目未知的语音信号瞬时混合的盲分离和信源数目检测算法。首先利用递归广义特征分解在线估计对应于最大广义特征值的广义特征向量, 利用向量相似度定义广义特征向量的平均相似度用以拟合信道互扰性能曲线, 并根据平均相似度曲线提取出“高相似度区间”近似“分离区间”。然后进一步提取分离性能较好的“高相似度区间”并剔除部分“混合区间”, 得到“高分离度区间”。然后对“高分离度区间”中的广义特征向量进行多阶段聚类确定源信号的数目, 在实现估计信号数目的同时完

成信号的盲分离。此外，所提算法可以克服盲提取问题中先提取出的信源性能好而后提取出的信源性能差的缺点。

4. 针对现有卷积混合盲源分离的频域算法存在的问题，即恢复出的信号是源信号和一个未知的滤波器的卷积，且由于部分频率点处盲源分离方法的失效和无法精确解决排列和尺度模糊问题，分离出的信号中会泄露进其余的信号分量，影响分离性能。提出一种基于多信道语音增强的频域盲源分离后处理方法以消除空间干扰和背景噪声。该方法有机结合盲源分离技术与阵列处理技术，既可充分利用空域信息，又无需增加其它先验信息。首先在频域盲源分离方法中利用语谱分裂技术得到  $M \times N$  个分离信号，将多输入多输出(MIMO)混合系统分裂成  $N$  个单输入多输出(SIMO)系统，泄露进该信号的其它信号分量可以视为干扰信号，然后分别用频域多信道语音增强方法重构出各个语音信号，以消除不同信源间的空间干扰和环境噪声。仿真结果表明该算法具有良好的性能。

关键词：盲源分离(BSS)，独立分量分析 (ICA)，瞬时混合，卷积混合，非平稳信号，非白信号，自然梯度，广义特征分解(GED)，多信道语音增强，变步长，变遗忘因子

## ABSTRACT

Blind source separation (BSS) aims to extract independent but unobserved source signals from their mixtures captured by a number of sensors without knowing the mixing coefficients. Over the past two decades, the problem of BSS has received much attention in various fields, such as speech and audio processing, image enhancement and biomedical signal processing. The main works can be summarized as follows:

1. An adaptive improved natural gradient algorithm for blind separation of instantaneous mixtures of independent sources is proposed. First, inspired by the well-known back-propagation algorithm, we incorporate a momentum term into the natural gradient learning process to accelerate the convergence rate and improve the stability. Then, an estimation function for the adaptation of the separation model is obtained to adaptively control a step-size parameter and a momentum factor. The proposed natural gradient algorithm with variable step-size parameter and variable momentum factor is therefore well suited to blind source separation in a time-varying environment. The expected improvement in the convergence speed, stability and tracking ability of the proposed algorithm is demonstrated by extensive simulation results in both time-invariant and time-varying environments. The ability of the proposed algorithm to separate extremely weak (or badly scaled) sources and many sources is also verified. In addition, a block recursive approach for blind source separation is presented. Firstly, based on natural gradient and nonlinear principal component analysis, a matrix equation is obtained by block recursive updating, and then the matrix equation is solved using QR factorization and back substitution to obtain the optimal separating matrix. Compared with other existing recursive-type BSS methods, the proposed algorithm is feasible to real-time processing, and the choice of the forgetting factor is simple. Compared with other block processing methods, the proposed algorithm has fast initial convergence speed.

2. An efficient variable forgetting factor recursive generalized eigen-decomposition algorithm is developed for blind separation of non-white sources when the mixing matrix changes abruptly. We derive a new recursive update equation for the multiplication of a cross-correlation matrix and the inverse of a covariance matrix with compact form and low computational complexity. The generalized eigenvectors are recursively estimated by using the approximated power method and the deflation procedure. Without additional priori information of the mixtures, we propose a novel

on-line decision rule to track the abrupt variations of the mixing matrix and then a variable forgetting factor recursive generalized eigen-decomposition algorithm for BSS is presented for the time-varying environments. The improved tracking ability and steady-state accuracy of the proposed algorithm are validated by the computer simulation results.

3. By exploiting of speech nonstationarity, a method for estimating the number of sources from instantaneous mixtures of speech signals with unknown source number is presented, and then the sources are extracted. The first dominant generalized eigenvector is estimated by recursive generalized eigen-decomposition. The mean similarity curve of the estimated generalized eigenvector is introduced to fit inter-channel interference curve, and then "High Similarity Intervals" are extracted to approach "Separation Interval". Moreover, "High Separation Intervals" are obtained by extracting "High Similarity Intervals" with better separation performance and eliminate "Mixtures Interval". Final, the number of the sources is estimated with multistage clustering techniques and the corresponding sources are extracted. The proposed algorithm can avoid suffering the error propagation of the deflation technique, which exists in all sequential algorithms.

4. A new post-processing method for convolutive mixtures blind speech separation is proposed. It utilizes multi-channel signal enhancement to suppress spatial interference and background noise. Due to imprecision for solving the permutation ambiguity problem, frequency domain blind source separation has its fundamental limitation in separation quality. To overcome that, by splitting spectrograms, the  $M \times N$  multi-input multi-output (MIMO) system will be converted into  $N$  single-input multi-output (SIMO) system in frequency domain blind speech separation system. Furthermore, to attenuate spatial interference from competing sources and background noise, the multi-channel signal enhancement method is exploited to reconstruct source signals from the  $N$  SIMO system respectively. The separation performance of the proposed algorithm is demonstrated by experiments.

**Key words:** Blind source separation (BSS), Independent Component Analysis (ICA), instantaneous mixtures, convolutive mixtures, nonstationary sources, non-white source, natural gradient, generalized eigen-decomposition (GED), multi-channel speech enhancement, variable step-size, variable forgetting factor

## 独创性声明

本人声明所呈交的论文是我个人在导师指导下进行的研究工作及取得的研究成果。尽我所知，除了本文中特别加以标注和致谢中所罗列的内容外，论文中不包含其他人已经发表或撰写过的研究成果；也不包含为获得西安电子科技大学或其它教育机构的学位或证书而使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已 在论文中做了明确的说明并表示了谢意。

本人签名： 刘建强

日期： 2009.6.8

## 关于论文使用授权的说明

本人完全了解西安电子科技大学有关保留和使用学位论文的规定，即：学校有权保留送交论文的复印件，允许查阅和借阅论文；学校可以公布论文的全部或部分内容；可以允许采用影印、缩印或其它复制手段保存论文。（保密的论文在解密后遵守此规定）

本人签名： 刘建强

日期： 2009.6.8

导师签名： 冯大政

日期： 2009.6.9

## 第一章 绪论

### § 1.1 引言

盲源分离(BSS: Blind Source Separation)是在源信号与传输信道未知的情况下仅根据源信号的统计独立性假设从观测信号(混合信号)中提取或恢复源信号。典型的例子是鸡尾酒会(cocktail party)问题,即在一个同时存在众多谈话者的环境中,由麦克风录取谈话者声音的混合信号,通过信号处理的方法,分离出各位谈话者或特定的谈话者的声音信号。这里术语“盲”有两重含义:第一,源信号不能被观测;第二,源信号是如何混合的是未知的。显然,当源信号与传输信道参数都未知时,盲源分离是一种很自然的选择。盲源分离的混合过程需要根据实际应用建模成不同的多输入多输出的数学模型,可分为无记忆和有记忆的,也可分别称为瞬时混合和卷积混合。

近二十多年来,盲源分离已成为信号处理领域最热门的研究课题之一,其研究正突飞猛进,已有诸多著作面世<sup>[1-9]</sup>。利用源信号或输入信号的部分量化信息(如源信号的相互统计独立性、恒模性、非白、非平稳和非高斯统计性等等),盲信号处理技术可以有效地实现系统辨识和源信号的恢复等。盲源分离技术已在语音识别、图像处理、生物信号处理、通信(特别是无线通信)信号处理以及源信号不易或不可能测量的自然信号处理等领域显示出十分诱人的应用前景。

### § 1.2 瞬时混合盲源分离研究发展与现状

盲源分离问题的早期研究集中于模型相对简单的瞬时混合情形。Jutten 和 Herault<sup>[10-12]</sup>最早在 1986 年提出了一种类神经盲源分离方法,该方法基于反馈神经网络,通过取奇次的非线性函数构成 Hebb 训练,从而达到盲源分离的目的。虽然该方法只能用于两个混合源信号的盲分离,非线性函数的选取具有随意性,并且缺乏理论解释,但是 Jutten 和 Herault 的工作作为盲源分离问题的研究开端,具有里程碑式的意义。Tong 等人<sup>[13]</sup>分析了盲源分离问题的可辨识性和不确定性,并给出一种基于高阶统计量的矩阵代数分解方法。Cardoso 等人<sup>[14]</sup>提出了一类基于高阶统计量的近似联合对角化盲源分离方法,并可用于盲波束形成。Comon<sup>[15]</sup>分析了瞬时混合信号的盲分离问题,明确了独立分量分析(ICA: Independent Component Analysis)的概念,利用可以测度源信号统计独立性的 Kullback-Leiber(K-L)准则作为对比函数(Contrast Function),并通过对概率密度函数的高阶近似,得出用于测度信号各分量统计独立的对比函数,并由此得出一类基于特征分解的独立分量分析

方法。这种方法需要对观测信号进行白化预处理,而且是一种批处理算法,运算量大,实时性差。需要指出的是独立分量分析虽然在概念上是一个简单的问题,但它可以引入许多有趣的技术,如高阶统计量、信息论、稀疏编码、投影跟踪和信息几何等等,并有着广泛的应用。Bell 等人<sup>[16]</sup>基于信息理论,通过最大化输出非线性节点的熵,得出一种最大信息传输的准则函数并由此导出一种自适应的盲源分离方法和盲反卷积方法,当该方法中非线性函数的选取逼近源信号的概率分布时,可以较好的恢复出源信号,但收敛慢,且需要矩阵求逆运算,稳定性也比较差。在从神经网络角度研究自适应信号分离的独立分量分析方法时,为克服随机梯度优化方法收敛比较慢的缺点,Cichocki 等人<sup>[17]</sup>提出了自然梯度算法,它的学习规则不仅容易计算,而且可以避免矩阵求逆运算。后来,Amari<sup>[18]</sup>从信息几何的角度诠释了自然梯度算法工作原理,并证明了自然梯度算法的有效性,对自然梯度算法的进一步研究参见文献[4]及其中文献,自然梯度算法的收敛性取决于步长的选取和初始分离矩阵的取值,如何选择这些参数以同时保证收敛速度、稳定性和跟踪能力较为关键。Cardoso 等人<sup>[19]</sup>则从自适应信号分离必须具有等变化性(equivariant)出发,相对于普通梯度(绝对梯度),定义了相对梯度,由之得到的算法称为相对梯度算法,可以证明自然梯度算法与相对梯度算法是等价的,因而自然梯度算法也具有等变化性。Yang 等人<sup>[20]</sup>已证明:熵最大化与独立分量分析是等价的。随着盲分离算法研究的深入,逐渐引入和产生了许多相关理论分析,如Cao 等人<sup>[21]</sup>的行可分解理论,Amari 等人<sup>[22]</sup>的半参数统计理论,Cardoso 等人<sup>[23-24]</sup>的局部稳定性理论等。

受独立分量分析的启发,Karhunen 等人<sup>[25-27]</sup>在主分量分析方法(PCA: Principal Component Analysis)中引入某种非线性,提出了一类非线性主分量分析的盲源分离算法。Pajunen 等人<sup>[26]</sup>通过结合非线性主分量分析(NPCA)和近似投影子空间跟踪(PAST)算法<sup>[28]</sup>提出了递归最小二乘(RLS)型的盲源分离算法,实验结果表明该算法的收敛速度优于已有的自适应盲源分离算法。由于RLS算法的推导是基于普通的随机梯度,而自然梯度算法明显优于普通的随机梯度算法,朱孝龙等人<sup>[29]</sup>将自然梯度算法与非线性主分量分析的递归最小二乘算法相结合,提出了基于自然梯度递归最小二乘算法。需要注意的是这两种递归算法<sup>[26,29]</sup>都需要对观测信号预白化。朱孝龙等人<sup>[30]</sup>随后又提出了无需对观测信号预白化的基于自然梯度和NPCA的递归最小二乘盲分离方法。

基于广义特征分解的盲源分离方法<sup>[31-33]</sup>是一种简单而有效的离线盲分离方法,它利用源信号的特性,如源信号的非白性、非高斯性或非平稳性,由混合信号的二阶或四阶统计量构造相应的矩阵束,通过对构成矩阵束的两个矩阵的精确联合对角化可求出矩阵束的广义特征向量矩阵,即为待求的分离矩阵。Cardoso 等人<sup>[34]</sup>提出了基于二阶统计量的多个不同时延相关矩阵的近似联合对角化算法

(SOBI), 通过对经白化预处理的观测信号构成不同延迟的相关矩阵, 通过求一个酉矩阵使得上述不同延迟的相关矩阵实现近似联合对角化, 从而可分离矩阵实现盲信号分离。基于非正交近似联合对角化的盲源分离算法<sup>[35-38]</sup>则可避免对观测信号进行白化预处理。

除了同步提取所有源信号的并行算法外, 盲源分离还可以采用一次只提取一个源信号的串行算法<sup>[39]</sup>, 比较典型的有定点(fixed-point)算法(也被称为FastICA算法)<sup>[40-45]</sup>, Hyvarinen等人<sup>[40-41]</sup>首先基于源信号非高斯性测度(或峭度), 给出一类定点训练算法, 该类算法能提取单个具有正或负峭度的源信号。该类算法由于计算复杂度小、分离性能好而受到广泛的关注。不同学者又相继提出了进一步改进或简化的FastICA算法及适用于不同情形的FastICA算法, 如复值信号的FastICA算法<sup>[42]</sup>, 复值非高斯非圆信号的FastICA算法<sup>[43]</sup>等。该类算法的收敛性能也得到了深入研究<sup>[44-45]</sup>。Li等人<sup>[46]</sup>提出了一种形式上类似于并行算法的盲提取算法, 该算法的特点是一步就可以提取出一个信号, 而不同于传统的盲信号提取算法提取一个源信号需要分提取和数据压缩两步完成。李细林等人<sup>[47]</sup>则提出了基于二阶统计量的盲信号提取算法, 通过对自相关矩阵的近似联合对角化实现信号的盲提取, 并且在部分先验信息给定的情况下可以提取特定的信号。

在现实世界中考虑一些问题时, 必须考虑源信号的非平稳性。所以近年来, 盲源分离问题中的非平稳性引起了大家广泛的关注。盲源分离中的非平稳性问题可以分为以下三类:

(1) 混合函数不时变, 源信号为非平稳的。研究盲分离中的非平稳性问题的大多数学者考虑的是这种情况, Souloumiac<sup>[48]</sup>对两个不同时刻协方差矩阵进行确切的联合对角化(这等价于解一个广义特征值问题), 由此得出的算法能实现非平稳源的盲分离。Matsuoka等人<sup>[49]</sup>考虑了一个本质上是高斯互信息的标准, 但他们对于其中的最优化问题只提出一个随机梯度问题。Pham等人<sup>[50]</sup>提出的拟牛顿型方法显然要优于前者。Choi等人<sup>[51]</sup>利用不同的时间窗数据框架结构得到多个时延相关矩阵, 然后用近似联合对角化方法来估计混合矩阵。但在实际环境中的非平稳信号大多不宜视为短时平稳的, 因此基于时频分布的非平稳信号的盲分离技术<sup>[52-54]</sup>已引起了广泛的专注。主要原因之一是它能够实现非平稳相关信源的分离, 而传统盲源分离方法的基本假设为信源是统计独立的。

(2) 混合系统时变、源信号为平稳的。混合系统时变也可分为混合矩阵慢变和突变, 这种情况通常需要在线算法跟踪混合矩阵的变化。Everson等<sup>[55]</sup>提出一种有效的跟踪混合矩阵同时估计源信号的方法来解决盲分离问题, 这里独立源的混合过程根据马尔可夫过程建模成进化过程。Ahmed<sup>[56]</sup>在混合系统时变, 源的密度能建模成有限个高斯混合的情况下, 用贝叶斯方法, 即著名的马尔可夫链蒙特卡罗(MCMC)方法来仿真实现以解决这种情况下的盲分离问题。Parga等人<sup>[57]</sup>也提出

一种混合系统缓慢平滑变化时的盲分离方法,在源信号平稳和各态历经时给出一种代数解法,但并没有给出仿真结果。

(3) 混合系统时变、源信号非平稳。这种情况最为复杂,研究相对较少,张贤达等人<sup>[58]</sup>提出了基于卡尔曼滤波的盲源分离算法,仿真实验表明该算法可以适用于混合矩阵缓慢变化、源信号是语音信号的情况。Pokharel 等人<sup>[59]</sup>提出了基于递归广义特征分解的盲源分离方法用以改善算法的收敛速度和跟踪性能,该文中分别推导了基于非白、非高斯和非平稳源信号的递归算法,仿真实验表明所提递归算法具有快速和稳定的收敛性能,且给出了混合矩阵慢变、源信号是语音信号的实验,但在该例中所用语音信号的样本太少(样本数为 2000),无法真正体现的语音信号的非平稳特性。该文也给出了信道慢变、源信号为 BPSK 信号的实验。

在实际应用上需要考虑混合矩阵突变的情形,混合矩阵发生突变时通常要求算法具有较快的跟踪性能。显然利用步长因子为常数的随机梯度的算法不适用于跟踪突变的环境,这是因为当步长因子太小时,算法的收敛速度太慢,而当步长因子太大时,失调太大可能会导致算法发散。而利用时变步长因子的在线算法来跟踪时变的环境通常能在快速收敛的同时得到更为精确的稳态性能。Douglas 等人<sup>[60]</sup>回顾了几种自适应步长选择方法,重点介绍了梯度自适应步长选择算法。Thomas 等人<sup>[61]</sup>给出了一种自调节步长选择算法,大大加快了收敛速度,跟踪能力也有很大提高。Chambers 等人<sup>[62]</sup>运用了一种梯度自适应步长因子的 ESAI 算法,将输入信号动态变化时的算法性能和混合矩阵的变化联系在一起。Yuan 等人<sup>[63]</sup>提出了变步长符号自然梯度算法,用于独立信号的在线盲分离,该算法的符号操作是通过推导归一化动态分离模型得到的。类似地,利用常数遗忘因子的递归盲源分离算法也不适用于跟踪突变的环境,这是因为当遗忘因子很接近于 1 时,算法的收敛速度太慢,而当遗忘因子太小时,失调太大可能会导致算法发散。而利用时变遗忘因子的递归算法来跟踪时变的环境通常能在快速收敛的同时得到更为精确的稳态性能。对于混合矩阵突变的情况,滑窗递归最小二乘算法比指数加权窗有好的跟踪性能<sup>[64]</sup>,基于 Papaodysseus 矩阵求逆引理<sup>[65]</sup>,利用观测数据通过滑窗加权求和的方式构造类似文献[59]的递归算法,但是这种方法需要根据实际情况选择滑窗的长度,滑窗太长则稳态性能较好而收敛速度较慢,滑窗较短可提高收敛速度但稳态性能较差,滑窗太短则易导致算法发散。

已有研究大多假设源信号个数是预先知道的,从而采用的分离矩阵行数等于信源个数,而实际中源信号的个数往往是未知的,甚至是随实际环境动态变化的,欠定盲信号分离只能成功分离部分源信号<sup>[66-67]</sup>,因此通常研究超定和适定情况下信源数目未知的盲源分离问题。Chichocki 等人<sup>[68]</sup>将适定盲源分离的自然梯度算法直接推广到源信号个数未知的超定盲源分离的新分离模型中,算法的输出信号中会有冗余信号,需要利用输出信号之间的相关性,检测并去除冗余输出同时确定

源信号个数。Ye 等人<sup>[69-70]</sup>研究了源信号数目未知与动态变化时的超定盲源分离可分离性及对比函数问题,推导了信源数目未知的盲源分离的自然梯度算法。Lv 等人<sup>[71]</sup>则给出了一种数目未知的稀疏信源的统一的盲源分离方法。

随着盲源分离技术应用范围的不断拓展及相应限制性条件的逐步放宽,不同的学者们提出了大量适用于特定环境的盲源分离算法,如基于非负矩阵分解和稀疏分量分析的欠定盲源分离算法<sup>[72-73]</sup>,在非线性混合情况下的盲源分离算法<sup>[74]</sup>等,这里不再一一赘述。

盲信号处理理论经过二十多年发展,无论在研究方向或应用领域都在不断扩张,除了上述已提及的部分国内学者的研究成果,其他国内学者也在瞬时混合盲信号处理理论和应用技术的各个方面都取得了重大进展<sup>[75-97]</sup>。凌燮亭等人<sup>[75-76]</sup>针对无线通信应用较早开展了盲源分离的研究。何振亚、刘璐和杨绿溪等人<sup>[77-81]</sup>基于特征分析和高阶谱的盲源分离,提出了一系列基于高阶统计量和信息理论的判据和算法,进一步研究了基于独立分量分析的数字水印技术。胡光锐等人<sup>[82]</sup>研究了基于统计独立的盲语音分离问题,并提出了基于混合高斯模型概率密度估计的语音分离方法。马远良等人<sup>[83-86]</sup>提出了水声信号的盲分离的实数和复数算法和性能评价准则,并提出了欠定情况下高阶累积量的复数混合矩阵盲估计算法。谢胜利等人<sup>[87-89]</sup>研究了盲源分离几何方法,盲信号分离的多路回波消除结构及稀疏分量分析。邱天爽等人<sup>[90]</sup>研究了基于分数低阶矩的脉冲信号的盲分离,并应用于水声和语音信号处理。殷福亮等人<sup>[91]</sup>研究了基于去相关和非平稳性的盲源分离方法。陈天平等人<sup>[92]</sup>研究了白化过程的动态行为和自然梯度算法的局部稳定性。王君等人<sup>[93-94]</sup>研究高阶统计量、径向基函数和遗传算法实现了非线性盲源分离,并提出了有交叉项抑制的时频分布基盲源分离技术。张丽清等人<sup>[95]</sup>研究有激活函数适应的自然梯度基盲源分离方法。张贤达等人<sup>[29-30, 47, 70-71, 96-99]</sup>对盲源分离进行了系统深入的研究,提出了基于递归最小二乘的盲源分离方法,信源数目未知和动态变化时的盲信号分离,学习率的自适应确定的分阶段方法与模糊集方法和基于二阶统计量的连续盲提取方法等。冯大政等人<sup>[100-101]</sup>研究了改进白化技术,利用矩阵的双正交性提出了一种信号盲分离问题的多阶段分解算法及矩阵组稳健代数方法从阵列信号中提取统计独立信号。此外,还有许多学者在瞬时混合盲源分离领域收获了卓著成果,这里不一一赘述。

### § 1.3 卷积混合盲源分离研究发展与现状

同瞬时混合模型相比,卷积混合是更接近真实状况的一种模型,对于解决实际问题具有极为重要的意义,例如,通信系统中的串扰问题,电话、电视会议系统中的回波对消问题,都可以看成是卷积混合信号的盲分离问题,因而卷积混合

的盲源分离问题在近几年来受到越来越多的关注。并且随着计算机的运算速度越来越快,该项技术可以广泛的应用在语音信号处理、图像信号处理、通讯信号处理、水声信号处理以及医学信号的检测、数据挖掘等领域,有着广泛的应用前景。目前很多卷积混合信号的盲分离算法是瞬时混合信号盲分离算法在卷积混合情况下的推广,而且这种推广既可以在时域内进行,又可以在频域内进行。因此卷积混合信号盲分离算法可以分为时域与频域算法两大类。

时域方法是较早提出的一类针对卷积混合信号盲分离的方法,可分为基于二阶统计量、高阶统计量和信息论等的算法。基于二阶统计量的算法有 Weinstein 算法<sup>[102]</sup>、Gerven 算法<sup>[103]</sup>及 Kawamoto 等人<sup>[104]</sup>基于源信号非平稳性的瞬时混合信号盲分离算法推广到卷积混合信号的盲分离算法,以及 Bousbia 等人<sup>[105]</sup>把 SOBI 算法推广得到的相关矩阵联合块状对角化算法等,基于二阶统计量算法的优点在于,只需要比较少的数据量就能获得较好的估计性能。基于高阶统计量的算法有 Yellin 算法<sup>[106]</sup>、推广的 EASI 算法<sup>[107]</sup>等,这类算法不仅能够提供系统的相位信息,而且可以有效地抑制高斯噪声。基于信息论的算法有 Choi 等人<sup>[108]</sup>的动态递归神经网络算法及 Douglas 等人<sup>[109]</sup>的自然梯度算法等。由于时域方法规避了频域算法中必须解决的排列模糊和尺度模糊问题,是卷积混合盲源分离的一个重要研究方向。但是,时域方法需要联合估计很多待定参数,计算复杂度高。

频域盲源分离方法<sup>[110-142]</sup>利用短时傅立叶变换把时域卷积混合盲分离问题转换为多个频域瞬时混合盲分离问题。此类方法可利用现有的瞬时混合盲源分离算法对各个频率分离问题进行独立处理<sup>[110]</sup>。一方面大大降低了运算量,另一方面便于参考借鉴性能优越的瞬时混合盲分离算法。然而要想得到满意的分离结果,所有的频域盲分离算法都必须解决盲源分离问题固有的排列和尺度模糊问题<sup>[111]</sup>。因为时域方法是联合处理,上述两种不确定性不影响其分离性能,而在频域方法中,各个频率分量上的排列不一致导致信号的重新混合,同时尺度不确定性也会导致合成信号频谱畸变。所以如何解决排列问题和尺度问题是研究各种频域盲分离问题的关键所在。目前用来解决频域排列问题的方法大致有以下几类,第一类方法基于频率平滑<sup>[112-113]</sup>,通过在时域限制滤波器长度来解决置换问题,这种方法可有效分离在没有混响的环境中录制的麦克风信号,但当混响存在时,该方法的性能严重下降;第二类方法基于阵列处理波达方向估计<sup>[114-115]</sup>,通过分析分离矩阵所形成的方向图来估计波达方向,并根据波达方向进行置换。第三类方法利用相邻频率间的相关性<sup>[116-117]</sup>来解决置换问题,即计算所有可能排列中相邻频率的频谱包络特性,相邻频率频谱包络相干性最大的排列被视为一致的排列。但由于该方法需要依次置换相邻频率,当某频率处盲源分离失效而得不到有效的分离矩阵时将导致误差累积,从而影响整个算法的性能;当各个频率处的盲分离性能都很好时,频率相关方法具有较高的精确度,而利用波达方向方法时一个频率上的失调不会

影响到其他频率,因此具有稳健性的优点。将两者相结合可以有效的解决排列模糊问题<sup>[118]</sup>。虽然各国学者针对频率置换问题在以上三大类方法的基础上已提出很多频率置换方法<sup>[119-122]</sup>,并且在特定的试验环境中能获得较满意的性能,但为了提高实用化水平,进一步研究能够在各种实际声音环境中通用的,既精确又稳健的频率置换问题依然是频域盲源分离问题的难点和热点。而现有的频率置换方法大多基于两输入多输出系统,即两个语音信号,两个或两个以上麦克风阵元模型,严重制约了盲源分离在语音增强等实际问题中的应用。在实际环境中常要考虑多个,甚至来自同一方向的多个说话者,而上述的频率置换方法大都不能直接推广到两个以上源信号卷积混合的问题中,因此研究多个语音信号卷积混合盲分离的频率置换问题是在盲源分离技术实用化的道路上需要攻克的主要难关之一。由于频域盲源分离方法便于实时处理,已有不同的学者给出了实时(或接近实时)处理算法和演示系统<sup>[123-125]</sup>。

在有混响的环境中,盲源分离的性能受到了很大的影响甚至无法分离出源信号,这是单纯的盲源分离技术无法克服的障碍。这是因为传统的盲源分离问题把多输入多输出声学系统当作一个“黑匣子”,在恢复源信号的过程中不考虑系统的内部结构,即不考虑空间声学的特性如独立语音源的方位,这样分离的信号都是单耳信号,这使得无法利用许多优越的多信道信号增强算法进一步抑制干扰和噪声信号。近来,获得分离信号的空间可辨性已经出现在立体声和多信道语音处理系统中。因此考虑把盲源分离和宽带阵列处理,如语音增强等技术有效的结合起来,真正的把盲源分离技术推向实用。同时,如何进一步利用阵列几何信息、语音信号固有特性以及现有自适应滤波方法等技术也引起了极大的关注。Parra等人<sup>[126]</sup>提出结合卷积盲源分离和几何波束形成的几何源分离方法,通过几何线性约束可以在一定程度上解决频域盲源分离问题中的内在的不确定性问题。Knaak等人<sup>[127]</sup>进一步提出了几何约束的独立分量分析算法,并详细的分析了算法的收敛性能。需要注意的是与基本的频域盲源分离方法相比,基于几何约束的盲源分离法<sup>[126-127]</sup>都需要一定的先验信息,如目标信号的方向角的大致估值等。一般频域盲源分离算法对每个频率作同样的处理,但是人耳对不同频率的敏感性不同<sup>[128]</sup>,因此Pan等人<sup>[129]</sup>提出一种基于E滤波器最小均方算法和人类听觉系统的感知卷积盲源分离算法,这种方法的基本思想是可以重点处理人耳比较敏感的频率而忽略人耳不敏感的频率。Low等人<sup>[130]</sup>提出一种卷积盲信号分离的后处理方法,该方法本质上是一种基于子带的语音增强方案,它基于盲源分离技术把目标信号和干扰信号进行空间分离,然后利用自适应噪声对消器抑制干扰。但是该方法主要适用于只有一个目标信号的情况。Huang等人<sup>[131]</sup>提出了一种在混响环境中基于盲信道辨识的两阶段算法实现语音信号的分离和去混响。Rivet等人<sup>[132]</sup>结合音视频语音处理和盲源分离技术从卷积混合中提取语音信号,利用音视频信息给出一种新的统计模型,

用以解决盲源分离问题中的置换和尺度的不确定性。此外,在欠定情况下卷积混合盲源分离算法也取得了深入研究<sup>[133-135]</sup>。

国内学者在卷积混合盲源分离各个方面的研究也取得了重大进展。何培宇等人<sup>[136-137]</sup>研究了语音信号的盲分离,对模型进行了简化,对多路声回波进行抑制。张安清等人<sup>[138]</sup>提出了频域卷积混合盲源分离的新算法。苏野平等<sup>[139]</sup>提出了一种改进的基于高阶累积量的语音盲分离算法,算法具有较快的速度。梅铁民等人<sup>[140]</sup>提出了一种基于在时域中优化频域独立性标准的盲源分离方法,可以避免频率置换问题。何昭水等人<sup>[141]</sup>基于稀疏性表示提出了一种基于自然梯度的频域盲源分离算法,并可应用于欠定盲源分离问题。张贤达等人<sup>[142]</sup>提出了基于斜投影的卷积信道盲信号分离。此外,还有许多学者在卷积混合盲源分离领域作出了卓著的成绩,这里不一一赘述。

## § 1.4 本文的研究内容

本论文对非平稳环境中的盲源分离问题进行了研究。这里“非平稳”有两重含义,第一,源信号是非平稳信号;第二传输信道是非平稳的,即混合矩阵是动态变化的。本文重点研究混合矩阵发生突变的情形,提出了不同的利用参数自适应调节的在线算法跟踪时变环境。此外还研究了数目未知的语音信号瞬时混合的盲分离问题,及语音信号卷积混合频域算法中的后处理方法。具体内容安排如下:

第二章重点介绍了盲源分离问题的基本知识。介绍了盲源分离的概念,研究了盲源分离问题不同的混合模型,对盲源分离问题的可分离性、分离原则及分离系统的设计进行了分析论证。

第三章首先提出一种变步长、变动量项因子的自然梯度算法实现源信号瞬时混合的盲分离。在后向传播算法的启发下,我们在自然梯度的学习过程中引入动量项以加快收敛速度,同时可以改善自然梯度算法的稳定性。然后,在分离模型中引入一个合适的测度函数来控制步长和动量项因子,由此得到的变步长、变动量项因子的自然梯度算法非常适合解决时变环境下的盲源分离问题。与经典的自然梯度算法及其它改进的自然梯度算法相比,即使在信源个数很多的情况下,自适应修正的自然梯度算法有更快的收敛速度和更好的稳态精确性,当混合矩阵突变或信号功率突变时自适应修正的自然梯度算法依然有较好的跟踪能力。此外,我们结合自然梯度和非线性主分量分析提出了一种块递归的盲源分离方法,构造出按块递归更新的矩阵方程,然后用 QR 分解和回代法求解该矩阵方程得到最优分离矩阵。与已有递归型盲源分离算法相比,所提方法适于实时处理且遗忘因子的选择相对简单,与其它块处理算法相比有较快的初始收敛速度。

第四章针对混合矩阵发生突变的情况,提出一种基于时变遗忘因子递归广义

特征分解的非盲源分离算法。首先给出一种协方差矩阵的逆矩阵和互相关矩阵乘积的递归更新方程, 具有较低的计算复杂度, 广义特征向量的估计经由近似幂迭代法和压缩处理在线实现, 得到的在线算法可以避免同时估计协方差矩阵及其逆矩阵。为跟踪混合矩阵的突然变化, 提出一种新颖的基于广义特征向量的在线决策规则, 可以判别混合矩阵是否发生突变, 从而利用变遗忘因子的递归广义特征分解算法改善在线盲源分离算法的性能, 利用生物信号进行仿真的实验结果表明所提在线决策规则可以准确跟踪混合矩阵的突然变化, 所用变遗忘因子的算法既能获得较好的跟踪能力又能得到精确的稳态性能。

第五章充分利用语音信号的非平稳特性, 对于数目未知的语音信号瞬时混合的盲分离问题, 在实现信号数目检测的同时完成信号的盲分离, 并可克服盲提取问题中先提取出的信源性能好而后提取出的信源性能差的缺点。首先利用递归广义特征分解在线估计的对应于最大广义特征值的广义特征向量, 由向量的相似度定义平均相似度用以拟合信道互扰性能曲线, 并根据平均相似度曲线提取出“高相似度区间”近似“分离区间”。然后进一步提取分离性能较好的“高相似度区间”并剔除部分“混合区间”, 得到“高分离度区间”。然后利用多阶段聚类的方法确定源信号的数目, 同时实现信号的盲分离或提取。仿真实现表明当信源数目适中时所提算法可以精确的估计出信源数目, 发生错误估计或漏估信号的概率较低, 并且用估计性能最好的广义特征向量作为分离向量, 从而获得分离性能最好的语音信号

第六章针对现有卷积混合盲源分离的频域方法存在的问题, 即恢复出的信号是源信号和一个未知的滤波器的卷积, 且由于部分频率点处盲源分离方法的失效和无法精确解决排列和尺度模糊问题, 分离出的信号中会泄露进其余的信号分量, 影响分离性能。提出一种基于多信道语音增强的频域盲源分离后处理方法以消除空间干扰和背景噪声。该方法有机结合盲源分离技术与阵列处理技术, 既可充分利用空域信息, 又无需增加其它先验假设。首先在频域盲源分离方法中利用语谱分裂技术得到  $M \times N$  个分离信号, 可视为多输入多输出(MIMO)混合系统分裂成  $N$  个单输入多输出(SIMO)系统, 泄露进该信号的其它信号分量可以视为干扰信号, 然后分别用频域多信道信号增强方法重构出各个语音信号, 以消除不同信源间的空间干扰和环境噪声。仿真结果表明基于多信道信号增强的后处理算法可以弥补由于排列模糊导致算法分离性能下降使得语音信号质量受损的情况。

第七章对全文的研究工作做了概括和总结, 对未来工作进行了展望, 给出了几个有待进一步研究的问题。

## 参考文献

- [1] Lee T W. Independent component analysis: theory and application. Dordrecht, pp. Kluwer Academic Publishers, 1998.
- [2] Haykin S. Unsupervised adaptive filtering. Vol 1: Blind source separation. New York, pp. Wiley, 2000.
- [3] Hyvarinen A, Karhunen J, Oja E. Independent Component Analysis. New York, pp. Wiley, 2001.
- [4] Cichocki A, Amari S. Adaptive Blind Signal and Image Processing. New York, pp. Wiley, 2003.
- [5] Stone J V. Independent Component Analysis, pp. A Tutorial Introduction. MIT Press, 2004.
- [6] Makino S, Lee T W, Sawada H. Blind speech separation, Springer, 2007.
- [7] 张贤达, 保铮. 通信信号处理. 国防工业出版社. 2002年.
- [8] 杨福生, 洪波. 独立分量分析的原理与应用. 清华大学出版社. 2006年.
- [9] 史习智. 盲信号处理-理论与实践. 上海交通大学出版社, 2008年.
- [10] Herault J, Jutten C. Space or time adaptive signal processing by neural network models. Neural Network for Computing: AIP Conference Proceeding, 151, New York, American Institute of Physics, 1986.
- [11] Jutten C, Herault J. Blind separation of sources, Part I: An adaptive algorithm based on neuromimetic architecture. Signal Processing, 1991, 24(1), pp. 1-10.
- [12] Comon P, Jutten C, Herault J. Blind separation of sources, Part II: Problems statement. Signal Processing, 1991, 24(1), pp. 11-20.
- [13] Tong L, Liu R, Soon V C, Huang J. F. Indeterminacy and identifiability of blind identification. IEEE Trans Signal Processing, 1991, 38(5), pp. 499-509.
- [14] Cardoso J F, Souloumiac A. Blind beamforming for non-Gaussian signals. IEE-Proceedings-F, 1993, 140 (6), pp. 362-370.
- [15] Comon P. Independent component analysis: a new concept? Signal Processing, 1994, 36 (3), pp. 287-314.
- [16] Bell A J, Sejnowski T J. An information maximization approach to blind separation and blind deconvolution. Neural Computation, 1995, 7(6), pp. 1129-1159.
- [17] Amari S, Cichocki A, Yang H. A new learning algorithm for blind signal separation. Advances in Neural Information Processing Systems, MIT Press: Cambridge, MA, 1996, 8, pp.757-763.

- 
- [18] Amari S. Natural gradient works efficiently in learning. *Neural Computation*, 1998, 10(1), pp. 251-276.
  - [19] Cardoso J F, Laheld B. Equivariant adaptive source separation. *IEEE Trans on Signal Processing*, 1996, 44(12), pp. 3017-3030.
  - [20] Yang H H, Amari S. Adaptive on-line learning algorithms for blind separation: maximum entropy and minimum mutual information. *Neural Computation*, 1997, 9(7), pp. 1457-1482.
  - [21] Cao X R, Liu R W. General approach to blind source separation. *IEEE Trans on Signal Processing*, 1996, 44(3), pp. 562-571.
  - [22] Amari S, Cardoso J F. Blind source separation: semi-parametric statistical approach. *IEEE Trans on Signal Processing*, 1997, 45(11), pp. 2692-2700.
  - [23] Cardoso J F. Blind signal separation: statistical principles. *Proceeding of IEEE*, 1998, 86(10), pp. 2009-2025.
  - [24] Cardoso J F. On the stability of source separation algorithms. *Journal of VLSI Signal Processing*. 2000, 26(1/2), pp. 7-23.
  - [25] Karhunen J, Joutsensalo J. Representation and separation of signals using nonlinear PCA type learning. *Neural Networks*, 1994, 7(1), pp. 113-127.
  - [26] Pajunen P, Karhunen J. Least-squares methods for blind source separation based on nonlinear PCA. *International Journal of Neural Systems*, 1998, 8(12), pp. 601-612.
  - [27] Karhunen J, Pajunen P, Oja E. The nonlinear PCA criterion in blind source separation: Relations with other approaches. *Neurocomputing*, 1998, 22, pp. 5-20.
  - [28] Yang B. Projection approximation subspace tracking. *IEEE Trans on Signal Processing*, 1995, 43(1), pp. 95-107.
  - [29] Zhu X L, Zhang X D. Adaptive RLS algorithm for blind source separation using a natural gradient. *IEEE Signal Processing Letter*, 2002, 9(12), pp. 432-435
  - [30] Zhu X L, Zhang X D, Ding Z Z, Jia Y. Adaptive nonlinear PCA algorithms for blind source separation without prewhitening. *IEEE Transactions on circuits and systems-I*, 2006, 53(3), pp. 745-753.
  - [31] Chang, Ding Z, Yau S, Chan F. A matrix-pencil approach to blind separation of nonwhite signals in white noise. in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Seattle, WA, May 1998, 4, pp. 2485-2488.
  - [32] Tome A M. Blind source separation using a matrix pencil. in *Proc. Int. Joint Conf. Neural Netw.*, Como, Italy, Jul. 2000, 3, pp. 339-342.
  - [33] Parra L, Sajda P. Blind Source Separation via Generalized Eigenvalue
-

- Decomposition. *Journal of Machine Learning Research*, 2003, 4(7-8), pp.1261-1269.
- [34] Belouchrani A, Abed-Meraim K, Cardoso J F, Moulines E. A blind source separation technique using second-order statistics. *IEEE Trans on Signal Processing*, 1997, 45(2), pp. 434-444.
- [35] Van der Veen A. J. Joint diagonalization via subspace fitting techniques. *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, 2001, pp. 2773-2776.
- [36] Yeredor A. Non-orthogonal joint diagonalization in the least-squares sense with application in blind source separation. *IEEE Trans on Signal Processing*, 2002, 50(7), pp. 1545-1553.
- [37] Ziehe A, Laskov P, Nolte G, Müller K R. A fast algorithm for joint diagonalization with non-orthogonal transformations and its application to blind source separation. *Journal of Machine Learning Research*, 2004, 5, pp. 777-800.
- [38] Li X L, Zhang X D. Nonorthogonal Joint Diagonalization Free of Degenerate Solution. *IEEE Trans on Signal Processing*, 2007, 55(5), pp. 1803-1814.
- [39] Delfosse N, Loubalon P. Adaptive blind separation of independent sources: A deflation approach. *Signal Processing*, 1995, 45(7), pp. 59-83.
- [40] Hyvarinen A, Oja E. A fast fixed-point algorithm for independent component analysis. *Neural Computation*, 1997, 9(7), pp. 1483-1492.
- [41] Hyvarinen A. Fast and robust fixed-point algorithms for independent component analysis. *IEEE Trans on Neural Networks*, 1999, 10(3), pp. 626-634.
- [42] Bingham E, Hyvarinen A. A fast fixed-point algorithm for independent component analysis of complex valued signals. *International Journal of Neural Systems*, 2000, 10(1), pp. 1-8.
- [43] Douglas S C. Fixed-point algorithms for the blind separation of arbitrary complex-valued non-Gaussians signal mixtures. *EURASIP Journal on Advances in Signal Processing*, 2007, 1(3), pp. 1-15.
- [44] Koldovsky Z, Tichavsky P, Oja E. Efficient variant of algorithm FastICA for independent component analysis attaining the Cramer-Rao lower bound. *IEEE Trans on Neural Networks*, 2006, 17(5), pp. 1265-1277.
- [45] Tichavsky P, Koldovsky Z, Oja E. Performance analysis of the FastICA algorithm and Cramer-Rao Bounds for linear independent component analysis. *IEEE Trans on Signal Processing*, 2006, 54(4), pp. 1189-1203.
- [46] Li Y Q, Wang J. Sequential blind extraction of instantaneously mixed sources. *IEEE Trans on Signal Processing*, 2002, 50(5), pp. 997-1006.

