

车载多位语音数据集

AISHELL-ASR0051



希尔贝壳 产品说明书 Copyright

A I S H E L L

目录

1 数据概述.....	2
2 录音语料.....	2
2.1 语料池的制作.....	2
2.1.1 录音文本的结果设计	2
2.1.2 语料池处理	3
3 发音人信息.....	3
3.1 发音人结构特征.....	3
3.1.1 性别比例	3
3.1.2 年龄比例	3
4 数据录制环境.....	4
4.1 录制环境.....	4
4.2 录制设备.....	4
4.3 录制方法.....	4
5 语音数据转写.....	5
6 数据文件目录.....	6
6.1 目录结构.....	6
6.2 命名规则.....	7
6.2.1 目录命名规则.....	7
6.2.2 设备命名规则.....	8
6.2.3 场景命名规则.....	8
7 版权声明.....	9

1 数据概述

AISHELL-ASR0051 智能驾舱语音交互数据集共计 3164 小时。录音语言为中文，录音地区为中国，共邀请 260 名录音人，年龄段为 7-55 岁，均在真实理想 one 车载环境内录制完成。

此数据库录音内容包含朗读（561H）、对话（2163H）和车载噪声（440H）三类，设备包含三类，共计 11 个点位：近讲为头戴（48kHz，16bit，210H），车机（16kHz，16bit，824H，4 拾音位），座位顶部（16kHz，16bit，846H，4 拾音位），车顶部（16kHz，16bit，424H），中控屏（16kHz，16bit，421H）。此数据库经过人工校验，文本句准确率在 95% 以上，可做语音识别、声纹识别等研究。

2 录音语料

2.1 语料池的制作

2.1.1 录音文本的结果设计

此数据库录音分朗读和对话，关窗场景每人分配 50 句，开窗场景每人分配 100 句，语料内容覆盖导航、媒体等领域。对话部分一组 2~4 人，其中主说话人 2 人，对话部分，一组对话约 10 分钟，对话内容主要为闲聊和自由交流。朗读和对话的细分类具体如下：

朗读：车控类，如空调。座椅等车载控件命令词

媒体音乐类：播放、查询音乐、听歌电台

联系人：拨打电话、查询联系人

导航类：POI

问询类：知识问答、语音搜索等

对话：人机闲聊、车内人自由对话

2.1.2 语料池处理

- 脱敏处理。删除政治敏感、个人隐私、色情暴力等内容。
- 删除 <, >, [,], ~, /, \, = 等符号，数字串部分除外。
- 删除含有中文和英文以外语言的内容。
- 统一格式。

3 发音人信息

3.1 发音人结构特征

3.1.1 性别比例

数据库总人数为 260 人，男 118 人，女 142 人。

性别	男性	女性	合计
比例	45.4%	54.6%	100%

图表 3-1-1

3.1.2 年龄比例

录音人年龄覆盖 7-55 岁，具体分布如下所示：A(7-17 岁)；B(18-25 岁)；C(26-40 岁)；D(41-55 岁)。

	年龄段	人数	比例
A	7-17 岁	25	9.6%
B	18-25 岁	122	46.9%
C	26-40 岁	47	18.1%
D	41-55 岁	66	25.4%
合计		260	100%

图表 3-1-2

4 数据录制环境

4.1 录制环境

真实车载环境，录音人白天以及夜晚，在停车、行驶以及开关窗等不同场景下录制，具体场景信息如下：

开/关窗场景	
序号	场景说明
1	停车，空调关闭、AEC 关闭
2	停车，空调中档、AEC 关闭
3	停车，空调高档、AEC 中档
4	非高速 (0~60)，空调关闭，AEC 关闭
5	非高速 (0~60)，空调中档，AEC 关闭
6	非高速 (0~60)，空调高档，AEC 中档
7	低速(0~40)，空调关闭，AEC 关闭
8	低速(0~40)，空调中档，AEC 关闭
9	低速(0~40)，空调高档，AEC 中档
10	中速(40~80)，空调关闭，AEC 关闭
11	中速(40~80)，空调中档，AEC 关闭
12	中速(40~80)，空调高档，AEC 中档
13	高速(80~120)，空调关闭，AEC 关闭
14	高速(80~120)，空调中档，AEC 关闭
15	高速(80~120)，空调高档，AEC 中档

表 4-1-1 场景说明

4.2 录制设备

录制设备覆盖头戴麦克风、驾舱内麦克风、间距为 5 厘米的线性阵列以及单麦。

4.3 录制方法

共计 11 个点位，分别包含录音人近讲，座位上方，中控屏以及车顶。设备采集位置示意图如下所以：



图 4-3-1 近讲头戴

录音人对话通过头戴麦克风进行录制，录音人的录音位置不固定，依据设定规则改变位置，以保证覆盖所有驾内的拾音区。



图 4-3-2 远讲录音点位分布

5 语音数据转写

数据转写人员根据所听到的音频写出内容，力求使文本内容与音频发音内容保持一致。

一般准则如下：

- 1) 转写的内容必须和听到的语音完全一致，不能多字、少字、错字。

2) 数字要转写为汉字形式, 如“一二三”, 而不是“123”。注意区分“一”和“幺”, “二”和“两”。

3) 音频中有英文发音的应写成相应的汉字或英文。具体分为以下几种情况:

