

# 婴儿哭声识别可行性分析

## 婴儿哭声识别可行性分析

- 一. 婴儿哭声需求调研
  - 1 用户调研
    - 1.1 用户调研
    - 1.2 竞品调研
- 二、科学合理性
  - 2.1 经验常识
  - 2.2 专业研究
  - 2.3 综合判断
- 三.问题与发现
- 四.总结

## 一. 婴儿哭声需求调研

婴儿哭声是婴儿监护中一个重要领域，哭闹是宝宝表达情感和寻求帮助的主要方式，也是一种健康的表现。就像在告诉父母：“我需要你！”，如果父母能够理解宝宝的需求并及时解决，会对宝宝的成长发育很有好处。现实中的年轻父母有“辅助破译哭声”的需求么？请看下图：

**吴建成问：宝宝怎么了？在固定的时间哭**

连着三天了，我家宝宝总在快八点到九点多一点的时候哭，哭的特别厉害，我都不知道她怎么了，要是什么地方..

光头嫻嫻答：我的宝宝是晚上哭，我去 给他看了，说是肚子涨，开了一些健脾的药，现在好了一些了，你不妨也带他去看看啊，或者也是哪里不舒服呢？他不会说的，所以要多留意的。

[查看其他10个回答](#)

**雪儿飞舞问：宝宝撕心裂肺的哭是怎么回事？**

我家宝宝三个月差几天，昨天晚上我给宝宝换了尿不湿准备睡觉，给她喂奶的时候开始哭，奶也不吃，直哭得嘶..

星星烁答：没事，看不哭时的饮食精神状态睡眠等情况，如果好，不要担心

[查看其他8个回答](#)

**周先生生问：我的宝宝整天哭**

我的宝宝整天哭怎么办呢？好心痛啊

新妍的脚印答：由于婴儿的肠胃发育还不完善，经常会因消化不良产生腹胀、胀气等，这些不舒服的症状最终导致婴儿的哭闹不安。按新妍妈妈的经验，解决的方法有两个：1、喂少量的益生菌类药物（如妈咪爱）帮助婴儿改善肠胃功能，..

[查看其他7个回答](#)