- 
- [47] Li X L, Zhang X D. Sequential Blind Extraction Adopting Second-Order Statistics. *IEEE Signal Processing Letters*, 2006, 14(1), pp. 58-61.
  - [48] Souloumiac A. Blind source detection and separation using second order non-stationarity. *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, 1995, pp. 1912-1915.
  - [49] Matsuoka K, Ohya M, Kawamoto M. A neural net for blind separation of nonstationary signals. *Neural Network*. 1995 8(3), pp. 411-419.
  - [50] Pham D T, Cardoso J F. Blind separation of instantaneous mixtures of nonstationary sources. *IEEE Trans on Signal Processing*, 49(9), pp. 1837-1848.
  - [51] Choi S, Cichocki A. Blind separation of nonstationary sources in noisy mixtures. *Electronics Letters*. 2000, 36(4), pp. 848-849.
  - [52] Belouchrani A, Amin M G. Blind source separation based on time-frequency signal representation. *IEEE Trans on Signal Processing*, 1998, 46(11), pp. 2888-2898.
  - [53] Fadaili E M, Moreau T N, and Moreau E. Non-orthogonal joint diagonalization/zero-diagonalization for source separation based on time-frequency distributions. *IEEE Trans on Signal Processing*, 2007, 55(5), pp. 1673-1687.
  - [54] Chabriel G, Barrere J, Moreau T N, Moreau E. Algebraic joint zero-diagonalization and blind sources separation. *IEEE Trans on Signal Processing*, 2008, 56(3), pp. 980-989.
  - [55] Everson R M, Roberts S. J. Bind source separation for non-stationary mixing *Journal of VLSI Signal Processing*. 2000, 26(1/2), pp. 5-23..
  - [56] Ahmed A. Signal separation. PhD. Thesis, Corpus Christi College, Cambridge University 2000.
  - [57] Parga N, Nadal J P. Blind source separation with time-dependent mixtures. *Signal Processing*, 2000, 80(10), pp. 2187-2194.
  - [58] Lv Q, Zhang X D, Jia Y. Kalman filtering algorithm for blind source separation. *IEEE International Conference Acoustics, Speech, Signal Processing*, 2005, 5, pp. 257-260.
  - [59] Pokharel P P, Ozertem U, Erdogmus D, Principe J C. Recursive complex BSS via generalized eigendecomposition and application in image rejection for BPSK. *Signal Processing*, 2008, 88(4), pp. 1368-1381.
  - [60] Douglas S C, Cichocki A. Adaptive step size techniques for decorrelation and blind source separation. *Signals, Systems & Computers Conference*, 1998, pp.1191-1195.
-

- 
- [61] Thomas P, Allen G, August N. Step-size control in blind source separation. International Workshop on Independent Component Analysis and Blind Source Separation, Helsinki, Finland, 2000, 509-514.
- [62] Chambers A, Jafari G, McLaughlin S. Variable step-size EASI algorithm for sequential blind source separation. *Electronics Letters*, 2004, 40(6), pp.393-394.
- [63] Yuan L X, Wang W W, Chambers A. Variable step-size sign natural gradient algorithm for sequential blind source separation. *IEEE Signal Processing Letters*, 2005, 12(8), pp.589-592.
- [64] Slock, D T M, Kailath T. Numerically stable fast transversal filters for recursive least squares adaptive filtering. *IEEE Trans on Signal Processing*, 1991, 39(1), pp.92-114.
- [65] Papaodysseus C.. A robust, parallelizable  $O(m)$ , a posteriori recursive least squares algorithm for efficient adaptive filtering. *IEEE Trans. on Signal Processing*, 1999, 47(9), pp. 2552-2557.
- [66] Amari S. Cichocki A. Adaptive Blind signal processing-neural network approaches. *Proceeding of IEEE*. 1998, 86(10), pp. 2026-2048.
- [67] Li Y Q, Wang J. Sequential blind extraction of instantaneously mixed sources. *IEEE Trans on Signal Processing*, 2002, 50(5), pp.997 – 1006.
- [68] Cichocki A, Karhunen J, Kasprzak W, Vigario R. Neural networks for blind separation with unknown number of sources. *Neurocomputing* 1999, 24, pp. 55-93.
- [69] 冶继民. 信源数目未知与变化时的盲信号分离方法研究. 西安电子科技大学, 博士学位论文, 2005.
- [70] Ye J M, Zhu X L, Zhang X D. Adaptive blind separation with an unknown number of sources. *Neural Computation*, 2004, 16(8), pp. 1641-1660.
- [71] Lv Q, Zhang X D. A Unified Method for Blind Separation of Sparse Sources With Unknown Source Number. *IEEE Signal Processing Letters*, 2006, 13(1), pp. 49-51.
- [72] Li Y Q, Amari S, Cichocki A, Ho D W C, Xie S L. Underdetermined Blind Source Separation Based on Sparse Representation. *IEEE Trans on Signal Processing*, 2006, 54(2), pp. 423-437.
- [73] Chen Z, Cichocki A, Rutkowski T M. Constrained Non-Negative Matrix Factorization Method for EEG Analysis in Early Detection of Alzheimer's Disease. *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2006, pp. 14-19.
-

- [74] Yang H H, Amari S, Cichocki A. Information-theoretic approach to for blind separation of sources from nonlinear mixture. *Signal Processing*, 1998, 64(3), pp. 291-300.
- [75] 凌燮亭. 近场宽带信号源的盲分离. *电子学报*. 1996, 24 (7), pp. 87 - 92.
- [76] 张昕, 胡波, 凌燮亭. 盲信号分离在数字无线通信中的一种应用. *通信学报*, 2000, 21(2), pp.73-77.
- [77] He Z Y, Liu J, Yang L X. Blind separation of mages using edge worth expansion based ICA algorithm. *Chinese Journal of Electronics*, 1999, 8(3), pp. 278-282.
- [78] He Z Y, Yang L X, Liu J. Blind source separation using cluster based multivariate density estimation algorithm. *IEEE Trans on Signal Processing*, 2000, 48 (2), pp. 575-579.
- [79] 刘琚, 何振亚. 盲源分离和盲反卷积. *电子学报* 2002, 30(4), pp. 570-576.
- [80] 刘琚, 聂开宝, 何振亚. 非线性混叠信号的可分离性及分离方法研究. *电子与信息学报*, 2003, 25(1), pp.54-61.
- [81] 刘琚, 孙建德. 基于ICA的数字水印方法. *电子学报*. 2004, 32(4), pp. 657-660.
- [82] 虞晓, 胡光锐. 基于统计估计的盲信号分离算法. *上海交通大学学报*, 1999, 33(5), pp. 566-569.
- [83] 倪晋平, 马远良, 孙超, 童立. 用独立成份分析算法实现水声信号盲分离. *声学学报*, 2002, 27(4), pp. 321-326.
- [84] 倪晋平, 马远良, 杨坤德. 水声信号盲分离的实数算法与复数算法. *西北工业大学学报*, 2002, 20(1), pp. 44-48.
- [85] 倪晋平, 马远良, 鄢社锋, 杨坤德. 水声信号盲分离算法性能评价准则. *探测与控制学报*, 2002, 24(2), pp. 40-44.
- [86] 倪晋平, 马远良, 鄢社锋. 基于高阶累积量的复数混合矩阵盲估计算法. *电子与信息学报*. 2002, 24(11), pp. 1506-1511.
- [87] 谢胜利, 王杰. 一个基于盲信号分离的多路回波消除结构. *电子学报*. 2004, 32(7), pp. 1124-1126.
- [88] 章晋龙, 何昭水, 谢胜利, 刘海林. 多个源信号混叠的盲分离几何算法. *计算机学报*, 2005, 28(9), pp. 1575-1580.
- [89] 何昭水, 谢胜利, 傅予力. 信号的稀疏性分析. *自然科学进展*. 2006, 16(9), pp. 1167-1173.
- [90] 张安清, 邱天爽, 章新华. 分数低阶矩的信号盲分离方法. *通信学报*, 2006, 27(3), pp.32-36.
- [91] Yin F L, Mei T M, Wang J. Blind source separation based on decorrelation and

- nonstationarity. *IEEE Trans on Circuits and Systems-I: Regular Papers*, 2007, 54(5), pp. 1150-1158.
- [92] Chen T P, Lin Q. Dynamic behavior of the whitening process Signal. *IEEE Processing Letters*, 1998, 5(1), pp.25-26.
- [93] Tan Y, Wang J, Zurada J M. Nonlinear blind source separation using a radial basis function network. *IEEE Trans on Neural Networks*, 2001, 12(1), pp.124-134.
- [94] Tan Y, Wang J. Nonlinear blind source separation using higher order statistics and a genetic algorithm. *IEEE Transactions on Evolutionary Computation*, 2001, 5(6), pp.600 – 612.
- [95] Zhang L Q, Cichocki A, Amari S. Self-adaptive blind source separation based on activation functions adaptation. *IEEE Transactions on Neural Networks*, 2004, 15(2), pp.233 – 244.
- [96] 张贤达, 保铮. 盲信号分离. *电子学报*, 2001, 29(12A), pp. 1766-1771.
- [97] 朱孝龙, 保铮, 张贤达. 基于分阶段学习的盲信号分离. *中国科学E辑*, 2002, 32(5), pp. 693-703.
- [98] 朱孝龙, 张贤达, 冶继民. 基于自然梯度的递归最小二乘盲信号分离. *中国科学, E辑*, 2003, 33(8), pp.741-748.
- [99] Lou S T Zhang X D. Fuzzy-based learning rate determination for blind source separation. *IEEE Trans on Fuzzy Systems*, 2003, 11(3), pp.375-383.
- [100] Feng D Z, Zhang X D, Bao Z. An efficient multistage decomposition approach for independent components. *Signal Processing*, 2003, 83(1), pp. 181-197.
- [101] Feng D Z, Zheng, W X, Cichocki A. Matrix-Group Algorithm via Improved Whitening Process for Extracting Statistically Independent Sources From Array Signals. *IEEE Trans on Signal Processing*, 2007, 55(3), pp. 962-977.
- [102] Weinstein E, Feder M, Oppenheim A V. Multi-channel signal separation by decorrelation. *IEEE Trans on Speech and Audio Processing*, 1993, 1(4), pp. 405-413.
- [103] Gerven V S, Compernelle V D. Signal separation by symmetric adaptive decorrelation stability, convergence and uniqueness. *IEEE Trans on Signal Processing*, 1995, 43(7), pp. 1602-1612
- [104] Kawamoto M, Matsucka K, Ohnishi N. A method of blind separation for convolved non-stationary signals. *Neurocomputing*, 1998, 22(1-3), pp. 157-171.
- [105] Bousbia S H, Belouchrani A, Abed M K. Blind separation of convolutive mixtures using joint block diagonalization. *IEEE Proceedings of International*

- Symposium on Signal Processing and its Applications, 2001, 1, pp. 13-16.
- [106] Yellin D, Weinstein E. Multichannel signal separation: methods and analysis. IEEE Trans on Signal Processing, 1996, 44 (1), pp. 106-118.
- [107] Sun X, Douglas S C. Multichannel blind deconvolution of arbitrary signals: adaptive algorithms and stability analysis. Proceedings of 34th Asilomar Conference on Signals, System, Computer, 2000, 2, pp. 1412-1416.
- [108] Choi S, Hong H, Glotin H, Berthommier F. Multichannel signal separation for cocktail party speech recognition: a dynamic recurrent network. Neurocomputing, 2002, 49(1-4), pp. 299-314.
- [109] Douglas S C, Sawada H, Makino S. Natural gradient multichannel blind deconvolution and speech separation using causal FIR filters. IEEE Trans on Speech and Audio Processing, 2005, 13(1), pp. 92-104.
- [110] Smaragdis P. Blind separation of convolved mixtures in the frequency domain. Neurocomputing, 1998, 22, pp. 21-34.
- [111] Araki S, Mukai R, Makino S. The fundamental limitation of frequency domain blind source separation for convolutive mixtures of speech. IEEE Trans on Speech and Audio Processing, 2003, 11(2), pp. 109-116.
- [112] Parra L, Spence C. Convolutive blind separation of nonstationary sources. IEEE Trans. Speech Audio Processing, 2000, 8(3), pp. 320-327.
- [113] Buchner H, Aichner R, Kellermann W. A generalization of blind source separation algorithms for convolutive mixtures based on second-order statistics. IEEE Trans. Speech and Audio Processing. 2005, 13(1), pp. 120-134.
- [114] Kurita S, Saruwatari H. Evaluation of blind signal separation method using directivity pattern under reverberant conditions. IEEE International Conference Acoustics, Speech, Signal Processing, 2000, pp. 3140-3143.
- [115] Ikram M Z, Morgan D R. A beamforming approach to permutation alignment for multichannel frequency-domain blind speech separation. IEEE International Conference Acoustics, Speech, Signal Processing, 2002, pp. 881-884.
- [116] Saruwatari H, Kawamura T, Nishikawa T, Lee A, Shikano K. Blind source separation based on a fast-convergence algorithm combining ICA and beamforming. IEEE Trans on Speech and Audio Processing, 2006, 14(2), pp. 666-678.
- [117] Murata N, Ikeda S, Ziehe A. An approach to blind source separation based on temporal structure of speech signals. Neurocomputing 2001, 41, pp. 1-24.
- [118] Asano F, Ikeda S, Ogawa M, Asoh H, Kitawaki N. A combined approach of array

- processing and independent component analysis for blind separation of acoustic signals. *IEEE Trans on Speech and Audio Processing*, 2003, 11(3), pp. 204–215.
- [119] Sawada H, Mukai R, Araki S, Makino S. A robust and precise method for solving the permutation problem of frequency-domain blind source separation. *IEEE Trans on Speech and Audio Processing*, 2004, 12(5), pp. 530–538.
- [120] Rahbar K, Reilly J P. A frequency domain method for blind source separation of convolutive audio mixtures. *IEEE Trans on Speech and Audio Processing*, 2005, 13(5), pp. 832–844.
- [121] Serviere C, Pham D T. A novel method for permutation correction in frequency-domain in blind separation of speech mixtures. in *Proc. Int. Conf. Independent Component Analysis and Blind Signal Separation*, 2004, pp. 807–815.
- [122] Wang W, Chambers J A, Sanei S. A novel hybrid approach to the permutation problem of frequency domain blind source separation. in *Proc. Int. Conf. Independent Component Analysis and Blind Signal Separation*, 2004, pp. 532–539.
- [123] Mukai R, Suwudu H, Araki S, Makino S. Robust real-time blind source separation for moving speakers in a room. *IEEE International Conference Acoustics, Speech, Signal Processing*, 2003, V, pp. 469–472.
- [124] Buchner H, Aichner R, Kellermann W. Real-time convolutive blind source separation based on a broadband approach. in *Proc. Int. Conf. Independent Component Analysis and Blind Signal Separation*, 2004, pp. 840–848.
- [125] Ding S X, Huang J, Wei D M, Cichocki A. A near real-time approach for convolutive blind source separation. *IEEE Trans on Circuits and Systems I: Regular Papers*, 2006, 53(1), pp. 114–128.
- [126] Parra L, Alvino C. Geometric source separation: Merging convolutive source separation with geometric beamforming. *IEEE Trans on Speech and Audio Processing*, 2002, 10(6), pp. 352–362.
- [127] Knaak M, Araki S, Makino S. Geometrically Constrained Independent Component Analysis. *IEEE Trans on Audio, Speech and Language Processing*, 2007, 15(2), pp. 715–726.
- [128] Zwicker E, Zwicker U T. Audio engineering and psychoacoustic: matching signals to the final receiver, the human auditory system. *Journal of the Audio Engineering Society*, 1991, 39(3), pp. 115–126.
- [129] Pan Q F, Aboulnasr, T. A new perceptual convolutive blind source separation

- algorithm for speech separation. Proceedings of International Conference on Signal Processing, 2004, pp. 323-326.
- [130] Low S Y, Nordholm S, Togneri R. Convolutional blind signal separation with post-processing. IEEE Trans on Speech and Audio Processing, 2004, 12(5), pp. 539-548.
- [131] Huang Y T, Benesty J, Chen J D. A blind channel identification-based two-stage approach to separation and dereverberation of speech signals in a reverberant environment. IEEE Trans on Speech and Audio Processing, 2005, 13(5), pp. 882-895.
- [132] Rivet B, Girin L, Jutten C. Mixing audiovisual speech processing and blind source separation for the extraction of speech signals from convolutional mixtures. IEEE Trans on Audio, Speech and Language Processing, 2007, 15(1), pp. 96-108.
- [133] Yilmaz O, Rickard S. Blind separation of speech mixtures via time-frequency masking. IEEE Trans on Signal Processing, 2004, 52(7), pp. 1830-1847.
- [134] Sawada H, Araki S, Mukai R, Makino S. Blind extraction of dominant target sources using ICA and time-frequency masking. IEEE Trans on Audio, Speech and Language Processing, 2006, 14(6), pp. 2165-2173.
- [135] Aissa A, Abed-Meraim K, Grenier Y. Blind separation of underdetermined convolutional mixtures using their time-frequency representation. IEEE Trans on Audio, Speech and Language Processing, 2007, 15(5), pp. 1540-1550.
- [136] 何培宇, 殷斌, Sommen P C W. 一种有效的语音盲信号分离简化混合模型. 电子学报, 2002, 30(10), pp. 1167-1173.
- [137] 何培宇, 周激流, 夏秀渝, 王永德, 赵刚. 一种基于二阶盲信号分离的多路声回波抑制模型. 电子学报, 2006, 34(11), pp. 2110-2114.
- [138] 张安清, 邱天爽, 章新华. 卷积混合信号频域盲分离技术研究. 大连理工大学学报. 2004, 44(5), pp. 723-728.
- [139] 苏野平, 何量, 杨荣震, 朱小刚. 一种改进的基于高阶累积量的语音盲分离算法. 电子学报, 2002, 30 (7), pp. 956-958.
- [140] Mei T M, Xi J T, Yin F L, Mertins A, Chicharo, J F. Blind source separation based on time-domain optimization of a frequency-domain independence criterion. IEEE Trans on Audio, Speech and Language Processing, 2006, 14(6), pp. 2075-2085.
- [141] He Z S, Xie S L, Ding S X, Cichocki A. Convolutional blind source separation in the frequency domain based on sparse representation. IEEE Trans on Audio, Speech and Language Processing, 2007, 15(5), pp. 1551-1563.

- [142] 彭春翌, 朱孝龙, 张贤达. 基于斜投影的卷积信道盲信号分离. 清华大学学报 (自然科学版), 2005, 45(4), pp. 517-520.

## 第二章 盲源分离的基础知识

### § 2.1 引言

盲源分离是在不知道源信号和传输信道参数的情况下,根据源信号的统计特性,仅由观测信号来分离或恢复出源信号的过程。盲源分离是信号处理领域中的一个基本问题,近二十年来已成为一个引人注目的研究热点。这是因为盲源分离问题有大量潜在的应用并且盲源分离本身也是有研究价值的问题。尽管现在已有不少盲源分离的文献,但依然有很多问题没有得到解决,因此盲源分离依然是一个活跃的研究领域,在多用户通信、语音处理、生物医学工程和图像处理等诸多领域有着非常重要的理论价值和研究意义。显然,当源信号与传输信道参数都未知时,盲源分离是一种很自然的选择。这里术语“盲”有两重含义:第一,源信号不能被观测;第二,源信号是如何混合的是未知的。文中把源信号的混合信号称为观测信号,源信号的混合模型称为混合系统。盲源分离的混合系统信号模型如图 2.1 所示:

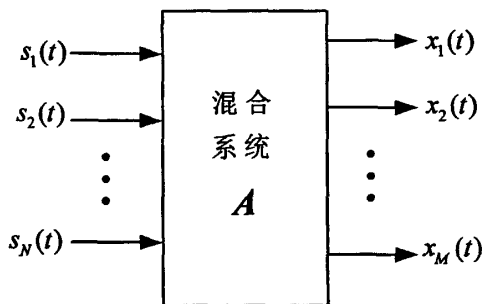


图 2.1 盲源分离的混合系统信号模型

图 2.1 中观测信号  $x_i(t)$  是源信号  $s_j(t), j=1,2,\dots,N$  的某一函数,这个把  $s_j(t)$  映射到  $x_i(t)$  的函数称为混合函数。观测者只能观测到传感器接收到的信号  $x_i(t), i=1,2,\dots,M$ , 而源信号和混合函数都是未知的;盲源分离的任务就是通过辨识混合函数来恢复源信号。由于上述系统的描述是一般性的,因此需要针对具体问题利用不同的标准对它进行分类,这些标准包括源信号的部分先验知识、混合函数的特性及它与源信号的关系。例如,根据观测信号和源信号个数的多少,盲源分离问题可以分为以下三种情况:

- ★ 适定盲源分离:观测信号和源信号的数目相等,也称为完备盲源分离;
- ★ 超定盲源分离:观测信号数目比源信号的数目多,也称为欠完备盲源分离;

★ 欠定盲源分离: 观测信号数目比源信号的数目少, 也称为过完备盲源分离。此外, 也可根据源信号的混合方式是否线性进行划分, 而混合过程可分为瞬时的或卷积的, 同样混合函数也可以是时不变的或时变的。大多情况下的研究对象是上述不同分类的组合形式: 如线性瞬时超定盲源分离问题、线性卷积混合超定盲源分离问题等。针对不同模型的盲源分离问题已经提出了大量不同类型的算法。例如, 根据工作方式的不同, 盲源分离算法可分为在线算法和离线算法, 前者是自适应算法, 后者是批处理算法。根据信号提取方式的不同, 盲源分离算法又可分为串行算法和并行算法, 前者逐个提取信号, 后者实现所有信号的同步提取。根据源信号的特性, 如非白、非高斯、非平稳性等, 盲源分离算法又可分为基于不同阶统计量的算法。与混合函数的分类类似, 大多数盲源分离算法也是上述不同分类的组合形式, 如在经典算法中, Cardoso 等人<sup>[1-2]</sup>的特征矩阵近似联合对角化算法(JADE 算法<sup>[1]</sup>和 SOBI 算法<sup>[2]</sup>)属于批处理并行算法, 其中 JADE 算法是基于四阶统计量, 而 SOBI 算法则基于二阶统计量; Hyvarinen 等人<sup>[3]</sup>的快速定点算法属于批处理串行算法; Amari 等人<sup>[4]</sup>的自然梯度算法和 Cardoso 等人<sup>[5]</sup>的等变化算法属于自适应并行算法。

近二十多年来, 盲源分离算法的研究和性能分析非常活跃, 所研究的信号模型及涉及到的基本理论各有不同, 本章主要针对线性瞬时混合模型和线性卷积混合模型介绍其所涉及到的基本理论。

## § 2.2 瞬时混合盲源分离

### 2.2.1 信号模型与假设

如图 2.2 所示, 假设  $N$  个统计独立的源信号  $s_j(t), j=1, 2, \dots, N$  经过线性瞬时混合后得到  $M$  个观测信号(或传感器信号)  $x_i(t), i=1, 2, \dots, M$ , 下面的方程对于线性时不变瞬时混合函数成立:

$$\begin{aligned} x_1(t) &= \sum_{j=1}^N a_{1j} s_j(t) \\ x_2(t) &= \sum_{j=1}^N a_{2j} s_j(t) \\ &\vdots \\ x_M(t) &= \sum_{j=1}^N a_{Mj} s_j(t) \end{aligned} \quad (2.1)$$

这里  $a_{ij} \in \mathbb{R}$ ,  $i=1, 2, \dots, M$ ,  $j=1, 2, \dots, N$  是混合参数。

式(2.1)能用向量和矩阵的形式来表示:

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) \quad (2.2)$$

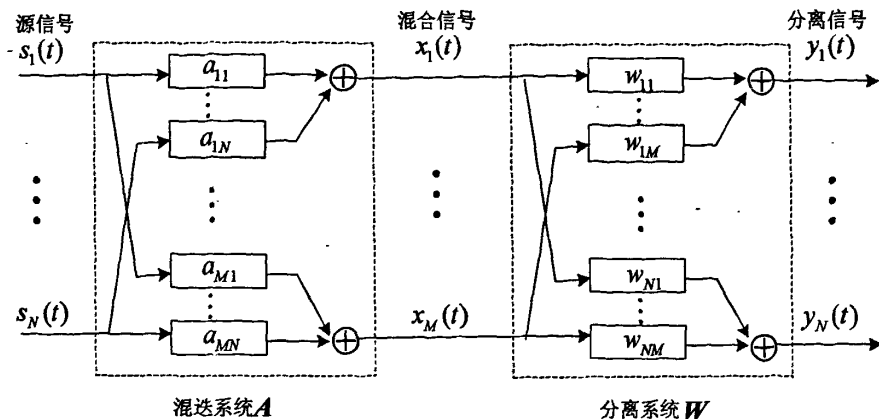


图2.2 盲源分离的线性瞬时混迭与分离信号模型

这里

$\mathbf{x}(t) = [x_1(t), x_2(t), \dots, x_M(t)]^T$  为  $M$  维观测信号向量；

$\mathbf{s}(t) = [s_1(t), s_2(t), \dots, s_N(t)]^T$  为  $N$  维源信号向量；

$\mathbf{A} = [a_{ij}]_{M \times N}$  是元素为  $a_{ij}$  的  $M \times N$  维矩阵，在盲源分离中称为混合矩阵。

$[\bullet]^T$  表示矩阵或向量的转置。

由于传输信道和传感器阵列包含加性噪声，因此在实际应用中必须考虑加性噪声，这样式(2.2)中的混合模型就变为

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t) \quad (2.3)$$

这里  $\mathbf{n}(t) = [n_1(t), n_2(t), \dots, n_M(t)]^T$  是加性噪声向量。

事实上，式(2.3)可写作

$$\mathbf{x}(t) = [\mathbf{A}, \mathbf{I}] \begin{bmatrix} \mathbf{s}(t) \\ \mathbf{n}(t) \end{bmatrix} \triangleq \tilde{\mathbf{A}}\tilde{\mathbf{s}}(t)$$

其中  $\mathbf{I}$  表示  $M$  维的单位矩阵。因此式(2.3)也属于式(2.2)的范畴。今后在线性瞬时混迭的情况下，如无特别说明，我们只讨论式(2.2)描述的不包含噪声的模型。

### 2.2.2 可分离性与分离原则

盲源分离问题的可分离性包括信号是否实际可分离和在什么条件下是可分离的。分离原则为在实现盲分离时所采用的原则。为了使用数学术语来陈述上面的问题，我们用下面的形式来描述盲源分离问题：

$$\mathbf{x}(t) = f(\mathbf{s}(t)) \quad (2.4)$$

和

$$\mathbf{y}(t) = g(\mathbf{x}(t)) \quad (2.5)$$

可分离性就是存在一个函数  $g(\cdot)$ ，使得每个  $y_i$  包含不同  $s_i$  的分量，即  $y_i = z_j(s_j)$ ，这里  $z_j$  是任意的滤波器算子。分离原则就是怎样来估计函数  $g(\cdot)$ 。我们这里的讨论限于线性时不变函数，同时假设源信号是独立的。

需要特别指出的是源信号的独立性假定是盲源分离问题中一个最基本的假定，从观测数据中分离独立的源信号是基础的而又具有挑战性的问题，近年来已受到广泛的关注。许多研究人员已经用不同的方法，例如神经网络、机器学习、高阶统计量、最小互信息、波束形成和自适应噪声对消等，不同程度上地解决了这个问题。尽管他们所用的方法不同，但是源信号的统计独立性假设是绝大部分解决方案中所共有的、最重要的假设。统计独立性意味着知道其中一个源信号，不能估计或描述任一别的源信号。独立性假定或由它导出的假定（如不相关）被上述大部分的解决方案用作分离标准。换句话说，若输出信号满足独立性标准，那么就可以认为分离是成功的。Yellin 等人<sup>[6]</sup>已证明若源信号是统计独立的，那么观测信号能分离的一个充分必要条件是分离系统的输出也是统计独立的。本文中给出的新算法也属于这一类。

源信号  $s_j$  是相互统计独立的当且仅当源信号的联合概率密度  $p_s(\mathbf{s})$  等于边缘密度  $p_{s_j}(s_j)$  的乘积，即

$$p_s(\mathbf{s}) = \prod_{j=1}^N p_{s_j}(s_j)。$$

Comon 等人<sup>[7]</sup>阐明了若输出是源信号的线性函数且分离后的信号是两两独立的，那么它们也是相互独立的。

Tong 等人<sup>[8]</sup>给出在两种情况下，也就是不相关源和独立源的可辨识的条件。他们在分析中假设：

- (1)  $f(\cdot)$  是一个线性时不变瞬时混合矩阵，即信道是线性的且无记忆能力；
- (2)  $f(\cdot)$  是列满秩的；
- (3) 源信号  $s_i$  是均值为零，协方差矩阵非奇异的广义平稳过程。

两种可分离情况分别为：

1. 不相关源  $s_j$  能被辨识，若满足下面的条件：

(i)  $s_j, j=1, \dots, N$  是不相关的；

(ii) 存在一个  $\tau > 0$  使得

$$\frac{E\{s_i(t)s_i(t-\tau)\}}{E\{s_i^2\}} \neq \frac{E\{s_j(t)s_j(t-\tau)\}}{E\{s_j^2\}}, \quad i \neq j。$$

2. 独立源  $s_j$  能被辨识, 若

(i)  $s_j, j=1, \dots, N$  是相互独立的;

(ii)  $\frac{E\{s_i(t)^4\}}{E\{s_i^2\}} \neq \frac{E\{s_j(t)^4\}}{E\{s_j^2\}}, i \neq j$ .

Tong 等人<sup>[9]</sup>给出了波形保持 (waveform preserving) 的概念, 他们认为分离后的输出的非唯一性的集合必须属于一个唯一的波形类别, 他们也证明了源信号也必须属于同一类。这就定义了解的可行族。在源信号的概率密度和传输信道的特性都未知的情况下不能唯一地辨识源信号。

在我们的假设下,  $f(\cdot)$  对应于混合矩阵  $A$ , 上面的分析表明为了尽可能好的恢复出源信号, 矩阵  $A$  必须是可逆的。这意味着在给定  $A$  依赖于源的物理的位置的情况下, 源信号必须是空间可分解的。进而, 源信号的数目需小于或等于传感器的数目 ( $N \leq M$ ), 从而使得系统不是欠定的。

在理想情况下, 当  $M = N$  时,

$$s(t) = A^{-1}x(t) = Wx(t) \quad (2.6)$$

这里  $W = A^{-1}$  是解混合矩阵, 或称为分离矩阵。

当  $M > N$  时,

$$s(t) = A^{\#}x(t) = Wx(t) \quad (2.7)$$

这里  $W = A^{\#}$  是解混合矩阵, 其中  $A^{\#}$  表示  $A$  的 Morre-Penrose 广义逆(伪逆)矩阵。

若混合系统是可逆的或存在伪逆, 这对应于  $M \geq N$  (传感器的数目大于或等于源信号的数目) 的情况, 在方阵的情况下混合矩阵是非奇异的。可以通过直接估计混合矩阵的逆或其伪逆处理这个问题。矩阵非奇异的要求在物理现实上可以解释为传感器信号是空间可分解的, 这也是盲源分离问题中最通常的要求。如图 2.2 所示, 对应于瞬时混合(见式(2.2))的解混合(分离)系统形式如下:

$$y(t) = Wx(t) = WAs(t) \quad (2.8)$$

这里  $W \in \mathbb{R}^{N \times M}$ ,  $y(t) \in \mathbb{R}^{N \times 1}$  是源信号的估计向量, 或称为源信号的拷贝。

然而, 在盲源分离问题中对混合矩阵  $A$  的完全辨识是不可能的, 即混合矩阵的辨识是一个病态问题。在盲源分离问题中有两个内在的模糊或不确定性:

- (i) 尽管可以将各信源正确分开, 但并不能确定其排列顺序, 这相当于同时交换源信号和混合矩阵与之对应的列的位置后, 所得到的观测向量是相同的;
- (ii) 一个信号和与之对应的混合矩阵  $A$  的列之间互换一固定比例因子, 对观测值不会产生影响, 即:

$$x(t) = As(t) = \sum_{i=1}^N \frac{a_i}{\alpha_i} \alpha_i s_i(t) \quad (2.9)$$

其中  $\alpha_i$  是任意复因子,  $a_i$  是  $A$  的列向量。

正是由于上述两个不确定性的存在,使得解混合系统和混合系统的乘积有以下形式:

$$WA = AP \quad (2.10)$$

这里  $A$  是对角矩阵,  $P$  是置换矩阵。

混合矩阵  $A$  的非完全辨识称为  $A$  的不确定性。既然  $A$  有不确定性,我们可以假定源信号有单位方差,即把源信号的复振幅的动态变化归并到混合矩阵  $A$  的相应的列中。由于源假定是不相关的,所以源信号方差的归一化结果为:

$$R_x(0) = I \quad (2.11)$$

从而有

$$R_x(0) = E\{x(t)x^H(t)\} = AA^H \quad (2.12)$$

归一化仍然使得混合矩阵  $A$  的各列的排列顺序和相位保持不确定性。为了解决这一不确定性问题,Cardoso 等人<sup>[1-2]</sup>将两个矩阵“本质相等”的概念引入盲源分离中。

**定义 2.1 (矩阵的本质相等)** 两个矩阵  $P$  和  $Q$  称为本质相等 (essentially equal), 并记作  $P \doteq Q$ , 若存在一个矩阵  $G$  使得  $P = QG$ , 其中  $G$  是一个广义交换矩阵 (即其每一行每一列只有一个非零元素), 并且其元素具有单位模。

由定义 2.1 知, 一个矩阵若其列向量交换位置与/或某列向量的所有元素都同乘以一个常数因子, 则所得矩阵是原矩阵的本质相等矩阵。

综上所述, 在盲源分离的讨论中, 应该将混合矩阵  $A$  的可辨识性可理解为确定一个与  $A$  本质相等的矩阵。显然, 这样的不确定性对源信号的分离不会有任何实质性的影响。此外, 对于一个不可逆的线性混合系统(它不存在左伪逆), 例如包含源的数目多于传感器的数目的情况, 这种情况必须联合估计源信号和混合系统。这要求关于源信号的更多的先验知识, 因此这种情况“盲”的程度要低一些。

### 2.2.3 分离系统与分离性能

若要设计一个分离算法, 就必须选择一个分离系统。其中最自然的选择就是利用混合系统的逆的结构, 另外一些通常选择的分离系统为前馈、后馈或反馈结构的神经网络。然后基于分离系统和它们相应的分离准则来估计参数。分离算法的形式虽多, 但是推导思想却大同小异。首先需要选择恰当的对比函数(contrast function), 然后采用某种优化方法来搜索对比函数的极值点, 即构造过程可以统一表示为<sup>[10]</sup>

$$\text{BSS 算法} = \text{对比函数} + \text{优化方法} \quad (2.13)$$

因此, 盲源分离算法的性能取决于两个方面。其中, 对比函数决定算法的统计性能, 包括一致性、渐进方差和稳健性等; 优化方法决定算法的性能, 包括收敛速度、运算量、内存需求和数值稳定性等。

盲源分离问题的对比函数是分离系统的输出  $y$  的分布的一个实值函数,它满足当且仅当  $y = Cs$  是源信号的一个拷贝时取到极大值,这里  $C$  表示一个广义交换矩阵。对比函数可以分为两大类:一类是对所有分离状态都等价的对称对比函数<sup>[11]</sup>,另一类是非对称的对比函数<sup>[12]</sup>。常用的对比函数有信息论对比函数<sup>[4,13-14]</sup>,其中包括最大熵对比函数<sup>[13]</sup>、最小互信息对比函数<sup>[4]</sup>和极大似然对比函数<sup>[14]</sup>。此外还有恒模对比函数<sup>[15]</sup>,非线性主分量分析对比函数<sup>[16,17]</sup>,高阶累积量对比函数<sup>[11,12]</sup>和最大去高斯性对比函数<sup>[18]</sup>等。

根据优化方法的特点,通常使用迭代算法或递归算法用来求分离矩阵,如梯度下降算法,自然梯度算法,牛顿型算法及递归最小二乘算法等。确定分离系统的另外一个选择就是用批处理算法还是用自适应算法。在批处理算法中,在其中任意一段数据中可以认为参数是不变的,它由这个时间段内的数据确定。由于我们不希望分离参数在短的时间周期内有显著的变化,为减少计算时间,可以用上一批数据的估计值作为下一个估计的开始点。若输入是非平稳的,那么批处理算法的性能通常会明显的偏差。这个影响可以通过相互重叠的分段批处理来减小,但不可避免的会增加计算时间。而在自适应算法中,参数会在处理每一个新的数据时得到更新,但这引入了新的参数,即步长或遗忘因子,它影响算法的收敛速度及跟踪性能。通常来说步长或遗忘因子取常数时算法的性能受限,需要采用变步长或变遗忘因子方法的方法在加快算法收敛速度的同时提高稳态精确性<sup>[17,19]</sup>。若不能正确的设置学习参数,该算法就不能正确的收敛甚至根本就不再收敛。

## § 2.3 卷积混合盲源分离

### 2.3.1 信号模型与假设

设在混响环境中用  $M$  个麦克风  $x_m(t)$  ( $m=1, \dots, M$ ) 接收  $N$  个统计独立的语音信号  $s_n(t)$  ( $n=1, \dots, N$ ) ( $M \geq N$ ), 这在数学上可以描述成一个  $M \times N$  多输入多输出 (MIMO) 有限脉冲响应 (FIR) 系统 (如图 2.3 所示), 第  $m$  个麦克风的第  $t$  个快拍可表示为

$$x_m(t) = \sum_{n=1}^N \sum_{l=0}^{L-1} a_{mn}(l) s_n(t-l), \quad m=1, \dots, M \quad (2.14)$$

这里  $a_{mn}(l)$  表示从第  $n$  个源信号到第  $m$  个麦克风的脉冲响应。

卷积混合模型也可用矩阵形式表示为

$$\mathbf{x}(t) = \sum_{l=0}^{L-1} \mathbf{A}(l) \mathbf{s}(t-l) \quad (2.15)$$

这里  $\mathbf{x}(t) = [x_1(t), \dots, x_M(t)]$  是观测向量,  $\mathbf{s}(t) = [s_1(t), \dots, s_N(t)]$  源信号向量,  $\mathbf{A}(t)$  为

