Vol 19 No. 1

文章编号: 1001-0920(2004)01-0073-04

基于小波域的非平稳卷积混合 语音信号的自适应盲分离

楼红伟, 胡光锐

(上海交通大学 电子工程系, 上海 200030)

摘 要: 一些卷积混合信号的盲分离算法是迭代性的,不适于实时应用 为此提出一种基于小波域的算法,用于卷积 混合信号的自适应盲分离 对基于小波域的算法进行仿真,并与频域盲信号分离算法进行对比,结果表明所提出的算 法能提高盲信号分离的性能

关键词: 语音信号分离; 神经网络; 二进小波变换; 非平稳卷积混合 中图分类号: TP391.42 文献标识码: A

Adaptive wavelet domain algorithm for blind separation of nonstationary convolved speech mixtures

LOU H ong w ei, HU Guang rui

(Department of Electronic Engineering, Shanghai Jiaotong University, Shanghai 200030, China Correspondent: LOU Hongwei, Email: louhw@sjtu edu cn)

Abstract A divide-and-conquer strategy is implemented by using the wavelet transform. Based on the wavelet domain algorithm, blind separation of non-stationary convolved speech mixtures is performed. And better performance is obtained compared with the frequency domain algorithm.

Key words: speech signal separation; neural network; dyadic wavelet transform; non-stationary convolved mixtures

1 引 言

盲信号分离是根据观测到的混合数据向量确定 一种变换,以恢复原始信号或信源 许多有关盲信号 分离的研究只对瞬时信号的线性混合进行简单处 理^[1,2].由于信号源与传感器之间存在不同的传输通 道,源信号之间的相对延迟就会发生,这种效应必须 用滤波器而不是标量矩阵来模型化 这是卷积混合 的情形,属于非线性混合模型下的非线性盲信号分 离 在这种情况下,大多数对策都假设信号源之间统 计独立,是基于瞬时混合的自适应算法的时域扩 展^[3].一些算法对此都有分组形式,但这种形式不能 严格处理实时闭环方面的问题 由于频域基分量是 正交的,在频域下优化分离滤波器较之时域更加容 易,计算收敛速度更快,且有更好的分离效果 但由 于具体实现时短时傅立叶变换的时频分辨率限制, 使它在非平稳信号分析中难以发挥更大的作用

为克服这种障碍, 人们努力寻找新的时频分布 形式 小波理论采取多分辨分析的思想, 非均匀地划 分时频空间, 使信号能在一组正交基上进行分解, 为 非平稳信号的分析提供了新的途径 本文提出一种 用于卷积混合语音信号自适应盲分离的小波域算 法 该方法比传统的方法有更好的性能, 更适合于实

收稿日期: 2002-08-25; 修回日期: 2002-10-16

基金项目: 国家自然科学基金资助项目(69672007).

作者简介: 楼红伟 (1976—), 男, 浙江永康人, 博士生, 从事语音信号处理与识别, 神经智能等研究; 胡光锐 (1938—), 男, 上海人, 教授, 博士生导师, 从事语音识别, 神经网络等研究

时应用

- 2 基于小波域的非平稳卷积混合盲信号 分离
- 2.1 卷积混合盲信号分离

卷积混合盲信号分离问题一般可表示为

$$x_{i}(n) = \int_{j=1}^{N} a_{ij}(\tau) s_{j}(n-\tau), \quad 1 \quad i \quad M.$$

(1)

其中:M 是传感器个数,N 是信号源个数,L 是窗函数的长度 各种信号源的叠加系数矩阵设为 $A = {a_{ij}}$.假定M = N, a_{ij} 为未知的有限冲激滤波器

对于白化的或标准化的数据矢量进行盲信号分 离,通常要求进行如下处理:

$$\widetilde{x} = V x.$$
 (2)

其中: $E[\tilde{xx}^{T}] = I; V$ 为白化矩阵, 可表示为

$$V = Q \Lambda^{-1/2} Q^{\mathrm{T}}, \qquad (3)$$

式中 Λ 和Q分别为观测信号x的协方差矩阵的特征 值和特征矢量矩阵

卷积混合盲信号分离问题的解决方法是设法找 到一滤波矩阵,对线性标准化的矢量 x 进行变换,得 到尽可能多的统计独立输出分量 y. 然而这里存在 与卷积混合信号有关的两个问题:

第 1 个问题是比例缩放和交换的不确定性, 在 所有的盲信号分离中都存在这种情况 如果这种卷 积混合问题是通过时域来处理, 交换不确定性就会 变得很严重, 因为时域输出信号要求包含所有的同 种信号源时域分量

许多卷积混合信号的盲分离算法是迭代性质 的,不适合于实时应用;此外,未知的混合源有一个 缓慢变化或扰动的传播媒介,在时域上有非平稳起 伏变化 因此,要求有一个象小波窗那样的滤波器来 降低这种非平稳影响,使在所确定的一段有限长滤 波长度内信号是分段稳态的 这种分割和解决的策 略可通过自适应小波变换方法予以实现 这里提供 一种基于小波域的算法,用于卷积混合信号的自适 应盲分离

$$x(t) 的离散二进小波变换(DW T) 为DW T (2t, k) = $\frac{1}{2^{t}}$ x(t) $\Psi^{*}\left(\frac{t-k}{2^{t}}\right)$ dt =
x(t) $\otimes \Psi_{2^{t}}(t)$. (4)$$

