

人耳对于不同频率成分分离知觉的定量分析

包永强, 周晓彦, 赵力, 邹采荣
(东南大学无线电工程系, 南京 210096)

摘要: 文章中研究了知觉性声源分离的最基本的问题, 即听觉对于不同频率成分特征声源的分离知觉。通过三组听觉辨别实验, 定量地分析了不同频率成分特征声源与产生分离知觉的关系, 得到了描述这种关系的近似函数, 提出了听觉对于不同频率成分特征的分离知觉模型。通过听觉辨别实验得出: 两种频率成分的 On-set 和 Off-set 的偏离是产生分离知觉的主要原因之一; 谐波频率成分之间频率发生的偏移也是产生分离知觉的要因之一; 要实现环境下的知觉性声源分离系统, 必须研究各个特征量相互作用时产生的综合结果。

关键词: 声场景分析; 声源分离; 频率成分; 知觉模型

中图分类号: TN912.3 文献标识码: A 文章编号: 1000-3630(2004)02-0106-03

Quantitative analysis on perceptibility of source separation of different characteristic frequencies

BAO Yong-qiang, ZHOU Xiao-yan, ZHAO Li, ZOU Cai-rong

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

Abstract: This paper studies the basic problem of acoustic source separation, the separation of different characteristic frequencies. An approximate function is obtained by analyzing the relationship between different characteristic frequencies and the creation of sense separation. A source separation model of characteristics of different frequencies is established. Experimental results show that both excursion of on-set and off-set frequency components and excursion of harmonic frequencies are important causes of auditory separation.

Key words: auditory scene analysis; acoustic source separation; frequency components; perceptual model

1 引言

声场景分析是由 Bregman 在 1990 年通过他的同名论著提出的一个新理念^[1], 目的是通过声音信息去理解和把握周围环境的状况。由于它结合心理学去研究人类在各种背景下获取听觉信息的机理, 并通过计算声场景分析去模拟和实现这些功能, 对于开发新型的人机接口, 建立高性能的语音信号处理技术具有重要意义。声源分离问题是在计算机上实现声场景分析的关键问题, 因此是声场景分析的研究重点^[2]。在讨论声源分离的问题时, 我们可以从两个方面去考虑声源和声源分离。一方面可以从

信号处理的角度, 认为声源即是实际的物理发音体, 声源分离是指从实际存在的多个声源中接收到的混合信号中恢复出某路或某几路声源信号的信号检测与处理过程。另一方面也可以从听觉感知的角度, 定义知觉性的声源是人们知觉成(或认识成)一个音的声音, 它可以来自几个物理性的声源^[3,4]。知觉性的声源分离, 即根据人的听觉知觉特性进行声源信息的分离和提取的过程。以往的声源分离主要指物理性声源, 它需要物理性声源的具体性质的信息。知觉性的声源分离对于发音体的具体性质的信息依赖不大, 即使存在未知的声源, 系统也能得到满足人们听觉效果的处理结果。因此, 对于知觉性的声源分离的研究也具有重要的实用价值。

要在计算机上实现知觉性的声源分离, 首先必须了解人们对于声谱是根据什么基准去知觉声音和提

取声音信息的, 而关于这个问题的最底层的部分是, 当不同的频率成分满足什么条件时被知觉成不同的音。本文研究了声源分离的最基本的问题, 即听觉对于不同频率成分特征的分离知觉, 定量地分析了不同频率成分时间关系和谐波关系特征对产生分离知觉的影响, 得到了描述这种关系的近似函数, 提出了听觉对于不同频率成分特征的分离知觉模型。

2 方法

试听对象共 10 名, 男 5 名, 女 5 名, 均为大专院校学生, 身体健康, 无耳病史, 无噪声接触史, 无耳病家族史, 年龄在 24~35 岁之间。

实验在安静的实验室中进行。用耳机把由计算机产生的各种频率成分特征的声音按随机的顺序发送给试听者。每次同样的过程, 让听者听 3 遍, 让其判断是两个以上的音(分离知觉) 还是一个音(融合知觉)。