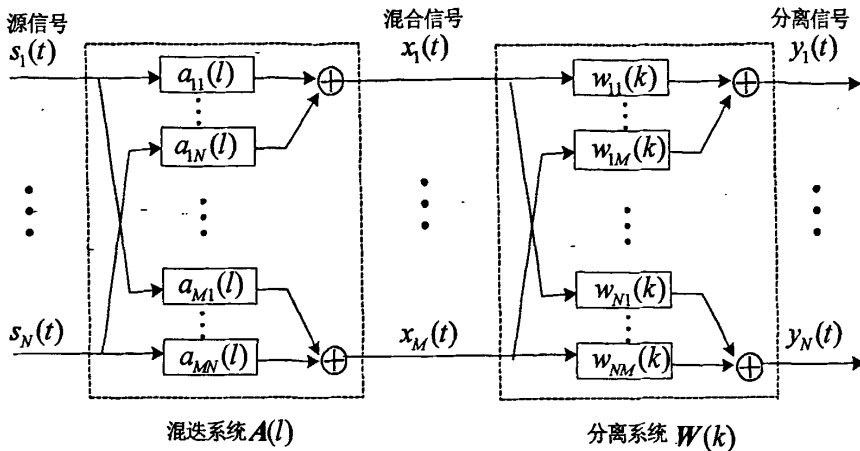


图2.3 盲源分离的卷积混迭与分离信号模型

滤波器混合矩阵，它的每个元素 $a_{mn}(l)$ 都是一个 $L$ 阶的滤波器。

此外，通过短时傅立叶变换（STFT）把卷积混合模型(2.15)中的观测信号 $\mathbf{x}(t)$ 变换到频域中，得到每个频率点上的瞬时混合

$$\mathbf{X}(f, t) = \mathbf{A}(f) \mathbf{S}(f, t) \quad (2.16)$$

这里  $\mathbf{X}(f, t) = [X_1(f, t), \dots, X_M(f, t)]^T$ ， $\mathbf{S}(f, t) = [S_1(f, t), \dots, S_N(f, t)]^T$ ，

$$\mathbf{A}(f) = \begin{bmatrix} A_{11}(f) & \cdots & A_{1N}(f) \\ \vdots & \ddots & \vdots \\ A_{M1}(f) & \cdots & A_{MN}(f) \end{bmatrix}, \text{ 其中 } S_n(f, t)、X_m(f, t) \text{ 和 } A_{mn}(f) \text{ 分别表示 } s_n(t)、$$

$x_m(t)$  和  $a_{mn}(l)$  的短时傅立叶变换。

### 2.3.2 可分离性与分离原则

业已证明<sup>[20,21]</sup>，瞬时混合和卷积混合模型的所有代数性质是等价的。瞬时盲源分离和多通道盲解卷积的关系和等价性参见文献[22]的表9.1。因此卷积混合盲源分离问题的可分离性与分离原则与瞬时盲源分离情形类似，这里不再进一步讨论。

需要注意的是上解介绍的卷积混合的盲源分离概念与多通道盲解卷积<sup>[20,23-25]</sup>有所不同，虽然两者的混合的数学模型相同，但是通常两者的前提假设条件和分离的结果却是不同的。在本节中介绍的卷积混合信号的盲分离，假设条件通常需要源信号之间是相互独立的或空间不相关的，其分离出的信号也不仅只有幅值的不确定性与排列的不确定性，而且是源信号经过一个传递函数后的滤波信号，即

$$y_m(t) = \sum_{l=0}^{\infty} d_n(l) s_n(t-l),$$

这里  $d_n(l)$  是一个未知的传递函数。而在多通道盲解卷积问题中, 其假设条件除了需要源信号之间是相互独立之外, 通常还需要每一个源信号都是一个独立同分布的过程。这类信号在现实中不易遇到, 但在通讯中经过调制后的信号大多是能满足这个条件的, 例如经过正交幅度调制后的QAM信号等, 同时, 盲解卷积后得到的信号与原始信号相比有着幅值的不确定性与延迟的不确定性, 即

$$y_m(t) = d_n s_n(t-\tau)$$

这里  $d_n$  是一个尺度因子,  $\tau$  表示延迟。显然, 多通道盲解卷积是卷积盲源分离的一种特殊情况, 它在系统输出端保持了源信号的时间结构<sup>[22]</sup>。

### 2.3.3 分离系统与分离性能

针对不同域的混合模型(2.15)和(2.16), 可以选择不同的分离系统实现源信号的盲分离。因此, 卷积混合盲源分离方法可以分为时域方法和频域方法。

时域方法是针对混合模型(2.15), 用一个长度为  $K$  的 FIR 滤波器  $w_{nm}(k)$  的集合成成分离系统来恢复源信号(如图 2.3 所示)

$$y_n(t) = \sum_{m=1}^M \sum_{k=0}^{K-1} w_{nm}(k) x_n(t-k), \quad n=1, \dots, N \quad (2.17)$$

首先提出在时域对卷积混合信号进行盲分离的方法<sup>[26]</sup>, 是利用独立分量分析的概念把瞬时混合时的标量混合矩阵扩展到卷积混合时的滤波器混合矩阵的方法。所有的时域分离方法都利用源信号的统计特性, 如源信号的非高斯性, 非平稳性和功率谱特性等<sup>[24-30]</sup>。从式(2.17)可见时域方法需要估计  $MNK$  个系数, 在滤波器长度较大时具有较高的计算复杂度, 此外收敛速度也是制约时域方法的一个重要问题。

频域方法(如图 2.4 所示)是针对混合模型(2.16), 由于在每个频率点观测信号都是复值信号的瞬时混合, 故可直接利用复值盲源分离算法求出分离矩阵  $W(f)$  以便重构源信号。每个频率点的分离信号可以表示为

$$Y(f, t) = W(f) X(f, t) \quad (2.18)$$

这里  $Y(f, t) = [Y_1(f, t), \dots, Y_N(f, t)]^T$ ,  $W(f) = \begin{bmatrix} W_{11}(f) & \dots & W_{1M}(f) \\ \vdots & \ddots & \vdots \\ W_{M1}(f) & \dots & W_{NM}(f) \end{bmatrix}$ , 其中

$Y_n(f, t)$  和  $W_{mn}(f)$  分别表示  $y_n(t)$  和  $w_{mn}(t)$  的短时傅立叶变换。

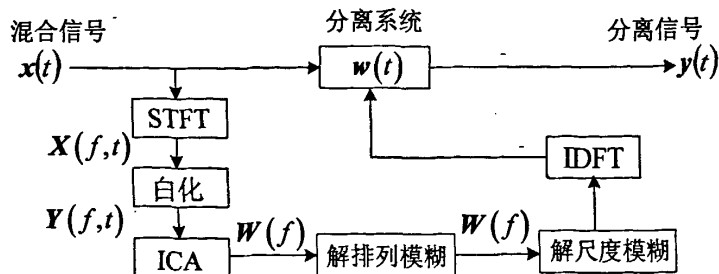


图 2.4 频域盲源分离流程

与卷积混合盲分离的时域算法相比，卷积混合分离的频域算法有着很大的优越性。在时域的盲分离算法中，当混合滤波器的阶数比较高时，分离滤波器每一阶的求解都要依靠滤波器中其余阶的求解，这会使得计算量变的非常大而且也会大大降低分离的效果。Chan<sup>[31]</sup>提出了分离矩阵对角常数化的算法，当混合滤波器的长度在 5 阶以下时，其算法可以很快地恢复出原始信号，当混合滤波器的长度变到 6 阶以上时，其分离速度就会非常地慢，而且分离的效果也会很差。相对而言，在频域中我们对每一个固定的频率用瞬时盲分离的方法进行分离，各个频率段的计算是相互独立的，因此，频域的卷积混合盲分离算法受到混合滤波器阶数的影响就较少，而且算法的计算量也较小。Smaragdis<sup>[32]</sup>用实验证明对于一个两输入两输出系统，当分离滤波器长度为 1024，采样频率为 44.1 kHz 时，用频域的计算法比用时域算法的计算量要减少了近 70 倍。

当然，频域算法也存在三个问题，主要是由于混合模型特性和瞬时盲源分离固有的两个不确定性所引起的。第一，在频域中求解卷积混合盲分离问题时，虽然从理论上讲可以利用已有复值瞬时混合盲源分离算法，但是在样本数目有限的情况下，在每个频率点面临的是小样本的盲源分离问题，一些在线算法在这种情况下分离性能较差或根本无法收敛<sup>[33]</sup>。第二，盲分离中排列的不确定性会导致在各个频段分离出的信号不能够正确的重构，从而可能使分离性能变差甚至失败<sup>[34]</sup>。第三，因为在盲源分离中有着幅值的不确定性，所以它会使得在各个频段分离的信号在频域上的幅值有偏差，从而会导致信号频谱的变形<sup>[35]</sup>。在语音信号的盲分离中就会使得语音信号的音质有所变化。针对上述三个问题，目前已有很多的算法<sup>[33-39]</sup>用以改进频率盲源分离方法的性能。

## § 2.4 本章小结

本章重点介绍了盲源分离问题的基本知识。介绍了盲源分离的概念,研究了盲源分离问题不同的混合模型,对盲源分离问题的可分离性、分离原则及分离系统的设计进行了分析论证。

### 参考文献

- [1] Cardoso J F, Souloumiac A. Blind beamforming for non-Gaussian signals. IEE-Proceedings-F, 1993, 140 (6), pp. 362-370.
- [2] Belouchrani A, Abed-Meraim K, Cardoso J F, Moulines E. A blind source separation technique using second-order statistics. IEEE Trans on Signal Processing, 1997, 45(2), pp. 434-444.
- [3] Hyvarinen A, Oja E. A fast fixed-point algorithm for independent component analysis. Neural Computation, 1997, 9(7), pp. 1483-1492.
- [4] Amari S, Cichocki A, Yang H. A new learning algorithm for blind signal separation. Advances in Neural Information Processing Systems, MIT Press: Cambridge, MA, 1996, 8, pp.757-763.
- [5] Cardoso J F, Laheld B. Equivariant adaptive source separation. IEEE Trans on Signal Processing, 1996, 44(12), pp. 3017-3030.
- [6] Yellin D, Weinstein E. Multichannel signal separation: methods and analysis. IEEE Trans on Signal Processing, 1996, 44 (1), pp. 106-118.
- [7] Comon P, Jutten C, Herault J. Blind separation of sources, Part II: Problems statement. Signal Processing, 1991, 24(1), pp. 11-20.
- [8] Tong L, Liu R, Soon V C, Huang J. F. Indeterminacy and identifiability of blind identification. IEEE Trans Signal Processing, 1991, 38(5), pp. 499-509.
- [9] Tong L, Inouye Y, Liu R. Waveform-preserving blind estimation of multiple independent sources. IEEE Trans Signal Processing, 1993, 41(7), pp. 2461-2470.
- [10] Hyvarinen A. Survey on independent component analysis. Neural Computing surveys, 1999, 2, pp. 94-128.
- [11] Comon P. Independent component analysis: a new concept? Signal Processing, 1994, 36 (3), pp. 287-314.
- [12] Moreau E, Moreau N T. nonsymmetrical contrasts for source separation. IEEE Trans Signal Processing, 1999, 47(8), pp. 2241-2252.

- [13] Bell A J, Sejnowski T J. An information maximization approach to blind separation and blind deconvolution. *Neural Computation*, 1995, 7(6), pp. 1129-1159.
- [14] Cardoso J F. Infomax and maximum likelihood in source separation. *IEEE Signal Processing Letter*, 1997, 4(4), pp. 112-114.
- [15] Cardoso J F. On the stability of source separation algorithms. *Journal of VLSI Signal Processing*. 2000, 26(1/2), pp. 7-23.
- [16] Karhunen J, Pajunen P, Oja E. The nonlinear PCA criterion in blind source separation: Relations with other approaches. *Neurocomputing*, 1998, 22, pp. 5-20.
- [17] Zhu X L, Zhang X D. Adaptive RLS algorithm for blind source separation using a natural gradient. *IEEE Signal Processing Letter*, 2002, 9(12), pp. 432-435
- [18] Hyvarinen A. Fast and robust fixed-point algorithms for independent component analysis. *IEEE Trans on Neural Networks*, 1999, 10(3), pp. 626-634.
- [19] Douglas S C, Cichocki A. Adaptive step size techniques for decorrelation and blind source separation. *Signals, Systems & Computers Conference*, 1998, pp.1191-1195.
- [20] Lambert R H. Multichannel blind deconvolution: FIR matrix algebra and separation of multipath mixtures. Ph.D. dissertation, Univ. Southern California, Los Angeles, 1996.
- [21] Sabala I, Cichocki A, Amari S. Relationships between instantaneous blind source separation and multichannel blind deconvolution. In *Proc. Int. Joint Conference on Neural Networks*, 1998, pp. 148-152.
- [22] Cichocki A, Amari S. *Adaptive Blind Signal and Image Processing*. New York, Wiley, 2003.
- [23] Amari S, Douglas S C, Cichocki A, Yang H H. Multichannel blind deconvolution and equalization using the natural gradient. *IEEE Proceedings of International Workshop Wireless Communication*, 1997, pp. 101-104.
- [24] Douglas S C, Sawada H, Makino S. Natural gradient multichannel blind deconvolution and speech separation using causal FIR filters. *IEEE Trans on Speech and Audio Processing*, 2005, 13(1), pp. 92-104.
- [25] Kokkinakis K, Nandi A K. Multichannel blind deconvolution for source separation in convolutive mixtures of speech. *IEEE Trans on Audio, Speech and Language Processing*, 2006, 14(1), pp. 200-212.
- [26] Torkkola K. Blind separation of convolved sources based on information

- maximization. IEEE Workshop on Neural Networks for Signal Processing, 1996, pp. 423-432.
- [27] Weinstein E, Feder M, Oppenheim A V. Multi-channel signal separation by decorrelation. IEEE Trans on Speech and Audio Processing, 1993, 1(4), pp. 405-413.
- [28] Gerven V S, Compernelle V D. Signal separation by symmetric adaptive decorrelation stability, convergence and uniqueness. IEEE Trans on Signal Processing, 1995, 43(7), pp. 1602-1612
- [29] Lindgren U A, Broman H. Source separation using a criterion based on second-order statistics. IEEE Trans on Signal Processing, 1998, 46(7), pp. 1837-1850
- [30] Kawamoto M, Matsucka K, Ohnishi N. A method of blind separation for convolved non-stationary signals. Neurocomputing, 1998, 22(1-3), pp. 157-171.
- [31] Chan D. Blind signal separation. PhD Thesis Trinity College, Cambridge University, 1997.
- [32] Smaragdis P. Blind separation of convolved mixtures in the frequency domain. Neurocomputing, 1998, 22, pp. 21-34.
- [33] Douglas S C, Gupta M. Scaled natural gradient algorithm for instantaneous and convolutive blind source separation. IEEE International Conference on Acoustics, Speech and Signal Processing, 2007, II, pp.637-640.
- [34] Araki S, Mukai R, Makino S. The fundamental limitation of frequency domain blind source separation for convolutive mixtures of speech. IEEE Trans on Speech and Audio Processing, 2003, 11(2), pp. 109-116.
- [35] Parra L, Spence C. Convolutive blind separation of nonstationary sources. IEEE Trans. Speech Audio Processing, 2000, 8(3), pp. 320-327.
- [36] Murata N, Ikeda S, Ziehe A. An approach to blind source separation based on temporal structure of speech signals. Neurocomputing 2001, 41, pp. 1-24.
- [37] Kurita S, Saruwatari H. Evaluation of blind signal separation method using directivity pattern under reverberant conditions. IEEE International Conference Acoustics, Speech, Signal Processing, 2000, pp. 3140-3143.
- [38] Ikram M Z, Morgan D R. A beamforming approach to permutation alignment for multichannel frequency-domain blind speech separation. IEEE International Conference Acoustics, Speech, Signal Processing, 2002, pp. 881-884.
- [39] Asano F, Ikeda S, Ogawa M, Asoh H, Kitawaki N. A combined approach of array

processing and independent component analysis for blind separation of acoustic signals. *IEEE Trans on Speech and Audio Processing*, 2003, 11(3), pp. 204–215.

- [40] Sawada H, Mukai R, Araki S, Makino S. A robust and precise method for solving the permutation problem of frequency-domain blind source separation. *IEEE Trans on Speech and Audio Processing*, 2004, 12(5), pp. 530–538.

## 第三章 改进的自然梯度算法

### § 3.1 引言

盲源分离是在不知道源信号和传输通道的先验信息的情况下,根据源信号的统计特性,仅由观测到的混合信号恢复出源信号的过程。在过去的二十几年中,盲源分离问题在声学和语音处理<sup>[1]</sup>图像增强以及生物信号处理<sup>[2]</sup>等诸多领域获得了广泛关注。为便于通过数字信号处理器实时实现这些应用,通常需要利用自适应学习的盲源分离算法,为此人们已经针对不同的盲源分离问题提出了大量的在线学习算法。自适应处理可以随着数据的陆续取得而逐步更新分离矩阵,使得分离信号各分量趋于相互独立,在线算法的计算一般比较简单,但收敛速度慢也是其固有缺点。

Cichocki 等人<sup>[3-5]</sup>提出了著名的自然梯度算法,既可以免除随机梯度算法<sup>[6]</sup>中矩阵求逆的问题,又保留了随机梯度算法的简单性和数值稳定性。1998年,Amari<sup>[7]</sup>利用微分几何学<sup>[8]</sup>证明了当用负对数似然函数作为损失函数时,自然梯度是 Fisher 有效的。自然梯度的性质以及扩展也获得了深入研究<sup>[9-16]</sup>。此外可以证明自然梯度算法实际上是与相对梯度<sup>[17]</sup>等效的,因此,自然梯度算法也具有等变化性。自然梯度算法的收敛性取决于步长和初始分离矩阵,如何选择这些参数以同时保证其收敛速度、稳定性和跟踪能力是很困难的。目前已经有很多选择步长的方法<sup>[18-22]</sup>。但是在混迭矩阵发生突变时,上述变步长盲源分离的跟踪能力还需进一步改善。

由于递归最小二乘(RLS)算法在收敛速度和跟踪能力方面要优于最小均方(LMS)算法<sup>[18]</sup>,Pajunen 等人<sup>[19]</sup>提出了基于非线性主分量分析的递归盲源分离方法。而自然梯度算法比一般梯度算法有更好的收敛性能和数值稳定性。Zhu 等人<sup>[20-21]</sup>结合非线性主分量分析和自然梯度,首先提出了一种 RLS 算法来解决盲源分离问题<sup>[20]</sup>,该方法的算法收敛速度快,数值稳定高。但要求分离矩阵为正交矩阵,若在递归学习中不能保持正交特性,则算法的性能会受到很大的影响,其仿真结果中正交性测度曲线呈间隔性起伏,从而会影响到算法的性能。Zhu 等人<sup>[21]</sup>随后又提无需对观测信号预白化的基于自然梯度和非线性主分量分析的递归最小二乘盲源分离方法。

本章在第 3.3 节中提出了一种用于盲源分离的自适应修正自然梯度算法。首先,在后向传播(BP)算法的启发下,我们在自然梯度的学习过程中引入动量项以加快收敛速度,同时改善自然梯度算法的稳定性。然后,在分离模型中引入一个合适的测度函数来控制步长和动量项因子,由此得到的变步长和变动量项因子的自

然梯度算法非常适合解决时变环境（例如混合矩阵突变或信号能量突变）下的盲源分离问题。在时不变和时变环境下的不同仿真实验表明本章所提算法在收敛速度、稳定性以及跟踪能力方面都有很大改善，并且当信号能量非常弱或信号能量比例严重失调时的分离效果仍然良好。该算法还适用于适定情况下的多个信号（例如：信号源的个数等于 10）分离。此外，为实现对观测数据按块处理时算法有快速收敛性能，在第 3.4 节中提出一种按块递归的盲源分离方法，基于文献[20]中用结合自然梯度和非线性主分量分析得到所求正交矩阵的最优解的方法，我们按块更新矩阵方程，且对其逐块求解来解决盲分离问题。在求解矩阵方程时利用 QR 分解和回代算法避免了直接求逆矩阵的运算，因此所提算法具有较好的数值稳定性，在迭代过程中保持了所求矩阵的正交特性。

### § 3.2 自适应改进的自然梯度算法

#### 3.2.1 自然梯度算法

在源信号统计独立的情况下，希望使用一神经网络从观测数据向量中实时恢复源信号，这一过程称为在线盲信号分离。图 3.1 画出了混合模型与盲源分离的前馈神经网络方框图。用  $M$  个传感器阵列接收  $N$  个统计独立的源信号  $\mathbf{s}(t) = [s_1(t), s_2(t), \dots, s_N(t)]^T$ ，混合矩阵根据不同的应用和环境会有不同的数学模型，本章主要讨论线性瞬时混合的情况，即观测信号  $x_j(t), j=1, \dots, M$  ( $M \geq N$ ) 是源信号  $s_i(t), i=1, \dots, N$  的线性瞬时混合信号。在这种情况下，信号模型通常表示为

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) \quad (3.1)$$

这里  $\mathbf{x}(t) = [x_1(t), x_2(t), \dots, x_M(t)]^T$  是阵列观测向量， $\mathbf{A} = [a_{ij}] \in \mathbb{R}^{M \times N}$  ( $M \geq N$ ) 是列满秩的未知混合矩阵。不失一般性，本章均考虑  $M = N$  的情形。

在线盲源分离的目的是用权矩阵  $\mathbf{W}(t) \in \mathbb{R}^{N \times N}$  从观测信号  $\mathbf{x}(t)$  中恢复源信号  $\mathbf{s}(t)$ ，即寻找矩阵  $\mathbf{B}$  使得  $\mathbf{B}\mathbf{A} = \mathbf{G}$ ，其中  $\mathbf{G}$  为广义交换矩阵。因此希望使用自适应学习算法，收敛后的权矩阵  $\mathbf{W}_\infty$  满足关系式

$$\mathbf{W}_\infty \mathbf{A} = \mathbf{G}$$

以使得输出向量

$$\mathbf{y}(t) = \mathbf{W}(t)\mathbf{x}(t) = \mathbf{W}_\infty \mathbf{A}\mathbf{s}(t) = \mathbf{G}\mathbf{s}(t) \quad (3.2)$$

为源信号的估计。因此，神经网络在时刻  $t$  的分离性能由合成矩阵  $\mathbf{T}(t) \triangleq \mathbf{W}(t)\mathbf{A}$  衡量，它描述混合-分离模型中信号被分离为独立分量的“精确性”。这里矩阵  $\mathbf{W}(t)$  在盲分离中称为分离矩阵或解混合矩阵，在神经计算和神经网络中称为突触权矩阵。

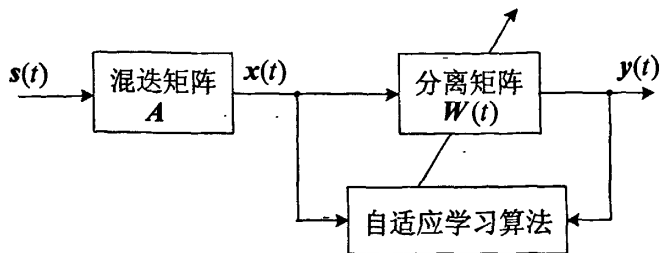


图 3.1 混合模型与盲源分离的前馈神经网络

下面从神经网络系统输出的互信息的最小化推导分离矩阵的更新公式。最小互信息(minimum mutual information, MMI)的基本思想是：选择神经网络的权矩阵  $W$ ，使输出  $y$  的各个分量之间的相依性最小化，即使得输出向量  $y(t)$  的各个分量尽可能相互统计独立。这一方法称为独立分量分析(ICA)，是 Comon 于 1994 年提出的<sup>[22]</sup>。现在，独立分量分析已成为信号处理、模式识别和神经计算等领域中的一种著名算法。

作为信号之间相依性的测度，通常采用输出向量  $y$  的联合概率密度函数与其边缘概率密度函数之间的 Kullback-Leibler(K-L)散度

$$\begin{aligned} D(W) &= K[p_y(y, W) \| q_y(y, W)] \\ &\triangleq \int p_y(y, W) \log \frac{p_y(y, W)}{q_y(y, W)} dy \end{aligned} \quad (3.3)$$

其中  $p_y(y, W)$  表示随机变量  $y$  联合概率密度函数， $q_y(y, W)$  是  $y$  的边缘概率密度函数的乘积，即  $q_y(y, W) = \prod_{i=1}^N q_i(y_i, W)$ 。

K-L 散度是两个概率分布差异的自然测度。因此  $D(W)$  给出了分布  $p_y(y, W)$  与分布  $q_y(y, W)$  之间的距离。显然，若各个输出分量  $y_i(t)$  相互独立，则  $p_y(y, W) = q_y(y, W)$ ，即  $D(W) = 0$ 。反之， $D(W) = 0$  意味着  $y_i(t)$  相互独立。

信号分量之间的 KL 散度与它们之间的互信息等价  $D(W) \equiv I(W)$ ，并且互信息为非负值，即  $I(W) \geq 0$ 。因此，当且仅当输出各分量独立时，互信息最小(等于零)，即

$$I(W) = 0 \quad \text{iff} \quad W = APA^{-1} \quad (3.4)$$

其中  $A$  为  $N \times N$  的对角矩阵， $P$  为  $N \times N$  的置换矩阵。

互信息也可以用熵表示为<sup>[5]</sup>

$$D(\mathbf{W}) = I(\mathbf{W}) = -H(\mathbf{y}, \mathbf{W}) + \sum_{i=1}^N H(y_i, \mathbf{W}) \quad (3.5)$$

其中  $H(\mathbf{y}, \mathbf{W}) = -\int p(\mathbf{y}, \mathbf{W}) \log p(\mathbf{y}, \mathbf{W}) d\mathbf{y}$  为联合熵,

$$H(y_i, \mathbf{W}) = -\int q_i(y_i, \mathbf{W}) \log q_i(y_i, \mathbf{W}) dy_i \text{ 为边缘熵.}$$

互信息最小的条件是: 熵  $H(\mathbf{y}, \mathbf{W})$  取最大值  $\sum_{i=1}^N H(y_i, \mathbf{W})$ 。即互信息最小与最大熵等价。

由  $\mathbf{y} = \mathbf{W}\mathbf{x}$ , 根据矩阵理论  $p(\mathbf{y}, \mathbf{W}) = p(\mathbf{x}) / \det(\mathbf{W})$ , 其中  $\det(\mathbf{W})$  表示矩阵  $\mathbf{W}$  的行列式, 微分熵可以表示为

$$H(\mathbf{y}) = H(\mathbf{x}) + \log |\det(\mathbf{W})| \quad (3.6)$$

对于  $q_i(y_i, \mathbf{W})$ , 可用类似的方法处理, 这样得到一个简单的代价(风险)函数

$$D(\mathbf{W}) = -H(\mathbf{x}) - \log |\det(\mathbf{W})| - \sum_{i=1}^N E\{\log q_i(y_i, \mathbf{W})\} \quad (3.7)$$

由于  $H(\mathbf{x})$  与分离矩阵  $\mathbf{W}$  无关, 可以把  $H(\mathbf{x})$  从式(3.7)中略去。

为使代价函数极小化, 一个简单的思想是用普通的随机梯度下降实时学习算法, 由下式给出

$$\Delta \mathbf{W}(t) = \mathbf{W}(t+1) - \mathbf{W}(t) = -\eta \frac{\partial D}{\partial \mathbf{W}} \quad (3.8)$$

这里  $\eta > 0$  为步长参数,  $\frac{\partial D}{\partial \mathbf{W}}$  是梯度矩阵。

由简单的微分矩阵计算, 可得

$$\Delta \mathbf{W}(t) = \eta \{ \mathbf{I} - f(\mathbf{y}(t)) \mathbf{y}^T(t) \} \mathbf{W}^{-T}(t) \quad (3.9)$$

其中  $\mathbf{I}$  为单位矩阵,  $\mathbf{W}^{-T}$  是  $\mathbf{W}$  逆阵的转置,  $f(\mathbf{y}) = [f(y_1), \dots, f(y_N)]^T$  中的激活函数  $f(y)$  为非线性奇函数。实际上, 大多算法中激活函数通常需要根据源信号的分布进行选择<sup>[5-6,17]</sup>。

上述随机梯度优化方法的主要缺点是收敛比较慢。因此, 希望有一种优化方法既能保留随机梯度方法的简单性和数值稳定性, 又能够得到很好的渐近收敛性。还希望其性能与混合矩阵无关, 以便即使当混合矩阵接近奇异矩阵时, 算法依然能够工作得很好。

当参数空间是欧几里德空间时, 梯度  $-\partial D / \partial \mathbf{W}$  代表函数  $D$  的最速下降方向。在现在的情形中, 参数空间由所有的非奇异  $N \times N$  矩阵  $\mathbf{W}$  组成。这是一个乘法群, 它的单位元是单位矩阵  $\mathbf{I}_N$ 。而且, 它是一个流形, 因此它形成一个李群。Amari

等<sup>[5,7,10]</sup>利用这个事实引入一个自然 Riemannian 测度到空间  $W$ ，他们证明了在黎曼空间中参数  $A$  的最陡下降方向不是沿梯度  $\frac{\partial D}{\partial W}$ ，而是沿自然梯度

$\frac{\partial D}{\partial W} W^T(t) W(t)$  的方向<sup>[7,10]</sup>。因此，自然梯度算法可以表示成在线形式：

$$\Delta W(t) = -\eta \frac{\partial D}{\partial W} W^T(t) W(t) = \eta \{I - f(y(t)) y^T(t)\} W(t) \quad (3.10a)$$

$$W(t+1) = W(t) + \Delta W(t) \quad (3.10b)$$

对应的块自然梯度算法可以表示为：

$$\Delta W(n) = \eta \left\{ I - \frac{1}{B} \sum_{t=B(n-1)+1}^{Bn} f(y(t)) y^T(t) \right\} W(n) \quad (3.11a)$$

$$W(n+1) = W(n) + \Delta W(n) \quad (3.11b)$$

其中  $B$  为块的大小。设置  $B=1$ ，可以很容易将块自然梯度算法转化为在线算法。

### 3.2.2 结合动量项的自然梯度算法

神经网络是一个由称为神经元的基本单元构成的并行分布传感器，构成神经网络的基本单元称为神经元。20 世纪 80 年代中期，Rumelhart 等人<sup>[23]</sup>提出了误差反向传播 (BP: back-propagation) 算法。从此，BP 算法被广泛应用于多层感知结构的训练。所谓多层感知器，就是单层感知器的推广，即在输入层和输出层之间加上隐含层（一层或多层）。这种由输入层、隐含层和输出层构成的神经网络称为多层前馈神经网络。

下面我们简单回顾一下 BP 算法。BP 学习过程可以描述如下：(1)工作信号正向传播：输入信号从输入层经隐单元，传向输出层，在输出端产生输出信号。在信号向前传递过程中网络的权值是固定不变的，如果在输出层不能得到期望的输出，则转入误差信号反向传播；(2)误差信号反向传播：网络的实际输出与期望输

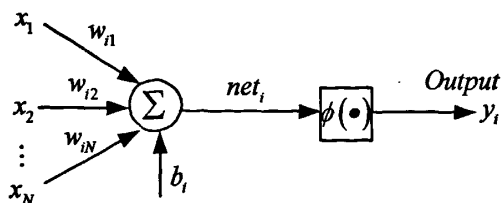


图 3.2 第  $i$  个神经元的简化模型

出之间的差值即为误差信号，误差信号由输出端开始逐层向前传播。在误差信号反向传播的过程中，网络的权值由误差反馈进行调节，通过权值的不断修正使网络的实际输出更接近期望输出。

图 3.2 为多层感知器结构中第  $i$  个神经元的简化模型，其中  $x_i, i=1, \dots, N$  表示观测信号， $w_{ij} (i=1, \dots, M, j=1, \dots, N)$  为权重， $net_i = \sum_{j=1}^N w_{ij} x_j$  为线性混合器的输出，激活函数通常定义为  $\phi(net_i) = \frac{1}{1 + \exp(-net_i)}$ 。

定义该网络的代价函数为均方差的累加和，即：

$$\xi(n) = \frac{1}{2} \sum_{i=1}^N e_i(n) e_i^*(n) = \frac{1}{2} \sum_{i=1}^N |e_i(n)|^2 \quad (3.12)$$

其中  $e_i(n) = d_i - y_i(n)$  为误差信号， $d_i$  为输出期望信号。

BP 算法就是通过不断调整权值，使代价函数  $\xi(n)$  达到极小值，由梯度下降法得到的权值更新公式为：

$$w_{ij}(n+1) = w_{ij}(n) + \Delta w_{ij}(n) \quad (3.13)$$

其中权值和负梯度成比例关系，所以更新项可以表示为：

$$\Delta w_{ij}(n) = -\mu \nabla_{w_{ij}} \xi(n) = -\mu \frac{\partial \xi(n)}{\partial w_{ij}} \quad (3.14)$$

这里  $\mu$  表示步长参数， $\nabla_{w_{ij}} \xi(n)$  为代价函数  $\xi(n)$  沿权值  $w_{ij}$  方向的梯度。

在式(3.14)中加入动量项可以加快 BP 算法的收敛速度<sup>[18, 23]</sup>，将更新项  $\Delta w_{ij}(n)$  修正为：

$$\Delta w_{ij}(n) = \alpha \Delta w_{ij}(n-1) - \mu \nabla_{w_{ij}} \xi(n) \quad (3.15)$$

其中， $\alpha$  称为动量项因子， $\Delta w_{ij}(n-1)$  为上一步的梯度。动量项为  $\Delta w_{ij}(n)$  增加了一个前馈循环，对 BP 算法的学习过程带来极大的好处，能够加快算法的收敛速度并且避免学习过程在陷入局部最小点而得不到最优解。

和 BP 算法类似，自然梯度算法也采用梯度下降法，所以，受 BP 算法启发，我们在自然梯度学习过程中引入动量项以加快其收敛速度，由此得到修正自然梯度算法：

$$\Delta \mathbf{W}(n) = -\eta \frac{\partial K}{\partial \mathbf{W}}(n) \mathbf{W}^T(n) \mathbf{W}(n) + \alpha \Delta \mathbf{W}(n-1) \quad (3.16a)$$

$$\mathbf{W}(n+1) = \mathbf{W}(n) + \Delta \mathbf{W}(n) \quad (3.16b)$$

其中  $\alpha \Delta \mathbf{W}(n-1)$  为动量项， $\alpha$  为动量项因子。

实际上，式(3.16a)可以看作  $\Delta \mathbf{W}(n)$  的一阶差分方程：

$$\Delta \mathbf{W}(n) = -\eta \sum_{k=0}^n \alpha^{n-k} \frac{\partial K}{\partial \mathbf{W}}(k) \mathbf{W}^T(k) \mathbf{W}(k) \quad (3.17)$$

$\Delta \mathbf{W}(n)$  是自然梯度序列的加权指数和, 或者说是式(3.16a)的平滑。因此, 在自然梯度算法中引入动量项也提高了学习过程的稳定性。

仿真结果表明结合动量项确实可以加快自然梯度算法的收敛速度并提高算法的稳定性, 但是, 在时变环境下的算法的跟踪能力及信号能量极弱时算法的分离性能有待提高。此外, 算法性能对  $\eta$  和  $\alpha$  很敏感, 而  $\eta$  和  $\alpha$  通常为经验值。下一节我们将提出一种自适应改进算法来消除自然梯度算法对参数的敏感性。

### 3.2.3 参数自适应控制算法

虽然在自然梯度算法中引入动量项可以加快收敛速度并提高学习过程的稳定性, 但是要选取最优的步长  $\eta$  和动量项因子  $\alpha$  很困难。一种简单的方法就是将  $\eta$  (或  $\alpha$ ) 定义为时间的指数下降函数, 但是时变环境下的跟踪能力并没有改善。自适应选择参数的方法<sup>[25-30]</sup>通常被认为是不错的选择, 特别是在时变环境下, 即能加快学习算法的收敛速度, 又能保证学习算法的稳态准确性。因此本节我们将提出一种自适应修正自然梯度算法。

自适应步长选择方法的原则是当被估计参数值与它的最优值之间的距离大时, 步长相应取较大的值, 反之, 步长应取较小的值<sup>[28]</sup>。在盲源分离问题中, 不可能直接利用实际误差来控制步长, 只能选择一种有效的替代测度来度量误差。Thomas 等人<sup>[28]</sup>定义了一种估计函数:

$$\mathbf{H}(n) = \mathbf{I} - \frac{1}{B} \sum_{t=B(n-1)+1}^{Bn} \mathbf{f}(\mathbf{y}(t)) \mathbf{y}^T(t) \quad (3.18)$$

令  $E\{\mathbf{H}(n)\}$  为  $\mathbf{H}(n)$  的期望, 用  $\mathbf{H}(n)$  的平滑形式  $\bar{\mathbf{H}}(n)$  来逼近期望值:

$$\bar{\mathbf{H}}(n) = (1 - \eta(n-1)) \bar{\mathbf{H}}(n-1) + \eta(n-1) \mathbf{H}(n) \quad (3.19)$$

有许多均方范数都可以作为合适的误差测度来控制步长, 我们采用了一种新的代价函数, 令  $\hat{H}_{ij}(n) = |\bar{H}_{ij}(n)|^2$ , 得到一个新的矩阵  $\hat{\mathbf{H}}(n)$ ,  $\hat{H}_{ij}(n)$  为第  $i$  行, 第  $j$  列元素。就可以用  $\hat{\mathbf{H}}(n)$  的 Frobenius 范数作为控制步长的代价函数, 即:

$$\Phi(n) = \|\hat{\mathbf{H}}(n)\|_F \quad (3.20)$$

步长  $\eta(n)$  根据  $\Phi(n)$  变化, 并且由下面自适应公式进行平滑:

$$\eta(n) = \beta \eta(n-1) + (1 - \beta) \rho \Phi(n) \quad (3.21)$$

其中  $\beta$  为遗忘因子, 一般取接近 1 的实数,  $\rho$  为比例因子。

由(3.16a)式知动量项因子  $\alpha$  对  $\Delta \mathbf{W}(n-1)$  的作用和步长  $\eta$  对  $\Delta \mathbf{W}(n)$  的作用相似, 即  $\alpha(n)$  由  $\Phi(n-1)$  控制, 得到动量项因子的自适应平滑形式:

$$\alpha(n) = \beta\alpha(n-1) + (1-\beta)\rho\Phi(n-1) \quad (3.22)$$

针对(3.21)和(3.22)的自适应形式，我们作如下注释：

**注释1：**通过大量仿真，我们发现要想在参数自适应调节的算法中用较小的运算复杂度得到最快的收敛速度和稳健性，选择一个合适的测度是关键所在。Thomas等人<sup>[28]</sup>定义代价函数为：

$$\Phi(n) = \max_{i,j} \bar{H}_{ij}^2(n) \quad (3.23)$$

其中 $\bar{H}_{ij}(n)$ 为 $\bar{\mathbf{H}}(n)$ 的第 $i$ 行，第 $j$ 列元素。此代价函数 $\Phi(n)$ 受限于对 $\mathbf{y}(t)$ 幅度的稳健性，尤其是当块大小 $B$ 很小时（如 $B=1$ ）。用 $\bar{\mathbf{H}}(n)$ 平方谱范数作为控制步长的代价函数更稳定，但是同时增加了运算复杂度。仿真结果表明(3.20)定义的代价函数具有较好的稳健性，运算量适中。

**注释2：**在式(3.21)和(3.22)的自适应形式中，合适的遗忘因子 $\beta$ 和比例因子 $\rho$ 能有效加快收敛速度，改善分离性能。Thomas等人<sup>[25]</sup>的自调节步长算法(记为SASS算法)选择 $\beta=0.998$ ， $\rho=0.25$ 。但是，SASS算法对 $\beta$ 和 $\rho$ 非常敏感。本文提出的算法在 $0.9 \leq \beta \leq 0.99$ ， $\rho=1$ 时效果都很好，说明我们的算法对 $\beta$ 敏感度降低，并且可以去掉比例因子 $\rho$ 。此外，为了使所有参数都可以自适应调节，遗忘因子 $\beta$ 也可以表示为自适应形式：

$$\beta(n) = (1-\eta(n-1))\beta(n-1) + \eta(n-1)(1-\Phi(n)) \quad (3.24)$$

由此得到式(3.21)和(3.22)的修正形式：

$$\eta(n) = \beta(n)\eta(n-1) + (1-\beta(n))\Phi(n) \quad (3.25)$$

$$\alpha(n) = \beta(n-1)\alpha(n-1) + (1-\beta(n-1))\Phi(n-1) \quad (3.26)$$

表3.1 自适应修正的盲源分离算法步骤

$\mathbf{H}(n) = \mathbf{I} - \frac{1}{B} \sum_{t=B(n-1)+1}^{Bn} \mathbf{f}(\mathbf{y}(t))\mathbf{y}^T(t)$
$\bar{\mathbf{H}}(n) = (1-\eta(n-1))\bar{\mathbf{H}}(n-1) + \eta(n-1)\mathbf{H}(n)$
$\Phi(n) = \ \hat{\mathbf{H}}(n)\ _F$
$\beta(n) = (1-\eta(n-1))\beta(n-1) + \eta(n-1)(1-\Phi(n))$
$\alpha(n) = \beta(n-1)\alpha(n-1) + (1-\beta(n-1))\Phi(n-1)$
$\eta(n) = \beta(n)\eta(n-1) + (1-\beta(n))\Phi(n)$
$\Delta\mathbf{W}(n) = \eta(n)\mathbf{H}(n)\mathbf{W}(n-1) + \alpha(n)\Delta\mathbf{W}(n-1)$
$\mathbf{W}(n+1) = \mathbf{W}(n) + \Delta\mathbf{W}(n)$