网址中包含的所有的字母或单词, 均为大写。例如: 发音内容为“www.abc.com”, 应转写为“三 W 点 A B C 点 com”

发音中包含的英文单词, 转写时全部为小写。

发音中包含的英文字母, 转写时全部为大写。

对于一些大写专有名词, 或者一些英文缩写全部大写加空格, 例如: CEO、CCTV 等。

4) 标注内容的完整性要与实际发音一致, 不得删减。

6 数据文件目录

6.1 目录结构

数据目录树	
数据目录结构	
AISHELL-ASR0051.pdf	(数据库简介)
└─DOC	(文本说明文件)
├─spkrinfo.xlsx	(录音人信息)
└─SPEECHDATA	(数据文件夹)
├─M	(关窗场景)
├─A1	(录音场景)
├─dialogue	(录音形式)
├─030	(对话编号)
	DA02.wav
	DA02.TextGrid
	DL01C01.wav
	DL01C01.TextGrid
	DX01C01.wav
	DX01C01.TextGrid
.....	
├─reading	(录音形式)
├─Audio	(语音文件)

─P0059	(录音人编号)
P0059_A02_0001.wav	(音频文件)
─Label	(文本文件夹)
─P0059.txt	(音频转写文本)
─N1	(开窗场景)
.....	
─N2	(开窗场景)
.....	
─N3	(开窗场景)
.....	

6.2 命名规则

6.2.1 目录命名规则

<USAGE>/<Window_info>/<Scene_ID>/<Recording_info>/<Dialogue_ID>/<wav_ID>

e. g. SPEECHDATA/M/A1/dialogue/030/DA02.wav

目录	内容	备注
USAGE	SPEECHDATA	文件夹名称
Window_info	M/N1/N2/N3	开关窗场景
Scene_ID	A~L	录制场景编号
Recording_info	dialogue	对话录音
Dialogue_ID	0030	对话编号
wav_ID	DA02.wav	音频文件

图表 6-2-1

<USAGE>/<Window_info>/<Scene_ID>/<Recording_info>/<File_type>/<Speaker_ID>/<wav_ID>

e. g. SPEECHDATA/M/A1/reading/Audio/P0059/P0059_A02_0001.wav

目录	内容	备注
USAGE	SPEECHDATA	文件夹名称
Window_info	M/N1/N2/N3	开关窗场景
Scene_ID	A~L	录制场景编号
Recording_info	Reading	朗读录音
File_type	Audio/Label	文件类型
Speaker_ID	P0059	录音人编号
wav_ID	P0059_A02_0001.wav	音频文件

图表 6-2-2

6.2.2 设备命名规则

录制涉及三类设备，共计 7 个点位，设备编号如下所示：

序号	编号	备注
1	A01~04、DA01~04	舒尔头戴、D 开头代表是对话
2	X01~06	车机，05 和 06 通道是回采信号
3	L01C01~C02	中控屏线阵，间距 50mm
4	L02C01~C02	车顶线阵，间距 50mm
5	S01~04	座位上方单麦

图表 6-2-3

6.2.3 场景命名规则

场景包含开关窗，车速、空调车机等设置，具体场景编号见下表：

N: 开窗场景 (三个开窗场景场景编号一致)	
场景编号	场景说明
A	停车，空调关闭、AEC 关闭
B	停车，空调中档、AEC 关闭
A C S	停车，空调高档、AEC 中档
D	非高速 (0~60)，空调关闭，AEC 关闭
E	非高速 (0~60)，空调中档，AEC 关闭
F	非高速 (0~60)，空调高档，AEC 中档

图表 6-2-4

N1: 主驾车窗开三分之一，天窗关闭

N2: 车窗关闭，天窗打开二分之一

N3: 主驾车窗开二分之一，天窗打开二分之一

M: 关窗	
场景编号	场景说明
A	停车，空调关闭、AEC 关闭
B	停车，空调中档、AEC 关闭
C	停车，空调高档、AEC 中档
D	低速(0~40)，空调关闭，AEC 关闭
E	低速(0~40)，空调中档，AEC 关闭

F	低速(0~40), 空调高档, AEC 中档
G	中速(40~80), 空调关闭, AEC 关闭
H	中速(40~80), 空调中档, AEC 关闭
I	中速(40~80), 空调高档, AEC 中档
J	高速(80~120), 空调关闭, AEC 关闭
K	高速(80~120), 空调中档, AEC 关闭
L	高速(80~120), 空调高档, AEC 中档

图表 6-2-4

注：场景编号后面的数字代表白夜，1 代表白天，2 代表夜晚

7 版权声明

本文内容禁止转载，AISHELL (北京希尔贝壳科技有限公司) 对本文拥有修改权、更新权及最终解释权。



希尔贝壳
A I S H E L L
A I S H E L L

Copyright