**真想有个哭声翻译器，宝宝不知道怎么了**

### 1 用户调研

#### 1.1 用户调研

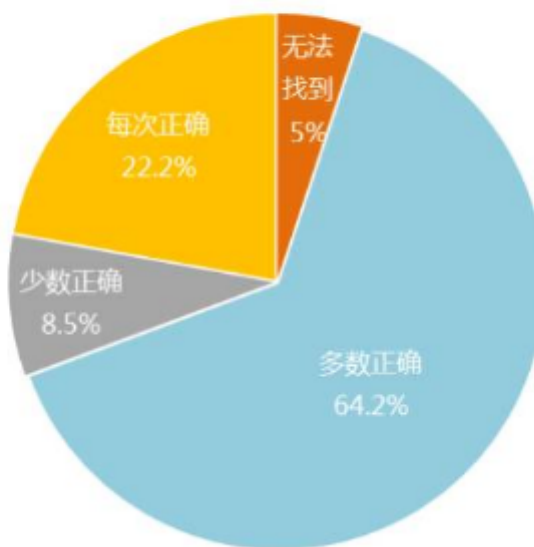
为了迅速获取更多信息，牺牲精度飙速度，做了一轮4道选择题的迷你调研。

目标人群是曾经带过0-6个月宝宝的家长，未要求“正在6个月中”这一苛刻条件。

共采集到176份样本，情况如下：

#### a) 父母判断啼哭原因的准确度

父母找到哭啼原因情况

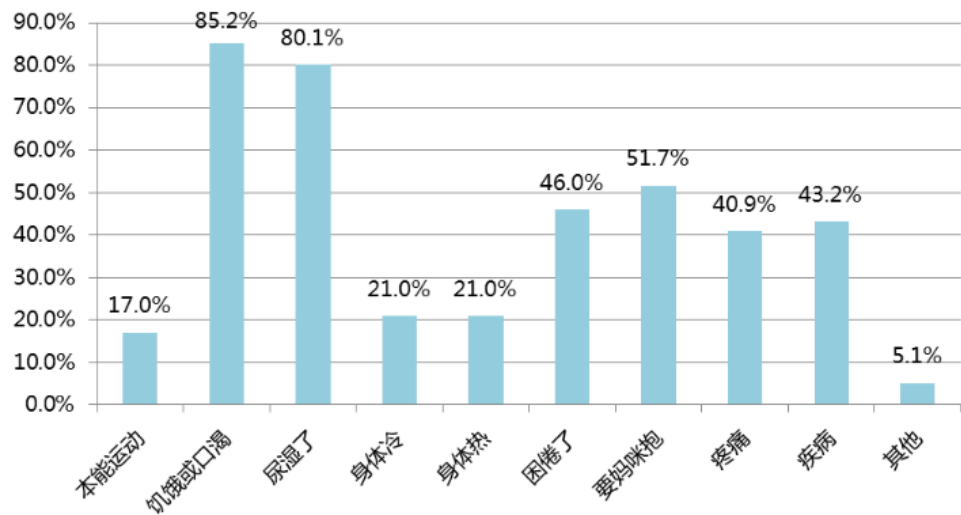


**全部正确（22%）**：这些父母拥有绝对的自信，几乎不需要借助工具。通过访谈，我了解到他们能够准确判断的主要原因有两点：经验丰富+宝宝让人省心。给他们打个标签叫“胸有成竹型”，后面会用到。

**多数正确（64.2%）**：大部分父母处于这种状态，打动他们需要一定的条件，卓越的用户体验是前提，同时要帮助他们提升诊断效率、准确率。给他们打个标签叫“经常找对型”。

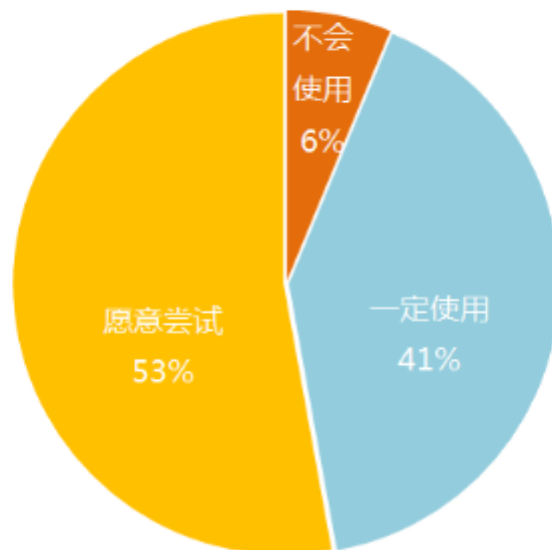
**很难找到（13.5%）**：有这么多父母面对宝宝啼哭不知所措，这里存在大量的机会可以挖掘。想想吧，中国人口基数这么大，13.5%的数量真的少么？他们的标签叫“困惑型”。

#### b) 常见的啼哭原因有哪些



### c) 父母对“破译啼哭APP”的兴趣程度

#### 愿意使用AI破译哭声APP么？



## 1.2 竞品调研

### a) 模拟声音类



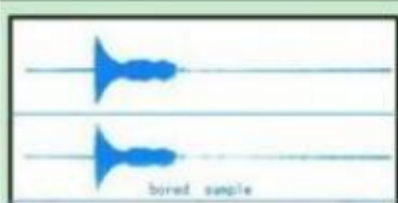



通过模拟胎儿在母体内听到的声音，来安抚宝宝。

这类产品的主要作用是安抚情绪，无法解决例如“生病”、“疼痛”、“尿布湿”等情况，治标不治本。

### b) 婴语翻译器

美国、中国台湾、西班牙、日本都有团队做过相关研究，程序通过分析哭声可给出肚子疼、尿布湿、想睡觉等原因。各团队都声明可覆盖95%以上的宝宝哭啼原因、准确度比人提高3倍等，但各团队找到的原因不一致。

此类产品目前未在大陆市场化，APP没找到可用的，独立设备倒是有，感兴趣的可以搜一下“贝客来婴儿哭声分析器”。这类产品已经是AI辨声的前辈了，其精准度、易用性、市场化均有提升空间。

	A 肚子饿：感到饥饿的哭声是精力充然开始可能声音不是很强烈，渐渐的就会越来越强烈和有节奏规律。	 <div>妈妈，我肚子饿饿 给我点吃的吧.....</div>
	B 枯燥无聊：孩子开始厌倦的时候可能是发出低低的呜咽声，如果没有人来理会他的时候就会变成尖叫。	 <div>爸爸，我好无聊 陪我玩玩吧....</div>
	C 不舒服(或生病)：这种哭叫是一种拉长声的哀诉，伴随着瞬间的呼吸停顿，宝贝短暂急促的喘息后，接着又开始了哀哭。	 <div>妈妈，我不舒服 帮我查查.....</div>

## 二、科学合理性

抛开那些竞品的干扰，推敲一下：【哭声特征】与【哭的原因】的关联到底可不可靠？如果二者之间具备真实的相关性，那是否意味着：无论映射逻辑多复杂，我们都有可能用机器学习拟合出预测模型？

### 2.1 经验常识

宝宝树这类育儿社区列举出大量听声音搞定哭闹的例子，百度经验中更是详细列举了15种哭声特点、原因以及解决办法，有经验的月嫂通过哭声能快速搞定婴儿——这些都说明：**靠声音特征判断很可能OK。**

### 2.2 专业研究

学术界对婴儿哭声含义的研究并不少，最前沿的甚至声称能诊断婴儿是否患有自闭症、脑损伤、神经类疾病和早产相关的疾病。

我觉得这个方向还是有戏的，值得继续研究，若感兴趣可以搜几份相关链接：

- 《美研发婴儿哭声诊断仪可检测神经类疾病》
- 《宝宝的哭声代表5种心情》

- 《宝宝哭声暗藏疾患线索》
- 《哭声分析仪，早期发现婴儿发育障碍》

## 2.3 综合判断

哭声不是将原因归类的唯一信号。我们可以将声音特征做主要线索优先预测一波，再用其他方法辅助判断。不苛求靠AI搞定100%的问题，只要能搞定一部分，提升解决效率，就可以发挥应用价值。