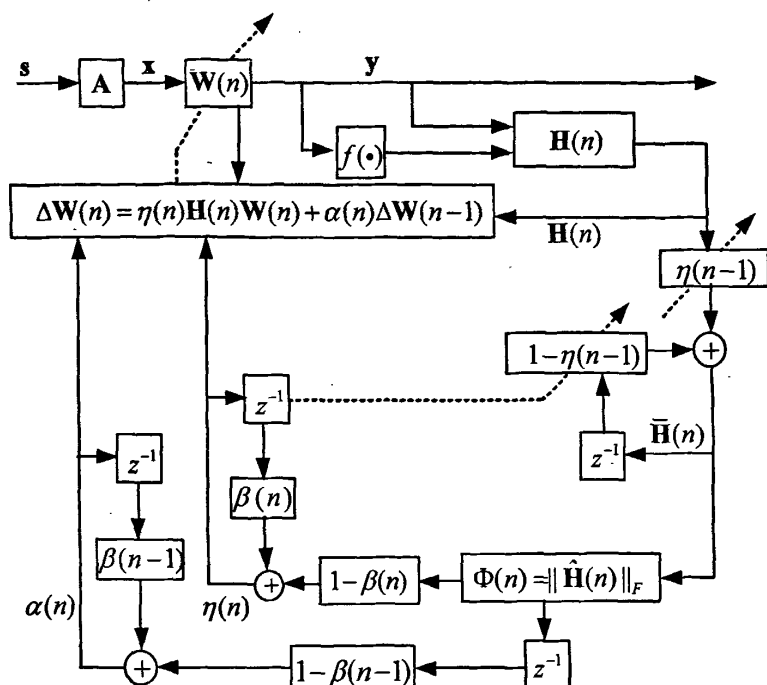


图 3.3 自适应修正的盲源分离算法流程图

**注释3:** 为了避免  $W(n)$ ,  $H(n)$ ,  $\eta(n)$ ,  $\alpha(n)$  和  $\beta(n)$  “溢出”, 我们约束  $\eta(n)$  的上限  $\eta_{up} = \frac{1}{2 \max \sum_i |H_{ij}(n)|}$ ,  $\alpha(n)$  的上限为 0.5。如果  $\beta(n) < 1 - \eta_{up}^2$  或  $\beta(n) > 0.99$ , 令  $\beta(n) = 1 - \eta_{up}^2$ 。

自适应修正的自然梯度算法 (记为 AI-NG) 框图如图 3.3 所示, 算法步骤见表 3.1。如果不加自适应动量项, 除了代价函数不同我们提出的算法本质上和 Thomas 等人<sup>[25]</sup>的自调节步长算法相同。仿真结果证明了新算法不仅可以加快收敛速度, 而且可以提高时变环境下 (混合矩阵突变或信号能量突变) 的跟踪能力。此外, 对极弱信号或信号比例严重失调的情况分离效果良好。

### 3.2.4 仿真实验

在本节中, 我们用计算机仿真结果来验证所提自适应修正的自然梯度算法的有效性。由于在后续的仿真实验中对功率不同的源信号进行分离, 为便于比较, 我们利用信干比(SIR)作为性能指标, 信干比率定义如下:

$$SIR = 10 \log_{10} \frac{1}{N} \sum_{i=1}^N \left( \frac{\max_p \left( |c_{ip}|^2 \text{var}(s_p) \right)}{\sum_{j=1}^N |c_{ij}|^2 \text{var}(s_p) - \max_p \left( |c_{ip}|^2 \text{var}(s_p) \right)} \right) \quad (3.27)$$

这里  $C = \{c_{ij}\} = WA$  是组合矩阵,  $\text{var}(s_p)$  表示信号  $s_p$  的方差。

在仿真实验中, 取十个复值信号作为源信号, 其中五个 4-QAM 信号, 五个 8-PSK 信号。每次仿真实验所用混合矩阵  $A = [a_{ij}]$  的各元素为服从复高斯分布的随机数, 且有  $E\{|a_{ij}|^2\} = 1$  和  $E\{a_{ij}^2\} = 0$ 。激活函数取为  $f(y) = |y|^2 y$ 。所有算法取同样的初始分离矩阵  $W(0) = 0.001I$ 。

实验一: 考虑在时不变环境下的两种情况:

(1) 首先比较结合动量项的自然梯度算法(记作 M-NG)和经典自然梯度算法(记作 NG)的性能。本例中动量项因子  $\alpha$  分别取不同的值( $\alpha = 0.2, 0.35, 0.5$ )。样本数  $T = 50000$ , 块长  $B = 10$ , 在两个算法中步长都取  $\eta = 0.01$ 。500 次独立运行的平均 SIR 性能曲线如图 3.4 所示它表明 M-NG 算法比 NG 算法有更快的收敛速度。动量项因子  $\alpha$  越大, 算法收敛越快, 但  $\alpha$  太大则可能导致算法发散。另一方面, 如果  $\alpha$  取值太小, 则会使动量项无法起到加快收敛速度的作用。因此对  $\alpha$  取值需要收敛速度和稳态性能之间折中。通常来说, 动量项的取值范围为  $0.2 \leq \alpha \leq 0.5$ 。

(2) 当只有小样本的观测数据可用时, 图 3.4 表明块处理的自然梯度算法无法收敛, 另一方面, 当块长太小时(如  $B = 1$ ), 在多信源的情况下(如  $n \geq 10$ )无法得到好的分离性能。在这种情况下, Douglas 等人<sup>[31]</sup>在每次迭代中重复利用所有观测数

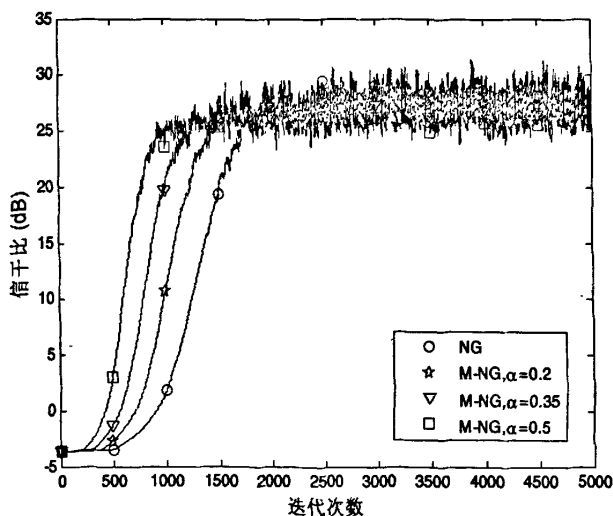


图 3.4 不同自然梯度算法经 500 独立运行的信干比均值图

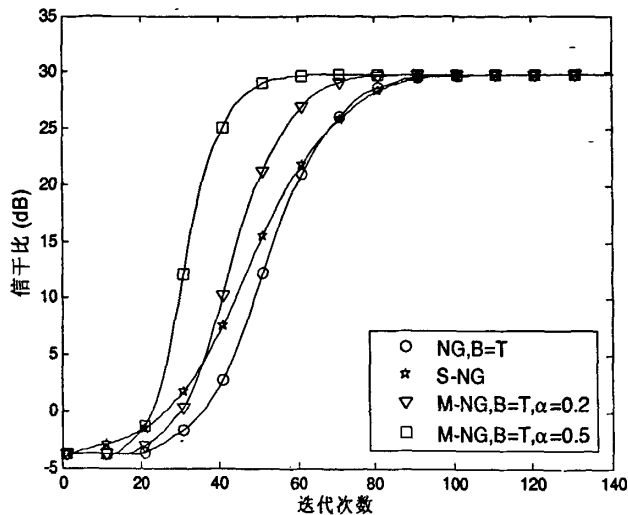


图 3.5 不同自然梯度算法在 1000 个样本  $B=T$  时经 500 独立运行的信干比均值图

据，这等价于在式(3.11)中令  $B=T$ ，对应算法记为 S-NG 算法。为比较算法性能，下面同时运行 NG 算法、M-NG 算法和 S-NG 算法，这里步长  $\eta=0.3$ ，样本数  $T=1000$ 。图 3.5 表明 M-NG 有较快的收敛性能，而三个算法的稳态精确性相同。

(3)为便于比较  $B=10$  时不同算法的分离性能，下面同时运行 NG 算法，M-NG 算法，SASS 算法<sup>[28]</sup>和我们提出的自适应修正自然梯度算法(记为 AI-NG)算法和没有结合动量项的 AI-NG 算法(记为 AI-NG-No-Mom)。在这五个算法中， $T=50000$ 。每个算法的其它参数见表 3.2。图 3.6(a)表明 AI-NG 算法和 AI-NG-No-Mom 算法的收敛速度和稳态精确性能明显优于其它三种算法，这也表明文中提出的测度(式(3.20))有更好的性能。

为进一步表明自适应动量项在 AI-NG 算法中的作用，AI-NG 算法和 AI-NG-No-Mom 算法的平均性能差如图 3.6(b)所示，这表明与自适应步长因子相

表3.2 实验一第三种情况不同算法中所用参数

NG: $\eta = 0.01$ .
M-NG: $\eta = 0.01, \alpha = 0.5$ .
SASS: initial of $\eta$ is 0.1, $\beta = 0.998, \rho = 0.25$ .
AI-NG-No-Mom: initial of $\eta$ is 0.1, initial of $\beta$ is 0.9.
AI-NG: initial of $\eta$ is 0.1, initial of $\alpha$ is 0.5, initial of $\beta$ is 0.9.

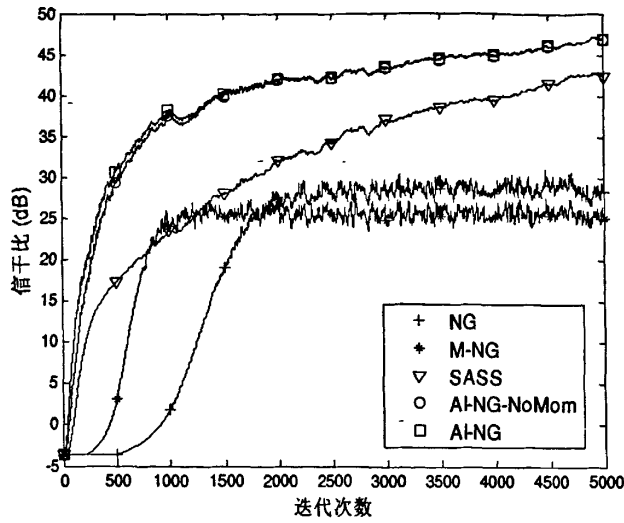


图 3.6(a): 五个不同自然梯度算法经 500 独立运行的信干比均值图

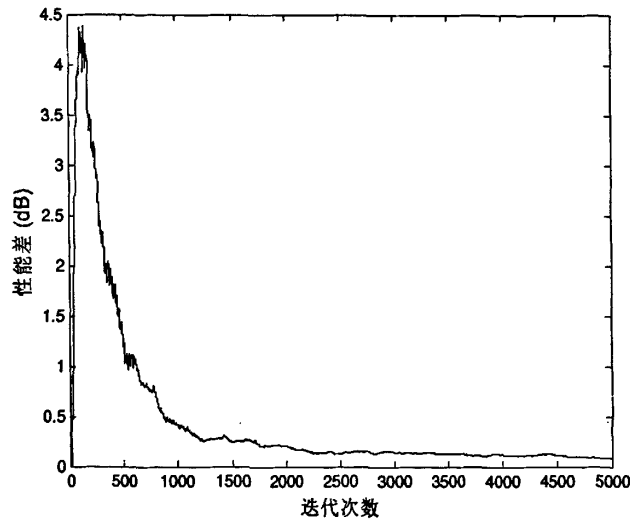


图 3.6(b): AI-NG 算法和 AI-NG-No-Mom 算法经 500 独立运行的平均性能差

比，自适应动量项因子对于 AI-NG 算法的分离性能起到一个辅助或次要的作用。这是因为自适应动量项因子  $\alpha(n)$  依赖于前一时刻的值  $\Phi(n-1)$ ，而自适应步长因子则依赖于当前值。此外，根据第 3.2.2 节中的分析，在自然梯度中结合动量项可以在学习过程中改善算法的稳定性。因此，自适应动量项对于 AI-NG 的性能有积极作用。

实验二：设瞬时混合的 10 个源信号中有 5 个幅度均为 1 的 8-PSK，5 个幅度均为  $10^{-9}$  的 8-PSK。其它参数的取值与实验一中第三种情况。分离性能曲线如图 3.7 所示，表明 NG 算法和 M-NG 算法收敛较慢，分离性能也比较差，SASS 算法和 AI-NG 算法收敛性能相对较好，但 AI-NG 算法比 SASS 算法有更快的收敛速度和更好的分离性能。

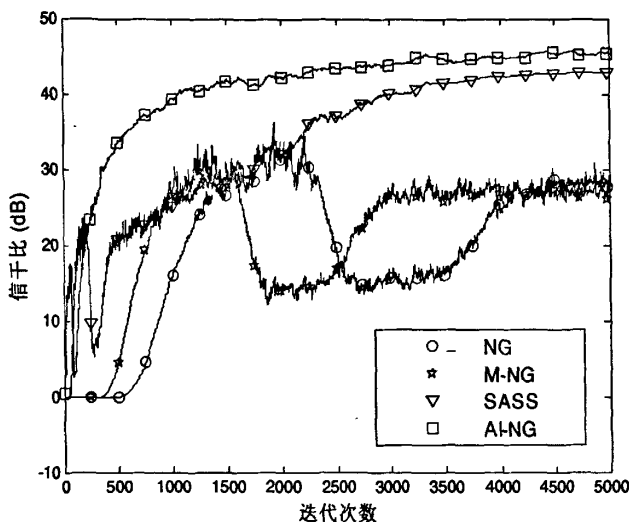


图 3.7 四个不同自然梯度算法经 500 独立运行的信干比均值图

实验三：为验证 AI-NG 算法在时变环境中的跟踪能力，我们在下面两种情况下同时运行参数自适应调节的 AI-NG 算法和 SASS 算法进行比较。

(1) 假设在样本数  $T=100000$  的情况下，混合矩阵每隔 25000 个样本发生一次突然变化，发生突变后的混合矩阵的元素同样服从复高斯分布，且有  $E\{a_{ij}^2\}=1$  和  $E\{a_{ij}^2\}=0$ 。其余参数取值与实验一第三种情况相同，分离性能曲线如图 3.8 所示。

(2) 假设在样本数  $T=50000$  的情况下，源信号功率隔 25000 个样本发生一次突然变化。为方便起见，假设在发生突变前的 10 个源信号中有 5 个幅度为  $10^{-9}$  的 4-QAM 信号，5 个幅度为 1 的 8-PSK 信号。在信号功率发生突变后，10 个源信号中有 5 个幅度为 1 的 4-QAM 信号，5 个幅度为 1 的 8-PSK 信号。其余参数的取值与实验一第三种情况相同，分离性能曲线如图 3.9 所示。

图 3.8 和图 3.9 表明本章提出的 AI-NG 算法在时变环境中有很好的跟踪能力。

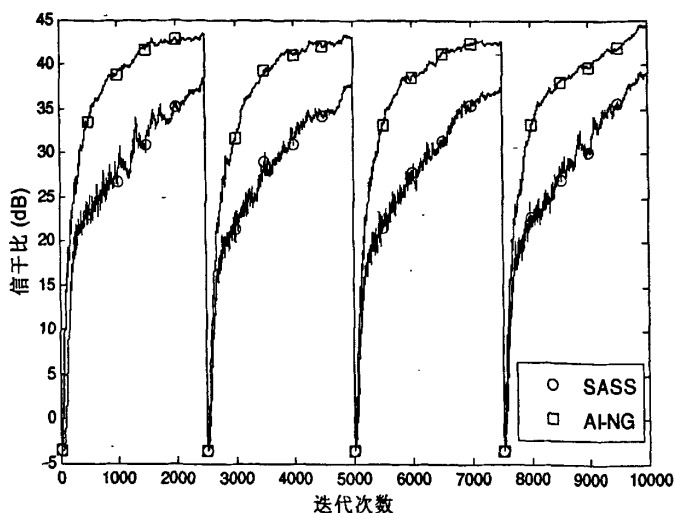


图 3.8 不同算法在混合矩阵发生突变时经 500 独立运行的信干比均值图

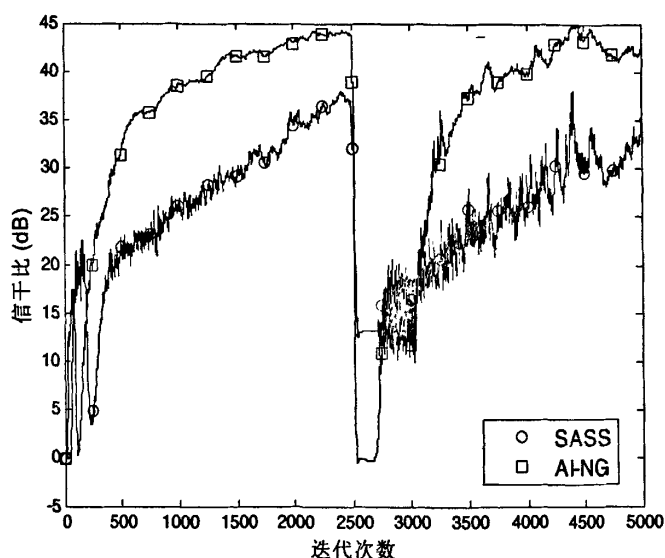


图 3.9 不同算法在信号功率发生突变时经 500 独立运行的信干比均值图

### § 3.3 基于自然梯度和非线性主分量分析的块递归盲源分离算法

本节考虑两阶段的盲分离方法,即首先线性瞬时混合模型(3.1)中的观测向量用矩阵  $\mathbf{V} \in \mathbb{R}^{N \times M}$  作预白化处理,然后经自适应学习求得正交矩阵  $\mathbf{W} \in \mathbb{R}^{N \times N}$ ,把源信号从经白化处理的观测数据中分离出来。因此,总体分离矩阵可记为  $\mathbf{B} = \mathbf{W}\mathbf{V}$ 。不

失一般性, 本节考虑  $M = N$  的情形。

白化矩阵  $V$  可以由离线或在线的方法近似计算<sup>[21]</sup>, 经预白化处理的观测向量

$$\mathbf{v}(t) = V\mathbf{x}(t) = V\mathbf{A}\mathbf{s}(t) = \tilde{Q}\mathbf{s}(t) \quad (3.28)$$

的协方差矩阵为单位矩阵, 即

$$E\{\mathbf{v}(t)\mathbf{v}(t)^T\} = VR_{xx}V^T = VAA^TV^T = \tilde{Q}\tilde{Q}^T = I \quad (3.29)$$

这里  $\tilde{Q} = VA$  为正交矩阵。

这样输出信号向量为

$$\mathbf{y}(t) = B\mathbf{x}(t) = W\mathbf{v}(t) = WV\mathbf{x}(t) \quad (3.30)$$

只需求出矩阵  $W$  即可分离出源信号, 可以证明<sup>[22]</sup>, 矩阵  $W$  的最优解为

$$W = PJQ^T \quad (3.31)$$

这里  $J$  为对角元素为  $\pm 1$  的对角矩阵, 易证  $W$  为正交矩阵。

### 3.3.1 递归最小二乘盲源分离算法

有很多准则可以用来确定正交矩阵  $W$ , 下面采用非线性主分量分析准则<sup>[35]</sup>:

$$J_1(W(t)) = E\left\{\left\|\mathbf{v}(t) - W^T(t)\mathbf{g}(W(t)\mathbf{v}(t))\right\|^2\right\} \quad (3.32)$$

其中  $\mathbf{g}(\mathbf{y}) = [g(y_1), \dots, g(y_N)]^T$ ,  $g(\bullet)$  为一非线性函数。

文献[19]提出了一种修正的最小二乘型准则, 它用指数加权求和代替式(3.32)中均方误差的期望, 用  $\mathbf{z}(t) = \mathbf{g}(W(t-1)\mathbf{v}(t))$  近似代替  $\mathbf{g}(W(t)\mathbf{v}(t))$ , 得到

$$J(W(t)) = \sum_{k=1}^t \beta^{t-k} \left\| \mathbf{v}(k) - W^T(t)\mathbf{z}(k) \right\|^2 \quad (3.33)$$

这里  $0 < \beta \leq 1$  是接近于1的遗忘因子。  $J(W(t))$  关于  $W(t)$  的梯度为

$$\nabla J(W(t)) = \sum_{k=1}^t \beta^{t-k} \left\{ -\mathbf{z}(k)\mathbf{v}^T(k) + \mathbf{z}(k)\mathbf{z}^T(k)W(t) \right\} \quad (3.34)$$

Amari<sup>[7]</sup>已经证明自然梯度学习比一般梯度学习更加有效, 自然梯度已广泛应用于盲源分离问题中。自然梯度  $\tilde{\nabla}J(W)$  和一般梯度  $\nabla J(W)$  的关系为

$$\tilde{\nabla}J(W) = WW^T \nabla J(W) - W[\nabla J(W)]^T W \quad (3.35)$$

Zhu等人<sup>[20]</sup>把式(3.34)代入式(3.35)经简化后得到

$$\tilde{\nabla}J(W(t)) = \sum_{k=1}^t \beta^{t-k} \left\{ -\mathbf{z}(k)\mathbf{v}^T(k) + \mathbf{y}(k)\mathbf{z}^T(k)W(t) \right\} \quad (3.36)$$

并由此得出基于自然梯度的递归最小二乘盲源分离方法。文献[20]中递归型算法得到  $W(t)$  在时刻  $t$  的最优值为

$$W_{opt}(t) = \left[ \sum_{k=1}^t \beta^{t-k} y(k) z^T(k) \right]^{-1} \left[ \sum_{k=1}^t \beta^{t-k} z(k) v^T(k) \right] = \Phi^{-1}(t) \Psi(t) \quad (3.37)$$

其中  $\Phi(t) = \sum_{k=1}^t \beta^{t-k} y(k) z^T(k)$ ,  $\Psi(t) = \sum_{k=1}^t \beta^{t-k} z(k) v^T(k)$ 。用递归的形式可分别写成

$$\Phi(t) = \beta \Phi(t-1) + y(t) z^T(t), \quad \Psi(t) = \beta \Psi(t-1) + z(t) v^T(t) \quad (3.38)$$

### 3.3.2 块递归的自适应盲源分离算法

#### 3.2.2.1 矩阵方程的求解

下面从矩阵方程求解的观点来考察式(3.37),  $W_{opt}(t)$  可以看成是矩阵方程

$$\Phi(t)W(t) = \Psi(t) \quad (3.39)$$

的解。为了讨论的方便, 下面对式(3.39)的分析省略了时间指标  $t$ 。

我们首先讨论矩阵方程

$$\Phi W = \Psi \quad (3.40)$$

的求解问题, 这里  $\Phi, W, \Psi$  均为  $N$  阶方阵, 且设  $\Phi$  可逆。我们知道, 在解线性方程组或矩阵方程时要遵循的基本原则是应当尽可能避免直接求矩阵的逆<sup>[32]</sup>。

因此在我们的问题中, 为避免在求解矩阵方程时直接求  $\Phi$  的逆, 首先对  $\Phi$  作 QR 分解

$$\Phi = QR \quad (3.41)$$

这里  $Q$  为正交矩阵,  $R$  为上三角矩阵。因为  $\Phi$  为可逆矩阵, 则  $R$  的对角线元素均为正数。对应的矩阵方程(3.40)变为

$$QRW = \Psi \quad (3.42)$$

这样在求解(3.42)时就可以充分利用对  $\Phi$  作 QR 分解得到的  $Q$  为正交矩阵这一优点, 即  $Q^{-1} = Q^T$ 。对(3.40)两边同乘  $Q^T$ , 且令  $\Gamma = Q^T \Psi$ , 得

$$RW = \Gamma \quad (3.43)$$

即得到一个下三角矩阵方程。对于下三角线性方程组, 可用回代方法<sup>[32]</sup>求解, 矩阵方程的解法与之类似, 但它是逐行求出  $W$ 。下三角矩阵方程的回代算法的伪代码见表3.3。

**注释一:** 在表3.3中,  $W(k,:)$  表示矩阵  $W$  的第  $k$  行,  $r_{jk}$  表示矩阵  $R$  的第  $(j,k)$  个

表3.3 回代算法的伪代码

---

```

W(n,:) = Γ(n,+)/rnn
for k = (n-1):(-1):1
    W(k,:) = [Γ(k,:) - ∑i=k+1n rkiW(i,:)] / wkk
end

```

---

元素。

**注释二：**在回代算法中，在计算 $W$ 的第 $k$ 行时，由于 $w_{kk}$ 作分母，通常需要判断其是否为零。如前所述， $R$ 的对角线元素均为正数，故可省略这一步骤。

### 3.3.2.2 块递归实现

式(3.37)中 $\Phi(t)$ 和 $\Psi(t)$ 可以递归计算，由上一节的分析对每个时刻 $t$ 我们可以通过两阶段的方法求矩阵方程的解 $W(t)$ 而不需要直接求 $\Phi(t)$ 的逆，这样的方法具有更好的数值稳定性。如前所述，为更精确的估计出 $W(t)$ ，下面采用按块更新矩阵、逐块求解矩阵方程的块递归方法。把所用数据 $y(k)$ ， $z(k)$ 和 $v(k)$ 分别分成长度为均为 $K$ 的块，把得到数据块分别求出互相关矩阵，再逐块相加来实现 $\Phi(t)$ 和 $\Psi(t)$ 按块递归更新。为此把(3.37)中指数加权求和的形式修正为

$$\begin{aligned}\Phi^b(t') &= \sum_{\tau=1}^{t'} \beta^{t'-\tau} \left( \frac{1}{K} \sum_{k=(\tau-1)K+1}^{\tau K} y(k) z^T(k) \right) \\ \Psi^b(t') &= \sum_{\tau=1}^{t'} \beta^{t'-\tau} \left( \frac{1}{K} \sum_{k=(\tau-1)K+1}^{\tau K} z(k) v^T(k) \right)\end{aligned}\quad (3.44)$$

其按块递归的形式分别为

$$\begin{aligned}\Phi^b(t') &= \beta \Phi^b(t'-1) + \frac{1}{K} \sum_{k=(t'-1)K+1}^{t'K} y(k) z^T(k) \\ \Psi^b(t') &= \beta \Psi^b(t'-1) + \frac{1}{K} \sum_{k=(t'-1)K+1}^{t'K} z(k) v^T(k)\end{aligned}\quad (3.45)$$

这里上标 $b$ 表示 $\Phi(t)$ 和 $\Psi(t)$ 经由块递归计算， $K$ 表示每一小块的长度，为与时刻 $t$ 区别，这里用 $t'$ 表示按块递归的时间指标，首先用求和近似计算各小块的互相关函数，然后再对不同的块用指数加权求和计算所有样本的互相关函数。若所处理的数据样本为 $T$ ，则相当于把 $T$ 个样本分成互不重叠的 $\lfloor T/K \rfloor + 1$ 个小块（这里 $\lfloor \cdot \rfloor$ 表示取下整数），即 $t' = 1, \dots, \lfloor T/K \rfloor + 1$ 。此外，在第 $\lfloor T/K \rfloor + 1$ 个小块中，只有 $\text{rem}(T, K)$ （表示 $T$ 除以 $K$ 的余数，它小于 $K$ ）个样本，此时只需对这 $\text{rem}(T, K)$ 个数据求互相关作为第 $\lfloor T/K \rfloor + 1$ 个小块的互相关矩阵。这样可由矩阵方程

$$\Phi^b(t') W(t') = \Psi^b(t'), \quad t' = 1, \dots, \lfloor T/K \rfloor + 1 \quad (3.46)$$

按第3.3.2.1节中的方法递归得出最优正交矩阵 $W_{opt}$ 。

### 3.3.3 仿真实验

为验证所提算法的有效性，这里做了不同的仿真实验，第3.2节定义的信干比(SIR)作为性能指标。在仿真实验中，我们使用文献[20]中的源信号：

$$s(t) = [\text{sign}(\cos(2\pi 155t)), \sin(2\pi 800t), \sin(2\pi 300t + 6\cos(2\pi 60t)), \sin(2\pi 90t), \omega(t)]^T$$

这里 $\omega(t)$ 是在 $[-1, 1]$ 上服从均匀分布的随机信号。源信号的采样率为 $10\text{kHz}$ ，源信号

与随机产生的方阵混合得到观测信号。

为表示的方便，把[19]中算法记为‘Pajunen’，[20] 中算法记为‘Zhu’，本文所提算法记为‘BR-BSS’。仿真中样本数 $T = 4000$ ， $\beta = 0.983$ ， $W(0) = I$ ，非线性函数  $g(t) = \tanh(t)$ 。

在BR-BSS算法中，为取得快速收敛并保持好的分离性能，前500个样本取  $K=1$ ，块迭代次数为  $t'=500$ 。从第501个样本起取  $K=10$ ，块迭代次数为  $t'=(T-500)/K=350$ 。BR-BSS算法中块长 $K$ 的取值越大，则算法运行一次所需时间越少，但算法的自适应性会减弱，因此需要对算法的运算时间和自适应性能折中。

三个算法按样本比较的平均性能指标见图3.10。Pajunen的算法是基于一般梯度的递归算法，在本例中出现发散，而Zhu的算法和BR-BSS算法由于基于自然梯度则不会发散。BR-BSS算法和Zhu的算法在前500次迭代的性能曲线基本重合，而在500次迭代之后BR-BSS算法的串音误差小于Zhu的算法。三种算法平均运行一次所需时间见表3.4，BR-BSS算法一次运行所需时间明显小于其它两种算法。Zhu的算法采用时变遗忘因子可以加快收敛速度和增强稳态性能，但BR-BSS算法可以在同时实现这两个优点的同时只需较少的运算时间，遗忘因子的选择相对简单。

表3.4 三种算法运行一次所需平均时间

算法	Pajunen	Zhu	BR-BSS
运行时间（秒）	0.484	0.437	0.156

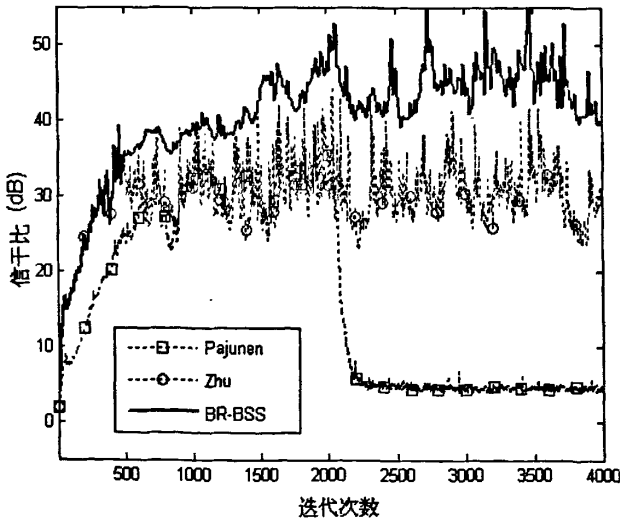


图3.10 不同递归算法500次独立运行的信干比均值图

为了与同样的按块处理的AI-NG算法的性能,下面同时运行BR-BSS算法和AI-NG算法,由于AI-NG算法中参数自适应控制,因此在行BR-BSS算法中使用时变遗忘因子<sup>[20]</sup>

$$\beta(t) = \begin{cases} 0.95 + 10^{-4}(t-1), & 1 \leq t \leq L \\ \beta_{\max}, & t > L \end{cases}$$

这里  $\beta_{\max} = 1$ ,  $L = (\beta_{\max} - 0.95)/10^{-4}$ 。

由图3.11可知, BR-BSS算法在初始迭代阶段收敛性能较快,但是稳态性能不如 AI-NG 算法且一次运行所需时间大于 AI-NG 算法。

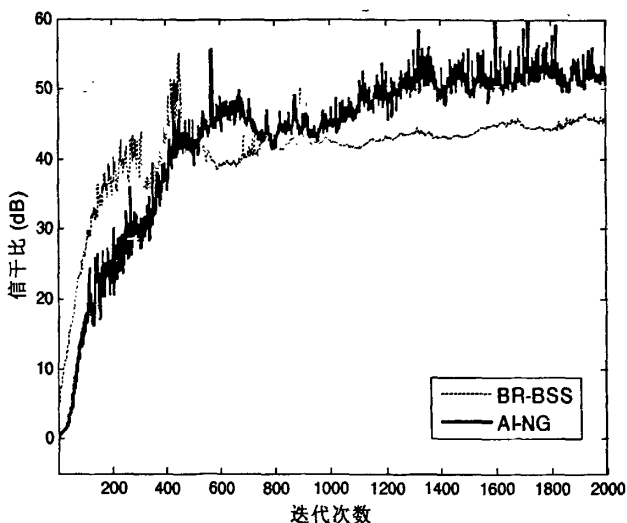


图 3.11 不同块处理算法 500 次独立运行信干比均值图

由于递归学习的矩阵  $\mathbf{W}$  是正交矩阵,因此需要考察在递归过程中  $\mathbf{W}$  是否保持正交性。同文献[31],我们用如下代价函数来衡量  $\mathbf{W}$  偏离正交矩阵的程度:

$\mathfrak{I} = \|\mathbf{W}\mathbf{W}^T - \mathbf{A}\|_F$ , 这里  $\mathbf{A} = \text{diag}(\mathbf{W}\mathbf{W}^T)$  表示由  $\mathbf{W}\mathbf{W}^T$  的对角元素构成的对角矩阵,

$\|\cdot\|_F$  表示Frobenius范数。如图3.12所示,显然本文所提算法得到矩阵  $\mathbf{W}$  的正交性能指标较小,说明在BR-BSS算法中  $\mathbf{W}$  在递归更新时更好的保持了正交性。

### § 3.4 本章小结

本章首先提出了一种变步长变动量相因子的自适应修正的自然梯度算法。首先,在后向传播(BP)算法的启发下,我们在自然梯度的学习过程中引入动量项以加快收敛速度,同时改善自然梯度算法的稳定性。然后,在分离模型中引入一个

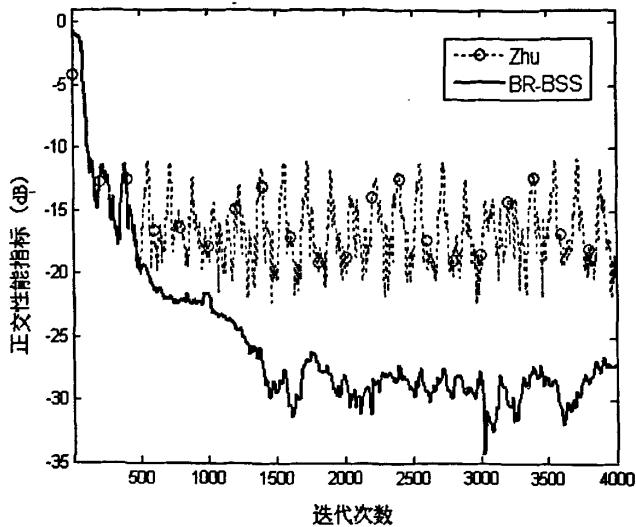


图 3.12 不同算法 500 次独立运行矩阵  $W$  的正交性能均值图

合适的测度函数来控制步长和动量项因子，由此得到的变步长和变动量项因子的自然梯度算法非常适合解决时变环境（例如混合矩阵突变或信号功率突变）下的盲源分离问题。对不同自然梯度的性能进行了详细的比较，实验表明与经典的自然梯度算法和其它改进的自然梯度算法相比，即使在信源个数很多的情况下，自适应修正的自然梯度算法有更快的收敛速度和更好的稳态精确性，当混合矩阵突变或信号功率突变时自适应修正的自然梯度算法依然有较好的跟踪能力。随后又提出一种按块递归的盲源分离方法，基于结合自然梯度和非线性主分量分析得到所求正交矩阵的最优解的方法，我们按块更新矩阵方程，且对其逐块求解来解决盲分离问题。在求解矩阵方程时利用QR分解和回代算法避免了直接求逆矩阵的运算，因此所提算法具有较好的数值稳定性，在迭代过程中保持了所求矩阵的正交特性。

### 参考文献

- [1] Brandstein M, Ward D. Microphone Arrays: Signal Processing Techniques and Applications (Digital Signal Processing), Springer-Verlag, 2001.
- [2] Cichocki A, Amari S. Adaptive Blind Signal and Image Processing. New York: Wiley, 2003.
- [3] Cichocki A, Unbehauen R, Rummert E. Robust learning algorithm for blind separation of signals. Electronics Letters, 1994, 30(17), 1386-1387.
- [4] Cichocki A, Unbehauen R. Robust neural networks with on-line learning for blind

- identification and blind separation of sources. *IEEE Trans. Circuits and Systems I: Fundamentals Theory and Applications*, 1996, 43(11), pp.894-906.
- [5] Amari S, Cichocki A, Yang H. A new learning algorithm for blind signal separation. *Advances in Neural Information Processing Systems*, MIT Press: Cambridge, MA, 1996, 8, pp.757-763.
- [6] Bell A, Sejnowski T. An information maximization approach to blind separation and blind deconvolution. *Neural Computation*, 1995, 7, pp.1129-1159.
- [7] Amari S. Natural gradient works efficiently. *Neural Computation*, 1998, 10, pp.251-276.
- [8] Amari S. *Differential-geometrical method in statistics*. Berlin: Springer-Verlag, 1985.
- [9] Amari S, Chen T, Cichocki A. Stability analysis of adaptive blind source separation. *Neural Networks*, 1997, 10(8), pp.1345-1351.
- [10] Yang H H, Amari A. Adaptive on-line learning algorithms for blind separation-maximum entropy and minimum mutual information. *Neural Computation*, 1997, 9, pp.1457-1482.
- [11] Rattray M, Saad D, Amari S. Natural gradient descent for on-line learning. *Physical Review Letters*, 1998, 81, pp.5461-5464.
- [12] Bottou L. Online algorithms and stochastic approximations. In D. Saad (Ed.), *Online learning and neural networks*. Cambridge: Cambridge University Press, 1998.
- [13] Amari S, Chen T, Cichocki A. Nonholonomic orthogonal learning algorithms for blind source separation. *Neural Computation*, 2000, 12, pp.1463-1484.
- [14] Amari S, Park H, Fukumizu K. Adaptive method of realizing natural gradient learning for multilayer perceptrons. *Neural Computation*, 2000, 12, pp.1399-1409.
- [15] Park H, Murata N, Amari S. Improving generalization performance of natural gradient learning using optimized regularization by NIC. *Neural Computation*, 2004, 16, pp.355-382.
- [16] Stefano S, Arcangeli A, Piazza F. Stability analysis of natural gradient learning rules in complete ICA: a unifying perspective. *IEEE Signal Processing Letters*, 2007, 14(1), pp.54-57.
- [17] Cardoso J F, Laheld B. Equivariant adaptive source separation. *IEEE Trans on Signal Processing*, 1996, 44(12), pp.3017-3030.
- [18] Haykin S. *Adaptive Filter Theory*, 3rd ed. Englewood Cliffs, NJ: Prentice-Hall, 1996.

