



Connected Vehicles Platform & Product

Product Requirements Document (PRD)

Voice recognition

Revision History

Version	Date	Writer	Note	status
1.0	2019/11/14	Wang Jingjing	Draft	Unreleased
2.0	2019/11/21	Wang Jingjing	1.修改测试工况和测试场景，增加噪音要求 2.删除“纠错”功能（百度不支持） 3.删除四川话（百度不支持） 4.删除用户自定义非全时免唤醒时长功能，增加非全时免唤醒退出功能 5.修改“退出”指令定义 6.车身控制功能待补充 7.语音控制多媒体逻辑待更新	Unreleased

Table of Contents

1 简介.....	4
2 语音识别原理.....	4
3 语音识别技术方案及关联件.....	4
3.1 技术方案.....	4
3.2 麦克风参数要求.....	4
4 性能要求.....	4
4.1 性能指标依赖场景.....	5
4.2 其他性能要求.....	5
5 功能清单.....	5
5.1 功能定义.....	6
5.1.1 语音交互.....	6
5.1.2 TTS.....	9
5.1.3 语音控制.....	9
5.2 按键及设置项.....	12
5.2.1 按键.....	12
5.2.2 设置项.....	12
5.3 附件.....	12

1 简介

语音识别是将自然语言转换成机械语言，从而实现车身及系统的控制和操作，以达到解放双手，实现方便和快捷操作的目的。

2 语音识别原理

TBD

3 语音识别技术方案及关联件

3.1 技术方案

双音区语音识别采用双麦克风阵列式分布方案，麦克风布置要求：

1. 麦克风布置在人前方 300~500mm
2. 双麦克风间距 40~200mm
3. 尽量远离空调出风口等噪音源
4. 麦克风对称布置在 Y0 位置

备注：以上可根据具体项目进行微调，建议布置在顶灯位置

麦克风布置示意图如下：

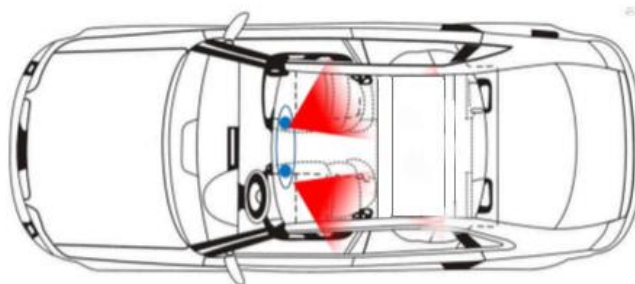


图 1 MIC 布置示意图

3.2 麦克风参数要求

麦克风要求：

1. 可采用双模拟或双数字麦克风，保证麦克风一致性
2. 尽量采用全指向麦克风
3. 麦克风信噪比 $\geq 60\text{dB}$
4. 麦克风灵敏度： $-3.5\pm 3\text{dB}$

备注：以上要求可根据具体项目进行微调。

4 性能要求

性能要求适用所有指令。

4.1 性能指标依赖场景

- 1. 基于云加端的混合识别模式进行评估；（须增加单独本地语音评估）
- 2. 网络环境 3G 及以上，上行不低于 100kbps，下行不低于 50kbps；
- 3. 百度提供测试集，验证指令不低于 2000 条，指令须覆盖全部垂类；
- 4. 测试人员口音基于标准普通话，正常说话状态，其中纯方言（粤语/四川话）和混合语言识别效果性能要求一致；
- 5. 百度提供测试报告，须加盖百度印章。

性能指标:

工况	唤醒率	误唤醒率	定位准确率	字识别率	响应时间
车速 0-40Km/h，空调关闭，安静环境	≥96%	≤3 次/81h	≥96%	≥96%	本地语音及在线语音（网络正常）≤1s
车速 40-80 Km/h，车窗半开	≥92%	≤4 次/81h	≥95%	≥92%	
车速 40-80 Km/h，空调低档，车窗关闭，噪音：70-80dB	≥90%	≤4 次/81h	≥95%	≥90%	
车速 80-120 Km/h，空调中档，车窗关闭，噪音：80-90dB	≥85%	≤5 次/81h	≥90%	≥85%	

