

周总结

2021.10.18-2021.10.24

钟舒桐¹

(1. 天津大学电气自动化与信息工程学院, 天津 30007)

摘要: 本周学习了 SPCE061A 编程软件 unsp IDE 中语音辨识的库函数使用方法, 完成了特定人语音辨识的代码编写和测试。在不同环境下, 不同的人群分别测试 50 次, 得到了特定人语音辨识的准确率

关键词: SPCE061A, 特定人语音辨识

1. 工作内容

1 特定人语音辨识

1.1. 语音识别库函数

(1) *BSR_DeleteSDGroup (0)*: 该函数的主要功能是初始化存储器 SRAM。参数“0”代表选择 SRAM, 并初始化。其返回值作为判断 SRAM 初始化成功的标志, 返回 0 代表初始化成功, 返回-1 代表初始化失败。

(2) *PlayRespond (RespondID)*: 该函数的主要功能是播放提示音, “RespondID”代表播放的语音段 (用计算机采集)。

(3) *TrainWord (int WordID, int TrainMode)*: 该函数的主要功能是语音训练, 其中“WordID”表示命令序号, “TrainMode”表示训练次数 (有 ONCE 和 TWICE 两种选择)。语音训练主要用 HMM 模型实现。

(4) *BSR_InitRecognizer (AudioSource)*: 该函数的主要功能是初始化语音识别器, “AudioSource”表示语音资源来源, 有 BSR_MIC (话筒输入) BSR_LINE_IN (模拟电压输出) 两种选择, 这里选择 BSR_MIC。

(5) *int BSR_GetResult (void)*: 该函数的主要功能是获取语音识别结果, 若识别成功返回命令序号, 否则返回 0 (无法识别)、-1 (识别器为未初始化) 或-2 (识别不合格)。该函

数的本质是一个比较、判断器，即将重新采集的语音标准特征矢量与训练产生的最佳标准特征矢量进行比较，并根据比较结果输出相应的控制信号。

(6) *BSR_PauseRecognizer(void)*: 暂停识别，但不释放中断等资源

(7) *BSR_ResumeRecognizer(void)*: 恢复被暂停的识别

(8) *BSR_GetRecognizerScore(void)*: 获得识别结果的可信度，返回值从-4096 到 4096，数值越大表示输入语音与特征模型的匹配度越高。

(9) *BSR_ExportSDWord(int CommandID)*: 使用函数库时，会自动创建一个 100 Word 的数组 *BSR_SDModel[100]*，可以把某条训练命令的特征模型数据导出到这个数组中。