- 
- [19] Pajunen P, Karhunen J. Least-squares methods for blind source separation based on nonlinear PCA. *Int. J. Neural Syst.*, 1998, 8(12), 601–612.
- [20] Zhu X L, Zhang X D. Adaptive RLS algorithm for blind source separation using a natural gradient. *IEEE Signal Processing Letter*, 2002, 9(12), pp. 432–435
- [21] Zhu X L, Zhang X D, Ding Z Z, Jia Y. Adaptive nonlinear PCA algorithms for blind source separation without prewhitening. *IEEE Transactions on circuits and systems-I*, 2006, 53(3), pp.745-753.
- [22] Comon P. Independent component analysis: a new concept? *Signal Processing*, 1994, 36(3), 287-314.
- [23] Rumelhart D E, Hinton G E, Williams. Learning representations by backpropogating errors. *Nature*, 1986, 323 pp. 533-536.
- [24] Haykin S. *Neural Networks: A Comprehensive Foundation*, Prentice-Hall: Upper Saddle River, NJ, 1999.
- [25] Douglas S C, Cichocki A. Adaptive step size techniques for decorrelation and blind source separation. *Signals, Systems & Computers Conference*, 1998, pp.1191-1195.
- [26] Jacobs, R. Increased rates of convergence through learning rate adaptation. *Neural Networks*, 1988, 1, pp.295-307.
- [27] Mathews V, Xie Z. A stochastic gradient adaptive filter with gradient adaptive step size. *IEEE Trans on Signal Processing*, 1993, 41(6), 2075-2087.
- [28] Thomas P, Allen G, August N. Step-size control in blind source separation. *International Workshop on Independent Component Analysis and Blind Source Separation*, Helsinki, Finland, 2000, 509-514.
- [29] Chambers A, Jafari G, McLaughlin S. Variable step-size EASI algorithm for sequential blind source separation. *Electronics Letters*, 2004, 40(6), pp.393-394.
- [30] Yuan L X, Wang W W, Chambers A. Variable step-size sign natural gradient algorithm for sequential blind source separation. *IEEE Signal Processing Letters*, 2005, 12(8), pp.589-592.
- [31] Douglas S C, Gupta M. Scaled natural gradient algorithm for instantaneous and convolutive blind source separation. *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2007, II, pp.637-640.
- [32] Moon T K, Stirling W. *Mathematical Methods and Algorithms for Signal Processing*. Prentice-Hall, 2000.
-

## 第四章 基于变遗忘因子递归广义特征分解的盲源分离算法

### § 4.1 引言

盲源分离在语音、无线通信, 图像, 雷达和声纳以及医学信号处理等诸多领域具有广阔的应用前景<sup>[1-4]</sup>。为便于通过数字信号处理器实现这些应用, 通常需要利用在线学习的盲源分离算法, 人们也已经针对不同的盲源分离问题提出了大量的在线学习算法<sup>[1-7]</sup>。

基于广义特征分解(GED)<sup>[8-9]</sup>的盲源分离方法<sup>[10-12]</sup>是一种简单而有效的盲分离方法, 它利用源信号的特性, 如非白性、非高斯性或非平稳性, 由混合信号的二阶或四阶统计量构造相应的矩阵束, 所构造矩阵束的广义特征向量矩阵即为待求的分离矩阵。为实现广义特征分解方法的在线处理, Rao 等人<sup>[13]</sup>提出了递归广义特征分解快速算法。基于递归广义特征分解的盲源分离方法<sup>[14-16]</sup>已经被用以改善算法的收敛速度和跟踪性能, 该文中分别推导了基于非白、非高斯和非平稳源信号的递归算法, 仿真实验表明所提递归算法具有快速和稳定的收敛性能, 且适用于信道缓慢变化, 即慢变混合矩阵的情形。但是文献[16]中没有涉及混合矩阵突变时算法跟踪性能的研究, 显然利用常数遗忘因子的递归算法不适用于跟踪突变的环境, 而利用时变遗忘因子的递归算法来跟踪时变的环境通常能在快速收敛的同时得到更为精确的稳态性能。对于基于随机梯度的盲源分离算法, 利用激活函数设计代价函数来自适应调节学习率可以加快算法的收敛并提高跟踪性能<sup>[17-18]</sup>。而对于基于广义特征分解的盲源分离方法而言, 一方面无需选择激活函数是它的优点, 因为激活函数的选择通常与信号的特性(如超高斯, 欠高斯)有关, 从而影响对应算法的普遍适用型; 另一方面则无法根据激活函数设计代价函数来自适应调节遗忘因子。此外, 在文献[16]中观测数据的协方差矩阵和它的逆矩阵的递归估计必须分别经由该文中更新公式(27)和定点算法(30)实现, 即同时使用两个独立的更新方程来估计观测数据的协方差矩阵及其逆矩阵。这种递归方式将会导致出现数值误差, 无法保证估计的协方差矩阵及其逆矩阵的乘积是单位矩阵, 在仿真实验中必须对协方差矩阵及其逆矩阵进行合适的初始化以避免算法发散。但是初始化通常需要根据所处理数据和反复实验选择, 因此文献[14-16]中的递归方式有待于进一步改进。

本章提出一种基于在线决策规则的时变遗忘因子递归广义特征分解方法实现非白源混迭的盲分离。首先给出一种新的协方差矩阵的逆矩阵和时延相关矩阵乘积的递归更新方程, 且具有较低的计算复杂度, 广义特征向量经由近似幂迭代法和压缩处理方法在线实现, 得到的在线算法可以避免同时估计协方差矩阵及其逆

矩阵。为进一步跟踪混合矩阵的突然变化,我们提出一种新颖的基于广义特征向量关于协方差矩阵正交化的在线决策规则,可以判定混合矩阵突变是否突变,当混迭矩阵被判定为发生突变时,然后利用变遗忘因子的递归广义特征分解算法改善在线盲源分离算法性能,既能获得较好的跟踪能力又能得到精确的稳态性能。

## § 4.2 基于广义特征分解的盲源分离算法

用  $M$  个传感器阵列接收  $N$  个统计独立的源信号  $\mathbf{s}(t) = [s_1(t), s_2(t), \dots, s_N(t)]^T$ , 其线性瞬时混合情形可建模为

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) \quad (4.1)$$

这里  $\mathbf{x}(t) = [x_1(t), x_2(t), \dots, x_M(t)]^T$  是阵列观测向量,  $\mathbf{A} = [\mathbf{a}_1, \dots, \mathbf{a}_N] \in \mathbb{R}^{M \times N}$  ( $M \geq N$ )

列满秩的未知混合矩阵。不失一般性,下面均考虑  $M = N$  的情形。

盲源分离的目的是用矩阵  $\mathbf{W} \in \mathbb{R}^{N \times N}$  从观测信号  $\mathbf{x}(t)$  中恢复源信号  $\mathbf{s}(t)$ , 以使输出向量

$$\mathbf{y}(t) = \mathbf{W}^T \mathbf{x}(t) \quad (4.2)$$

为源信号的拷贝。若不考虑排列和尺度的不确定性,则有  $\mathbf{W}^T = \mathbf{A}^{-1}$  是最优分离矩阵。

由文献[12]中分析可知,求分离矩阵的过程可转化为求解两个协方差矩阵的广义特征分解问题。根据源信号特性的不同(如非平稳源,非高斯源,非白源),协方差矩阵的选择有所不同。本章针对非白平稳源信号,下面给出详细的推导。

观测信号的时延分别为 0 和  $\tau$  的两个协方差矩阵为

$$\mathbf{R}^0 = E\{\mathbf{x}(t)\mathbf{x}(t)^T\} = \mathbf{A}E\{\mathbf{s}(t)\mathbf{s}(t)^T\}\mathbf{A}^T = \mathbf{A}\mathbf{A}_s^0\mathbf{A}^T \quad (4.3)$$

$$\mathbf{R}^\tau = E\{\mathbf{x}(t)\mathbf{x}(t+\tau)^T\} = \mathbf{A}E\{\mathbf{s}(t)\mathbf{s}(t+\tau)^T\}\mathbf{A}^T = \mathbf{A}\mathbf{A}_s^\tau\mathbf{A}^T \quad (4.4)$$

其中  $E\{\cdot\}$  表示数学期望。由于假设源信相互统计独立,故  $\mathbf{A}_s^0$  和  $\mathbf{A}_s^\tau$  皆为对角矩阵。

在实际计算中为满足矩阵的对称性要求,用  $\hat{\mathbf{R}}^\tau = (\mathbf{R}^\tau + (\mathbf{R}^\tau)^T)/2$  代替  $\mathbf{R}^\tau$ , 则有

$$\hat{\mathbf{R}}^\tau = \mathbf{A}\tilde{\mathbf{A}}_s^\tau\mathbf{A}^T \quad (4.5)$$

这里  $\tilde{\mathbf{A}}_s^\tau$  是对角矩阵。

用矩阵  $\mathbf{W}$  分别右乘(4.3)、(4.5)两个等式的两边,得到

$$\mathbf{R}^0\mathbf{W} = \mathbf{A}\mathbf{A}_s^0\mathbf{A}^T\mathbf{W}, \quad \hat{\mathbf{R}}^\tau\mathbf{W} = \mathbf{A}\tilde{\mathbf{A}}_s^\tau\mathbf{A}^T\mathbf{W} \quad (4.6)$$

由于  $\mathbf{A}_s^0$  是源信号的自相关矩阵,故可逆,则由(4.6)可得

$$\hat{\mathbf{R}}^T \mathbf{W} = \mathbf{R}^0 \mathbf{W} \mathbf{A} \quad (4.7)$$

其中  $\mathbf{A} = (\mathbf{A}_i^0)^{-1} \mathbf{A}_i^T$  为对角矩阵。这样构成一个广义特征方程, 即矩阵束  $(\hat{\mathbf{R}}^T, \mathbf{R}^0)$  的广义特征值为对角矩阵  $\mathbf{A}$  的对角元, 对应的广义特征向量矩阵为分离矩阵的转置矩阵。因此求分离矩阵的问题就转化为求矩阵束  $(\hat{\mathbf{R}}^T, \mathbf{R}^0)$  的广义特征向量矩阵。进一步, 式(4.7)中  $\mathbf{R}^0$  为对称正定矩阵, 故可逆, 因此式(4.7)可写为

$$\mathbf{D} \triangleq (\mathbf{R}^0)^{-1} \hat{\mathbf{R}}^T = \mathbf{W} \mathbf{A} \mathbf{W}^T \quad (4.8)$$

即矩阵  $\mathbf{D}$  的特征向量矩阵  $\mathbf{W}$  为分离矩阵的转置矩阵。求解特征值问题时, 只要特征值和其对应的特征向量在矩阵  $\mathbf{A}$  和  $\mathbf{W}$  中位置相对应, 就能确保式(4.8)成立。因此求出的特征向量矩阵  $\mathbf{W}$  存在排列和尺度的不确定性, 从而导致分离出的源信号之间也有排列和尺度的不确定性。

广义特征值分解的基本性质<sup>[8,9]</sup>如下: 设矩阵  $\mathbf{R}$  和  $\mathbf{Q}$  都是  $N \times N$  的实对称矩阵, 且  $\mathbf{R}$  是正定矩阵, 那么矩阵束  $(\mathbf{Q}, \mathbf{R})$  有实的广义特征值  $\lambda_i, i=1, \dots, N$ , 对应“ $\mathbf{R}$ -正交化”的广义特征向量

$$\begin{aligned} \mathbf{Q} \mathbf{w}_i &= \lambda_i \mathbf{R} \mathbf{w}_i \\ \mathbf{w}_i^T \mathbf{R} \mathbf{w}_j &= \delta_{ij}, \quad i, j \in \{1, \dots, N\} \end{aligned} \quad (4.9)$$

这里  $\delta_{ij}$  是 Kronecker delta 函数。易证广义特征向量也具有“ $\mathbf{Q}$ -正交化”性质:

$$\mathbf{w}_i^T \mathbf{Q} \mathbf{w}_j = \lambda_i \delta_{ij}, \quad i, j \in \{1, \dots, N\} \quad (4.10)$$

对于盲源分离问题来说, 对由观测向量的二阶和四阶统计量构成的矩阵束进行同时(联合)对角化就等价于广义特征值问题的求解, 矩阵束的广义特征向量构成的矩阵即为所求分离矩阵。

### § 4.3 递归广义特征分解的盲源分离算法

#### 4.3.1 递归更新

为满足实际应用的需要通常需对观测数据进行自适应在线处理, 以获得精确的跟踪能力。为此本章用指数加权求和代替式(4.3)、(4.4)中相关函数的数学期望, 即用观测信号的指数加权求和来估计相关函数:

$$\mathbf{R}^0(t) = \sum_{i=1}^t \beta^{t-i} \mathbf{x}(i) \mathbf{x}^T(i) \quad (4.11)$$

$$\hat{\mathbf{R}}^T(t) = \sum_{i=\tau+1}^t \beta^{t-i} (\mathbf{x}(i) \mathbf{x}^T(i-\tau) + \mathbf{x}(i-\tau) \mathbf{x}^T(i)) / 2 \quad (4.12)$$

这里  $\beta$  为遗忘因子, 通常取小于但趋于 1 的数。

把式 (4.11)、(4.12) 写成递归的形式如下:

$$\mathbf{R}^0(t) = \beta \mathbf{R}^0(t-1) + \mathbf{x}(t)\mathbf{x}^T(t) \quad (4.13)$$

$$\hat{\mathbf{R}}^r(t) = \beta \hat{\mathbf{R}}^r(t-1) + \left( \mathbf{x}(t)\mathbf{x}^T(t-\tau) + \mathbf{x}(t-\tau)\mathbf{x}^T(t) \right) / 2 \quad (4.14)$$

令  $\mathbf{P}(t) = (\mathbf{R}^0(t))^{-1}$ , 由矩阵求逆引理<sup>[19]</sup>得到  $\mathbf{P}(t)$  递归形式为

$$\mathbf{P}(t) = \frac{1}{\beta} \mathbf{P}(t-1) - \frac{1}{\beta} \mathbf{Q}(t)\mathbf{x}(t)\mathbf{x}(t)^T \mathbf{P}(t-1) \quad (4.15)$$

这里  $\mathbf{Q}(t) = \frac{\mathbf{P}(t-1)}{\beta + \mathbf{x}(t)^T \mathbf{P}(t-1) \mathbf{x}(t)}$ 。

由式(4.8)、(4.14) 和 (4.15), 可以递归地求出

$$\mathbf{D}(t) = \mathbf{D}(t-1) + \left[ \left( \mathbf{P}(t)\mathbf{x}(t)\mathbf{x}^T(t-\tau) + \mathbf{P}(t)\mathbf{x}(t-\tau)\mathbf{x}^T(t) \right) / 2 - \mathbf{Q}(t)\mathbf{x}(t)\mathbf{x}(t)^T \mathbf{D}(t-1) \right] \quad (4.16)$$

由式(4.8)可知, 矩阵  $\mathbf{D}$  的特征向量矩阵对应于分离矩阵, 这样递归求矩阵束  $(\hat{\mathbf{R}}^r(t), \mathbf{R}^0(t))$  的广义特征向量的问题就转化为递归求解矩阵  $\mathbf{D}(t)$  的特征向量。

### 4.3.2 近似幂迭代方法

现在考虑在递归更新矩阵  $\mathbf{D}(t)$  的同时求其特征向量以实现盲源分离。求给定矩阵的特征向量可根据实际问题的需要选择不同的方法。本章中对信源的平稳性假设意味着对于固定的时延  $\tau$ , 当  $t$  的值很大时  $\mathbf{Q}(t)$  是确定的矩阵。根据文献[20], 当  $t$  的值很大时协方差矩阵的逆阵  $\mathbf{P}(t)$  具有近似确定(或近近平稳)的特点。故而  $\mathbf{D}(t) = \mathbf{P}(t)\mathbf{Q}(t)$  对于固定的时延和大的  $t$  值也是近似确定的。因此为实现在线处理以及出于算法复杂度的考虑, 且避免同时用不同的更新方程估计协方差矩阵  $\mathbf{R}(t)$  及其逆矩阵  $\mathbf{P}(t)$ 。我们这里对  $\mathbf{D}(t)$  使用近似幂迭代方法, 或称为自然幂迭代方法<sup>[21]</sup>

$$\begin{cases} \mathbf{v}(t) = \mathbf{D}(t)\mathbf{w}(t-1) \\ \mathbf{w}(t) = \mathbf{v}(t) / \|\mathbf{v}(t)\|_2, \quad t = 1, 2, \dots \end{cases} \quad (4.17)$$

这里  $\|\cdot\|_2$  表示向量的 Euclidean 范数。式(4.17)表示递归估计  $\mathbf{D}(t)$  的最大特征值对应的特征向量, 即对  $\mathbf{D}(t)$  每更新一次, 同时利用幂迭代法对  $\mathbf{D}(t)$  迭代一次, 直到算法收敛。

Hua 等人<sup>[21]</sup>中证明了自然幂迭代法具有全局和指数收敛性能。这是自然幂迭代法与其它方法相比最大的优点。此外, 由于本章中矩阵  $\mathbf{D}(t)$  的特征向量需要逐个估计, 因此  $\mathbf{D}(t)$  的近似平稳特性对于有效利用自然幂迭代方法是至关重要的。

下面考虑运算量, 近似幂迭代方法(4.17)需要  $(N^2 + 2N)N$  次运算, 而在文献[16]中对非白源信号的递归盲源分离算法(记为 RBSS-NW)中的迭代规则需要  $(3N^2 + 3N + 1)N$  次运算。显然, 本文所用迭代法的运算量小于对应于 RBSS-NW 算法中的运算量, 特别是当  $N$  的取值较大时。此外如前所述, 近似幂迭代法可以避免同时估计协方差矩阵  $\mathbf{R}_i$  和它的逆阵  $\mathbf{P}_i$ 。

#### 4.3.3 压缩处理

下面对  $\mathbf{D}(t)$  进行压缩处理, 以便估计  $\mathbf{D}(t)$  的其余特征向量。

为清楚地说明压缩过程, 把  $\mathbf{R}^0(t)$ 、 $\hat{\mathbf{R}}^r(t)$  和  $\mathbf{D}(t)$  分别记为  $\mathbf{R}_1^0(t)$ 、 $\hat{\mathbf{R}}_1^r(t)$  和  $\mathbf{D}_1(t)$ ,  $\mathbf{D}_1(t)$  的最大特征值对应的特征向量记为  $\mathbf{w}_1(t)$ , 则经  $k-1$  次压缩后得到矩阵记为  $\mathbf{R}_k^0(t)$ 、 $\hat{\mathbf{R}}_k^r(t)$  和  $\mathbf{D}_k(t)$ , 其对应的特征向量记为  $\mathbf{w}_k(t)$ 。

由压缩处理规则<sup>[13]</sup>, 对于  $\hat{\mathbf{R}}_k^r(t)$ , 令

$$\hat{\mathbf{R}}_{k+1}^r(t) \triangleq \hat{\mathbf{R}}_k^r(t) - \frac{\hat{\mathbf{R}}_k^r(t) \mathbf{w}_k(t) \mathbf{w}_k^T(t) \hat{\mathbf{R}}_k^r(t)}{\mathbf{w}_k^T(t) \hat{\mathbf{R}}_k^r(t) \mathbf{w}_k(t)} \quad (4.18)$$

对于  $\mathbf{R}_k^0(t)$ , 令  $\mathbf{R}_{k+1}^0(t) = \mathbf{R}_k^0(t)$ , 则有

$$\mathbf{P}_{k+1}(t) = \mathbf{P}_k(t) \quad (4.19)$$

由式(4.8)、(4.18)和(4.19)有

$$\mathbf{D}_{k+1}(t) = \mathbf{D}_k(t) - \frac{\mathbf{D}_k(t) \mathbf{w}_k(t) \mathbf{w}_k^T(t) \hat{\mathbf{R}}_k^r(t)}{\mathbf{w}_k^T(t) \hat{\mathbf{R}}_k^r(t) \mathbf{w}_k(t)} \quad (4.20)$$

由式(4.20)可得

$$\begin{cases} \mathbf{D}_{k+1}(t) \mathbf{w}_k(t) = 0, \\ \mathbf{D}_{k+1}(t) \mathbf{w}_j(t) = \lambda_j(t) \mathbf{w}_j(t), j \neq k \end{cases}$$

从上式可知把  $\mathbf{D}_k(t)$  压缩为  $\mathbf{D}_{k+1}(t)$  的过程就是把已求出的对应于  $\mathbf{w}_k(t)$  的特征值置零, 而其余的特征值和它们对应的特征向量保持不变, 即压缩过程使得  $\mathbf{D}_k(t)$  的最大特征值为零, 除此之外  $\mathbf{D}_k(t)$  和  $\mathbf{D}_{k+1}(t)$  具有相同的非零广义特征对。这样经过对然后  $\mathbf{D}_k(t)$  压缩处理后再用近似幂迭代法求  $\mathbf{D}_{k+1}(t)$  的最大特征值对应的特征向量, 依次类推, 经  $n-1$  步后就可以依次估计出  $\mathbf{D}(t)$  所有特征向量。

组合式(4.16)、(4.17)和(4.24)即可得到一种新的基于递归广义特征分解的盲源分离算法,该方法可视为 RBSS-NW 算法的改进算法。为方便且与 RBSS-NW 算法区别,上述利用常数遗忘因子的本节所提基于递归广义特征分解的盲源分离算法简记为 CFF-RGED。显然, CFF-RGED 方法的收敛速率取决于近似幂迭代方法。

#### § 4.4 混合矩阵突变时基于时变遗忘因子的递归盲源分离算法

在时变环境中合理选择遗忘因子对于递归算法的收敛性能是至关重要的,常数遗忘因子限制了算法的收敛性能,且跟踪性能较差。通常需要利用有效代价函数自适应调节遗忘因子。下面我们提出一种在线决策规则来跟踪混合矩阵的突然变化,同时结合时变遗忘因子来改善第4.3节中递归算法的跟踪性能。

##### 4.4.1 时变遗忘因子

当混合矩阵不发生变化时,如下的时变遗忘因子<sup>[22]</sup>

$$\beta(t) = \begin{cases} 0.95 + 10^{-4}(t-1), & 1 \leq t \leq L \\ \beta_{\max}, & t > L \end{cases} \quad (4.21)$$

对于获得快速收敛性能和好的稳态精确性是一个简单而有效的选择,这里  $0.95 < \beta_{\max} \leq 1$ ,  $L = (\beta_{\max} - 0.95)/10^{-4}$ 。但是这样的时变遗忘因子在  $\beta(t)$  已经增加到  $\beta_{\max}$  后,混合矩阵发生突变时就相当于常数遗忘因子,跟踪性能较差。事实上,在盲源分离问题中,发生突变的时刻通常是未知的,因此无法直接利用该遗忘因子改善算法的跟踪性能。

为验证本节下面提出的时变遗忘因子递归盲源分离算法在混合矩阵发生突变时的跟踪性能。这里把混合矩阵发生突变的时刻  $t'$  已知时,上节所提基于递归广义特征分解的盲源分离方法结合式(4.21)中的时变遗忘因子作为一个参考算法,记为 VFF-RGED-Ref。这样在 VFF-RGED-Ref 算法中,对应于式(4.21)的遗忘因子可以表示成

$$\beta_t = \begin{cases} 0.95 + 10^{-4}(t-t'-1), & t'+1 \leq t \leq t'+L \\ \beta_{\max}, & t-t' > L \end{cases} \quad (4.22)$$

其中  $t'$  表示混合矩阵发生突变的时刻。本节假设混合矩阵发生相邻两次突变的时间间隔  $L'$  大于式(4.22)中  $L$  的值。因为即使  $\beta_{\max}$  取最大值,即  $\beta_{\max} = 1$ ,对应于  $L$  的最大值为500,因此上述假设可以视为一个合理的假设。而在所有时刻混合矩阵都发

生剧烈动态变化的情况则不在本章考虑范围之内。

#### 4.4.2 在线决策规则

在盲源分离问题中,发生突变的时刻 $t'$ 通常是未知的,而上节所提基于递归广义特征分解的盲源分离方法并没有利用激活函数,因此也不能象第三章中那样利用激活函数控制时变参数。为此下面利用广义特征向量本身的“ $\mathbf{Q}$ -正交化”特性来提出在线决策规则以跟踪混合矩阵的突然变化,需要指出的是这里无需使用混合矩阵的额外的先验信息。需要注意的是这里没有利用“ $\mathbf{R}$ -正交化”特性是为了避免同时更新协方差矩阵 $\mathbf{R}(t)$ 和它的逆阵 $\mathbf{P}(t)$ 。

为减小计算量,且提高决策规则对于混合矩阵变化的敏感性。下面只利用主广义特征向量 $\mathbf{w}_1$ 和 $\mathbf{w}_2$ 来设计代价函数。广义特征向量 $\mathbf{w}_1$ 和 $\mathbf{w}_2$ 之间在不同时刻的“ $\mathbf{Q}$ -正交化”可用如下表达式

$$c_p(t) = \left( \left| \mathbf{w}_2^T(t-p)\mathbf{Q}(t)\mathbf{w}_1(t) \right| + \left| \mathbf{w}_1^T(t-p)\mathbf{Q}(t)\mathbf{w}_2(t) \right| \right) / 2, p=1, \dots, P \quad (4.23)$$

来衡量,这里 $\mathbf{w}_k(t)$ 和 $\mathbf{w}_k(t-p)$ ,  $k=1,2, p=1, \dots, P$ 分别表示主广义特征向量 $\mathbf{w}_k$ 的当前时刻和过去时刻的估计量。用向量的形式可表示为

$$\mathbf{c}(t) = [c_1(t), c_2(t), \dots, c_p(t)] \quad (4.24)$$

由于在式(4.16)中 $\mathbf{D}(t)$ 的值随随着每次样本更新的变化而变化,且在式(4.17)中用近似幂迭代方法来估计 $\mathbf{D}(t)$ 的特征向量,因此即使混合矩阵不变且已获得良好的分离性能的情况下,相邻时刻的估计的特征向量间也有差异。因此, $\mathbf{c}(t)$ 的分量的最大值和最小值之间的差异也是明显的。

另一方面,当混合矩阵在某一时刻 $t'$ 发生突然变化后,由于估计的特征向量不再是混合矩阵没发生变化时矩阵束的好的估计量,因此广义特征向量的当前时刻估计量 $\mathbf{w}_i(t')$ 和过去时刻估计量 $\mathbf{w}_j(t'-p)$  ( $i \neq j, p=1, \dots, P$ )之间的“ $\mathbf{Q}$ -正交化”特性就会失效,仿真结果表明当混合矩阵发生突变时 $\mathbf{c}(t)$ 的分量的最大值和最小值之间的差异变得非常之小。换句话说, $\mathbf{c}(t)$ 的分量的最大值和最小值的差异大小可以用来判定混合矩阵是否发生突变,因此,定义如下比值

$$CF_t = \frac{\max(\mathbf{c}(t))}{\max(\mathbf{c}(t)) + \min(\mathbf{c}(t))} \quad (4.25)$$

作为代价函数来表明混合矩阵的变化情况,这里 $\max(\mathbf{c}(t))$ 和 $\min(\mathbf{c}(t))$ 分别表示 $\mathbf{c}(t)$ 的分量的最大值和最小值,显然,当 $\max(\mathbf{c}(t))$ 和 $\min(\mathbf{c}(t))$ 的差值很小时, $CF_t$ 接近于0.5。如仿真实验中图4.3所示,当混合矩阵不发生突变时 $CF_t$ 的在区间

$[0.8, 1)$  内, 而当混合矩阵发生突变时  $CF_t$  的在区间  $(0.5, 0.6]$  内。

下面用假设  $H_p$  和  $H_a$  分别表示混合矩阵发生和没有发生突然变化。设  $\gamma$  为一适当的门限值, 则在线决策规则为

$$CF_t \underset{H_p}{\overset{H_a}{\gtrless}} \gamma \quad (4.26)$$

即当  $CF_t < \gamma$  时, 判定混合矩阵发生突然变化, 否则判定混合矩阵没有发生突变。

一旦混合矩阵在时刻  $t'$  被判定为发生突变, 就可从时刻  $t'+1$  开始在上节所提基于递归广义特征分解的盲源分离算法中用式(4.22)中的时变遗忘因子来实施跟踪。这里把上节所提基于递归广义特征分解的盲源分离法, 时变遗忘因子(4.22)和在线决策规则(4.26)相结合得到的在线跟踪算法记为VFF-RGED-DR, 它可用以改善CFF-RGED算法的跟踪性能。需要注意的是在VFF-RGED-DR算法中, 混合矩阵发生突变时刻  $t'$  是经过在线决策规则(4.26)判定的, 由于样本更新的累积效应, 所判定的混合矩阵突变时刻  $t'$  可能比混合矩阵真正发生突变的时刻有所延迟, 仿真实验中图4.6也证实了这一点。

在  $L' > L$  的假设下, 一旦  $H_p$  在时刻  $t'$  为真, 即混合矩阵被判定为发生突变, 那么在区间  $(t'+1, t'+L]$  中混合矩阵则不会发生再次突变, 即在区间  $(t'+1, t'+L]$  中  $H_a$  为真, 因此在  $(t'+1, t'+L]$  中无需再计算  $CF_t$  的值。只在  $(t'+1, t'+L]$  中用时变遗忘因子(4.22)来跟踪发生突变后的混合矩阵。当  $t = t' + L + 1$  时, 重新开始计算  $CF_t$  的值用以判定混合矩阵是否再次发生突变。

此外, 为了评估所提在线决策规则(4.26)的性能, 下面引入正确判决混合矩阵突变的概率

$$P_\alpha = \frac{\text{正确判决的次数}}{\text{实验总次数}} \quad (4.27)$$

当  $P_\alpha$  的值越大, 表明在线决策规则(4.26)越有效。

关于向量  $\mathbf{c}(t)$ , 参数  $\beta_{\max}$  和  $CF_t$ , 有如下注释:

**注释一:** 本节中  $\mathbf{c}(t)$  的取值对于判定混合矩阵是否发生突变是至关重要的。由于压缩处理所致的误差累积使得所提RGED算法中估计的特征向量的精确性依次降低。与利用所有特征向量设计代价函数相比, 我们这里只利用主特征向量  $\mathbf{w}_1(t)$  和  $\mathbf{w}_2(t)$  来设计代价函数对混合矩阵的变化情况更为敏感, 从而具有更为有效的跟踪能力, 此外也具有较小的计算复杂度。

**注释二:** 在时变遗忘因子(4.22)中,  $\beta_{\max}$  的值对于VFF-RGED-DR算法的稳态分离性能有很大影响。 $\beta_{\max}$  的值越大, 则稳态精确性越好, 然而, 当  $\beta_{\max}$  的值太大,

如  $\beta_{\max}=1$  时, VFF-RGED-DR 算法虽然能得到较好的分离性能, 但是如果混合矩阵发生缓慢变化, 此时, VFF-RGED-DR 算法的跟踪性能则会变差。因此为使所提算法更具普遍性, 应使  $\beta_{\max}$  的值小于 1。

**注释三:** 我们使用如下方法在仿真实验中利用式(4.25)和式(4.26)判别混合矩阵是否发生突变。在区间  $[1, L]$  中令  $CF_t = \eta_{\max}$ , 这里  $\eta_{\max}$  是大于  $\gamma$  的常数。当  $t > L$  时, 开始计算根据式(4.25)计算  $CF_t$  的值并利用式(4.26)作出判决, 一旦  $H_p$  在某一时刻  $t'$  为真, 在  $t'+1$  开始利用式(4.22)跟踪混合矩阵的突变, 又因为在混合矩阵发生突变后, 在  $L' > L$  的假设下在区间  $(t'+1, t'+L)$  内  $H_o$  为真, 因此可以停止计算  $CF_t$ , 可将其设为  $CF_t = \eta_{\max}$ , 这里  $\eta_{\max}$  为某一接近于 1 的常数。当  $t = t' + L$  时, 重新开始式(4.25)计算  $CF_t$  的值并利用式(4.26)作出判决。这样在递归过程中  $CF_t$  取值为如下分段函数

$$CF_t = \begin{cases} \frac{\max(c(t))}{\max(c(t)) + \min(c(t))}, & t \leq t' \text{ 或 } t > t' + L \\ \eta_{\max}, & 1 \leq t \leq L \text{ 和 } t' < t \leq t' + L \end{cases} \quad (4.28)$$

## § 4.5 仿真实验

在本节中, 我们用仿真实验来验证所提算法的有效性。我们考虑分离ICALAB 工具箱中<sup>[1]</sup>中ABio7.mat文件包含的四个典型的生物信号  $s_1, \dots, s_4$ , 信号曲线如图4.1所示。在下面的不同算法中, 取  $\tau=1$ 。所提算法的分离性能由平均信干比 (SIR)(dB)<sup>[20, 30]</sup>来衡量, 这里SIR的值越大, 说明算法的性能越好。

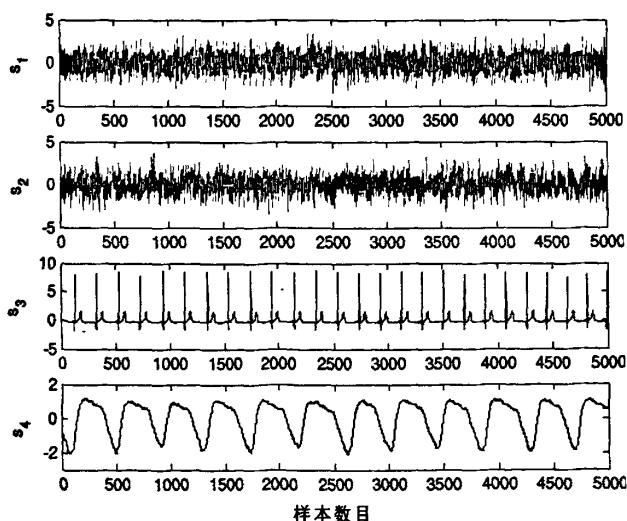


图 4.1 仿真实验中所用源信号波形

**实验一：**为便于比较，我们同时运行文献[17]中的RBSS-NW算法和我们提出的CFF-RGED算法和基于变遗忘因子(12)的RGED算法。混合矩阵 $A$ 的各元素为在 $[-1,1]$ 区间服从均匀分布的随机数。三个算法的初始分离矩阵的列向量均为由随机产生的单位列向量，遗忘因子分别等于0.996和1。矩阵 $Q$ 和 $P$ 被分别初始化为 $\delta I$ 和 $\delta^{-1}I$ 。在RBSS-NW算法中 $R$ 的初始化矩阵为 $\delta I$ 。令 $\delta = 10^{-4}$ ，对三种算法均由1000次蒙特卡罗实验得到算法的平均性能。

图4.2表明这两种算法对于相同的遗忘因子有几乎相同的性能曲线。此外，在初始迭代阶段，对于同一算法即使遗忘因子不同也可得到大致相同的收敛速率。这是因为本文所提算法中近似幂迭代方法和RBSS-NW算法中定点算法对于收敛速度有重要作用。 $\beta=1$ 能得到比 $\beta=0.996$ 更好的稳态精确性能。然而，由于用两个独立的方程来更新 $R(t)$ 和 $P(t)$ ，如果取不到合适的初始化参数 $\delta$ ，RBSS-NW算法有潜在的失败风险。在我们的试验中如果令 $\delta=1$ ，则每运行1000次RBSS-NW算法，有大致4-5次算法会出现发散的情况，而我们所提的算法则不会出现这种情况，这是由于本文算法避免了同时更新 $R(t)$ 和 $P(t)$ ，且更新公式相对简单，计算复杂度低。因此本文所提CFF-RGED算法进一步改进了RBSS-NW算法。

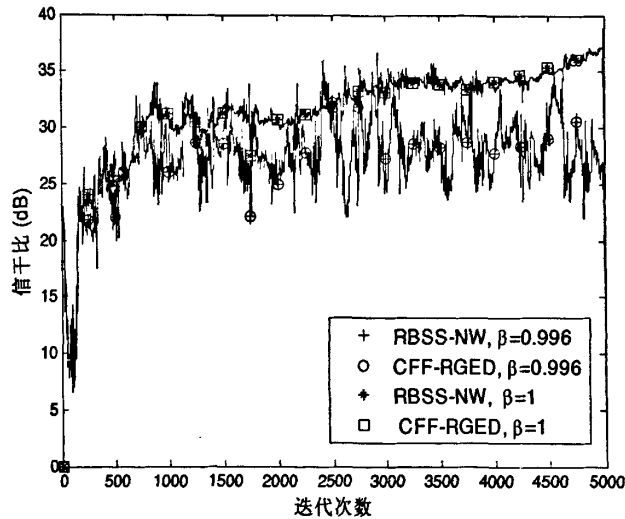
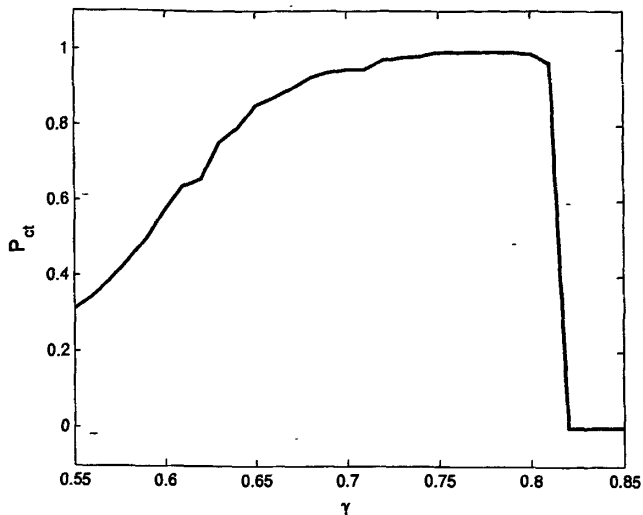
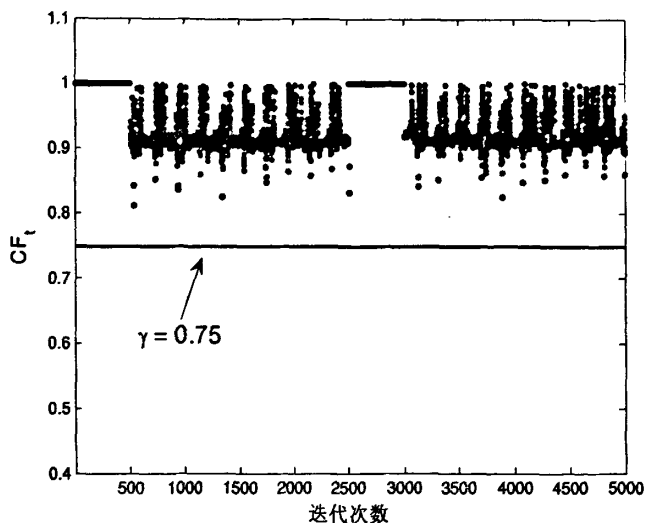


图 4.2 两种递归算法对于不同常数遗忘因子 1000 次平均的信干比均值图

**实验二：**下面我们考虑盲源分离问题中混合矩阵发生的突然变化的情况。混合矩阵 $A$ 的各元素为在 $[-1,1]$ 区间服从均匀分布的随机数且每隔2500个样本发生一次突然变化。用判决混合矩阵突变的正确的概率 $P_d$ 来评估所提在线决策规则(4.26)的性能。VFF-RGED-DR算法中矩阵 $Q$ ， $P$ 及初始分离矩阵的初始化与实验一中的CFF-RGED算法中参数的初始化相同。在VFF-RGED-DR算法中，取 $P=10$ ，

图4.3 混合矩阵发生突变时正确判决概率  $P_\alpha$ -门限  $\gamma$  性能指标图图4.4 混合矩阵发生突变的情况下  $CF_i$  在1000次实验中的值

$\beta_{\max}=1$ ,  $\eta_{\max}=1$ 。VFF-RGED-DR算法对不同的门限  $\gamma$  运行1000次蒙特卡罗实验。

图4.3中画出了对应于门限值  $\gamma$  的概率值  $P_\alpha$ ，它表明在区间  $[0.7, 0.81]$  正确判决混合矩阵发生突变的概率  $P_\alpha \geq 0.95$ 。图4.4给出了当  $\gamma = 0.75$  时一次实验在线估计的  $CF_i$  的值，除了算法初始收敛阶段和混合矩阵发生突变后的时间间隔  $L$  外，混合矩阵不发生突变时  $CF_i$  的值都大于0.8而在混合矩阵发生突变后  $CF_i$  随之变为0.58。虽然每次实验的结果不尽相同，但可以说明式(4.24)中  $c(t)$  的分量最大值与最小值在混合矩阵发生突变时近似相同，在混合矩阵不发生突变时差异较大。从而表明所

提在线决策规则的有效性。图4.3和图4.4表明本文所提出的在线决策准则能够有效跟踪混合矩阵的突然变化。

**实验三：**为进一步表明所提算法的跟踪能力，我们比较VFF-RGED-DR算法、VFF-RGED-Ref算法和CFF-RGED算法在混合矩阵发生突变时的跟踪能力。对于CFF-RGED算法，常数遗忘因子 $\beta$ 分别取0.996和1。其它参数同实验一。VFF-RGED-Ref算法中的参数除变遗忘因子外其余参数同实验一中的CFF-RGED算法参数的取值。在变遗忘因子(式(4.22))中， $\beta_{\max}=1$ 。对于VFF-RGED-DR算法，参数同实验二中所用的参数。对三个算法均运行1000次蒙特卡罗实验，图4.5给出了当混合矩阵发生突变时三种算法的平均性能曲线，表明CFF-RGED算法在 $\beta=1$ 没有跟踪能力，而在 $\beta=0.996$ 时具有跟踪能力，但跟踪能力依然较差，而VFF-RGED-DR算法由于能检测到混合矩阵的突变时刻，从而可以利用变遗忘因子实现快速跟踪并可得到较好的稳态精确性。图4.5和图4.6表明VFF-RGED-DR的算法性能趋于VFF-RGED-Ref算法的性能，两者性能差异如图4.6所示，在混合矩阵发生突变后的短时间内，VFF-RGED-Ref算法的性能优于VFF-RGED-DR算法，这是由于样本更新的累积效应，所判定的混合矩阵突变时刻 $t'$ 可能比混合矩阵真正发生突变的时刻有所延迟，因此性能分离相应略差，之后则趋于相同。图4.5和图4.6表明VFF-RGED-DR算法适合于盲源分离问题中混合矩阵发生突然变化的情况，并对需要在线处理的盲分离任务有潜在的应用。混合矩阵发生突变时在线提取的信号如图4.7所示。

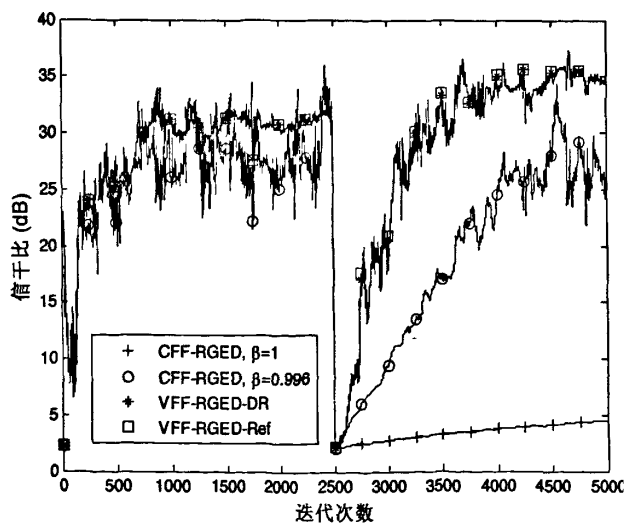


图4.5 三种递归算法在混合矩阵突变的情况下1000次平均的信干比均值图

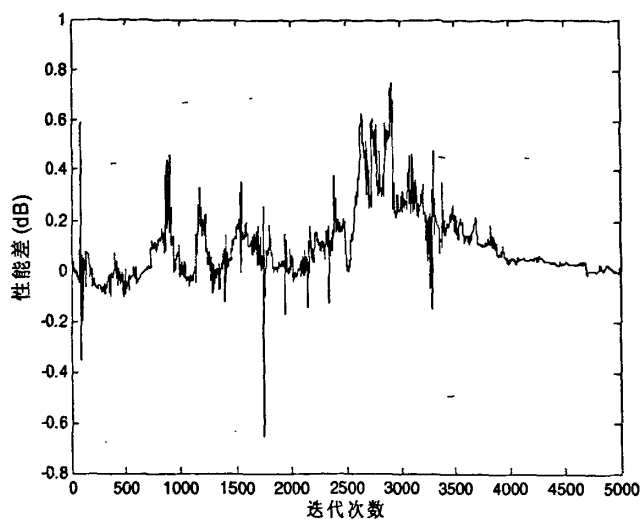


图4.6 VFF-RGED-DR 算法与VFF-RGED-Ref 算法之间的信干比性能差

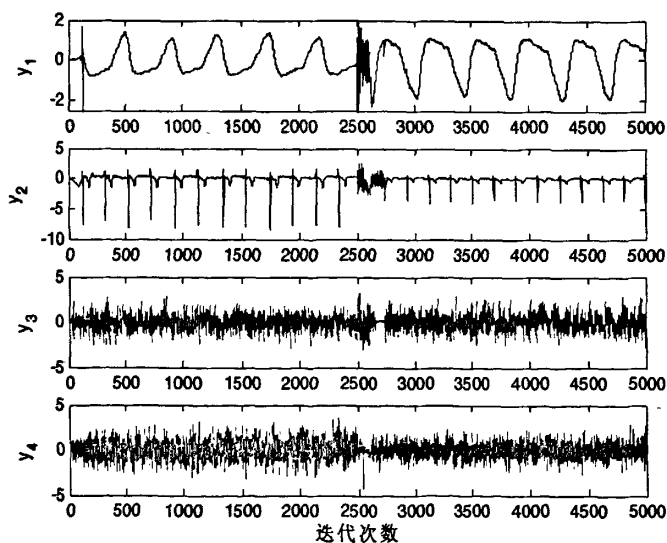


图 4.7 混合矩阵突变的情况下在线提取出的信号

## § 4.6 本章小结

针对盲源分离问题中混合矩阵发生突变的情况,本章提出一种基于时变遗忘因子递归广义特征分解的盲源分离算法。已有递归广义特征分解算法且在递归过程中利用不同的公式同时在线估计协方差矩阵及其逆矩阵,容易导致数值误差,

算法的性能对初始分离矩阵的取值比较敏感,易导致算法发散,缺乏应用的普遍性。此外,已有递归广义特征分解算法基于常数遗忘因子,在混合矩阵发生突变时跟踪能力较差,有待进一步改进。

本章首先给出一种协方差矩阵的逆矩阵和互相关矩阵乘积的递归更新方程,且具有较低的计算复杂度,广义特征向量经由近似幂迭代法和压缩映象方法在线实现,得到的在线算法可以避免同时估计协方差矩阵及其逆矩阵。为跟踪混合矩阵的突然变化,我们提出一种新颖的基于广义特征向量的“ $Q$ -正交化”特性的在线决策规则,可以判别混合矩阵是否发生突变,从而利用变遗忘因子的递归广义特征分解算法改善在线盲源分离算法性能,既能获得较好的跟踪能力又能得到精确的稳态性能。

### 参考文献

- [1] Haykin S. Unsupervised adaptive filtering. Volume1: Blind source separation. New York: Wiley, 2000.
- [2] Hyvarinen A, Karhunen J, Oja E. Independent Component Analysis. New York, Wiley, 2001.
- [3] Cichocki A, Amari S. Adaptive Blind Signal and Image Processing. New York: Wiley, 2003.
- [4] Makino S, Lee T W, Sawada H. Blind speech separation, Springer, 2007.
- [5] Bell A, Sejnowski T. An information-maximization approach to blind separation and blind deconvolution, Neural Computation, 1995, 7, pp. 1129–1159.
- [6] Amari S, Cichocki A, Yang H. A new learning algorithm for blind signal separation. Advances in Neural Information Processing Systems. Cambridge, MA: MIT Press, 1996, 8, pp. 757-763.
- [7] Cardoso J F, Laheld B. Equivariant adaptive source separation. IEEE Trans on Signal Processing, 1996, 44(12), pp.3017-3030.
- [8] Horn R A, Johnson C R. Matrix Analysis. Cambridge: Cambridge University Press, 1986.
- [9] Golub G H, Loan C. Matrix Computations. John Hopkins University Press, Baltimore, MD, 1991.
- [10] Chang, Ding Z, Yau S, Chan F. A matrix-pencil approach to blind separation of nonwhite signals in white noise. in Proc. IEEE Int. Conf. Acoust., Speech, Signal Process., Seattle, WA, May 1998, 4, pp. 2485-2488.