4.2 其他性能要求

- 1. 不允许存在回声问题，即 TTS 播报音被收录的情况
- 2. 网络正常时，文字转写不允许存在延时问题（百度技术评估）

5 功能清单

序号	一级功能	二级功能	功能描述	备注
1	语音交互	语音唤醒词	用户可通过唤醒词唤醒语音，定制唤醒词“你好，福特”等，林肯系列：“你好，林肯”	
2		自定义唤醒词	用户可通过设置项手动输入，或者通过语音自定义唤醒词，例：“你好福特，以后你叫大白吧”	
3		非全时免唤醒	非全时免唤醒：唤醒后一段时间内无需再次唤醒	待确认垂类
4		情景化免唤醒	在特定场景下无需唤醒，直接实现语音控制	详见指令清单
5		全局免唤醒	在所有界面下无需唤醒，直接实现语音控制	详见指令清单
6		one-shot	唤醒+命令词，无需等待，例：“你好福特，播放林俊杰的歌”	
7		多轮交互	支持多轮交互	详见指令清单

8		上下文交互	支持上下文理解，例：“我想听林俊杰的歌”，“我想听他的江南”	垂类待完善
9		可见即可说	所有界面可看见的文字，均可语音控制，包含 APP 内部操作	须开放两个字的指令
10		全双工语音/打断	在语音助理进行播报时，用户可随时进行进一步的交互或说出新的命令，让助理执行相关操作，而无需等待播报结束；支持双方同时说，互不干扰	
11		纠错	支持纠错，例： “我想听周杰伦，不对，林俊杰的歌”	
12		语音退出	可通过语音“退下吧”等，暂停语音服务	
13	声源定位	/	抢麦功能，判定用户方位，排除车内其他干扰、显示定位方向	
14			本地/在线语音识别，将用户语音识别为文字	
15	自然语音识别	/	粤语/四川话方言识别，所有指令均支持	百度提供垂类
16			混合语言识别（中英文混合识别）（周三）	百度提供垂类和指标
17	语音播报	/	TTS 播报音选择	
18			交互信息、提示信息、百度百科、电子说明书等语音播报	
19	语音控制	/	语音控制主机功能（导航、音乐、电话、设置等）、空调等车身控制，语音小游戏等	
20	语音查询指令	/	生活服务、信息服务、股票、天气等信息语音查询（具体根据信源调整）	
21	智能家居	/	远程控制智能家居	

备注：功能支持情况及计划见附表。未特殊标注均为全局指令适用。

5.1 功能定义

5.1.1 语音交互

语音交互是语音识别的基础功能，除特殊要求外，所有通过语音实现的功能均应满足该要求。

1.唤醒

用户可以通过三种方式唤醒语音系统：

- a.硬按键，若整车配置有语音硬按键，则可通过硬按键唤醒语音系统
- b.软按键，用户可通过显示屏常驻软按键唤醒语音，软按键设计需经过福特确认通过
- c.唤醒词，用户可通过唤醒词唤醒语音系统

唤醒语音须遵循以下规则：

- a.无特殊要求时，语音系统唤醒后，须有提示音或提示语反馈用户，同时弹出对话框
- b.无特殊要求时，语音系统唤醒后，若 3s 内无声音输入，则自动退出语音系统，并有提示音或提示语

反馈用户，提示音后或者 TTS 播报完毕后立即关闭对话框

c.语音唤醒后，用户的任何非语音动作均可打断语音交互，例如操作方向盘控制按键，操作控制面板硬按键，显示屏软按键，调出空调界面，接打蓝牙电话、倒车等

语音唤醒：

用户可以通过唤醒词唤醒语音功能。

福特定制唤醒词：你好，福特/福特，你好/哈喽，福特/福特，哈喽/嗨，福特/福特，嗨（林肯系列：你好，林肯）。此外，DuerOS 系统也可通过“小度小度”唤醒语音。