这与智能手环的绿光测心率的道理类似，光电法的准确度比不上心电信号法（医疗级别），只靠心率也很难预测心脑血管疾病。但手环的天然优势就是佩戴方便+实时监控+主动提醒。

若使用者结合经验、血压等指标综合判断，就有改善生活习惯、预测风险的可能。

这类产品符合早预防、早治疗的养生理念，经常能取得较好的市场反馈。

按照综合按断的思路，梳理出如下表格，训练模型就是要预测出其中高亮的“哭声原因”。

哭声性质	哭声原因	辅助判断	方法建议
生理性	本能运动	哭而无泪，吃奶、睡眠、玩耍都很好	轻轻地触摸几下，或对宝宝笑笑，或是把两只小手放在腹部轻轻摇一摇，小宝贝很快就会停止啼哭
需要性	饥饿或口渴	宝宝时不时用小舌头舔嘴唇，并嘴唇发干，神情有些不耐烦，就是渴了；如果没有，就是饿了	喂奶或喂水，宝宝马上就会停止啼哭
	尿湿了	没多少眼泪，小腿一个劲儿地蹬被，大多发生在睡醒或吃奶后	换掉湿衣被，宝宝很快就会停止啼哭
	身体冷	胳膊和腿的活动较少，小手小脚摸上去有些发凉，嘴唇也有些紫	给宝宝加上被子、衣物，或放到暖和的地方，宝宝就不再啼哭
	身体热	神情有些烦躁，身上摸上去也有些热，皮肤潮红，脖子上出汗，尿少	减掉衣被或安放在凉爽处时，宝宝就会马上安静下来
	困倦了	身体运动减少，眼睛半闭半睁，目光不灵活，有时眼皮出现闪动，脸上没什么表情	把宝宝放在一个舒适、安静的地方，让小宝贝安静下来，不然容易引起夜里啼哭
	要妈咪抱	头不停地左右转动，一付“左顾右盼”的样子；或把小脸转向妈咪，手抓住不放	抱抱小宝贝，或让小手抚着妈咪的身体，这样就会很快安静下来
	疼痛	仔细检查小宝贝的被褥、衣物及皮肤，有无异物或虫咬伤	赶快去除引起宝宝疼痛的异物
疾病性	疾病	夜里总醒，哭一会睡一会，很不安宁，如同受了惊吓	妈咪应该赶快带去看医生，进行补钙
		哭声无力，无精打彩，易出现烦躁	营养不良，需请医生指导喂养
		非母乳喂养，吃完奶后哭闹不止，并在哭闹时可听见小宝贝的肚子内咕咕做响	可能是肠痉挛，改吃低乳糖、无乳糖奶粉或豆浆、米面制品后，症状很快好转
		注意观察，上腹部是否可摸到有压痛的香蕉	可能是肠管迭套，妈咪应该赶快带小宝贝去医院就诊
		一阵阵的青紫、脸上及四肢肌肉抖动，多在出生时有窒息史或产伤	可能颅内压过高，赶快去医院就医，否则病情很危险
		哭时总晃动着头，或一碰到耳朵就哭，耳朵里有脓性物流出	可能患了中耳炎，需要马上找医生治疗
		赶紧观察阴囊肿块，是否平卧时也不会消失？有无压痛感？	可能患了疝气嵌，赶紧带小宝贝到外科就医

这就引出了我对该工具的定位：**通过正确判断+有效提示，帮助父母尽快掌握读懂宝宝的技能，用完即走不粘人。**

### 三.问题与发现

经过前面的需求和判断分析后，提出以下内容。

问题：

1.当前哭声病理数据集并不多，只有大约450条。

解决方法：计划先暂时搁置病理分析，因为目前来看国际上相关数据集大都不开源，我在和一个上交大以婴儿哭声检测为毕业设计的研究生交流过程中知道，他自己找了一段时间，但是和我一样也没有什么进展，有的数据集需要商业购买。

新发现:

#### 1.找到视频数据集

在查找数据集的时候发现了开源的视频数据集[https://research.google.com/audioset///dataset/baby\\_cry\\_infant\\_cry.html](https://research.google.com/audioset///dataset/baby_cry_infant_cry.html)，这也和大创思路相同，通过视觉辅助音频做哭声处理。不过这里的数据并没有标记哭声的原因。

## 四.总结

---

基于上面的调研，我们组计划前期视频数据拿到手进行清洗，之后把常用的几种识别方法都用在这个数据集中进行比较。预期结果是视觉加音频精度更高一些。虽然视频消耗的资源多，不过随着5G来临，并且硬件性能提升，视频分析不会受网络带宽和算力限制。