- [11] Tome A M. Blind source separation using a matrix pencil. in Proc. IEEE-INNS-ENNS Int. Joint Conf. Neural Netw., Como, Italy, Jul. 2000, 3, pp. 339-342.
- [12] Parra L, Sajda P. Blind Source Separation via Generalized Eigenvalue Decomposition. Journal of Machine Learning Research, 2003, 4(7-8), pp.1261-1269.
- [13] Rao Y N, Principe J C, Wong T F. Fast RLS-like algorithm for generalized eigendecomposition and its applications. Journal of VLSI Signal Processing, 2004, 37, pp. 333-344.
- [14] Ozertem U, Erdogmus D. Recursive generalized eigendecomposition for independent component analysis. Proceedings of ICA 2006, Berlin Heidelberg: Springer-Verlag, 2006, 198-205.
- [15] Pokharel P P, Ozertem U, Erdogmus D, Principe Jose C. Recursive complex blind source separation via eigendecomposition of cumulant matrices. in Proc. IEEE Int. Conf. Acoust., Speech, Signal Process., Seattle, WA, May 2007, II, pp. 645-648.
- [16] Pokharel P P, Ozertem U, Erdogmus D, Principe J C. Recursive complex BSS via generalized eigendecomposition and application in image rejection for BPSK. Signal Processing, 2008, 88(4), pp. 1368-1381.
- [17] Douglas S C, Cichocki A. Adaptive step size techniques for decorrelation and blind source separation. Signals, Systems & Computers Conference, 1998, pp.1191-1195.
- [18] Thomas, P, Allen G, August N. Step-size control in blind source separation. International Workshop on Independent Component Analysis and Blind Source Separation, Helsinki, Finland, 2000, pp. 509-514.
- [19] Haykin S. Adaptive Filter Theory, 3rd ed. Englewood Cliffs, NJ: Prentice-Hall, 1996.
- [20] Eleftheriou E., Falconer D D. Tracking properties and steady-state performance of RLS adaptive filter algorithms. IEEE Trans. Acoustic Speech, Signal Processing, 1986, 34(10), pp. 1097-1109.
- [21] Hua Y, Xiang Y, Chen T, Abed-Meraim K, Miao Y. A new look at the power method for fast subspace tracking. Digital Signal Processing, 1999, 9(10), pp.297-314.

- [22] Lv Q, Zhang X D, Jia Y.. Kalman filtering algorithm for blind source separation. IEEE International Conference Acoustics, Speech, Signal Processing, 2005, pp. 257–260.
- [23] Hild K, Erdogmus D and Principe J C. Blind source separation using Renyi's mutual information. IEEE Signal Processing Letter, 2001, 8(6), pp. 174–176.

## 第五章 数目未知的语音信号瞬时混合的

### 盲分离与信源数目检测算法

#### § 5.1 引言

对于盲源分离问题及其在不同领域的应用,人们已提出了很多算法<sup>[1-2]</sup>。在语音处理中盲源分离技术的一个典型应用是“鸡尾酒会”问题,即在多人同时说话的情况下,把混杂在其中的一个或几个人的声音提取或分离出来。在实际环境中,由于混响的存在,接收到的信号通常为源信号的延迟和滤波信号。这种情况下的混合方式可用卷积混合近模型近似。卷积混合盲语音分离问题已经引起了广泛的兴趣<sup>[3-4]</sup>,主要处理方法可分为时域方法和频域方法,但是时域方法通常需要很大的计算量<sup>[5]</sup>,频域方法则存在频率置换问题<sup>[6-10]</sup>。由于麦克风阵列中不同麦克风接收信号时存在时延,瞬时混合的盲源分离方法通常不适用于分离麦克风阵列接收信号。但是当麦克风阵列中的麦克风位置近似相同,且对不同的方向有不同的增益,源信号的卷积混合可用瞬时模型来近似<sup>[11-13]</sup>。因此对于瞬时混合的如语音信号的盲分离具有研究意义和实用价值。

源信号的非平稳特性对于解决盲源分离问题是非常有用的信息, Matsuoka 等人<sup>[14]</sup>基于神经网络提出了适用于非平稳过程的盲分离方法。利用语音信号的短时平稳特性, Choi 等人<sup>[15-16]</sup>对经过预白化处理的观测信号通过不同的时间滑窗得到多个不同时刻、不同时延的相关矩阵,这些相关矩阵利用近似联合对角化方法<sup>[17]</sup>估计混合矩阵。Pham 等人<sup>[18]</sup>利用多个正定矩阵的近似联合对角化方法提出了非平稳源瞬时混合的盲分离方法。Ying 等人<sup>[19]</sup>利用非白和非平稳特性,提出了基于解相关的代价函数,并分别给出了批处理算法和在线算法。此外基于广义特征分解的盲源分离方法<sup>[20-22]</sup>是一种有效的盲分离方法,它利用源信号的特性,如非平稳性,非白或非高斯性,由观测数据构造不同的矩阵束,所构造矩阵束的广义特征矩阵即为所求分离矩阵。对于非平稳信号,其相关矩阵是时变的,因此对观测信号的在线处理可以更好的反映其时变特性。

已有研究大多假设源信号数目是预先知道的,从而采用的分离矩阵行数等于信源数目,而实际中源信号的数目往往是未知的,甚至是随实际情况动态变化的,由于欠定盲信号分离只能成功分离部分源信号<sup>[23-24]</sup>,因此我们本章主要研究语音信号数目未知的超定盲信号分离。Ye 等人<sup>[25]</sup>研究了信源数目未知与变化时的盲信号分离方法。但是针对数目未知的语音信号瞬时混合盲分离的研究尚不多见。

对于信源数目已知时语音信号瞬时混合盲分离的算法已有较为成熟的研究,

但大都不能适用于信源数目未知的盲分离问题,无法分离出所有的源信号,虽然盲提取算法可以逐个提取源信号,并通过一定的判决准则,确定是否把所有的信号提取出来,但是大多数的盲提取算法在提取出一个源信号后需要对数据作压缩处理,把已提取出的信号的信息从混合信号或所用相关矩阵中消去,这样逐个提取的过程会产生误差累积,导致所提取信号的质量依次降低,在信源数目较多的情况下甚至无法提取出个别信源。

本章充分利用人类发音的特点和语音信号的非平稳特性,对于数目未知的语音信号瞬时混的盲分离问题,通过对在线估计的广义特征向量进行聚类处理,在实现信号数目检测的同时完成信号的盲分离,并可克服盲提取问题中先提取出的信源性能好而后提取出的信源性能差的缺点。虽然 Pokharel 等人<sup>[26]</sup>提出了基于递归广义特征分解的非平稳源的盲分离方法,并以语音信号为例研究了在信道(即混合矩阵)慢变情况下的盲分离问题,但该文在仿真试验中只用了 2000 个样本,无法真实体现语音信号的非平稳特性,以语音信号的采样率为 16 KHz 为例,如果录制一段时间长度为 3 秒的对话,则相应应有 48000 个样本。由于人类对话交流的特点,说话人在说一句话时在词或词组间可能会有停顿,也可能在一定的时间间隔内保持沉默,那么在多个说话人同时说话的情况下(经典的“鸡尾酒会问题”),当构成矩阵束的两个协方差矩阵的时延较大时,在线估计的对应的最大广义特征值的广义特征向量,是随着实际语音环境的变化而变化,并不对应于特定的说话人,而是对应不同说话人,这样提取出的信号中不同的时间间隔内对应着不同的说话人,如果实际语音环境的动态变化较大,则分离出信号听起来只是无法辨别的“噪音”。

但也正是语音信号的上述特性给我们解决数目未知的语音信号盲分离问题提供了一种新的解决思路,由于所用近似幂迭代方法的快速收敛特性,在线估计的最大广义特征值对应的广义特征向量在某些时间间隔内能得到很好的估计性能,且对应于某一个说话人。在线估计的最大特征值对应的特征向量在某些时间间隔内能得到很好的估计性能,这样在某个区间内估计的特征向量就可以有效的提取出一个源信号。那么对于每个说话人都会在某个或某些时间间隔内有在线估计的广义特征向量与之相对应,这样如果我们能够得到所有不同的广义特征向量,则可以确定信源的数目并利用得到的广义特征向量提取出所有的源信号。为此本节利用向量的相似度和广义特征分解的特性提出有效的代价函数,并利用提取“高分离度区间集”和对“高分离度区间集”中估计的广义特征向量迭代聚类的方法,在完成信源数目的检测的同时实现源信号的盲分离。

## § 5.2 基于广义特征分解的非平稳信号的盲提取算法

### 5.2.1 非平稳信号构成矩阵束的广义特征分解

用  $M$  个传感器阵列接收  $N$  个统计独立的源信号  $\mathbf{s}(t) = [s_1(t), s_2(t), \dots, s_N(t)]^T$ ,

其线性瞬时混合情形可建模为

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) \quad (5.1)$$

这里  $\mathbf{x}(t) = [x_1(t), x_2(t), \dots, x_M(t)]^T$  是阵列观测向量,  $\mathbf{A} = [\mathbf{a}_1, \dots, \mathbf{a}_N] \in \mathbb{R}^{M \times N} (M \geq N)$

列满秩的未知混合矩阵。

盲源分离的目的是用分离矩阵从观测信号  $\mathbf{x}(t)$  中恢复源信号  $\mathbf{s}(t)$ , 以使输出向量

$$\mathbf{y}(t) = \mathbf{W}^T \mathbf{x}(t) \quad (5.2)$$

作为源信号的拷贝, 这里  $\mathbf{W} \in \mathbb{R}^{N \times M}$ 。若不考虑排列和尺度的不确定性, 则有  $\mathbf{W}^T = \mathbf{A}^{-1}$  是最优分离矩阵。

由文献[21]中分析可知, 求分离矩阵的过程可转化为求解两个协方差矩阵的广义特征分解问题。根据源信号特性的不同(如非平稳源, 非高斯源, 非白源), 协方差矩阵的选择有所不同,

本章考虑的语音信号为非平稳信号, 则观测信号的协方差矩阵

$$\mathbf{R}_x(t) = E\{\mathbf{x}(t)\mathbf{x}(t)^T\} = \mathbf{A}E\{\mathbf{s}(t)\mathbf{s}(t)^T\}\mathbf{A}^T = \mathbf{A}\mathbf{R}_s(t)\mathbf{A}^T \quad (5.3)$$

随时间  $t$  的变化而变化, 由假设源信号相互统计独立, 故  $\mathbf{R}_s(t)$  为对角矩阵。

我们选择两个不同时刻  $t_1, t_2$ , 则

$$\begin{aligned} \mathbf{R}_x(t_1) &= \mathbf{A}\mathbf{R}_s(t_1)\mathbf{A}^T \\ \mathbf{R}_x(t_2) &= \mathbf{A}\mathbf{R}_s(t_2)\mathbf{A}^T \end{aligned} \quad (5.4)$$

用矩阵  $\mathbf{W}$  分别右乘式(5.4)中两个等式的两边, 得到

$$\begin{aligned} \mathbf{R}_x(t_1)\mathbf{W} &= \mathbf{A}\mathbf{R}_s(t_1)\mathbf{A}^T\mathbf{W} \\ \mathbf{R}_x(t_2)\mathbf{W} &= \mathbf{A}\mathbf{R}_s(t_2)\mathbf{A}^T\mathbf{W} \end{aligned} \quad (5.5)$$

由于  $\mathbf{R}_s(t_2)$  是源信号的自相关函数, 故其对角元均为非零元素, 则由式(5.5)可得

$$\mathbf{R}_x(t_1)\mathbf{W} = \mathbf{R}_x(t_2)\mathbf{W}\mathbf{A} \quad (5.6)$$

其中  $\mathbf{A} = \mathbf{R}_s(t_2)^{-1}\mathbf{R}_s(t_1)$  为对角矩阵。这样构成一个广义特征方程, 即矩阵束  $(\mathbf{R}_x(t_1), \mathbf{R}_x(t_2))$  的广义特征值为  $\mathbf{A}$  的对角元素,  $\mathbf{W}$  的列为对应的广义特征向量。因此求分离矩阵的问题就转化为求矩阵束  $(\mathbf{R}_x(t_1), \mathbf{R}_x(t_2))$  的广义特征矩阵, 进一步式(5.6)中  $\mathbf{R}_x(t_1), \mathbf{R}_x(t_2)$  中均为对称正定矩阵, 故均为可逆矩阵, 因此式(5.6)可写为

$$\mathbf{D} \triangleq \mathbf{R}_x(t_2)^{-1} \mathbf{R}_x(t_1) = \mathbf{W} \mathbf{A} \mathbf{W}^T \quad (5.7)$$

即需求矩阵  $\mathbf{D}$  的特征向量矩阵  $\mathbf{W}$ 。求解特征值问题时, 只要特征值和和其对应的特征向量在矩阵  $\mathbf{A}$  和  $\mathbf{W}$  中位置相对应, 就能确保式(5.7)成立, 因此求出的特征向量矩阵  $\mathbf{W}$  存在排列和尺度的不确定性。从而导致分离出的源信号之间也有排列和尺度的不确定性。

## 5.2.2 基于递归广义特征分解的盲提取算法

### 5.2.2.1 协方差矩阵的递归实现

由于语音信号为非平稳信号, 其统计特性随时间的变化而动态的变化。因此通常需对观测数据进行自适应处理, 以获得更精确的跟踪能力以满足由于实际应用的需要。为实现递归自适应运算, 用指数加权求和代替式(5.4)中相关函数的数学期望, 得到观测信号的指数加权的相关函数:

$$\begin{aligned} \mathbf{R}_x(t_1) &= \sum_{i=1}^{t_1} \lambda^{t_1-i} \mathbf{x}(i) \mathbf{x}^T(i) \\ \mathbf{R}_x(t_2) &= \sum_{i=1}^{t_2} \lambda^{t_2-i} \mathbf{x}(i) \mathbf{x}^T(i) \end{aligned} \quad (5.8)$$

这里  $\lambda$  为遗忘因子, 通常取小于但趋于 1 的数。

把式(5.8)写成递归的形式

$$\begin{aligned} \mathbf{R}_x(t_1) &= \lambda \mathbf{R}_x(t_1-1) + \mathbf{x}(t_1) \mathbf{x}^T(t_1) \\ \mathbf{R}_x(t_2) &= \lambda \mathbf{R}_x(t_2-1) + \mathbf{x}(t_2) \mathbf{x}^T(t_2) \end{aligned} \quad (5.9)$$

由矩阵求逆引理<sup>[27]</sup>,  $[\mathbf{R}_x(t_2)]^{-1}$  递归形式为

$$\mathbf{R}_x^{-1}(t_2) = \frac{1}{\lambda} \mathbf{R}_x^{-1}(t_2-1) - \frac{1}{\lambda} \mathbf{Q}(t_2) \mathbf{x}(t_2) \mathbf{x}(t_2)^T \mathbf{R}_x^{-1}(t_2-1) \quad (5.10)$$

这里  $\mathbf{Q}(t_2) = \frac{\mathbf{R}_x^{-1}(t_2-1)}{\lambda + \mathbf{x}(t_2)^T \mathbf{R}_x^{-1}(t_2-1) \mathbf{x}(t_2)}$ 。其中  $t_1, t_2$  的取值可根据递归时间指标  $t$  确定,

下面取  $t_2 = t$ ,  $t_1 = t - \kappa$  其中  $\kappa > 0$  为不同时刻  $t_1, t_2$  之间的时间间隔。

令  $\mathbf{P}(t) = \mathbf{R}_x^{-1}(t)$ , 由式(5.7)、(5.9) 和(5.10), 可以递归地求出

$$\begin{aligned} \mathbf{D}(t) &= \mathbf{P}(t) \mathbf{R}_x(t - \kappa) \\ &= \mathbf{D}(t-1) + [\mathbf{P}(t) \mathbf{x}(t - \kappa) \mathbf{x}(t - \kappa)^T - \mathbf{Q}(t) \mathbf{x}(t) \mathbf{x}(t)^T \mathbf{D}(t-1)] \end{aligned} \quad (5.11)$$

为实现在线处理以及出于算法复杂度的考虑, 本章用近似幂迭代法<sup>[28]</sup>估计  $\mathbf{D}(t)$  的最大特征值对应的特征向量。由于(5.11)式中  $\mathbf{D}(t)$  为递归更新, 幂迭代法的

计算格式是：取分离矩阵  $W(0)$  的每一列都为任意的单位向量，其第一列  $w_1(0)$  作为初始向量，然后进行迭代

$$\begin{cases} v(t) = D(t)w_1(t-1) \\ w_1(t) = v(t) / \|v(t)\|_2 \\ \lambda_1(t) = w_1^T(t)D(t)w_1(t) \end{cases}, t = 1, 2, \dots \quad (5.12)$$

即  $D(t)$  由(5.11)式更新的同时，利用幂迭代法由(5.12)式对  $D(t)$  迭代一次得到  $w_1(t)$ ，直到算法收敛。由幂迭代法的收敛性分析<sup>[29]</sup>可知：当  $D(t)$  的特征值按模互不相等时， $\lambda_1(t)$  收敛于  $D(t)$  的按模最大的特征值， $w_1(t)$  按方向收敛于  $D(t)$  的最大特征值对应的特征向量。

### §5.3 数目未知的语音信号瞬时混合盲分离与信源数目检测算法

#### 5.3.1 语音信号在线提取特性分析

为便于分析语音信号在线分离的特性，本节在假设混合矩阵  $A$  已知的情况下来分析对应于最大广义特征值的广义特征向量  $w_1$  的时变特性，为此我们定义如下的信道互扰(Inter-channel interference)<sup>[30]</sup>作为性能指标

$$ICI_t = \frac{\sum_{i=1}^N |c_i(t)|^2 - \max_i |c_i(t)|^2}{\max_i |c_i(t)|^2} \quad (5.13)$$

这里  $[c_1(t), c_2(t), \dots, c_N(t)] = w_1^T(t)A$  是全局混合-解混向量， $y(t) = w_1^T(t)As(t)$  即为对应于  $w_1$  提取出的源信号。 $ICI_t$  的值越小，说明算法的分离性能越好，下面通过两个简单的例子说明  $w_1$  的时变特性。

**例 1.** 假设用两个传感器接收两个语音信号(即  $M = N = 2$ )，语音信号的采样率为 16 KHz，信号波形分别见图 5.1(a)和图 5.1(b)。混合矩阵  $A$  的各元素为在  $[-1, 1]$  区间服从均匀分布的随机数。图 5.1(c)给出了其中一个混合信号。利用第 5.2 节中基于递归广义特征分解的盲提取方法在线计算对应于最大广义特征值的广义特征向量。取初始向量  $w_1(0)$  为任意的单位向量。下面研究遗忘因子小于 1 且时间间隔  $\kappa$  取不同值时的在线提取的信号波形及  $ICI_t$  性能曲线。

图 5.1(d)给出了遗忘因子  $\beta = 0.9996$  且  $\kappa = 3000$  时递归算法一次运行在线提取出的信号波形。对应的  $ICI_t$  性能曲线(见图 5.1(e))呈现间隔性起伏。从图 5.1 中标注的四个从左到右虚线矩形框中可以直观地看出：图 5.1(d)中在线提取的信号波形中第一个和第三个虚线框中的信号属于第一个源信号(见图 5.1(a))，第二个和第四

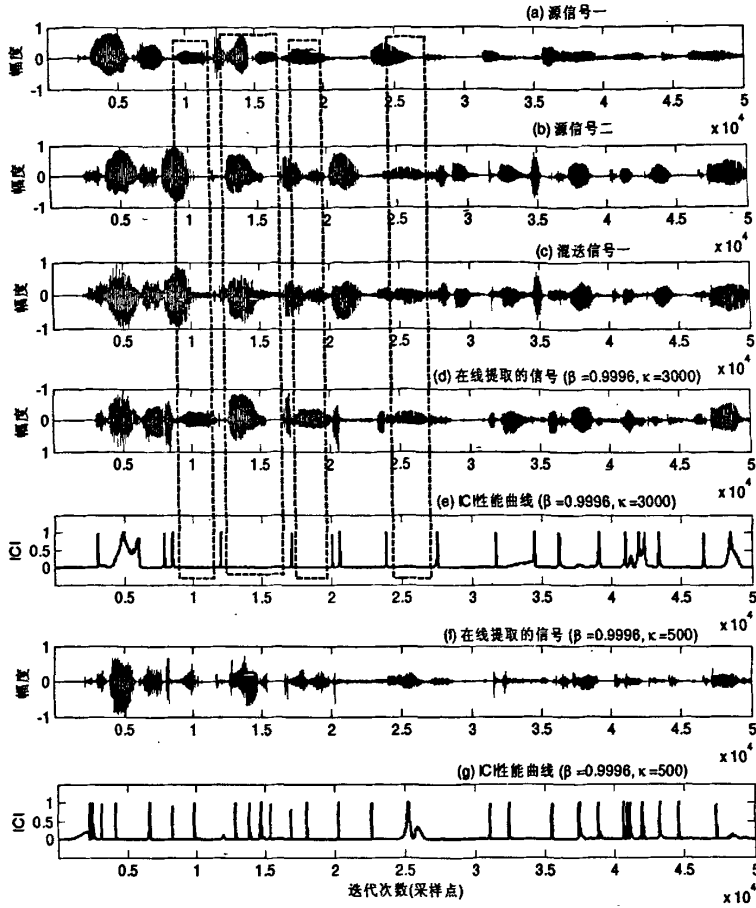


图 5.1 两个源信号混合的在线提取语音信号特性图

个虚线框中的信号属第二个源信号(见图 5.1(b)), 并且图 5.1(e)中对应于四个虚线框的  $ICI_t$  的值都接近于 0, 表明此时有很好的分离性能。图 5.1(f)给出了遗忘因子  $\beta = 0.9996$  且  $\kappa = 500$  时在线提取出的信号波对应的  $ICI_t$  性能曲线见图 1(g)。与图 1(e)相比, 图 1(g)中  $ICI_t$  性能曲线的间隔性起伏较为密集但分离性能较好的采样区间长度较短, 对应的在线提取的信号波形则不易直观看出与源信号的对应关系。

在实验中, 混合矩阵  $A$  是不变的, 因此  $ICI_t$  性能曲线(见图 5.1(e))呈现的间隔性起伏是由于  $w_1$  的在线估计向量的时变特性所致。为便于表述, 我们把  $w_1$  的估计向量能取得良好分离性能的采样区间简称为“分离区间”, 所有“分离区间”组成的集合简称为“分离区间集”, 记为

$$\{[t', t''] | ICI_t < \varepsilon, t \in [t', t''], \varepsilon > 0\} \quad (5.13)$$

这里  $\varepsilon$  取趋于零的正常数。与之相对应, 区间  $[1, T]$  与“分离区间集”差集合

$$[1, T] - \{[t', t''] | ICI_t < \varepsilon, t \in [t', t''], \varepsilon > 0\} \quad (5.14)$$

称为“混合区间集”，包含在其中的区间称为“混合区间”。显然，算法在混合区间中的分离性能较差甚至无法实现分离。

通过例 1 可以直观看到在线提取的信号、源信号与部分“分离区间”的对应关系，下面重点研究多个传感器接收多个语音信号的超定盲提取问题中各参数的时变特性。

**例2.** 假设用八个传感器接收四个语音信号(即  $M=8, N=4$ )， $\kappa=1000$ ，其它参数同例 1。考察在线提取的信号波形， $ICI_t$  性能曲线和如下参数

$$PI_\lambda(t) \triangleq \mathbf{w}_1^T(t) \hat{\mathbf{R}}_x(t_1) \mathbf{w}_1(t) \quad (5.15)$$

的时变特性，这里  $\hat{\mathbf{R}}_x(t_1) = \mathbf{R}_x(t_1) / \|\mathbf{R}_x(t_1)\|_F$ 。

图 2(a)-(d)给出了四个语音信号的波形，图 2(e)给出了其中一个混合信号的波形。图 2(f)给出在线提取的信号的波形，但不易看出在线提取的信号、源信号与部分“分离区间”的对应关系。与例 1 中  $ICI_t$  性能曲线(图 1(e))相比，图 2(g)表明在信源较多的情况下  $ICI_t$  性能曲线的间隔性起伏更为密集。相应的“分离区间”也更多。图 2(h)给出了在线估计的参数  $PI_\lambda(t)$ ，可见在大部分的“分离区间”内  $PI_\lambda(t)$  的值接近于零。这是由于按照广义特征值分解的性质(见第 4.2 节)， $\lambda(t) \triangleq \mathbf{w}_1^T(t) \mathbf{R}_x(t_1) \mathbf{w}_1(t)$  为在线估计的矩阵束  $(\mathbf{R}_x(t_1), \mathbf{R}_x(t_2))$  的广义特征值。在递归算法中  $\mathbf{R}_x(t_1)$  按式(5.8)指数加权求和，从而具有较大的范数值，故式(5.15)对  $\mathbf{R}_x(t_1)$  作归一化处理后所得参数  $PI_\lambda(t)$  的值较小，接近于零。下节将利用  $PI_\lambda(t)$  的时变特性作为一个性能指标确定信源数目。

从例 1 和例 2 的实验结果中容易验证，在同一“分离区间”中不同采样点的对应的向量  $\mathbf{w}_1(t)$  都是同一广义特征向量的估计值，而不同“分离区间”中的向量  $\mathbf{w}_1(t)$  则可能对应同一广义特征向量，也可能对应不同广义特征向量。对于这种现象有如下解释：由于语音的非平稳特性和说话人在发音时有停顿，因此，在线估计的对应于最大广义特征值的广义特征向量随着实际语音环境的变化而变化，并不总对应于特定的说话人，而是对应不同说话人，这样提取出的信号中不同的采样间隔内对应着不同的说话人。当对应于最大广义特征值的广义特征向量发生突然变化时就会导致算法分离性能下降，而保持不变时则可取得良好的分离性能，从而使得性能曲线  $ICI_t$  呈现不同采样间隔的起伏现象，导致所提取的信号在不同的

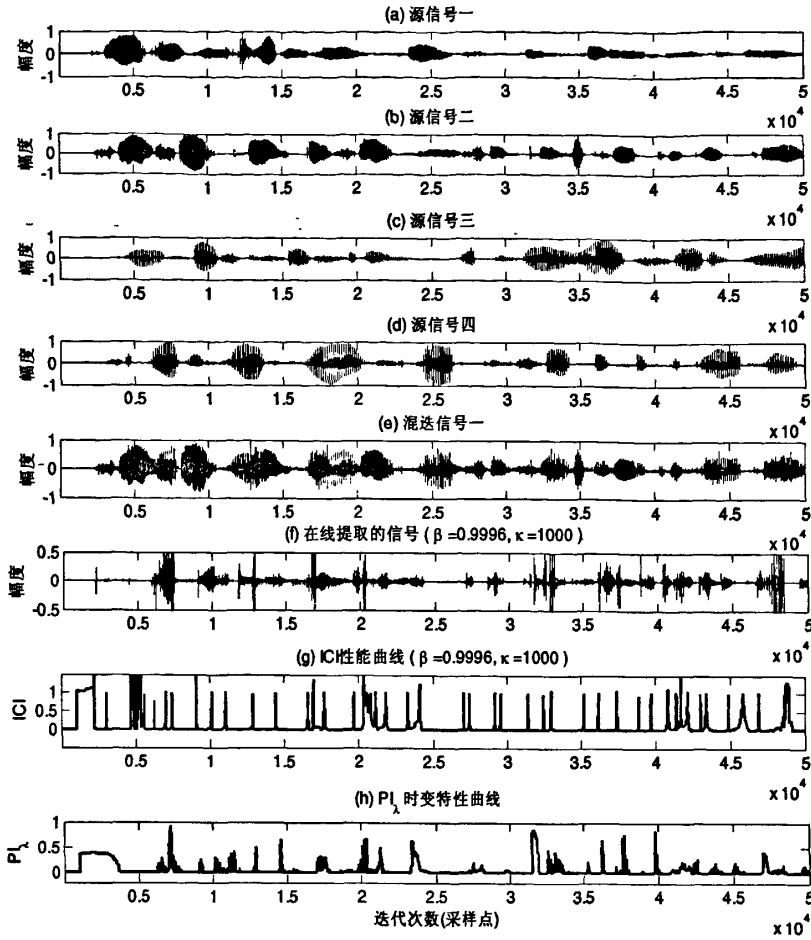


图 5.2 四个源信号混合的在线提取语音信号特性图

“分离区间”可能对应于不同的源信号，在“混合区间”依然是源信号的混合。总体而言无法获得单独提取出单个信源，如果实际语音环境的动态变化较大时则提取出信号听起来只是无法辨别的“噪音”。另外，“分离区间”的数目与信源数目、时间间隔  $\kappa$ ，和门限值  $\varepsilon$  有关。

### 5.3.2 信源数目检测与分离算法

对应于不同广义特征值的广义特征向量的数目正是信源的数目，从对第 5.3.1 节中例 1 和例 2 的分析可知，不同的“分离区间”对应不同的广义特征向量的估计值，正是这种现象为我们解决信源未知或动态变化时的瞬时混合语音信号盲分离和信源数目检测问题提供了重要的研究思路。若对“分离区间”中向量  $w_i(t)$  进行聚类，向量聚类的数目即为信源的数目。但是在盲源分离问题中，混合矩阵  $A$  是未知的，因此实际无法给出  $ICI_i$  性能曲线，从而无法直接得到“分离区间”。

本节在混合矩阵  $A$  未知的情况下充分利用第 5.2 节中递归盲提取算法得到的分离向量序列  $w_1(t)(t=\kappa+1, \dots, T)$ , 语音的非平稳特性、 $PI_\lambda(t)$  的时变特性及向量相似度 (similarity) 等信息通过拟合“分离区间”和多阶段聚类算法实现信源数目的检测并同时完成信源的盲分离。

要想在混合矩阵  $A$  未知的情况下正确检测出信源数目, 需对在线估计的向量序列  $w_1(t)(t=\kappa+1, \dots, T)$  进行分类, 并把一定意义下分离性能最好的向量作为广义特征向量的估计, 同时根据估计的对应于不同广义特征值的广义特征向量的数目确定信源的数目。显然, 分离性能最好的向量一定是上节定义的“分离区间集”中采样时刻对应的向量。但由于混合矩阵未知, 所以“分离区间”不能直接得到。因此需要根据已知信息“拟合”(或近似)出“分离区间”, 并对近似的“分离区间”中采样时刻对应的向量进行聚类或分类, 从而根据聚类数目确定信源的数目。

### 5.3.2.1 提取“高分离度区间”

本节分别利用两种性能指标拟合  $ICI_t$  性能曲线, 提取出两类不同特性的区间集合, 进一步得到“高分离度区间集”, 属于其中的区间称为“高分离度区间”。

利用向量的相似度或相异度 (dissimilarity) 对向量进行分类是一种简单的选择, 针对要解决实际问题, 向量的相似度可以有不同定义: 如 Euclidean 距离, Mahalanobis 距离, 夹角余弦、Tanimoto 测度以及目标-概念距离<sup>[31]</sup>等。本节利用两个  $M$  维向量  $u_1$  和  $u_2$  的夹角余弦函数

$$\Gamma(u_1, u_2) = \cos \theta = \frac{|u_1^T u_2|}{\|u_1\|_2 \|u_2\|_2} \quad (5.16)$$

作为向量相似性的测度, 显然  $0 \leq \Gamma(u_1, u_2) \leq 1$ ,  $\Gamma(u_1, u_2)$  的值越大则表示两个向量的相似度越高,  $\Gamma(u_1, u_2) = 1$  则表示两个向量只相差一个比例系数, 即  $u_1 = \xi u_2$ , 这里  $\xi$  为非零常数。

为了拟合  $ICI_t$  性能曲线, 我们把当前分离向量  $w_1(t)$  与过去分离向量  $w_1(t-j), j=1, \dots, L$  之间的平均相似度定义为

$$MS_t = \begin{cases} 1 - \frac{1}{t-\kappa-1} \sum_{j=1}^{t-\kappa-1} \Gamma(w_1(t), w_1(t-j)), & t-\kappa \leq L \\ 1 - \frac{1}{L} \sum_{j=1}^{t-L} \Gamma(w_1(t), w_1(t-j)), & t-\kappa > L \end{cases} \quad (5.17)$$

这里用近似幂迭代法(5.12)在线估计的向量  $\mathbf{w}_1(t), t = \kappa+1, \dots, T$  都是单位向量, 故上式中  $\Gamma(\mathbf{w}_1(t), \mathbf{w}_1(t-j)) = |\mathbf{w}_1^T(t) \mathbf{w}_1(t-j)|$ 。

从式(5.17)可知,  $MS_i$  的值越小, 当前时刻向量  $\mathbf{w}_1(t)$  与其临近向量的相似程度越高。由于在同一“分离区间”中不同采样时刻估计的向量是同一广义特征向量的估计, 因此这些向量具有很高的相似度, 对应  $MS_i$  的值接近于零。换言之, 在  $MS_i$  的值都很小(小于某个门限)的采样区间内的向量间具有很高的向量相似度, 这样的区间称为“高相似度区间”。所有“高相似度区间”的组成的集合称为“高相似度区间集”, 记为

$$\{\{\tilde{t}, \hat{t}\} \mid MS_i < \gamma, t \in [\tilde{t}, \hat{t}], 0 < \gamma < 1\} \quad (5.18)$$

其中  $\gamma$  为趋于 0 的正数。

由于  $ICI_i$  性能曲线和平均相似度  $MS_i$  曲线在“分离区间”中的值都是趋于零, 因此我们利用平均相似度  $MS_i$  曲线来拟合  $ICI_i$  性能曲线, 再进一步利用平均相似度  $MS_i$  曲线提取出“高相似度区间集”以拟合“分离区间集”。

下面利用平均相似度  $MS_i$  曲线采用如下决策规则在线提取“高相似度区间”。其基本思想是对于给定的门限  $\gamma$ , 比较  $MS(t)$  的值与  $MS(t-1)$  的值判定当前采样时刻  $t$  是否属于某一“高相似度区间”, 若属于一个“高相似度区间”, 进一步判定该点是否为区间的端点, 若该点不是区间端点, 则该“高相似度区间”的长度加 1, 若当前时刻  $t$  被判定为一个“高相似度区间”的起点, 则把该点设为区间起点, 并把“高相似度区间”的数目加 1, 若当前时刻  $t$  被判定是一个“高相似度区间”外的第一个点, 则把前一时刻  $t-1$  设为区间终点结束该区间。基于 Matlab 语言的“高相似度区间”的在线搜索算法见表 5.1。

需要指出的是在同一个“分离区间”中的不同向量都是同一个广义特征向量的近似, 因此这些向量与其临近向量有很高的平均相似度。但是向量间具有很高的平均相似度的区间并不一定是“分离区间”, “混合区间”中的向量间也可能会有很高的平均相似度。换句话说, “高相似度区间集”中既包含所有“分离区间”, 也可能包含部分“混合区间”。因此还需要进一步作后续处理以剔除“混合区间”。此外, 对应于同一广义特征向量的不同“分离区间”中的分离向量的分离性能也有差异。

根据上述分析, 为正确检测信源数目, 还需要对“高相似度区间集”进行两个步骤的处理:

- (i) 从提取出的“高相似度区间集”中剔除“混合区间”;
- (ii) 剔除分离性能较差的“高相似度区间”。

表5.1基于Matlab语言的“高相似度区间”的在线搜索算法

```

Flag_number = 0; % 用于统计“高相似度区间”的个数
Flag_length = 0; % 用于统计每个“高相似度区间”的长度
High_similarity_interval = zeros(T,3);
% 该矩阵的每一行分别记录一个“高相似度区间”的起点、终点
% 和区间长度
if MS(t) <  $\gamma$ 
    if MS(t-1) >  $\gamma$ 
        Flag_number = Flag_number + 1;
        High_similarity_interval(Flag_number,1) = t;
        Flag_length = 1;
    else
        Flag_length = Flag_length + 1;
    end
else
    if MS(t-1) <  $\gamma$ 
        High_similarity_interval(Flag_number,2) = t-1;
        High_similarity_interval(Flag_number,3) = Flag_length;
    end
end
High_similarity_interval = High_similarity_interval(1:Flag_number,:);