在未自定义唤醒词的情况下，其余唤醒词均不能唤醒语音功能。

性能要求见 4-性能要求中性能指标。

2.自定义唤醒词

用户可以通过设置，自定义唤醒词。

自定义唤醒词方式共两种：

a.设置项

用户可通过设置项手动输入自定义唤醒词；同时可通过设置项清除自定义唤醒词，清除后自定义唤醒词开关自动切换至关闭状态。

b.语音设置

用户可通过语音设置唤醒词，例，语音唤醒后，指令“以后你就叫大白吧”，即可设置成功，具体指令可由供应商定义，指令集须释放福特确认。

自定义唤醒词须满足以下规则：

a.无论是否存在自定义唤醒词，定制唤醒词均可唤醒语音；

b.仅能存在一个自定义唤醒词，即用户自定义一个唤醒词后，除定制唤醒词外，须覆盖前一个自定义唤醒词。例，用户现有自定义唤醒词“大白”，重新定义唤醒词“小黑”，则“小黑，你好”可以唤醒语音，“大白，你好”不能唤醒语音；

c.自定义唤醒词最高支持 8 个汉字，不支持特殊字符、标点符号、敏感词汇、英文名称、数字、叠字。

3.非全时免唤醒

用户唤醒一次后，在一定时间内，用户可以随时说任何语音指令，无需再次唤醒，系统可正确响应。

非全时免唤醒语音须遵循以下规则：

a.非全时免唤醒功能可通过设置项设置打开和关闭

b.非全时免唤醒时长~~用户可自定义：30s、60s、90s~~，默认为 30s，即用户唤醒一次后，从唤醒开始计时，在 30s 内无需再次唤醒，可直接说语音指令

c.非全时免唤醒对话流满足语音退出功能，即语音退出动作可以终止语音服务

d.在免唤醒时间内，用户语音须转写文字并显示：

1) 接收到有效指令后，转写文字并直接执行相应动作

- 2) 接收无效指令后，转写文字，若满行则最前面文字前移消除，从接收到的最后一个字开始计时，1s 后无输入则文字自动消除
- 3) 接收无效指令后，从接收到的最后一个字开始计时，1s 内有输入，则在当前行继续转写文字，满行后，最前面文字前移消除

4.情景化免唤醒

在特定场景下，即仅在当前功能界面内，部分指令可直接执行，无需唤醒。福特 DuerOS 系统在导航界面（其余功能待补充），有关导航的部分指令支持免唤醒，详细指令清单见附表。（百度已提供，待更新确认）

5.全局免唤醒

无论当前是任何界面，指令无需唤醒词，可直接执行。其中部分指令需要该功能运行时才能执行，DuerOS 系统要求有关多媒体的部分指令支持免唤醒，详细指令清单见附表。（百度已提供，待更新确认）

6.全双工语音/打断

在语音播报的同时，监听用户指令并响应，即 TTS 播报时，用户可以进一步交流，直接打断播报，无需等待播报结束。

7.语音退出

可通过以下方式退出语音服务：

- a.硬按键：当语音在唤醒状态时，用户按下方向盘语音按键（短按），即可终止语音服务；
- b.软按键：当语音在对唤醒状态时，用户触摸显示屏常驻软按键，即可终止语音服务；
- c.语音指令：当语音在监听用户指令时，用户可通过语音指令终止语音服务，包含：
 - 1) 退出
 - 2) 退下吧
 - 3) 拜拜/再见

如有其他指令可以终止语音服务，须与福特确认。

8.特殊指令要求：

返回

无特殊要求时，系统接收到“返回”指令时，仅界面返回至上一层，直至返回到主界面，APP 退至后台继续运行；若有界面涉及音源切换，则直接切换至当前界面的音源继续播放；

取消

无特殊要求时，仅在对话流中需要用户选择“取消”时，该指令有效，其他界面该指令不做响应；

退出

~~若在 APP 内部，则退出 APP，同时关闭语音对话；当前无 APP 打开，则直接关闭语音对话；~~
仅退出语音服务；

5.1.2 TTS

TTS 有四种可选：标准女声，标准男声，情感女声，明星声音，默认标准女声。TTS 播报音需经过福特评审确认。

控制类指令 TTS 反馈须与控制指令完全匹配，例：“打开地图”，TTS：“地图已打开”；“打开导航”，TTS：“导航已打开”。

英文播报发音标准，包含中英文混合播报。

语音控制响应失败时：

- 1. 若是车身控制指令，例空调、座椅等的控制，则根据具体要求进行 TTS 反馈；
- 2. 若是 IVI 系统内部控制，例导航、随心听等的控制，除特殊要求外，不做 TTS 反馈。