3 实验和结果

3.1 实验 1: 频率成分的时间偏离和分离知觉的关系

本实验的目的是测定几个频率成分的声音的起始上升时刻(On-set) 的偏离和结束下降时刻(Off-set) 的偏离和分离知觉的关系。

测试使用的音料是两个频率组成的音, 如图 1 所示。 f_1 是 250Hz, f_2 是 500Hz, 两种频率固定, 振幅相等。On-set 和 Off-set 的偏离分两种情况考虑, 一种是 On-set 偏离, Off-set 一致; 另一种是 On-set 一致, Off-set 偏离。实验时, On-set 或 Off-set 偏离按 5ms 的步幅变化, 偏移的方向如图 1 所示。

试听时, 音料有声部分的长度 V 和中间间隔无声部分的长度 S 分别取为: a) $V = 300 \text{ ms}$, $S = 300 \text{ ms}$; b) $V = 500 \text{ ms}$, $S = 500 \text{ ms}$; c) $V = 1000 \text{ ms}$, $S = 500 \text{ ms}$ 。

实验结果如表 1 所示。结果表明, 如果 On-set 和 Off-set 没有偏离, 则始终感觉是融合知觉。当 On-set 或 Off-set 偏离增大, 融合知觉的程度减少, 分离知觉的程度增加, 当超过阈值以后, 基本上呈现分离知觉, 而且 On-set 的偏离对分离知觉的影响比 Off-set 偏离的影响要大。另外从实验结果中也可以看出, 声音自身的长度和无声间隔的长度对分离

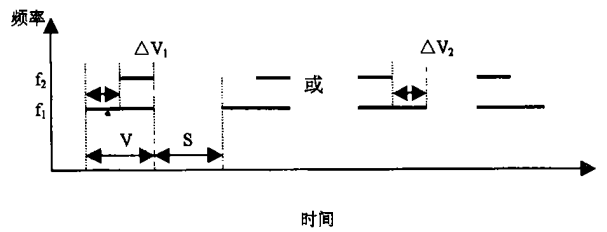


图 1 实验 1 的音料概念图

知觉的影响不大。为了建立融合和分离两境界点之间偏离和分离知觉的关系, 我们定义了判断分离知觉程度的 10 个等级, 等级越高分离知觉程度就越大。并通过利用最小二乘算法, 对落入这一区域的点, 进行曲线拟合估计可知, 它们基本上符合二次曲线。图 2 和图 3 是当 $V = 300 \text{ ms}$, $S = 300 \text{ ms}$ 时分别由 On-set 时间差 Δ_1 和 Off-set 时间差 Δ_2 引起的分离知觉实验的分离知觉模型曲线, 从这些曲线中我们可以得到描述这种关系的近似函数。

表 1 融合知觉和分离知觉的境界点

	On-set 偏离 Off-set 一致		On-set 一致 Off-set 偏离	
	$V = 300 \text{ ms}$ $S = 300 \text{ ms}$	On-set 时间差 Δ_1	融合 $\leq 8 \text{ ms}$ 分离 $\geq 18 \text{ ms}$	Off-set 时间差 Δ_2
$V = 500 \text{ ms}$ $S = 500 \text{ ms}$	On-set 时间差 Δ_1	融合 $\leq 8 \text{ ms}$ 分离 $\geq 27 \text{ ms}$	Off-set 时间差 Δ_2	融合 $\leq 34 \text{ ms}$ 分离 $\geq 47 \text{ ms}$
$V = 1000 \text{ ms}$ $S = 500 \text{ ms}$	On-set 时间差 Δ_1	融合 $\leq 9 \text{ ms}$ 分离 $\geq 19 \text{ ms}$	Off-set 时间差 Δ_2	融合 $\leq 29 \text{ ms}$ 分离 $\geq 58 \text{ ms}$