```

下面利用在线估计的参数  $PI_{\lambda}(t)$  的时变特性提取  $PI_{\lambda}(t)$  的值很小的区间(称为“ $PI_{\lambda}(t)$ 趋零区间集”)以得到“高分离度区间”。如 5.3.1 节中图 2(h)所示,分离区间中对应的  $PI_{\lambda}(t)$  的值接近于零。把  $PI_{\lambda}(t)$  的值都小于某一给定门限  $\rho$  (如  $\rho=0.01$ ) 的采样区间称为“ $PI_{\lambda}(t)$ 趋零区间”,类似“高相似度区间”的搜索算法,我们也可以提取出“ $PI_{\lambda}(t)$ 趋零区间”。“ $PI_{\lambda}(t)$ 趋零区间”对应于“分离区间”中分离性能较好的部分,因此我们可以取“ $PI_{\lambda}(t)$ 趋零区间”与“高相似度区间”的交集作为“高分离度区间”。一方面可以提取出分离性能较好的分离区间,另一方面则可剔除部分“混合区间”。

用于存储“高分离度区间集”的矩阵记为 *High\_separation\_interval*, 该矩阵的行数即为交集的数目,该矩阵列数设为四,其中前三列用于存储每个“高分离度区间”的起点、终点和区间长度。注意到“高分离度区间”是“ $PI_{\lambda}(t)$ 趋零区间”与“高相似度区间”的交集,即每一个“高分离度区间”都分别是某一“高相似度区间”的子区间,因此 *High\_separation\_interval* 第四列存储该“高分离度区间”

在矩阵 *High\_similarity\_interval* 中所在行的行数, 用于标识该“高分离度区间”是某一“高相似度区间”的子区间, 以便下一阶段的进一步处理。

同提取“高相似度区间”遇到的问题一样, “ $PI_{\lambda}(t)$  趋零区间”也不一定是“分离区间”, 因此, 根据“ $PI_{\lambda}(t)$  趋零区间”无法完全剔除“混合区间”, 需在下一阶段进一步处理。需要指出的是“高相似度区间集”与“ $PI_{\lambda}(t)$  趋零区间集”都是在估计向量  $w_1(t)$  的同时在线提取的。

### 5.3.2.2 “高分离度区间”聚类

“高分离度区间”的区间长度越大说明算法在该区间内的收敛性能越稳定, 从而该区间中的向量为广义特征向量的精确估计值的概率越大。因此本节利用“高分离度区间”的长度和属于不同“高分离度区间”中的向量间的相似度对“高分离度区间集”中的“高分离度区间”进行迭代聚类, 并初步确定信源数目。

迭代步骤如下:

初始化: 令估计的源信号数目  $\hat{N}=0$ ;

第1步: 首先从“高分离度区间集”中找到区间长度最大的“分离区间”, 即

$$[Length\_max, Number\_row] = \max(High\_separation\_interval(:, 3));$$

这里等式右边表示求 *High\_separation\_interval* 中第三列的最大值, *Length\_max* 表示区间长度最大值, *Number\_row* 表示区间长度最大值所在的矩阵行数。用如下符号

$$tmp\_i \triangleq High\_separation\_interval(Number\_row, i), i=1, 2$$

表示该矩阵的第 *Number\_row* 行第 *i* 列元素。

把该“高分离度区间”  $[tmp\_1, tmp\_2]$  作为对应于提取第一个源信号的广义特征向量的精确估计值区间, 并判定源信号数目为 1, 然后在该区间  $[tmp\_1, tmp\_2]$  中求在  $MS_i$  值最小的采样时刻, 即

$$[MS\_min, Sample\_instant] = \min(MS(tmp\_1:tmp\_2))$$

这里  $MS\_min$  表示在区间  $[tmp\_1, tmp\_2]$  内  $MS_i$  的最小值, *Sample\_instant* 表示在区间  $[tmp\_1, tmp\_2]$  内的第 *Sample\_instant* 个点处取得  $MS_i$  的最小值, 并在对应的采样时刻 ( $\hat{t} = tmp\_1 + Sample\_instant - 1$ ) 估计的向量  $w(:, \hat{t})$  作为一个广义特征向量的估计值, 相应可以提取出一个源信号。

同时存储与已判定出的广义特征向量的相似度大于给定门限  $\eta$  (如  $\eta=0.99$ ) 的

“高分离度区间”，这些区间内的向量均视为同一广义特征向量的估计，可以提取出同一源信号，这些区间全体称为“同源聚类区间集”。为避免重复提取同一个源信号，需要把“同源聚类区间集”从“高分离度区间集”中剔除。

第 $k$ 步( $k=2,3,\dots$ ): 如果在剔除掉 $k-1$ 个“同源聚类区间集”的“高分离度区间集”中与对应于已判定为源信号的广义特征向量的相似度小于 $\eta$ 的区间数目大于零；则按第一步的方法提取下一个广义特征向量，得到相应的分离信号，并把检测出的信源数目加一( $\hat{N}:=\hat{N}+1$ )，直到与对应于已判定为源信号的广义特征向量的相似度小于 $\eta$ 的区间数目为零。

### 5.2.2.3 剔除“孤立区间”

由第 5.2.2.1 节中分析可知，根据“高相似度区间”和“ $PI_{\lambda}(t)$ 趋零区间”无法完全剔除“混合区间”，因此经过第 5.2.2.2 节的聚类处理后部分“混合区间”也被判定为对应于源信号的“分离区间”被保留下来，这样检测出的信源数目可能大于真实的信源数目，因此需要进一步处理。

考察对应于被判定为信源的不同“同源聚类区间集”。由于“同源聚类区间集”中区间个数一般较多，且“混合区间”中的估计的广义特征向量与它向量的相似性较小，因此我们考察“同源聚类区间集”中区间的数目。若区间集中区间数目只有一个，则该区间判定为“混合区间”，这样的区间称为“孤立区间”，将其从“高分离度区间集”中剔除。若一个“同源聚类区间集”中区间数目大于一个，但 these 区间属于“高相似度区间集”同一个区间或相邻的两个区间，也将其视为“混合区间”，这可能是由于“高相似度区间集”中的区间本是“混合区间”，和“ $PI_{\lambda}(t)$ 趋零区间集”取交集后被分割成的若干子区间，它们还是“混合区间”，这样的区间也视为“孤立区间”。把“孤立区间”剔除后保留的“同源聚类区间集”的集合数目即为信源数目。

## §5.4 仿真实验

以下实验假设用传感器阵列接收信源数目未知的语音信号，为验证所提算法的性能，首先以传感器数目 $M=8$ ，信源数目分别 $N=4$ 为例阐述算法的提取“高分离度区间”和“高分离度区间”聚类的过程，然后在相同的传感器数目( $M=8$ )接收数目不同( $N=3,4,5,6$ )的源信号的情况下检测信源的数目，在下面不同实验中所用到的所有源信号波形见图 5.3。为体现信源数目未知的特点，我们在试验中

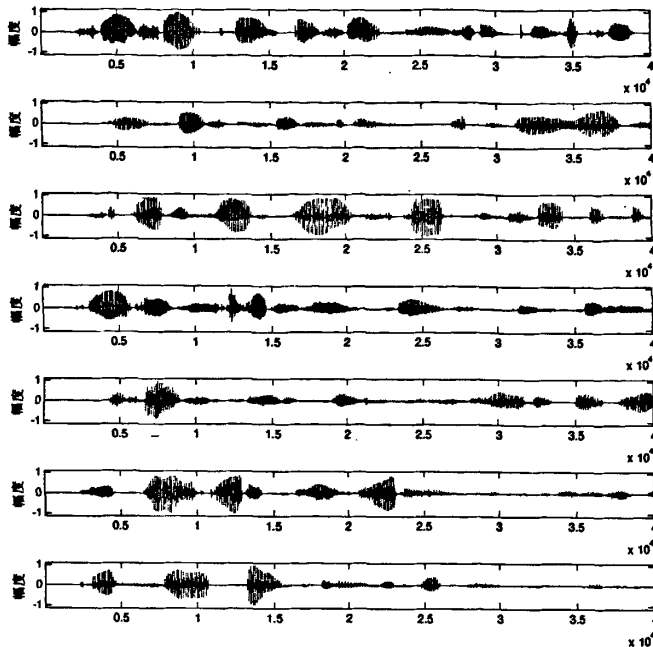
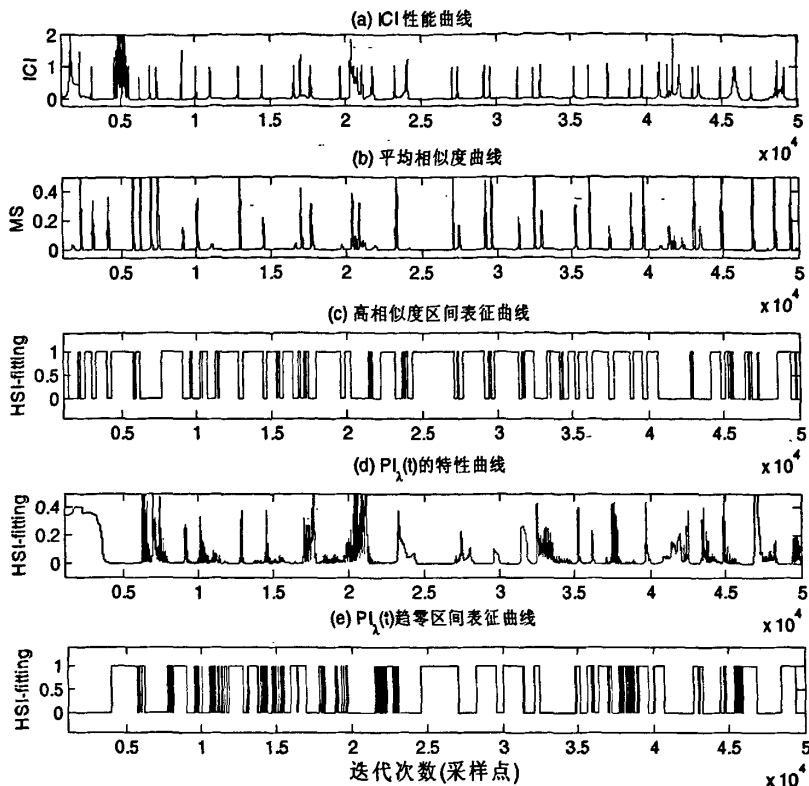


图 5.3 仿真试验中所用源信号

用一组相同的参数来分别检测信源数目不同时的信源数目。混合矩阵  $\mathbf{A}$  的各元素为在  $[-1,1]$  区间服从均匀分布的随机数。仿真实验中所用的参数取值分别为  $\kappa=1000$ ,  $\gamma=0.0002$ ,  $L=100$ ,  $\rho=0.004$ ,  $\eta=0.99$ ,  $\beta=0.9996$ 。

实验一：在假设用八个传感器接收四个语音信号(即  $M=8, N=4$ )的情况下利用第 5.3 节中的方法检测信源数目。所用源信号为图 5.3 中的前四个信号。

图 5.4(a)给出  $ICI_t$  性能曲线以检验所提性能参数(式(5.15)和式(5.17))拟合  $ICI_t$  性能( $ICI_t$  性能曲线是在混合矩阵已知的情況下得到的,但在我们实验中并不利用混合矩阵的信息)。首先利用向量相似度(相异度)拟合“ $ICI_t$  性能曲线”。由式(5.17)给出的平均相似度曲线见图 5.4(b),与图 5.4(a)中  $ICI_t$  性能曲线相比较,除去部分采样区间(如  $[4500,5500]$ 、 $[45500,46500]$  等),平均相似度曲线在接近于零时可以较好的拟合  $ICI_t$  性能曲线。这表明我们可以从平均相似度曲线中提取“高相似度区间集”来近似“分离区间集”,但需要从中剔除相似度较高的“混合区间”,如  $[4500,5500]$ 、 $[45500,46500]$  等。由于长度太小的区间的“高相似度区间”的可信度不高,在从平均相似度曲线搜索“高相似度区间集”时只保留区间长度大于 50 的“高相似度区间”。为直观表示“高相似度区间集”,这里用 Matlab 语言的表示法,令  $HSI\_fitting = \text{zeros}(1,T)$ , 这里  $HSI\_fitting$  表示一个元素全为零的  $1 \times T$  行向量,

图 5.4 拟合  $ICI_t$  性能曲线与其区间表征曲线

把与“高相似度区间”中采样时刻相对应的  $HSI\_fitting$  中的区间内的点都赋值为 1，得到的折线如图 5.4 (c) 所示，即折线上为 1 的部分对应的采样区间为“高相似度区间”，折线上为 0 的部分判定为“混合区间”。下面用类似的折线直观表示所提取的各种区间。为进一步提取“高分离度区间集”并剔除部分“混合区间”，图 5.4(d) 给出了  $PI_\lambda(t)$  特性曲线，根据  $PI_\lambda(t)$  特性曲线提取得到的“ $PI_\lambda(t)$  趋零区间”表征曲线如图 4(e) 所示，这时提取出的区间个数较多，说明在给定的门限下  $PI_\lambda(t)$  随时间变化比较剧烈。

图 5.5(a) 给出了图 5.4 (c) 中所示的“高相似度区间集”和图 5.4 (e) 所示的“ $PI_\lambda(t)$  趋零区间”的交集，即“高分离度区间集”的表征曲线。只包含在一种区间集合中的部分“混合区间” (如区间 [4500, 5500] 等) 已经取交集后剔除，另外部分分离性能不好的“高相似度区间”也在取交集后剔除。

经过对“高分离度区间集”中区间利用第 5.3 节中迭代聚类 ( $\eta = 0.99$ ) 并剔除“孤立区间”后的区间集合表示曲线如图 5.5(b) 所示，这里区间中元素都为 1 的区

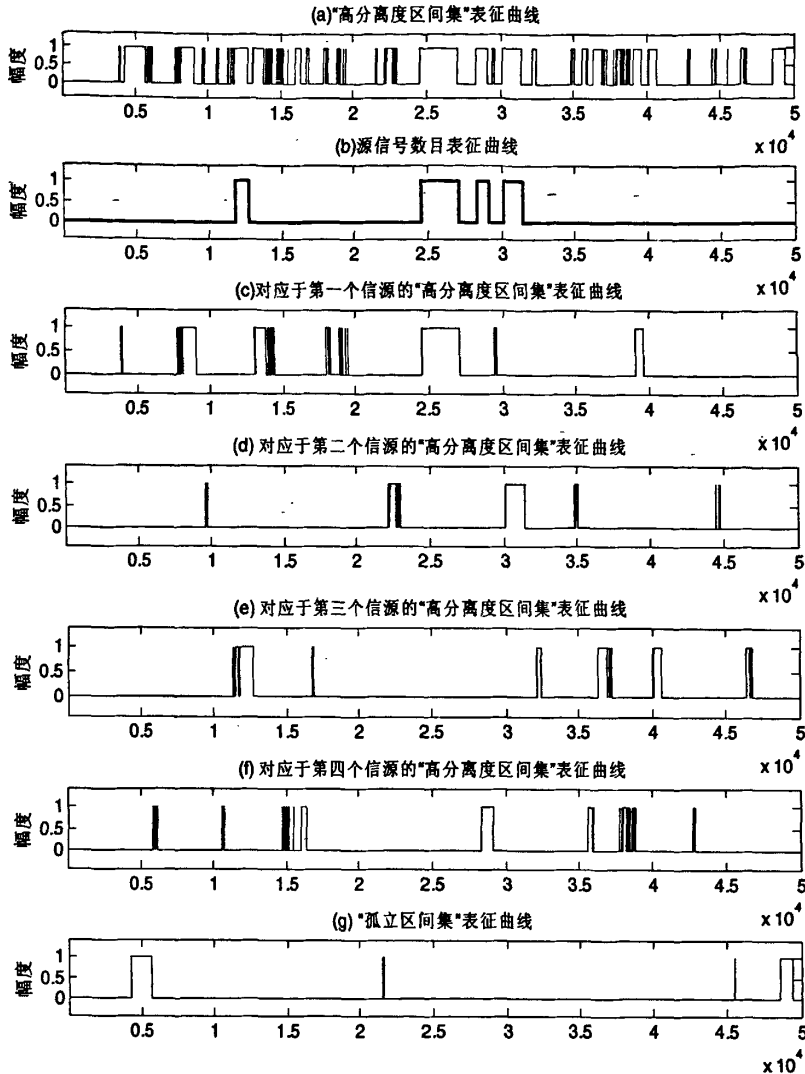


图 5.5 “高分离度区间”、信源数目、“同源区间”及“孤立区间”表征曲线

间的数目则为经检测得到的信源的数目。图 5.5 (c)-(f)分别给出了对图 5.5 (a)中的区间的对应于四个源信号的“同源聚类区间集”表征曲线，由此可见对应于同一源信号的广义特征向量可在多个“聚类区间”中出现，这种现象表明漏检某一源信号的概率较小。图 5.5(g) 则给出了对应于“混合区间”的“孤立区间”表征曲线，这些区间中对应的向量与其它区间中向量的相似度较小。

为避免估计出的信源数目虽然正确的，但对应的某个广义特征向量却可能是一“混合区间”中的一个估计向量，而另外一个对应于源信号广义特征向量估计被漏检，这样必然有一个源信号无法正确的分离出来。为此需要进一步验证算法对于信源数目的检测的可靠性，我们分别对数目不同的信源的混合的情形进行如

下的 Mento-Carlo 实验。假设检测出对应于信源的广义特征向量存储在  $\hat{N} \times M$  矩阵  $\hat{\mathbf{W}}$  中, 这里  $\hat{N}$  是实际检测的信源的数目。由于每次仿真混合矩阵的随机性, 每次实验中检测出信源的顺序具有不确定性, 为验证每次仿真实验的检测性能, 我们定义类似于式(5.13)的信道互扰性能指标

$$ICI_t = \sum_{i=1}^{\hat{N}} \left( \sum_{l=1}^N \frac{|\bar{b}_{il}(t)|^2}{\max_p |b_{pl}(t)|^2} \right) - \hat{N} \quad (5.19)$$

这里  $\mathbf{B}(t) = \hat{\mathbf{W}}^T(t)\mathbf{A}$  是全局混合-解混矩阵, 其中  $\mathbf{B}(t) = [b_{ij}(t)]_{\hat{N} \times N}$ 。

**实验二:** 用数目相同的传感器阵列 ( $M=8$ ) 接收数目不同 ( $N=3, \dots, 7$ ) 的源信号的情况下检测信源的数目, 为体现信源数目未知的特点, 其余参数同实验一。

表 5.2 是用所提信源数目检测算法对于信源数目不同时分别进行 1000 次重复实验检测出信源数目正确率。当信源数目为 3 时检测正确率为 100%, 当信源数目增多时, 正确率随之下降, 这主要是因为信源数目不同时对应的性能指标  $ICI_t$  (式(5.13))有所不同,  $ICI_t$  的值随着信源数目的增加而变大, 而实验中为体现信源数目

表 5.2 1000 次仿真实验检测信源数目的正确识别率

信源数目	$N=3$	$N=4$	$N=5$	$N=6$	$N=7$
正确识别率	100 %	99.4 %	97.7 %	94.8 %	86.3 %

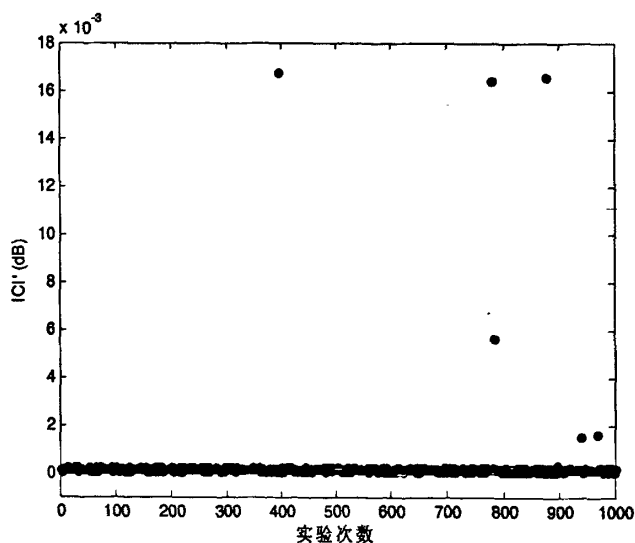


图 5.6 信源数目为 4 时 1000 实验的性能指标  $ICI_t$  的值

未知的特点,用一组相同的参数来分别检测信源数目不同时的信源数目。因此参数的设置需要一定程度的折中。

图 5.6 是信源数目为 4 时,1000 次不同实验的  $ICI_t'$  性能指标,每个点表示一次实验的  $ICI_t'$  的值。从图 5.6 中可以看到性能指标值明显偏大的点的数目为 6 个,与表 5.2 中信源数目为 4 时错误检测结果的次数相同,这表明所提信源数目检测算法可以较为准确的检测出信源数目,并对应于真实的信源。此外,一般的盲提取算法通常会在提取的过程中进行数据压缩导致误差累积,从而提取出的信源的质量依次下降,而在本文恢复出的信号都对应于估计性能最好的广义特征向量,因此都具有很好的分离性能。

## §5.5 本章小结

本章充分利用人类发音的特点和语音信号的非平稳特性,对于数目未知的语音信号瞬时混的盲分离问题,在实现信号数目检测的同时完成信号的盲分离,并可克服盲提取问题中先提取出的信源性能好而后提取出的信源性能差的缺点。由于人类对话交流的特点,在线估计的对应的最大广义特征值的特征向量,是随着实际语音环境的变化而变化,并不对应于特定的说话人,由于所用近似幂迭代方法的快速收敛特性,在线估计的最大广义特征值对应的广义特征向量在某些时间间隔内能得到很好的估计性能,且对应于某一个说话人。这样可以利用估计性能良好的不同广义特征向量的数目确定信源的数目,并利用得到的广义特征向量提取出所有的源信号。

本章首先利用向量的相似度定义了平均相似度  $MS_t$ ,用以拟合  $ICI_t'$  性能曲线,并从  $MS_t$  曲线中提取出“高相似度区间”拟合“分离区间”。然后利用“ $PI_\lambda(t)$ 趋零区间”进一步提取分离性能较好的“高相似度区间”并剔除部分“混合区间”,得到“高分离度区间”。然后利用多阶段聚类的方法确定源信号的数目,并可在完成信源数目的检测的同时实现信号的盲分离或提取。仿真实验表明本章提出的信源数目检测算法具有良好的性能。

## 参考文献

- [1] Cichocki A, Amari S. Adaptive Blind Signal and Image Processing. New York: Wiley, 2003.
- [2] Haykin S. Unsupervised adaptive filtering. Volume1: Blind source separation. New York: Wiley, 2000
- [3] Benesty J, Huang Y. Adaptive Signal Processing: Applications to Real-World Problems. Berlin, Germany: Springer, 2003.

- [4] Makino S, Lee T W, Sawada H. Blind speech separation, Springer, 2007.
- [5] Torkkola K. Blind separation of convolved sources based on information maximization. IEEE Workshop on Neural Networks and Signal Processing, Kyoto, Japan, 1996.
- [6] Smaragdis P. Blind separation of convolved mixtures in the frequency domain, Neurocomputing, 1998, 22(2), pp. 21-34,.
- [7] Parra L, Spence C. Convolutional blind separation of nonstationary sources. IEEE Trans on Speech Audio Processing, 2000, 8(3), pp. 320-327.
- [8] Murata N, Ikeda S, Ziehe A. An approach to blind source separation based on temporal structure of speech signals. Neurocomputing 2001, 41, pp. 1-24.
- [9] Asano F, Ikeda S, Ogawa M, Asoh H, Kitawaki N. A combined approach of array processing and independent component analysis for blind separation of acoustic signals. IEEE Trans on Speech and Audio Processing, 2003, 11(3), pp. 204-215.
- [10] Araki S, Mukai R, Makino S. The fundamental limitation of frequency domain blind source separation for convolutional mixtures of speech. IEEE Trans on Speech and Audio Processing, 2003, 11(2), pp. 109-116.
- [11] Ito M, Takeuchi Y, Matsumoto T. Moving-source separation using directional microphones. Proc. Int. Symp. Signal Process. Inf. Theory, Dec. 2002, pp. 523-526.
- [12] Cauwenberghs G, Stanacevic M, G. Zweig. Blind broadband source localization and separation in miniature sensor arrays. Proc. IEEE Int. Symp. Circuits Syst., 2001, 3, pp. 193-196.
- [13] Pedersen M S, Wang D L, Larsen J, Kjems U. Two-microphone separation of speech mixtures. IEEE Trans on Neural Network, 2008, 19(3), pp. 475-492.
- [14] Matsuoka K, Ohya M, Kawamoto M. A neural net for blind separation of nonstationary signals. Neural Network, 1995, 8(3), pp. 411-419.
- [15] Choi S, Cichocki A. Blind separation of nonstationary sources in noisy mixtures. Electronics Letters, 2000, 36(4), pp. 848-849.
- [16] Choi S, Cichocki A. Robust whitening procedure in blind source separation context. Electronics Letter, 2000, 36(24), pp. 2050-2051.
- [17] Belouchrani A, Abed-Meraim K, Cardoso J F, Moulines E. A blind source separation technique using second-order statistics. IEEE Trans Signal Processing, 1997, 45(2), pp. 434-444.
- [18] Pham D T, Cardoso J F. Blind separation of instantaneous mixtures of nonstationary sources. IEEE Trans on Signal Processing 2001, 49(9), pp.

1837-1848.

- [19] Yin F L, Mei T M, Wang J. Blind source separation based on decorrelation and nonstationarity. *IEEE Trans. Circuits and Systems-I:Regular Papers*, 2007, 54(5), pp. 1150-1158.
- [20] Chang C, Ding Z, Yau S, Chan F. A matrix-pencil approach to blind separation of colored nonstationary signals. *IEEE Trans on Signal Processing*, 2000, 48(3), pp. 900-907.
- [21] Parra L, Sajda P. Blind source separation via generalized eigenvalue decomposition. *Journal of Machine Learning Research*, 2003, 4(7-8): 1261-1269
- [22] Pokharel P P, Ozertem U, Erdogmus D, Principe J C. Recursive complex BSS via generalized eigendecomposition and application in image rejection for BPSK. *Signal Processing*, 2008, 88(4), pp. 1368-1381.
- [23] Li Y Q, Wang J, and Zurada J M. Blind extraction of singularly mixed sources signals. *IEEE Trans on Neural Network*, 2000, 11, pp.1413-1422.
- [24] Li Y Q, Wang J. Sequential blind extraction of instantaneously mixed sources. *IEEE Trans. Signal Processing*, 2002, 46(7), pp. 997-1006.
- [25] Ye J M, Zhu X L, Zhang X D. Adaptive blind separation with an unknown number of sources. *Neural Computation*, 2004, 16(8), pp. 1641-1660.
- [26] Pokharel P P, Ozertem U, Erdogmus D, Principe J C. Recursive complex BSS via generalized eigendecomposition and application in image rejection for BPSK. *Signal Processing*, 2008, 88(4), pp. 1368-1381.
- [27] Haykin S. *Adaptive Filter Theory*, 3rd ed. Englewood Cliffs, NJ: Prentice-Hall, 1996.
- [28] Hua Y, Xiang Y, Chen T, Abed-Meraim K, Miao Y. A new look at the power method for fast subspace tracking. *Digital Signal Processing*, 1999, 9(10) , pp.297-314.
- [29] Golub G H, Loan C. *Matrix Computations*. John Hopkins University Press, Baltimore, MD, 1991.
- [30] Douglas S C, Gupta M. Scaled natural gradient algorithm for instantaneous and convolutive blind source separation. *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2007, II, pp.637-640.
- [31] 张贤达. 矩阵分析与应用. 北京: 清华大学出版社, 2004.

## 第六章 基于多信道信号增强的卷积混合语音

### 信号盲分离的后处理算法

#### § 6.1 引言

近年来, 麦克风阵列语音信号处理已成为信号处理领域的研究热点<sup>[1-5]</sup>。在典型的语音通信环境(如音视频会议、车载免提应答系统、声控系统和助听器等)中, 用户通常希望在通话时能够自由活动且无需佩戴或手持麦克风。但由麦克风阵列接收的信号中往往掺杂了背景噪声和室内混响等会影响用户之间顺畅的交流。因此如何在有噪声和混响的声学环境中高质量地跟踪、增强和分离目标语音信号是一个很有实用价值且极具挑战性的课题。

在语音增强技术研究的早期, 人们利用单传感器接收信号并由观测信号的谱信息实现单信道语音增强<sup>[6]</sup>。为进一步提高受损声音和语音信号的质量, 基于麦克风阵列的多信道信号增强技术由于可以同时利用空域信息和谱信息, 其性能比传统的单传感器处理技术有显著提高。因此目前的语音增强研究已由单信道模型转向立体声和多信道情形<sup>[1]</sup>。利用阵列的空域信息使多信道语音信号处理算法得到更多自由度, 利用波束形成技术<sup>[7-8]</sup>在增强目标信号的同时抑制干扰和噪声, 有利于提高算法性能以进一步推进语音通信系统的商业化。然而传统波束形成技术虽充分利用了麦克风阵列流形(如传感器和信源的位置), 但无法避免由于信道误差导致的互扰和泄漏问题; 并且为了获得较好的性能常需较多的阵元, 但阵元太多会增加设备造价, 不利于多信道信号增强技术的实用化发展。

在一些非合作的声学环境(如鸡尾酒会问题)中, 往往没有信道或信源的先验信息可以利用。在这种情况下, 传统的波束形成技术的有效应用受到了一定的制约。而盲源分离技术<sup>[9-10]</sup>正是在没有任何先验知识的条件下从观测信号中恢复或提取出源信号。由于盲源分离技术无需已知阵列的几何流形和物理特性, 所以能有效地克服信道误差, 增强阵列处理的性能。即使信源波达方向未知, 盲源分离抑制干扰的程度也已接近甚至超过传统的自适应波束形成, 其潜在应用目前是一个很活跃的研究课题, 因此盲信号分离的技术已在声学 and 信号处理领域引起了极为广泛的关注。

盲源分离的混合过程通常被建模成多输入多输出(MIMO)的数学模型, 这种模

型可分为无记忆和有记忆的,也可分别称为瞬时混合和卷积混合。盲源分离问题的早期研究集中于相对简单的瞬时混合情形<sup>[9-10]</sup>,但在典型的声学环境中(如著名的“鸡尾酒会”问题),由于声音传播的多径效应,源信号的混合方式建模为卷积混合是更为现实<sup>[9]</sup>。而语音信号以卷积的方式混合会使问题的解决变得相当困难。通常需要考虑用有限冲激响应(FIR)滤波器矩阵来分离卷积混合的信号。

目前已有许多方法用来处理基于卷积混合的盲源分离问题,这些方法大致可分为两类,第一类是时域方法<sup>[11-13]</sup>,若所用算法收敛,则能获得较好的性能,但通常现有算法具有复杂度高和收敛速度慢的缺点。第二类是频域算法<sup>[14-20]</sup>,即通过短时傅立叶变换把时域卷积混合盲分离问题转换为频域多个瞬时混合盲分离问题,可在整个频带范围内,利用现有的瞬时混合盲源分离算法对各个频率分离问题进行独立处理。这样一方面大大降低了运算量,另一方面可利用现有的研究成果,便于参考借鉴性能优越的瞬时混合盲分离算法。卷积混合的盲语音分离问题的频域算法由于便于实时处理已被广泛研究,国内外学者在其特定的实验环境中已取得了较好的结果,且开发了不同的盲源分离实时处理系统。然而要想得到理想的分离结果,所有的频域盲分离算法都必须解决由于盲源分离问题本身存在的置换模糊和尺度模糊问题。因为时域方法是联合处理,上述两种不确定性不影响其分离性能,而在频域方法中,各个频率分量之间分别处理,易出现各个频率分量上的排列不一致问题,进而重构时会导致信号频谱的重新混合,各个频率分量上尺度模糊也会导致合成信号频谱产生畸变。因此,若不能有效解决置换模糊问题和尺度模糊问题,则会导致分离性能的下降甚至无法分离。虽然各国学者针对频率置换问题已提出很多频率置换方法,并且在特定的试验环境中都能获得较满意的性能,但是在真实的语音环境中有噪声和混响的存在,无法得到真正既精确又稳健的频率置换解决方法。此外由于上述的不确定性,大多频域盲源分离算法得到的分离信号是源信号的卷积信号,在强混响的声学环境中,盲源分离算法的分离性能会严重降低。

本章针对现有频域盲源分离方法存在的问题,即恢复出的信号是源信号和一个未知的滤波器的卷积,且由于部分频率点处盲源分离方法的失效和无法精确解决排列和尺度模糊问题,分离出的信号中会泄露进其余的信号分量,影响分离性能。本章提出一种基于多信道语音增强的频域盲源分离后处理方法。该方法有机结合盲源分离技术与阵列处理技术,既可充分利用空域信息,又无需增加其它先验假设。首先在频域盲源分离方法中利用语谱分裂技术得到 $M \times N$ 个分离信号,可

视为多输入多输出 (MIMO) 混合系统分裂成  $N$  个单输入多输出 (SIMO) 系统, 泄露进该信号的其它信号分量可以视为干扰信号, 然后分别独立用频域多信道语音增强方法重构出各个语音信号, 计算机仿真结果表明本章所提算法具有良好的性能。

## § 6.2 频域盲源分离算法

### 6.2.1 卷积混合盲源分离问题的数学模型

设用  $M$  个麦克风接收  $N (M \geq N)$  个统计独立的语音信号  $s_n(t) (n=1, \dots, N)$  得到观测信号  $x_m(t) (m=1, \dots, M)$ , 这在数学上可以描述成一个  $M \times N$  多输入多输出 (MIMO) 有限冲激响应 (FIR) 系统 (如图 6.1 所示), 第  $m$  个麦克风的第  $t$  个快拍可表示为

$$x_m(t) = \sum_{n=1}^N \sum_{l=0}^{L-1} a_{mn}(l) s_n(t-l) + e_m(t), \quad m=1, \dots, M \quad (6.1)$$

这里  $a_{mn}(l)$  表示从第  $n$  个源信号到第  $m$  个麦克风的脉冲响应。  $e_m(t)$  为均值为零方差为  $\sigma^2$  的白高斯噪声。

卷积混合模型也可用矩阵形式表示为

$$\mathbf{x}(t) = \mathbf{A}(t) * \mathbf{s}(t) + \mathbf{e}(t) \quad (6.2)$$

这里  $\mathbf{x}(t) = [x_1(t), \dots, x_M(t)]$  是观测矢量,  $\mathbf{s}(t) = [s_1(t), \dots, s_N(t)]$  源信号矢量,  $\mathbf{A}(t)$

为滤波器混合矩阵,  $\mathbf{e}(t) = [e_1(t), \dots, e_M(t)]$  是噪声矢量。

卷积混合信号盲分离的目的是通过一个分离系统来重构源信号。重构的信号

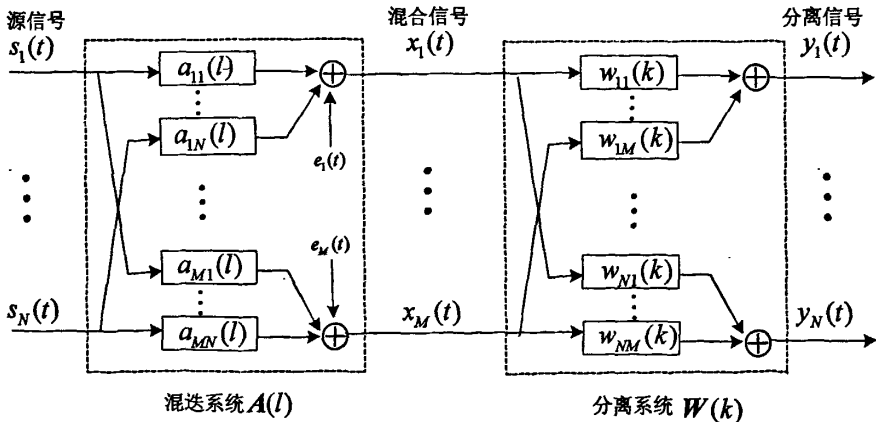


图6.1 盲源分离的卷积混迭与分离信号模型

是源信号的滤波, 且重构信号与源信号存在排列和尺度的不确定性模糊<sup>[14]</sup>。时域方法是用一个长度为  $K$  的 FIR 滤波器  $w_{nm}(k)$  的集合构成分离系统来产生分离信号 (如图 6.1 所示)

$$y_n(t) = \sum_{m=1}^M \sum_{k=0}^{K-1} w_{nm}(k) x_n(t-k), \quad n=1, \dots, N \quad (6.3)$$

从式 (6.3) 可见卷积混合盲源分离问题的时域方法需要估计的系数较多, 具有较大的计算复杂度。

### 6.2.2 频域盲源分离算法

频域盲源分离的流形如图 6.2 所示。通过短时傅立叶变换 (STFT) 把卷积混合的观测信号变换到频域中, 得到每个频率点上的瞬时混合

$$X(f, t) = A(f)S(f, t) + E(f, t) \quad (6.4)$$

这里  $X(f, t) = [X_1(f, t), \dots, X_M(f, t)]^T$ ,  $S(f, t) = [S_1(f, t), \dots, S_N(f, t)]^T$ ,

$$A(f) = \begin{bmatrix} A_{11}(f) & \dots & A_{1N}(f) \\ \vdots & \ddots & \vdots \\ A_{M1}(f) & \dots & A_{MN}(f) \end{bmatrix}, \quad E(f, t) = [E_1(f, t), \dots, E_M(f, t)]^T, \quad \text{其中 } S_n(f, t)、$$

$X_m(f, t)$ 、 $E_m(f, t)$  和  $A_{mn}(f)$  分别表示  $s_n(t)$ 、 $x_m(t)$ 、 $e_m(t)$  和  $a_{mn}(t)$  的短时傅立叶变换。

由于在每个频率点都是复值信号的瞬时混合, 故可分别用复值瞬时盲源分离算法求出分离矩阵  $W(f)$  以分离出源信号。分离信号可以表示为

$$Y(f, t) = W(f)X(f, t) \quad (6.5)$$

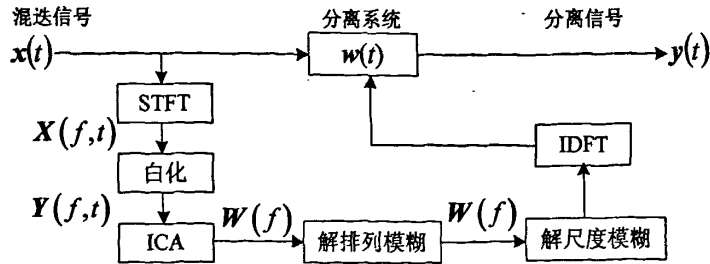


图 6.2 频域盲源分离流形

$$\text{这里 } \mathbf{Y}(f, t) = [Y_1(f, t), \dots, Y_N(f, t)]^T, \quad \mathbf{W}(f) = \begin{bmatrix} W_{11}(f) & \cdots & W_{1M}(f) \\ \vdots & \ddots & \vdots \\ W_{N1}(f) & \cdots & W_{NM}(f) \end{bmatrix}.$$

### 6.2.2.1 复值瞬时盲源分离

瞬时混合盲源分离问题的分离矩阵可通过一个两阶段的分离系统来求得, 首先对观测信号预白化,  $N \times M$  白化矩阵可表示为  $\mathbf{V}(f) = \Lambda_s^{-1/2}(f) \mathbf{E}_s^H(f)$ , 然后用独立分量分析 (ICA) 方法求得一个  $N \times N$  酉矩阵  $\mathbf{U}(f)$  实现源信号的盲分离。因此, 总体分离矩阵为  $\mathbf{W}(f) = \mathbf{U}(f) \mathbf{V}(f)$ 。

为获得稳定和快速的收敛, Araki 等人<sup>[20]</sup>用如下的复值 ICA 的批处理自然梯度学习规则

$$\begin{aligned} \mathbf{U}(f, t+1) &= \mathbf{U}(f, t) + \Delta \mathbf{U}(f, t) \\ \Delta \mathbf{U}(f, t) &= \mu_1 \left[ \text{diag} \left( \frac{1}{B} \sum_{t=B(N-1)+1}^{BN} \Phi(Y(f, t)) Y(f, t)^H \right) \right. \\ &\quad \left. - \frac{1}{B} \sum_{t=B(N-1)+1}^{BN} \Phi(Y(f, t)) Y(f, t)^H \right] \mathbf{U}(f, t) \end{aligned} \quad (6.6)$$