5.1.3 语音控制

可通过语音控制系统内及车辆功能，包含多媒体、导航、蓝牙、空调等，详细指令列表见附件。（百度提供）。

1 语音控制 APP

所有的 APP 均可通过语音控制打开和关闭（退出）；支持深度控制的 APP 须单独表表明，具体控制范围及指令列表见附件。

序号	一级功能	二级功能	功能描述	备注
1	语音控制	随心听	在线多媒体：QQ 音乐、有声、新闻 本地多媒体：收音机、蓝牙音乐、USB 音乐 具体控制范围见指令列表清单	
2		导航	控制范围见指令列表清单	
3		爱奇艺	具体控制范围见指令列表清单	
4		外卖	支持语音控制支付，具体控制范围见指令列表清单	
5	查询及购买	电影票	控制范围见指令列表清单	
6	语音查询	/	见指令列表清单	仅支持查询

备注：

- 1) 用户通过语音搜索在线音乐时，须检测 QQ 音乐账户登录情况：若 QQ 音乐账号未登录时，不支持语音搜索歌曲，须直接反馈用户登陆账号，并跳转至账号登录界面；若 QQ 音乐账号已登录，则直接进行搜索；
- 2) 随心听仅代表 QQ 音乐，例：“打开随心听”，若当前是非 QQ 音乐，须切换至 QQ 音乐并随机播放；若是 QQ 音乐，则直接打开界面，并继续播放；

- 3) 在线多媒体及本地多媒体中的模块可通过语音单独控制打开和关闭，例：“打开有声”，则打开有声书并随机播放有声；

2 语音控制音乐逻辑

语音指令 当前状态	播放音乐、我要听歌等	播放蓝牙音乐等	播放 U 盘音乐等
非音乐播放	打开在线音乐并播放	打开蓝牙音乐并播放	打开 U 盘音乐并播放
在线音乐	打开在线音乐界面并继续播放	打开蓝牙音乐并播放	打开 U 盘音乐并播放
U 盘音乐	打开在线音乐并播放	打开蓝牙音乐并播放	打开 U 盘音乐并继续播放
蓝牙音乐	打开在线音乐并播放	打开蓝牙音乐并继续播放	打开 U 盘音乐并播放

备注：

- 1) 蓝牙未连接时，语音控制播放蓝牙音乐，直接跳转至蓝牙连接界面，并暂停当前音源；
- 2) 用户未明确音乐搜索路径时，默认搜索在线音乐；
- 3) 用户明确音乐路径时，则按照搜索路径，记忆上次播放状态进行播放。例：“我要听 U 盘音乐”，则 U 盘中断点记忆位置继续播放；
- 4) 当前非多媒体界面时
 - a) 用户明确音源信息时，包含收音机、在线音乐、有声、新闻，直接后台播放指定音源并 TTS 播报，界面保持当前界面不变。其中，蓝牙音乐和 U 盘音乐不支持音乐名称搜索，统一按照未明确音源信息处理。例：“我想听明朝哪些事”则后台播放有声“明朝那些事”，并 TTS 播报；
 - b) 用户未明确音源信息时，包含收音机、在线音乐、有声、新闻，则跳转至指定音源界面，直接播放。例：“我想听收音机”，则界面跳转至收音界面并播放；**FO 提出与音源 source 切换按键逻辑保持一致，该逻辑需要确认是否修改，福特内部同步后输出；**
- 5) 当前是多媒体界面时

无论用户是否明确音源信息，均需跳转至指定音源界面并播放。例：当前在收音机界面，“我想听歌”，则界面跳转至在线音乐界面并播放；
- 6) 以上音源切换须满足断点记忆要求，即当音源状态未发生改变时，语音切换音源且未明确音源信息，那么音源之间切换，须在上次记忆位置继续播放，例：当前播放在线音乐，语音控制打开爱奇艺并播放视频，此时语音指令“我想听歌”，响应结果：记忆上次在线音乐断点位置并继续播放。（与音源管理 FO 确认断点记忆）