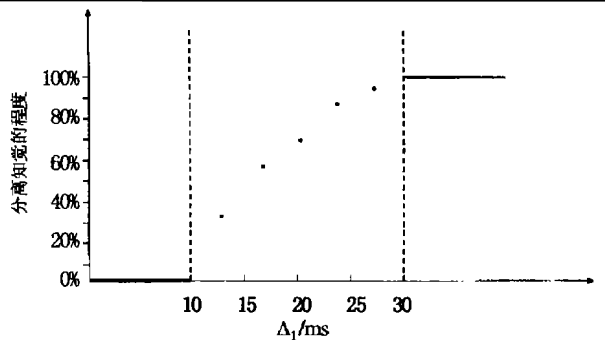


图 2 On-set 时间差 Δ_1 引起分离知觉实验结果

3.2 实验 2: 频率成分的谐波频率偏移和分离知觉的关系

本实验的目的是测定几个谐波频率成分之间, 当谐波频率发生偏移时, 偏离量和分离知觉的关系。测试使用的音料是 3 个频率成分组成的音。基

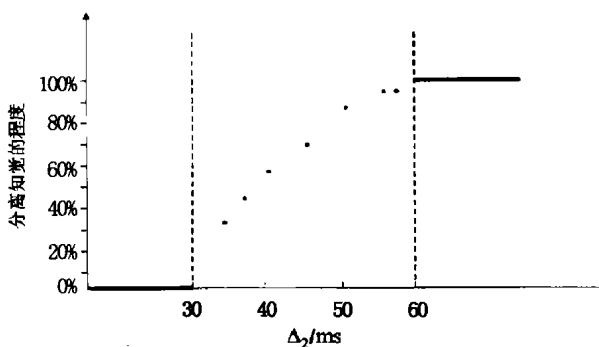


图 3 off-set 时间差 Δ_2 引起分离知觉实验结果

频 f_1 是 250Hz 以及 2 次谐波 f_2 和 3 次谐波 f_3 , 实验时基频 f_1 固定, f_2 和 f_3 以每次 1.25Hz 的步幅, 按以下 3 种形式变化: 1) f_1 、 f_2 一定, f_3 的变化范围为: $750 \pm 20\text{Hz}$; 2) f_1 、 f_2 一定, f_3 的变化范围为 $500 \pm 20\text{Hz}$; 3) f_1 一定, f_2 和 f_3 的变化范围分别以 $500\text{Hz} \pm 20\text{Hz}$ 和 $750\text{Hz} \pm 20\text{Hz}$ 。并且变化的形式为 4 种: a) f_2 、 f_3 同时向上变化; b) f_2 、 f_3 同时向下变化; c) f_2 向上, f_3 向下变化; d) f_3 向上, f_2 向下变化。

3 个频率成分的振幅相等, On-set 和 Off-set 相同, 持续时间均为 3000ms。为了说明实验结果, 首先定义如下两个参数 v_2 和 v_3 。

$$v_2 = \frac{f_2 - 2 * f_1}{f_1} \quad (1)$$

$$v_3 = \frac{f_3 - 3 * f_1}{f_1} \quad (2)$$

通过实验可知, 由基频以及 2 次谐波和 3 次谐波组成的声音, 若谐波频率成分之间没有频率偏移, 则始终是融合知觉。若有频率偏移, 则偏离增大, 融合知觉的程度减少, 分离知觉的程度增加。只有当 f_2 偏移且偏离量满足 $|v_2| \geq 0.03$ 时, 基本上呈现分离知觉。当只有 f_3 偏移且偏离量满足 $|v_3| \geq 0.04$ 时, 基本上呈现分离知觉。 f_2 和 f_3 同时偏移时, 偏移的方向不同, 则分离知觉的境界点不同。同方向偏移时间, 偏离量满足 $|v_2| \geq 0.03$ 或 $|v_3| \geq 0.04$, 逆方向偏移时间, 偏离量满足 $|v_2| \geq 0.015$ 或 $|v_3| \geq 0.02$, 基本上呈现分离知觉。整个分离知觉的分布区域如图 4 所示。另外, 从实验中 2 次谐波的偏移比 3 次谐波的偏移对分离知觉的影响大可知, 低次谐波的偏移对分离知觉的影响比高次谐波要大。

