

# 声纹语音数据库说明书

**AISHELL-VPR0061**

北京希尔贝壳科技有限公司

**Beijing Shell Shell Technology Co.,Ltd**

Add: Room 3-621, 6F, Zhongguancun Lifangting No. 1, Shanyuan Road, Haidian District, Beijing 100080, P.R.China

Tel: +86 10 80225006 E-mail: [bd@aishelldata.com](mailto:bd@aishelldata.com)

目录

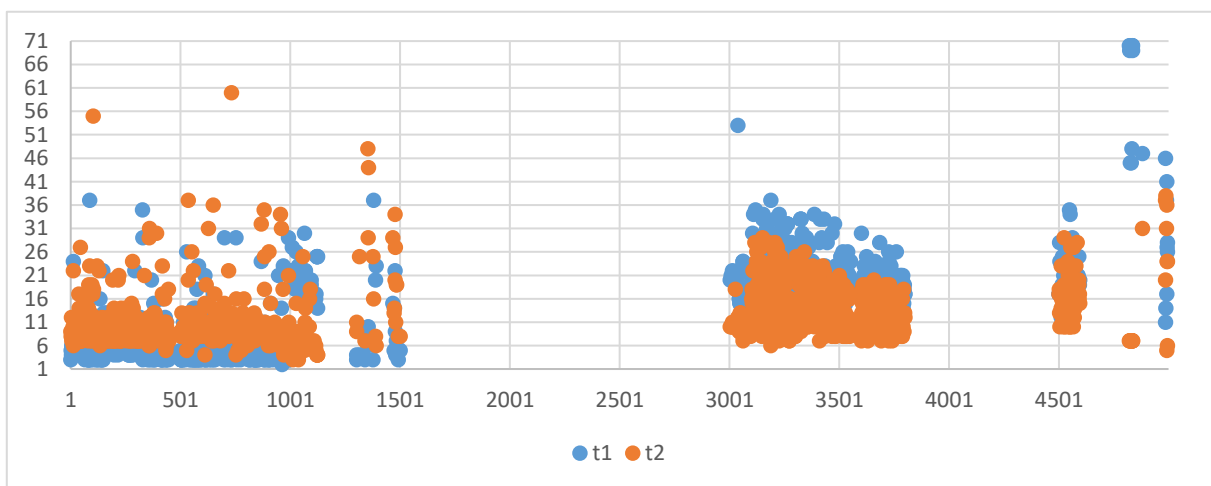
1 产品概述.....	3
2 产品目录结构.....	3
2.1 目录结构.....	3
2.2 命名规则.....	4
3 文本设计.....	4
4 版权声明.....	4

# 1 产品概述

AISHELL-VPR0061 语音数据库共 410 小时。录音语言为中文；录音地区为中国。共邀请 1300 名来自中国不同口音区域的发言人，每人在不同日期的早、午、晚三个时段分别录制 3 次语音，时间跨度为 7-78 天完成一个人的录制。录音人的年龄段为 9-65 岁，使用手机录制语音，覆盖常用手机机型。

此数据库可用于声纹识别等研究。

1300 人声纹三次录制时间间隔表，其中 t1 表示第二次和第一次录制的的时间间隔，t2 表示第三次和第二次录制的的时间间隔。



# 2 产品目录结构

## 2.1 目录结构

数据目录结构	
数据目录结构	
AISHELL-VPR0061.pdf	(数据库简介)
spk_info.xlsx	(发言人信息)
content.txt	(语音文本内容)
└─SPEECHDATA	(数据文件夹)
└─wav	(音频文件夹)

图表 2-1-1 数据目录结构

## 2.2 命名规则

AISHELL-VPR0061\SPEECHDATA\wav\

e.g. AISHELL-VPR0061\SPEECHDATA\wav\0001\0001\_F0001.wav

名称	内容	备注
<speakerID>	0001-4595	发言人编号
<wavID>	<speakerID>_<F/S/T><textID>.wav	音频名称
<F/S/T>	F-第一次录制, S-第二次录制, T-第三次录制	录制次数

图表 2-2-1 命名规则

## 3 文本设计

AISHELL-VPR0061 语音数据考虑到多场景工业用途, 录音分为三次收集, 分别为 F、S、T。三次的文本设计如下:

F 文本: 第一次录制。

每人 130 句, 内容为数字串, 例如“零幺一二三四五六七八九”, 每句 5~15 个数字, 平均 10 个数字。

S 文本: 第二次录制。

每人 160 句, 0001~0025 句内容为语控词+唤醒词, 0026~0050 句内容为覆盖安防、医药等敏感词汇自由文本, 0051~0160 句内容为广域自由文本。

T 文本: 第三次录制。

每人 130 句, 内容全部为广域自由文本。

## 4 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

**Copyright**