3 音量调节

详见附件。



Voice control
volume spec-2011

4 车身控制

可通过语音控制空调、座椅加热、座椅通风、座椅按摩、氛围灯、天窗，除特殊要求外，车身控制功能需按照原有功能既定逻辑执行，即仅作控制，不做逻辑判断。（待福特提供列表清单）

功能	描述	备注
空调	所有可通过 soft button 控制的空调功能，均可通过语音控制	
座椅加热	可通过语音控制座椅加热开关	仅控制开关，不做具体功能调节
座椅通风	可通过语音控制座椅通风开关	仅控制开关，不做具体功能调节
座椅按摩	可通过语音控制座椅按摩开关	仅控制开关，不做具体功能调节
氛围灯	可通过语音调节氛围灯	
天窗	可通过语音控制天窗	
方向盘加热	可通过语音控制防线盘加热开关	仅控制开关
熏香模式	待确认	
待确认		

车身控制须根据车身控制需求更新，须做在线配置。

备注：

➤ 空调控制要求：

- 1) 语音调节空调后，须有对应的 TTS 反馈，弹框提示须按照空调既定逻辑执行，不做特殊要求；
- 2) 风向调节可单独控制打开或关闭，也可两两组合同时控制打开或关闭；
- 3) 可通过语音直接调节到指定风量等级；未指定风量等级时，以 1 个风量等级为单位进行步进调节，直至最大/最小值；可通过语音直接调节风量增大/减小等级，例：当前风量为 1 挡，语音“风量调大 3 挡”，则调节后风量应为 4 挡；
- 4) 可通过语音直接调节到指定温度值；未指定温度值时，以 2 个温度值为单位进行步进，直至最大/最小值；可通过语音直接调节温度增大/减小值，例：当前温度为 18℃，语音“温度增大 3℃”，则调节后温度应为 21℃；
- 5) 温度调节：

- 同步模式打开时：

若用户指定调节区域，则仅调节指定区域温度，同时关闭空调同步模式。例：若当前双区温度为 18℃，指令“主驾温度调到 22℃”，则调节主驾温度到 22℃，副驾温度保持不变；

若用户未指定调节区域，则同步调节。例：即若当前双区温度为 18℃，指令“温度调到 22℃”，则主副驾同时调节到 22℃；

- 同步模式关闭时：

若用户指定调节区域，则调节指定区域温度。例：主驾发出指令“副驾温度调节到 22℃”，则仅调节副驾温度到 22℃；

若用户未指定调节区域，则关联声源定位，调节发声区域温度。例：主驾发出指令“温度调到 22℃”，则仅调节主驾温度到 22℃

- 天窗控制
主机通过 LIN 与天窗模块交互，具体控制逻辑见《voice control sunroof & sunshade spec》（仅 CD542 项目，其余做在线配置区分处理，整理后输出至百度）。
- 香薰控制
具体控制逻辑整理后发出

5.2 按键及设置项

5.2.1 按键

	操作	响应
方向盘按键	短按	唤醒/停止语音服务
	长按	无效

5.2.2 设置项

	描述	备注
允许语音唤醒	语音唤醒开关，打开后可通过定制唤醒词唤醒语音，默认开	
自定义唤醒词	自定义唤醒词开关，打开后可通过自定义唤醒词唤醒语音，且可编辑自定义唤醒词；关闭后自定义唤醒词置灰，且无法编辑，默认关	
语音播报	可切换 TTS 播报音	
非全时免唤醒	非全时免唤醒开关，打开后可使用非全时免唤醒功能，并可设置免唤醒时长： 30s、60s、90s ，默认关，打开后默认时长 30s	
One-shot	One-shot 开关，打开后用户可以直接唤醒+指令，无需等待 TTS 反馈，默认开	

备注：

语音设置项默认值均为出厂设置值，用户自定义语音设置项后，打开和关闭设置项开关以及主机休眠开机，不影响用户自定义状态。

5.3 附件

语音功能及指令清单（非最终状态，须更新）。