3.3 实验 3: 频率成分的时间偏移和谐波频率偏移共同存在和分离知觉关系

当时间的偏离和谐波的偏移同时存在时, 对分离知觉将产生什么影响是本实验调查的目的。从实

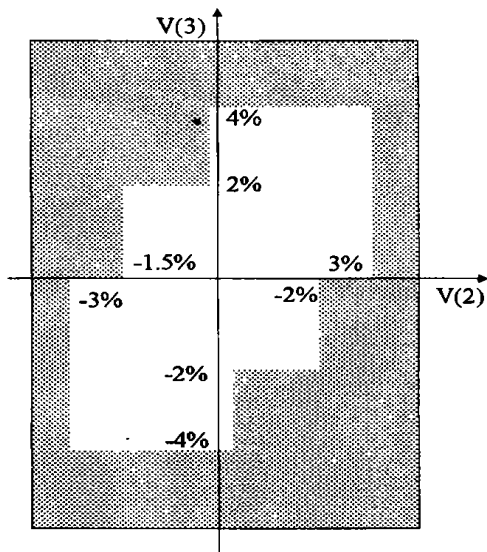


图 4 谐波频率偏移和分离知觉的关系

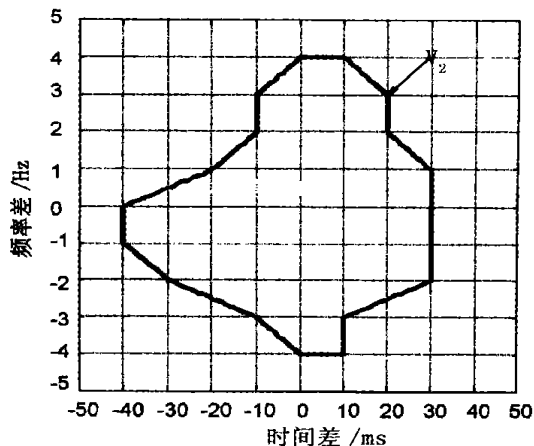


图 5 时间和谐波频率偏移共存和分离知觉的关系

验 1 可知, 和 Off-set 的偏离相比, On-set 的偏离对分离知觉的影响较大。所以, 本实验仅测定 On-set 的偏离和谐波的偏移共同存在时和分离知觉的关系。

测试使用的音料是两个频率成分组成的音。 f_1 是 250Hz, f_2 以 500Hz 为中心, 按 2.5Hz 的步幅, 在 $500 \pm 20\text{Hz}$ 范围内变化, f_2 的 On-set 时刻, 以 f_1 为基准, 按 5ms 的步幅, 前后最大的变化是 $\pm 50\text{ms}$ 。 f_1 、 f_2 的 Off-set 时刻相同。

实验的结果如图 5 所示。图中的 v_2 即式 (1) 定义的 f_2 的偏离量。图中结果的闭曲线的外围部分表示基本上呈现分离知觉的区域。

从图中可以看出, 在实验 2 中不呈现分离知觉的一部分音, 由于受到 On-set 偏离的影响而呈现分离知觉。例如: v_2 在 0.01 附近应呈现融合知觉, 但在 On-set 偏离量的影响下, 在 -20 ms 以下或 30 ms

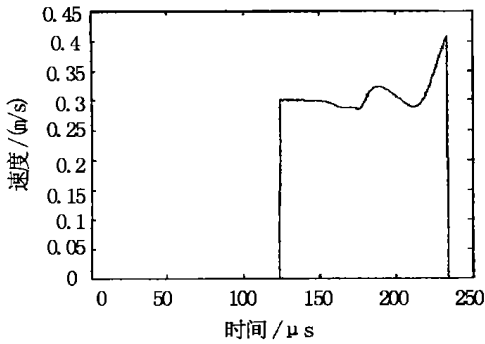


图9 速度测量结果(情形二)

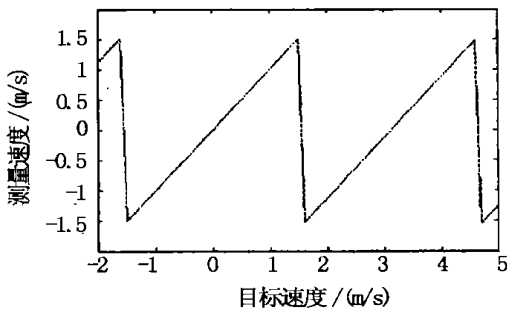


图10 测量速度曲线

4 结束语

(上接 108 页)