其中: 2['] 是尺度因子, $\Psi'(t)$ 是小波函数 $\Psi(t)$ 的复共 轭 通过取不同尺度 *l*, 可进行母小波 $\Psi(\bullet)$ 的扩伸 从信号处理的角度看, DW T 小波变换可认为是一 个有着恒定的相对带宽, 即常 Q 的滤波器的输出结 果 每个滤波器的冲激响应为 $2^{-i}\Psi(2^{-i}t)$, 其带宽和 中心频率正比于 $1/2^{l}$. 原信号可重构如下:

$$x(t) = \mathbf{DW} \mathbf{T}(2^{t}, k) =$$

$$\int_{|z|=2^{t}} \frac{1}{2^{t}} \mathbf{DW} \mathbf{T}(2^{t}, k) \Psi \left(\frac{t-k}{2^{t}}\right) dk.$$
(5)

如果滤波是在小波域中进行,分离标准(独立和非线性)是在时域中完成,则交换不确定性和缩放比例的问题就不会存在基于小波域的分离算法如图 1 所示

在图 1 中, 令 $\tilde{x}_n(l,k)$ 和 $Y_m(l,k)$ 分别为尺度 l= 1, ..., L - 1 下输入和分离信号的小波变换细节 系数, 令 $\tilde{x}_n(L,k)$ 和 $Y_m(L,k)$ 分别为尺度 l= 1, ..., L - 1 下输入和分离信号的小波变换近似系数

第 2 个问题是白化影响,这种现象可通过循环 神经网络在小波域的应用得以避免 当*M* = 2 维时, 小波域递归神经网络模型中每个尺度下的输入输出 关系为

$$\begin{cases} Y_{1}(l,k) = \widetilde{X_{1}}(l,k) - W_{12}(l,k)Y_{2}(l,k), \\ Y_{2}(l,k) = \widetilde{X_{2}}(l,k) - W_{21}(l,k)Y_{1}(l,k). \end{cases}$$
(6)

$$\diamondsuit \quad W(l,k) = \begin{bmatrix} 1 & W_{12}(l,k) \\ W_{21}(l,k) & 1 \end{bmatrix},$$



图 1 小波域分离算法

© 1994-2010 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

$$Y(l,k) = \begin{bmatrix} Y_1(l,k) \\ Y_2(l,k) \end{bmatrix}, \ \widetilde{X}(l,k) = \begin{bmatrix} \widetilde{X}_1(l,k) \\ \widetilde{X}_2(l,k) \end{bmatrix},$$

则式(6) 等效为

$$Y(l,k) = (W(l,k))^{-1} X^{\widetilde{}}(l,k).$$
(7)

其中: Y 和 \tilde{X} 中的元素分别为y 和 \tilde{X} 的小波变换, l 为 尺度 分离矩阵W 可通过下面介绍的训练规则加以 修正

2 2 基于小波变换的分离算法

语音信号采用以帧为信号处理单位进行。假设 第*m* 帧第 *i* 个传感器收到的信号为 *x_i*(*n*), 第 *i* 个输 出的分离信号为 *y_i*(*n*), 令式(7) 中分离矩阵的修正 系数为

 $\Delta W_{ij}(l,k) = \Theta(l,k,y_j)\Gamma(l,k,y_i).$ (8) 其中 y_i 和 $\overline{y_i}$ 为该帧下的分离系数 $Y_i(l,k)$ 和 $Y_j^*(l,k)$ k) 的离散二进小波逆变换, 这里的 $Y_j^*(l,k)$ 为 $Y_j(l,k)$ k) 的共轭

在计算小波逆变换时, 会涉及到近似和细节分 量的重构问题 对于每个分离矩阵的修正, 可以简单 地写成如下形式:

 $W_{ij}(l,k) = W_{ij}(l,k) + \Delta W_{ij}(l,k).$ (9) 非线性函数 $\Theta(\bullet)$ 和 $\Gamma(\bullet)$ 可表示为

$$\begin{cases} \Theta(l, k, \overline{y_i}) = \text{DW T}(\theta(\overline{y_i})), \\ \Gamma(l, k, y_i) = \text{DW T}(Y(y_i)). \end{cases}$$
(10)

离散二进小波变换提供了一种有效处理非平 稳混合信号的数学工具 只要源信号有偶次的概率 分布函数,且非线性函数Θ(•)和Γ(•)为奇次,学习 方程(8)的收敛性便可得到保证 同时,以下列方式 对学习方程进行修正,可以显著提高分离性能

$$\begin{cases} \Delta W_{ij}(l,k) = \Theta(l,k,\overline{y_j}) \Gamma(l,k,y_i),\\ \Delta W_{ji}(l,k) = \Theta(l,k,y_i) \Gamma(l,k,\overline{y_j}). \end{cases}$$
(11)