这里  $\mu_1$  是步长,  $B$  为块的大小。非线性矢量函数  $\Phi(\bullet)$  定义为

$$\Phi(Y(f, t)) = [\Phi(Y_1(f, t)), \dots, \Phi(Y_N(f, t))]^T \quad (6.7)$$

其中  $\Phi(Y_n(f, t)) = [1 + \exp(-Y_n^{(R)}(f, t))]^{-1} + j * [1 + \exp(-Y_n^{(I)}(f, t))]^{-1}$ , 这里  $Y_n^{(R)}(f, t)$  和  $Y_n^{(I)}(f, t)$  分别表示  $Y_n(f, t)$  的实部和虚部, 文献[14]表明这种类型的非线性函数适宜处理频域语音信号分离。

### 6.2.2.2 解排列模糊

目前用来解决排列问题的方法大致有以下几类, Parra 等人<sup>[15]</sup>基于频率平滑, 通过在时域限制滤波器长度来解决置换问题, 这种方法可有效分离在没有混响的环境中录制的麦克风信号, 但当混响存在时, 该方法的性能严重下降; Kurita 等人<sup>[16]</sup>和 Morgan 等人<sup>[17]</sup>基于阵列处理波达方向估计, 通过分析分离矩阵所形成的方

向图来估计波达方向,并根据波达方向进行置换;Ikeda 等人<sup>[18-19]</sup>利用相邻频率间的相关性来解决置换问题,即计算所有可能排列中相邻频率的频谱包络特性,使得相邻频率频谱包络相关性最大的排列被视为一致的排列进行置换;但由于该方法需要依次置换相邻频率,当某频率处盲源分离失效而得不到有效的分离矩阵时将导致误差累积,从而影响整个算法的性能。当各个频率处的盲分离性能都很好时,频率相关方法具有较高的精确度,而利用波达方向方法时一个频率上的失调不会影响到其他频率,因此具有稳健性的优点。虽然各国学者针对频率置换问题在以上三大类方法的基础上已提出很多频率置换方法,并且在特定的试验环境中都能获得较满意的性能,但在没有阵列流形等先验知识的情形下,很难既精确又稳健的解决频率置换问题,使分离出的信号中含有其它干扰信号分量,影响分离信号的质量。

用上述不同方法均可求得排列矩阵  $P(f)$ , 然后按下式修正分离矩阵

$$W(f) \leftarrow P(f)W(f) \quad (6.8)$$

### 6.2.2.3 解尺度模糊

解决尺度模糊问题的一个有效的方法是分裂语谱方法<sup>[18]</sup>,用该方法解决尺度模糊问题,对于  $M$  个传感器,  $N$  个源的情况会得到  $M \times N$  个分离信号,即每个传感器上得到  $N$  个分离信号,由推导可知分离信号是源信号与一未知滤波器的卷积。文献[18]中没有对其做进一步处理,直接作为源信号的估计。

Matsuoka<sup>[22]</sup>提出了最小畸变原理,尺度模糊可由下式解决:

$$W(f) \leftarrow \text{diag}(W^\dagger(f))W(f) \quad (6.9)$$

由此解尺度模糊后得到分离信号为

$$Z_n(f, t) = A_m(f)S_n(f, t), \quad n=1, \dots, N \quad (6.10)$$

由式(6.10)得到  $N$  个分离信号是用分裂语谱方法得到的分离信号中的一部分,即分离性能较好的主路径信号。且无需把每个传感器  $N$  个分离结果都求出来,与文献[18]相比减少了计算量。但两者的共同的缺点是得到的分离信号是源信号的滤波信号,且分离出的信号中可能会泄露进其它信号。

### § 6.3 基于多信道信号增强的后处理算法

#### 6.3.1 分裂的语谱与多信道信号增强模型的关系

为实现基于多信道信号增强的后处理,下面通过分裂语谱法<sup>[18]</sup>得到给出一种直观的推导。简单起见,考虑 $M=N=2$ 的无噪情形,对于 $M \geq N$ 的情形有类似结果,只是式(6.12)中 $W(f)$ 的逆阵相应变为伪逆。

假设排列模糊问题已解决,则分离信号为

$$Y(f,t) = W(f)X(f,t) = W(f)A(f)S(f,t) = D(f)S(f,t) \quad (6.11)$$

其中 $W(f)A(f) = D(f) = \text{diag}\{k_1(f), k_2(f)\}$ 表示对角的尺度矩阵,这里 $k_1(f)$ 和 $k_2(f)$ 是依赖于频率 $f$ 的标量。由(6.11)可得

$$W^{-1}(f) = A(f)D^{-1}(f) = \begin{bmatrix} A_{11}(f)/k_1(f) & A_{12}(f)/k_2(f) \\ A_{21}(f)/k_1(f) & A_{22}(f)/k_2(f) \end{bmatrix} \quad (6.12)$$

把以 $W^{-1}(f)$ 第一行为元素构成对角矩阵记做

$$A_1(f) = \text{diag}\{A_{11}(f)/k_1(f), A_{12}(f)/k_2(f)\} \quad (6.13)$$

则有

$$\hat{Y}_1(f,t) = A_1(f)Y(f,t) = \begin{bmatrix} A_{11}(f)S_1(f,t) \\ A_{12}(f)S_2(f,t) \end{bmatrix} \quad (6.14)$$

其中 $\hat{Y}_1(f,t) = [\hat{Y}_{11}(f,t), \hat{Y}_{12}(f,t)]^T$ ,  $\hat{Y}_i(f,t)$ 表示第 $i$ 个传感器接收到第 $j$ 个源信号。式(6.14)相当于把(6.4)中第一个传感器上接收的两个语音信号的混合分裂为两个分量,这个过程称为分裂语谱。同理有

$$\hat{Y}_2(f,t) = A_2(f)Y(f,t) = \begin{bmatrix} A_{21}(f)S_1(f,t) \\ A_{22}(f)S_2(f,t) \end{bmatrix} \quad (6.15)$$

其中 $\hat{Y}_2(f,t) = [\hat{Y}_{21}(f,t), \hat{Y}_{22}(f,t)]^T$ 。

由于解决尺度模糊后得到的分离信号是源信号的滤波,此外排列问题不可能做到完全精确,会有部分信号分量泄露进其它信号中,两者将引起频域盲源分离问题的分离性能的下降甚至无法分离,而一般的卷积混合盲源分离技术没有对这个问题做进一步处理。为提高分离质量,本文充分利用空域信息对该问题做进一

步处理,在增强源信号的同时消除泄露进该信号的其余信号分量和噪声。为此,把式(6.14)和(6.15)改写为

$$\begin{bmatrix} \hat{Y}_{11}(f,t) \\ \hat{Y}_{21}(f,t) \end{bmatrix} = \begin{bmatrix} A_{11}(f) \\ A_{21}(f) \end{bmatrix} S_1(f,t) \quad (6.16a)$$

$$\begin{bmatrix} \hat{Y}_{12}(f,t) \\ \hat{Y}_{22}(f,t) \end{bmatrix} = \begin{bmatrix} A_{12}(f) \\ A_{22}(f) \end{bmatrix} S_2(f,t) \quad (6.16b)$$

从式(6.16)我们可以看到还有空域信息可以利用,如(6.16)式第一个方程可视为用阵列接收源信号  $s_1(t)$ ,再把泄露进  $s_1(t)$  的另一信号  $s_2(t)$  的部分分量作为干扰信号,则为典型的多信道信号增强问题<sup>[23]</sup>,可以通过多信道信号增强算法分别重构出源信号。即通过分裂语谱,可以把多输入多输出(MIMO)的盲源分离问题分裂成两个独立的单输入多输出(SIMO)的多信道信号增强问题。

### 6.3.2 基于多信道信号增强的后处理算法

把式(6.16)推广到  $M \geq N$  的一般情况,则可以把  $N$  个输入  $M$  个输出的盲源分离问题分裂成  $N$  个独立的单输入  $M$  个输出的多信道增强问题,即

$$\hat{Y}^{(n)}(f,t) = \mathbf{a}_n(f) S_n(f,t) + \mathbf{Q}^{(n)}(f,t) + \mathbf{E}^{(n)}(f,t), n=1, \dots, N \quad (6.17)$$

这里用  $\hat{Y}^{(n)}(f,t)$  表示第  $n$  个源信号对应的阵列接收信号,同理下面用上标  $(n)$  表示处理第  $n$  个源信号。 $\mathbf{a}_n(f)$  表示混合矩阵  $\mathbf{A}(f)$  的第  $n$  列,  $S_n(f,t)$  表示第  $n$  个源信号的时频域表示,  $\mathbf{Q}^{(n)}(f,t)$  表示泄露进信号  $S_n(f,t)$  的其它残余信号分量,视为干扰信号,  $\mathbf{E}^{(n)}(f,t)$  表示加性白噪声。

下面分别解决这  $N$  个 SIMO 多信道信号增强问题,增强语音信号的同时消除干扰和噪声。利用 Gannot 等人<sup>[23]</sup>提出的传递函数-广义旁瓣相消(TF-GSC)的多信道信号音增强算法,该算法不需要直接估计传递函数,而是提出了传递函数比的概念,利用语音信号的非平稳特性,用次优的算法来估计传递函数比。然后利用频域广义旁瓣相消结构来消除空间干扰和噪声以实现信号增强;具体的算法推导详见文献[23]。由此得到的频域盲源分离解混及多信道信号增强后处理的流程如图 6.3 所示。

显然式(6.17)和 TF-GSC 算法的模型相吻合,有以下注释:

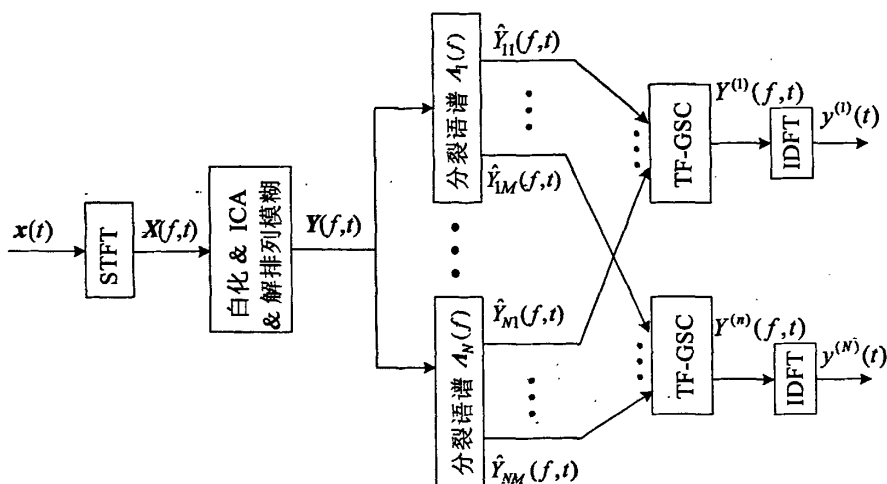


图 6.3 频域盲源分离混及多信道信号增强后处理的流程图

**注释1:** 式(6.17)中  $\mathbf{a}_n(f)$  与 TF-GSC 算法的传递函数相对应。TF-GSC 算法的传递函数是任意的，不要求特定的阵列流形，在求解过程中无需具体求出传递函数，只需求传递函数比；而盲源分离问题中只要能重构出源信号，可以不求出混合矩阵，即无需求出  $\mathbf{a}_n(f)$  的值。两者对传递函数的要求是一致的。这样既可充分利用空域信息，又不用增加新的先验信息，如参考信号等。对应于式(6.17)的传递函数比定义为  $\mathbf{H}^{(n)}(f) = \mathbf{a}_n(f)/a_{1n}(f)$ ，即用  $\mathbf{A}(f)$  的第  $n$  列元素除以该列第一行的元素；

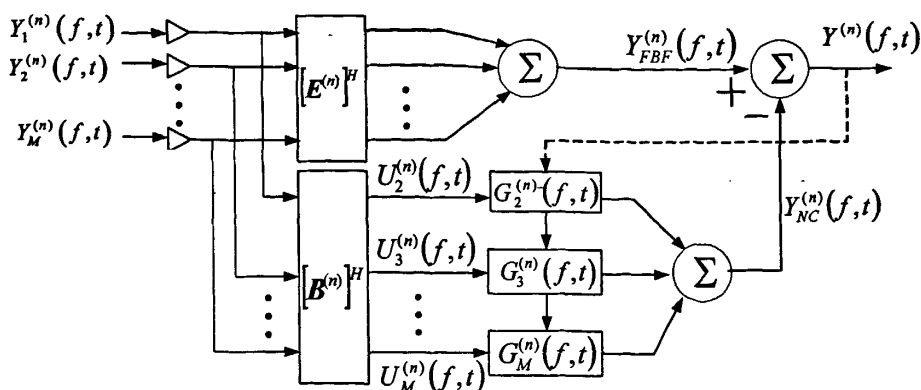
**注释2:** 文献[23]中的干扰信号既包括有方向的噪声分量，也包括环境噪声，式(6.17)中把泄露进该信号的其余信号的残余信号  $\mathbf{Q}^{(n)}(f,t)$  考虑为有向噪声，同时考虑环境噪声存在。

由上述分析可知，式(6.17)和 TF-GSC 算法的模型相吻合，因此对式(6.17)可用多信道信号增强算法作进一步处理，改进盲源分离问题中分离信号的性能。

对式(6.17)表示的 SIMO 的多信道增强问题利用 TF-GSC 算法重构第  $n(n=1, \dots, N)$  个源信号结构如图 6.4 所示。

TF-GSC 算法的算法步骤如下：

- (1) 固定波束形成器：  $\mathbf{Y}_{FBF}^{(n)}(f,t) = [\mathbf{H}^{(n)}(f)]^H \hat{\mathbf{Y}}^{(n)}(f,t) / \|\mathbf{H}^{(n)}(f)\|^2$ ；
- (2) 噪声参考信号：  $\mathbf{V}^{(n)}(f,t) = [\mathbf{B}^{(n)}(f)]^H \hat{\mathbf{Y}}^{(n)}(f,t)$ ；

图6.4 对应于第 $n$ 个源信号的TF-GSC结构

其中

$$\mathbf{B}^{(n)}(f) = \begin{bmatrix} -\frac{a_{2n}^*(f)}{a_{1n}^*(f)} & -\frac{a_{3n}^*(f)}{a_{1n}^*(f)} & \dots & -\frac{a_{Mn}^*(f)}{a_{1n}^*(f)} \\ 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ & & \dots & \ddots \\ 0 & 0 & \dots & 1 \end{bmatrix};$$

$$(3) \quad \text{输出信号: } Y^{(n)}(f, t) = Y_{FBB}^{(n)}(f, t) - [\mathbf{G}^{(n)}(f, t)]^H \mathbf{V}^{(n)}(f, t)$$

$$(4) \quad \text{滤波器更新: } \tilde{G}_m^{(n)}(f, t+1) = G_m^{(n)}(f, t) + \mu_2 V_m^{(n)}(f, t) [Y_n(f, t)]^* / P_{est}^{(n)}(f, t)$$

$$G_m^{(n)}(f, t+1) \xleftarrow{FIR} \tilde{G}_m^{(n)}(f, t+1)$$

$$\text{这里 } P_{est}^{(n)}(f, t) = \rho P_{est}^{(n)}(f, t-1) + (1-\rho) \sum_m |\hat{Y}_m^{(n)}(f, t)|^2$$

$$(5) \quad \text{利用叠接相加 (overlap-add) 法重构源信号 } y^{(n)}(t), n=1, \dots, N。$$

## §6.4 仿真实验

仿真实验用三个阵元接收两个统计独立的语音信号，即用语音信号与抽头为8的FIR滤波器卷积混合。为体现盲处理特性，本文仿真对阵列流形没有特别约束，这样在解决排列问题时需用[18]中不依赖于阵列流形的方法，这里分别取脉冲响应滤波器为服从均值为零方差为一的均匀分布。实验所用参数值见表6.1。

表6.1 仿真参数

参数	取值
语音信号采样率	16KHz
DFT 的长度	512
帧移	128
步长	$\mu_1 = 0.001 \mu_2 = 0.005$
遗忘因子	$\rho = 0.96$

本章用如下的性能指标来度量所提后处理算法的性能。盲源分离问题的基本的假设是源信号统计独立, 因此可以利用重构的信号矢量和源信号矢量的相似度(或相异度)<sup>[24]</sup>来测度分离性能, 理想的盲分离的结果要求第  $i$  个源信号与其重构信号有较高相似度, 与第  $j$  个源信号的重构信号有较强的相异度。这里把两个信号矢量夹角的余弦函数

$$d_{ij} = D(s_i(t), z_j(t)) = \frac{|s_i^T(t) z_j(t)|}{\|s_i(t)\|_2 \|z_j(t)\|_2}, i, j = 1, \dots, N \quad (6.18)$$

作为两个向量相似度(或相异度)的度量。本文所用两个信源号的相异度为 0.0024, 说明两信源可视为独立的。若能精确的分离出源信号, 则以  $d_{ij}$  为元素的矩阵  $D$  为广义置换矩阵。因此可用[12]中的性能指标(PI)

$$PI = \sum_{i=1}^N \left( \sum_{j=1}^N \frac{|d_{ij}|}{\max_k |d_{ik}|} - 1 \right) + \sum_{j=1}^N \left( \sum_{i=1}^N \frac{|d_{ij}|}{\max_k |d_{jk}|} - 1 \right) \quad (6.19)$$

来刻画算法的性能, PI 越小, 即第  $i$  个源信号与其重构信号的相似性越强, 与第  $j$  个源信号的重构信号的相似性越弱, 说明该算法的性能越好。

下面给出所提算法与不同的解尺度模糊算法的性能比较, 第一个算法是基于最小畸变原理<sup>[22]</sup>的方法, 第二个算法是分裂语谱<sup>[17]</sup>方法, 由于分裂语谱方法把每个阵列都分离出两个信号, 故在每个阵列上得到不同性能指标, 本文所提算法是在分裂语谱法的基础上, 利用多信道信号增强方法重构源信号, 图 6.5 是源信号、混合信号与用后处理增强方法重构的信号语谱图(不影响直观效果, 这里只给出其中两个混合信号的语谱图)。

图 6.6 是用分裂语谱方法得到的 6 个分离信号的语谱图, 如图 6.6 中箭头所示,

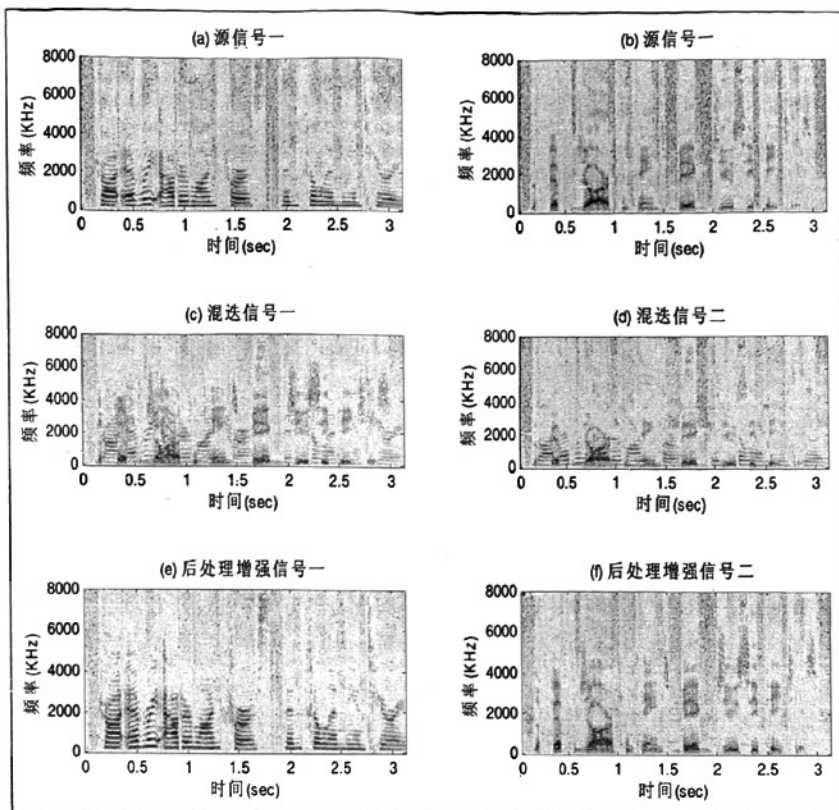


图 6.5 源信号、混合信号与后处理方法重构的信号的话谱图

很显然由于排列模糊导致一个信号语谱中含有另一信号的部分语谱,而图 6.5 中用后处理方法增强的语音信号则没有出现这个问题,说明本文提出的后处理方法明显改善分离结果。另外通过主观听取重构的语音信号,同样会发现未经后处理得到的分离信号中还会混杂着另一源信号的声音,而经后处理得到的分离信号则是较为纯净的语音。从图 6.7 可见,在不同信噪比下,本章所提算法的性能指标明显优于上述两种算法。

## §6.5 本章小结

频域盲源分离方法利用短时傅立叶变换把时域卷积混合盲分离问题转换为多个频域瞬时混合盲分离问题。此类方法可利用现有的瞬时混合盲源分离算法对各个频率分离问题进行独立处理。一方面大大降低了运算量,另一方面便于参考借鉴性能优越的瞬时混合盲分离算法。然而要想得到满意的分离结果,所有的频域盲分离算法都必须解决盲源分离问题固有主要问题:1)分离出的信号是源信号的滤

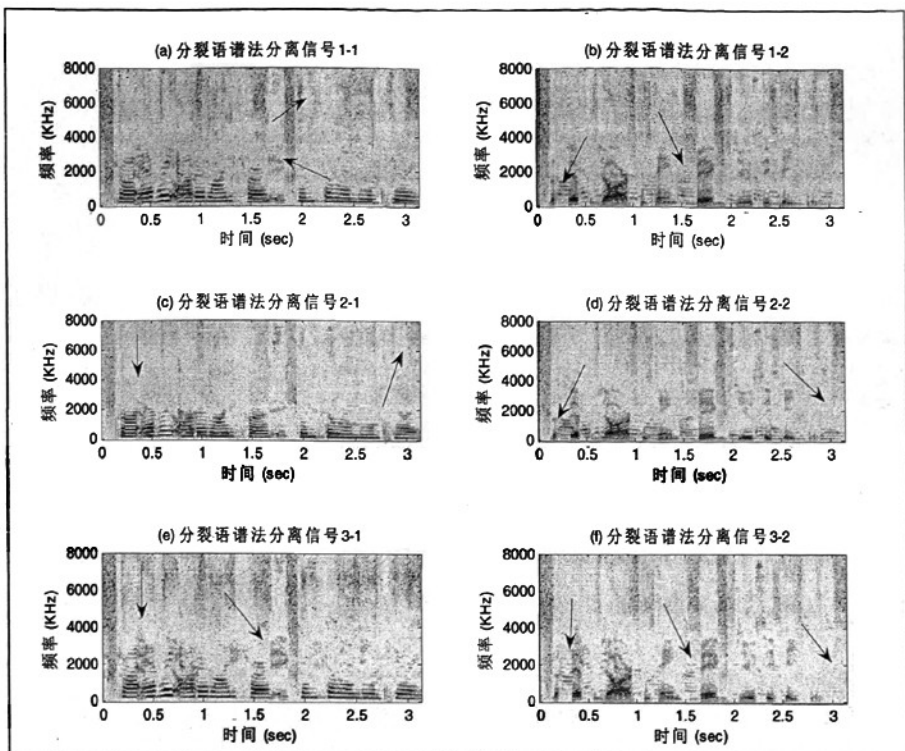


图 6.6 分裂语谱法得到分离信号的语谱图

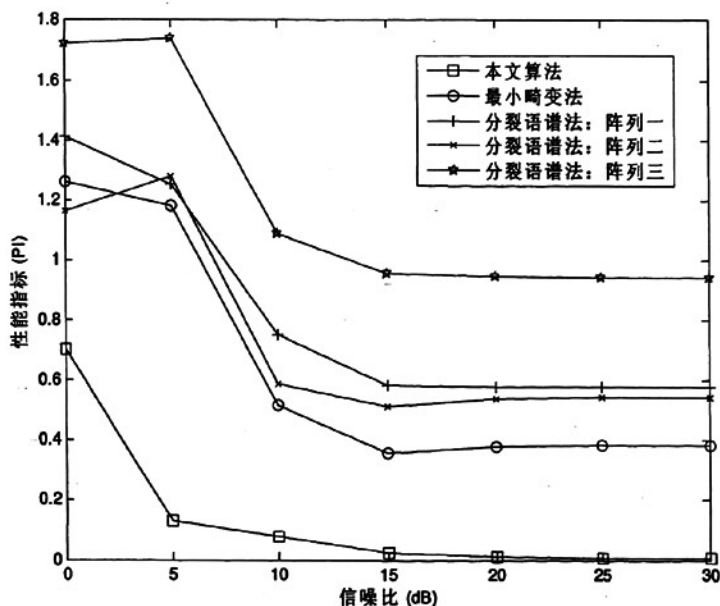


图 6.7 不同信噪比下的性能指标

波；2) 部分频率点处盲源分离方法失效；3) 无法精确解决排列和尺度模糊问题，这样分离出的信号中会泄露进其余的信号分量，影响分离性能。

本章提出一种基于多信道信号增强的频域后处理方法。本该方法有机结合盲源分离技术与阵列处理技术，既可充分利用空域信息，又无需增加其它先验假设。首先在频域盲源分离方法中利用语谱分裂技术得到  $M \times N$  个分离信号，可视为得到  $N$  个单输入多输出(SIMO)系统，泄露进该信号的其它信号分量可以视为干扰信号，然后分别用基于传递函数比和广义旁瓣相消结构频域多信道语音增强方法重构出各个语音信号，仿真结果表明该算法具有良好的性能，由语谱图可以说明本章提出的基于多信道信号增强的频域盲源分离后处理方法明显改善了排列不确定问题。另外通过主观听取重构的语音信号，同样会发现未经后处理得到的分离信号中还会混杂着另一源信号的声音。此外，其它有效的多信道后处理方法有待于进一步研究，如根据噪声分量和干扰分量的强弱分别进行噪声对消或干扰抑制等。

### 参考文献

- [1] Brandstein M, Ward D. Microphone Arrays: Signal Processing Techniques and Applications (Digital Signal Processing), Springer-Verlag, 2001.
- [2] Benesty J, Huang Y. Adaptive Signal Processing: Applications to Real-World Problems. Berlin, Germany: Springer, 2003.
- [3] Special Issue on for Signal Processing for Acoustic Communication Systems. EURASIP Journal on applied signal processing. 2003,11(5).
- [4] Special Issue on Multichannel Signal Processing for Audio and Acoustics Applications. IEEE Trans. Speech and Audio Processing. 2004, 12(5).
- [5] Special section on multi-channel acoustic signal processing. IEICE Trans. Fundamentals, 2005, 88(7).
- [6] Berouti M, Schwartz R, Makhoul J. Enhancement of speech corrupted by additive noise. in Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing, 1979, pp. 208-211.
- [7] Van Veen B D, Buckley K M. Beamforming: A versatile approach to spatial filtering. IEEE Trans on Acoust, Speech, Signal Processing, 1998, 5(2), pp. 4-24.
- [8] 张贤达, 保铮. 通信信号处理. 国防工业出版社. 2002年.
- [9] Haykin S. Unsupervised adaptive filtering. Volume1: Blind source separation John Wiley & Sons, Inc., 2000.

- [10] Cichocki A, Amari S. Adaptive Blind Signal and Image Processing. New York: Wiley, 2003.
- [11] Amari S, Douglas S, Cichocki A, Yang H. Multichannel blind deconvolution and equalization using the natural gradient. IEEE Workshop Signal Processing Advanced Wireless Communication, 1997, pp. 101–104.
- [12] Lambert R H, Bell A J. Blind separation of multiple speakers in a multipath environment. IEEE Int. Conf. Acoustic, Speech, Signal Processing, 1997, 1, pp. 423–426.
- [13] Torkkola K. Blind separation of convolved sources based on information maximization. IEEE Workshop on Neural Networks and Signal Processing, Kyoto, Japan, 1996, pp.423-432.
- [14] Smaragdīs P. Blind separation of convolved mixtures in the frequency domain. Neurocomputing, 1998, 22, pp. 21–34.
- [15] Parra L, Spence C. Convolutional blind separation of nonstationary sources. IEEE Trans. Speech Audio Processing, vol. 8, no. 3, pp. 20–327, May 2000.
- [16] Kurita S, Saruwatari H. Evaluation of blind signal separation method using directivity pattern under reverberant conditions. IEEE International Conference Acoustics, Speech, Signal Processing, 2000, pp. 3140–3143.
- [17] Ikram M Z, Morgan D R. A beamforming approach to permutation alignment for multichannel frequency-domain blind speech separation. IEEE International Conference Acoustics, Speech, Signal Processing, 2002, pp. 881–884.
- [18] Murata N, Ikeda S, Ziehe A. An approach to blind source separation based on temporal structure of speech signals. Neurocomputing 2001, 41, pp. 1–24.
- [19] Asano F, Ikeda S, Ogawa M, Asoh H, Kitawaki N. A combined approach of array processing and independent component analysis for blind separation of acoustic signals. IEEE Trans. Speech Audio Processing, 2003, 11(3), pp. 204–215.
- [20] Araki S, Mukai R, Makino S. The fundamental limitation of frequency domain blind source separation for convolutional mixtures of speech. IEEE Trans on Speech Audio Processing, 2003, 11(2), pp. 109–116.
- [21] Douglas S C, Gupta M. Scaled natural gradient algorithm for instantaneous and convolutional blind source separation. IEEE International Conference on Acoustics, Speech and Signal Processing, 2007, II, pp.637-640.
- [22] Matsuoka K. Independent component analysis and its application to sound signal separation. IWAENC2003, Kyoto, Japan, pp. 15-18, 2003.
- [23] Gannot S, Burshtein D, Weinstein E. Signal enhancement using beamforming

and nonstationarity with applications to speech. IEEE Trans on Signal Processing, 2001, 49(8), pp. 1614–1626.

[24] 张贤达 矩阵分析与应用 清华大学出版社 2004年9月.

## 第七章 结束语

### § 7.1 本文内容总结

本论文对非平稳环境中的盲源分离问题进行了探索性研究。这里“非平稳”有两重含义,第一,源信号是非平稳信号;第二传输信道是非平稳的,即混合矩阵是动态变化的。本文重点研究混合矩阵发生突变的情形,提出了利用自适应参数的不同在线算法跟踪时变环境。此外还研究了数目未知的语音信号瞬时混合的盲源分离问题及语音信号卷积混合频域算法中的后处理方法,主要内容可概括如下:

1. 首先提出一种变步长、变动量项因子的自然梯度算法实现源信号瞬时混合的盲分离。在后向传播(BP)算法的启发下,我们在自然梯度的学习过程中引入动量项以加快收敛速度,同时可以改善自然梯度算法的稳定性。然后,在分离模型中引入一个合适的测度函数来控制步长和动量项因子,由此得到的变步长、变动量项因子的自然梯度算法非常适合解决时变环境下的盲源分离问题。与经典的自然梯度算法及其它改进的自然梯度算法相比,即使在信源个数很多的情况下,自适应修正的自然梯度算法有更快的收敛速度和更好的稳态精确性,当混合矩阵突变或信号功率突变时自适应修正的自然梯度算法依然有较好的跟踪能力。此外,我们结合自然梯度和非线性主分量分析提出了一种块递归的盲源分离方法,构造出按块递归更新的矩阵方程,然后用QR分解和回代法求解该矩阵方程得到最优分离矩阵。与已有递归型盲源分离算法相比,所提方法适于实时处理且遗忘因子的选择相对简单,与其它块处理算法相比有较快的初始收敛速度。

2. 针对混合矩阵发生突变的情况,提出一种基于时变遗忘因子递归广义特征分解的非白源盲分离算法。首先给出一种协方差矩阵的逆矩阵和互相关矩阵乘积的递归更新方程,具有较低的计算复杂度,广义特征向量的估计经由近似幂迭代法和压缩处理在线实现,得到的在线算法可以避免同时估计协方差矩阵及其逆矩阵。为跟踪混合矩阵的突然变化,提出一种新颖的基于广义特征向量的在线决策规则,可以判别混合矩阵是否发生突变,从而利用变遗忘因子的递归广义特征分解算法改善在线盲源分离算法的性能,利用生物信号进行仿真的实验结果表明所提在线决策规则可以准确跟踪混合矩阵的突然变化,所用变遗忘因子的算法既能获得较好的跟踪能力又能得到精确的稳态性能。

3. 充分利用语音信号的非平稳特性,对于数目未知的语音信号瞬时混合的盲分离问题,在实现信号数目检测的同时完成信号的盲分离,并可克服盲提取问题

中先提取出的信源性能好而后提取出的信源性能差的缺点。首先利用递归广义特征分解在线估计的对应于最大广义特征值的广义特征向量,由向量的相似度定义平均相似度用以拟合信道互扰性能曲线,并根据平均相似度曲线提取出“高相似度区间”近似“分离区间”。然后进一步提取分离性能较好的“高相似度区间”并剔除部分“混合区间”,得到“高分离度区间”。然后利用多阶段聚类的方法确定源信号的数目,同时实现信号的盲分离或提取。仿真实现表明当信源数目适中时所提算法可以精确的估计出信源数目,发生错误估计或漏估信号的概率较低,并且用估计性能最好的广义特征向量作为分离向量,从而获得分离性能最好的语音信号。

4. 针对现有卷积混合盲源分离的频域方法存在的问题,即恢复出的信号是源信号和一个未知的滤波器的卷积,且由于部分频率点处盲源分离方法的失效和无法精确解决排列和尺度模糊问题,分离出的信号中会泄露进其余的信号分量,影响分离性能。提出一种基于多信道语音增强的频域盲源分离后处理方法以消除空间干扰和背景噪声。该方法有机结合盲源分离技术与阵列处理技术,既可充分利用空域信息,又无需增加其它先验假设。首先在频域盲源分离方法中利用语谱分裂技术得到  $M \times N$  个分离信号,可视为多输入多输出(MIMO)混合系统分裂成  $N$  个单输入多输出(SIMO)系统,泄露进该信号的其它信号分量可以视为干扰信号,然后分别用频域多信道信号增强方法重构出各个语音信号,以消除不同信源间的空间干扰和环境噪声。仿真结果表明基于多信道信号增强的后处理算法可以弥补由于排列模糊导致算法分离性能下降使得语音信号质量受损的情况。

## § 7.2 工作展望

在过去的近30年内,盲源分离问题的研究已经取得了长足的进步,在理论研究方面已经得到了有意义的结果,并且提出了许多有效的算法。但是,盲分离技术在实际的信号分离问题中的应用,则还需要进行大量艰苦的深入研究,也只有在盲源分离算法能够在信号处理领域得到较广泛的实际应用之后,才能实现盲源分离问题研究的真正意义。作者拟提出以下几个方面作为未来的研究方向以供探讨:

### 1. 弱信号检测和分离研究问题

在复杂电磁背景中,观测信号为大量信号的混合,且信源数目未知,而弱信号往往携带我们需要的信息。在信源数目很多的情况下,目前的并行盲源分离算法性能会严重退化。而在盲提取问题中,所提取信号的质量由于压缩处理的误差累积往往随着提取顺序下降。此外通常采用观测数据矩阵的奇异值分解和协方差矩阵的特征值分解来估计信源数目。在动态的环境中,需要在线算法跟踪环境

的变化。因此,在信源数目未知的信号分离问题中,对弱信号的检测与分离就成为我们需要解决的一个难题。

## 2. 低信噪比情况下的在线盲源分离问题

当前的在线盲源分离算法通常不考虑噪声,或者假定噪声非常小,以致可以忽略不计。但在实际应用中,观测信号总要受噪声的污染,某些情况下噪声还比较大,甚至强于信号。因此对噪声,尤其是对色噪声的抑制方法的研究是必须考虑的又一个难题,这个问题不解决,在线盲源分离的应用范围将受很大的影响。

## 3. 多信源卷积混合盲源分离问题

频率置换问题是解决频域盲源分离问题的最大障碍,且目前大多算法是在两信源情况下才能取得好的分离性能,因此研究两个以上语音信号卷积混合盲分离的频率置换问题,该问题被认为是盲源分离技术实用化道路上所必须完成的艰巨任务。为了在混响环境中分离出混合的语音信号,必须克服盲源分离技术和宽带阵列处理技术各自固有的缺点,将盲源分离技术和宽带阵列处理技术的优点有机结合起来,研究适合于非平稳噪声环境的语音增强方案,实现多个语音信号的分离,降低所提出算法的计算复杂性,增强算法的稳健性和实用性。

## 4. 欠定卷积混合盲源分离问题

目前研究欠定盲语音分离问题的算法,大都是利用语音信号在时频域中的稀疏特性,基于稀疏分量分析和聚类算法实现信号的分离,但是在实际环境中,语音信号在时频域中的稀疏特性并不是总是成立,在某些时频点出并不是只有一个语音信号存在,因此需要进一步利用语音特性、人类听觉感知特性及先进的聚类技术设计有效的欠定盲源分离算法并尽可能消除音乐噪声。

## 致谢

四年多的博士学习生活即将结束，回忆起这段美好的时光真的让人难忘！我的毕业论文能够顺利的完成，离不开许多良师益友的指导帮助、家人的鼓励和支持。在此，我向他们表示衷心的感谢！

首先，我要感谢我的导师冯大政教授。论文的完成离不开冯老师对我在研究方向上的指引、研究方法上的指导和研究内容上的指正。冯老师以其渊博的知识、敏捷的思维和非凡的洞察力，引导我逐渐踏上了科学研究之路，并指导我对研究内容逐步深入和细致，同时还倡导我们不断扩大知识面。冯老师不仅教给我从事科学研究和思考问题的方法，而且以其正直的人品、严谨的治学态度、实事求是的工作作风和忘我的工作精神为我们树立了一位优秀科研工作者的榜样；他的许多优秀品质将是我取之不尽、用之不竭的宝贵财富。

感谢中国科学院院士保铮教授，有幸加入由保老师领导并壮大的雷达信号处理国家重点实验室这个集体，这个优秀的团队深深吸引了我，也全方位地培养了我。保老师学术水平渊博，造诣极深。从他身上我感受到了我国老一辈科学家的崇高的思想境界以及对科学事业的无私奉献。他对科学事业的不懈追求，将是我一生学习的榜样！

感谢雷达信号处理国家重点实验室提供的良好学习和工作环境。感谢吴顺君教授、刘宏伟教授、水鹏朗教授、杨克虎教授、邢孟道教授给予我的指导和帮助。感谢实验室陈瑞林老师、李兰老师、白洁老师、张淑红老师、刘华锐老师给予的各方面的帮助和支持。

特别感谢我的师母张海琴副教授在生活上对我的无微不至的关心和照顾及教学工作上的大力帮助。

感谢同门师兄姐妹们：张斌、周祎、郑勇、武楠、刘宝泉、陈霖、吕雁、聂卫科、张华、张莉、李晓明、徐先峰、李涛、夏宇垠、吕辉、和洁、向聪、薛海伟、曹杨等，与他们的朝夕相处和学术上的交流讨论，都将成为我珍藏的回忆。

感谢我深爱的父母多年来对我的教育、支持和无限的关爱，感谢三位姐姐及她们全家这么多年对我的关心、鼓励和帮助。

最后，感谢我的爱人张伟伟女士对我真挚的爱和精神上的支持与鼓励，我能顺利完成学业，都是与她对家庭的无私奉献与分不开的。也感谢岳父母对我的关心、理解和支持。

作者： 刘建强

二零零九年四月于西安

## 作者在读期间的研究成果

1. **Jian-qiang Liu**, Da-zheng Feng and Wei-wei Zhang. Adaptive improved natural gradient algorithm for blind source separation. *Neural Computation*, 2009, 21(3), pp. 872-889. (SCI: 000264022900011)
2. 刘建强, 冯大政, 周炜. 基于多信道信号增强的卷积混迭语音信号盲分离的后处理方法. *电子学报*, 2007, 35(12), pp. 2389-2393. (EI: 20080411058653)
3. 刘建强, 冯大政. 一种按块递归的盲源分离方法. *西安电子科技大学学报*, 2008, 35(2), pp. 233-236. (EI: 20082111272494)
4. **Jian-qiang Liu**, Da-zheng Feng and Tao Li. Variable forgetting factor recursive generalized eigen-decomposition algorithm for blind source separation. Submitted to *Signal Processing* (欧洲).
5. 刘建强, 冯大政. 数目未知的语音信号瞬时混合盲分离与信源数目检测算法. 已投声学学报.
6. Yi Zhou, Da-zheng Feng, **Jian-qiang Liu**. A novel algorithm for two-dimensional frequency estimation. *Signal Processing*, 2007, 87 (1), pp. 1-12.
7. Yi Zhou, Da-zheng Feng, **Jian-qiang Liu**. A novel approach for DOA estimation in unknown correlated noise fields. *Journal of Electronics (China)*, 电子与信息学报 (英文版), 2006, 23(1), pp. 44-47.
8. 周炜, 冯大政, 刘建强. 一种新颖的 DOA 估计算法. *电子与信息学报*, 2005, 27(10), pp. 1513-1516.
9. 周炜, 冯大政, 刘建强. 二维频率估计的多阶段分解算法. *自然科学进展*, 2006, 16(1), pp. 79-85.