以上的区域却呈现分离知觉。同样在实验 1 中, $O_{\text{H-set}}$ 偏离量是 10ms 时, 应该呈现融和知觉, 但在 v_2 为 -0.03 以下或 0.04 以上的部分却呈现分离知觉。总之, 时间的偏离和谐波的偏移共同作用时, 比各自单独作用对分离知觉的影响要大。

4 讨论

根据听觉心理学的研究结果, 不同频率成分的时间关系和谐波关系是产生分离知觉的主要特征。本文通过听觉辨别实验定量地分析了它们对分离知觉的影响。

首先, 单从不同频率成分的时间关系来看, 两种频率成分的 $O_{\text{H-set}}$ 和 $O_{\text{F-set}}$ 的偏离是产生分离知觉的主要原因之一, 当偏离超过一定阈值以后, 基本上呈现分离知觉, 而且 $O_{\text{H-set}}$ 的偏离对分离知觉的影响比 $O_{\text{F-set}}$ 偏离的影响要大。在我们的实验中, 没有考虑 $O_{\text{H-set}}$ 和 $O_{\text{F-set}}$ 的倾斜对分离知觉的产生的影响, 而且只考虑了一种 $O_{\text{H-set}}$ 和 $O_{\text{F-set}}$ 的偏离方向, 这些因素对分离知觉的影响将成为我们今后的研究课题之一。同时, 谐波频率成分之间频率

计算机仿真的结果较好地验证了本文对 BBA-DCP 的分析研究。但这个模型尚未考虑散射信号的统计特性, 还不能进行 BBADCP 测量误差的分析, 不少问题有待进一步研究。

参 考 文 献

- [1] Michael R Simpson. Discharge measurements using a broadband acoustic doppler current profiler [R]. Open-file report OF-1, California: United States: Geological Survey: 2001. 1-33.
- [2] Brumley. Broadband acoustic doppler current profiler [P]. United States Patent Number: 5,208,785, May 4, 1993
- [3] 沈允春. 扩谱技术 [M]. 北京: 国防工业出版社, 1995. 1-54, 284-288.
- [4] 张道平. 宽带多普勒测流仪信号特性分析 [J]. 海洋技术, 2001, 20(1): 78-82.
- [5] James Tsui 著, 杨小牛, 陆安南译. 宽带数字接收机 [M]. 北京: 电子工业出版社, 2002. 165-240.
- [6] 李建新, 刘乃安, 刘继平. 现代通信系统分析与仿真 [M]. 西安: 西安电子科技大学出版社, 2000. 86-115.
- [7] 鄢社锋, 马远良, 倪继平, 等. 一种海洋宽带噪声场数值模拟方法 [J]. 声学技术, 2003, 22(1): 30-32.

发生的偏移也是产生分离知觉的要因之一。频率的偏移量以及偏移的方向都是影响分离知觉的特征。各种分离知觉特征同时作用时, 比各自单独作用时对分离知觉影响要大。在一种特征作用下讨论另一种特征的作用阈值时, 往往要比单独作用时的阈值减小或降低许多。所以, 要实现实际环境下的知觉性声源分离系统, 必须研究各个特征量相互作用时产生的综合结果, 确定合理的分离规则, 这也是我们今后要着重研究的课题之一。另外在听觉辨别实验中, 实验环境、受试者的精神状态和判断标准等都会对实验结果产生影响。

参 考 文 献

- [1] Bregman A S. Auditory Scene Analysis [M]. USA: MIT Press, 1990
- [2] 吴锡扬, 张子喻, 李想, 赵力. 听觉场景分析的研究进展 [J]. 电路与系统学报, 2001, 6(2): 68-73.
- [3] Abe M, Ando S. Auditory scene analysis based on time-frequency integration of shared FM and AM [C]. ICASSP'98 (Vol. 4), 1998. 1580-1591.
- [4] XU J. Basic Investigation for Scene Analysis by Acoustic Signals [M]. Kyoto: Kyoto Institute of Technology, Japan, 2000.