其中: *i*= 1, ...,*N* - 1; *j* = *i*+ 1, ...,*N*. 根据修正的 学习算法(11), 有

$$\begin{cases} \Delta W_{12}(l,k) = \Theta(l,k,\overline{y_2}) \Gamma(l,k,y_1), \\ \Delta W_{21}(l,k) = \Theta(l,k,y_1) \Gamma(l,k,\overline{y_2}). \end{cases}$$
(12)

对于瞬时混合信号, 如果 Θ(•) 是线性函数, Γ(•) 是 某种奇次非线性函数, 则这种修正分离算法仍是很 有效的, 并且学习算法可简化为下列方式^[4]:

$$\Delta W_{12}(l,k) = \Theta(l,k,\overline{y_2}) Y_1(l,k),$$

$$\Delta W_{21}(l,k) = \Theta(l,k,\overline{y_1}) Y_2^*(l,k).$$
(13)

根据最大似然准则和最大熵准则,非线性函数 θ(•)的最优选择可表示为

7

$$\boldsymbol{\theta}(\mathbf{y}_i) = - \frac{1}{p(\mathbf{y}_i)} \frac{\mathrm{d}p(\mathbf{y}_i)}{\mathrm{d}y_i}, \qquad (14)$$

其中p(y)是输出数据的概率密度函数 尽管这与所研究的盲信号出发点相违背,因为这里盲信号分离本身并不知信号的概率密度函数,但只要数据的概率密度分布函数是偶函数, $\theta(\bullet)$ 是非线性奇函数,这种学习算法仍具有高效性 超高斯的信号对该函数的选择为 $\theta(y) = \tanh(y)$,而亚高斯的信号则选为 $\theta(y) = 2y + \operatorname{sign}(y)y^2$.本文在算法验证时取为亚高斯信号,并令Y(y) = y.

2.3 计算复杂性

考虑本文介绍的卷积混合模型,假设输入信号 和输出信号的个数相同,即有两个传感器,两个模型 输出,一帧取N个(2的幂次方)采样值 该算法的计 算量是由模型的输出和分离矩阵的更新决定的 假 设在进行输入信号的离散二进小波变换时取最大尺 度为L,则运用M allat 算法计算时要进行 8V(1- 2^{-L})次乘法计算 运用自回归神经网络进行加权优 化时循环一次有(6- 2^{2-L})N次乘法,最后进行二 进逆小波变换时需 4(1- 2^{1-L})N次乘法 如果进行 K次循环,则共需乘法次数为

 $12N - 2^{4-L}N + (6 - 2^{2-L})N K.$

如果有M 个传感器,则共需乘法次数为

 $6(M^2 - M)N + 2^{3-L}(M^2 - M)N +$

 $(3 - 2^{1-L}) (M^2 - M) N K.$

这表明该算法的计算量不大,具有实时应用性

3 仿真结果

对于本文提出的算法进行了计算机仿真,实验 采用Biorthogonal小波系 其主要特征体现在具有 线性相位性,该方法在信号重构中得到广泛的应用

首先利用前述算法进行盲信号分离比较实验 实验是在室内有背景音乐的情况下,让一个人说汉 语的数字"1"到"10"话筒之间、说话人之间以及话 筒与说话人之间相距都是 50 cm,四点形成一个正 方形.采样率为 16 kHz,分辨率为 16 bit,窗函数的 长度N 取为 256 对比实验采用文献[5]提出的基于 频域的卷积混合盲信号分离方法,比较结果如图 2 所示 从图 2(a)可以看出,用频域分离的信号有明 显的非相关信号干扰

其次进行小波域盲信号分离实验,采用与上一 实验相同布局的两个话筒同时录音,其中一个发数 字"1"到"5"的男音,另一个发数字"6"到"10"的女 音.用该算法得到的实验结果如图3所示可以看出 分离的效果较好,用耳机听取分离的两段语音,可以 感受到各自纯净的声音,背景噪声很小



4 结 论

本文提出一种基于小波域的非平稳卷积混合语 音信号的自适应盲分离算法 为消除交换不确定性 和缩放比例问题, 把滤波放在小波域中进行, 而分离 标准在时域中完成 为避免信号的白化效果, 在算法 优化中采用了自回归神经网络模型 计算机仿真表 明该算法适于实时应用, 同基于频域的盲信号分离 相比, 具有更好的分离性能

参考文献(References):

- 2

[1] Amari S, Cardoso J F. Blind source separation semiparametric statistical approach [J] IEEE Trans on Signal Processing, 1997, 45(5): 2692-2700

- [2] Pham D T. B lind separation of mixtures of independent sources through a quasimaximum likelihood approach
 [J] IEEE T rans on S ignal P rocessing, 1997, 45 (2): 722-725
- [3] Nguyen Thi L, Jutten C Blind source separation for convolutive mixtures [J] Signal Processing, 1995, 45 (1): 209-229.
- [4] 杨福生 小波变换的工程分析与应用[M] 北京: 科学 出版社, 2000
- [5] Back A D, Tsoi A C. Blind deconvolution of signals using a complex recurrent network [A] Proc of IEEE Workshop — N eural N ew orks for Signal Processing IV[C] IEEE Press, 1994 565-574

76