

# 模型生成中文情感合成语音数据库

AISHELL-TTS0073



希尔贝壳  
A I S H E L L

产品说明书

Copyright

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.....	3
3 发音人信息.....	4
4 版权声明.....	4



Copyright

# 1 产品概述

AISHELL-TTS0073 模型生成中文情感合成语音数据库共计 46.2 万条，共计 92 个发音人数据。合成引擎来自讯飞、火山、阿里和微软：讯飞（16kHz，16bit）；火山（24kHz，16bit）；阿里（16kHz/24kHz/48kHz，16bit）；微软（48kHz，16bit）。

## 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-TTS0073.pdf	(数据库简介)
└─spker-info.xlsx	(音色信息)
└─SPEECHDATA	(数据文件夹)
├─XF	(合成平台)
├─关山	(音色名称)
├─标准	(情感/风格分类)
0000001.pcm	(音频文件)
0000001.txt	(语音文本内容)
└─HS	(合成平台)

图表 2-1-1 数据目录结构

### 2.2 命名规则

CORPUS/USAGE/PLATFORM/SPEAKER/STYLE/SPEECH\_ID

e.g.AISHELL-TTS0073/SPEECHDATA/XF/关山/标准/0000001.pcm

目录名称	内容	备注
CORPUS	AISHELL-TTS0073	语音数据库编号
USAGE	SPEECHDATA	文件夹名称
PLATFORM	XF	合成接口平台
SPEAKER	关山	音色名称
STYLE	标准	风格或者情感分类

SPEECH_ID	0000001.pcm	WAV 文件
-----------	-------------	--------

图表 2-2-1 命名规则

### 3 发音人信息

发音人信息性别如下表格所示，共包含 92 个音色，覆盖各个平台基础之外，共计 30 个情感风格。每人每个情感风格 2000 条，每句不低于 5 秒。

性别	人数	占比
男	31	33.7%
女	54	58.7%
儿童	7	7.6%
合计	92	100%

图表 3-1-1 人员分布

### 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。