(10) *BSR_ImportSDWord(int CommandID)*: 可以把 *BSR_SDModel* 数组中的数据导入为某条语音命的特征模型。

1.2 特定人语音识别程序

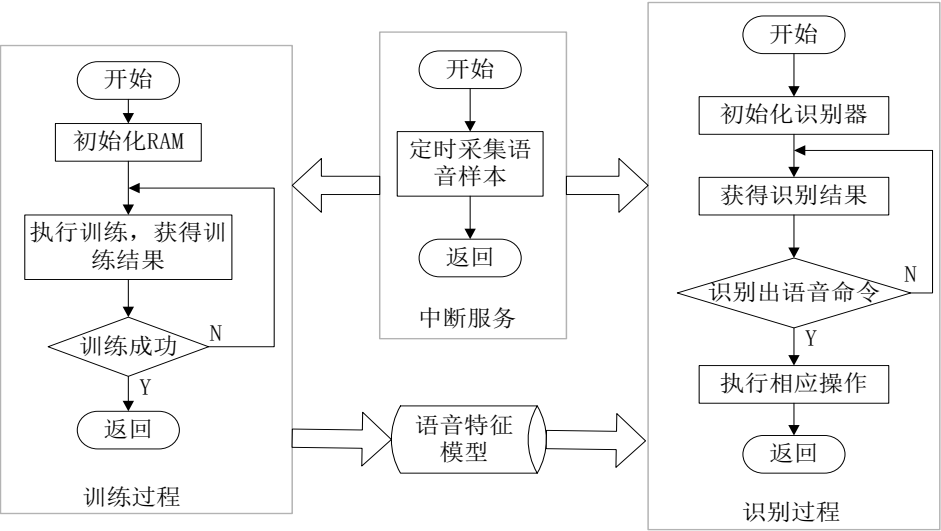


图 1 程序流程图

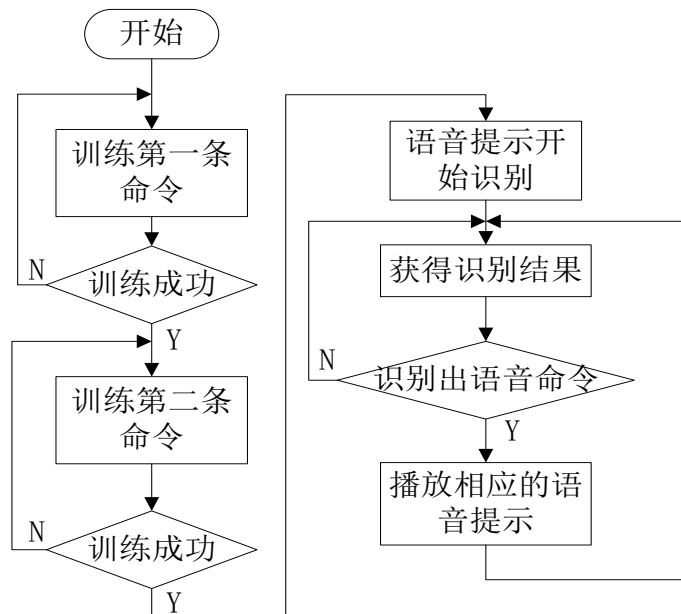


图 2 程序流程图

```

int main(void)
{
    unsigned int uiTemp=0,Loop_Counter=0;
    int res;
    ReTrainBS(); //调用语音识别的训练程序，在Sound.c中定义
    Play_Sound(12);
    //.....开始识别命令.....
    BSR_InitRecognizer(BSR_MIC); //辨识器初始化

    while(1)
    {
        res = BSR_GetResult(); //获取语音识别结果，在库中定义
        switch(res) //根据识别的结果进行分支跳转
        {
            case 0x100: //name ID
                Play_Sound(7); //播放: lalalalala
                Loop_Counter = 0; //计数清零
                break;
            case 0x101: //命令一 ID
                Play_Sound(5); //播放: I'm fine
                Loop_Counter = 0; //计数清零
                break;
            case 0x102: //命令二 ID
                Play_Sound(8); //播放: Happy new year
                Loop_Counter = 0; //计数清零
                break;
            case 0x103: //命令三 ID
                Play_Sound(6); //播放: I love you too
                Loop_Counter = 0; //计数清零
                break;
            case 0x104: //命令四 ID
                Play_Sound(9); //播放: zzzzzz (打呼噜声)
                Loop_Counter = 0; //计数清零
                break;
            default:
                break;
        }
        Loop_Counter++; //循环计数器自加一，用于程序中的状态判断
        if(Loop_Counter>=600) //如在循环次数里没有命令，则显示初始图片
        {
            Loop_Counter = 0;
        }
        *P_Watchdog_Clear = 0x0001; //清看门狗
    }
}

```

图 3 主函数文件

训练过程：

- (1) 听到提示“请输入第一条命令”后，说出一句语音命令。
- (2) 如果输入成功，会提示“请再说一遍”，重复刚才的语音命令。
- (3) 如果由于某种原因导致输入失败，会提示“没有听到任何声音”、“环境噪声太大”，如果两次训练的语音差别过大，会提示“两次输入不相同”。
- (4) 第一条命令训练成功后，用同样的办法训练第二条语音命令。

1.3 实验数据

实验测试时，选取了安静的房间，播放音乐的环境，嘈杂的大街三种环境，训练时的音色为男生，识别时选取本人音色，女生音色，另外一个男生的音色进行测试，测试距离为1m，音调为正常说话的声音。每组测试 50 次，实验识别率和有效距离如表 1.1 所示，其中有效距离代表识别正确率在 50%以上的最大距离。

音色 环境	本人	女音	男音	有效距离
安静的房间	96%	12%	38%	2.6m
播放音乐的环境	86%	0%	12%	1.8m
嘈杂的大街	88%	0%	10%	1.9m

表 1.1 实验识别率和有效距离

可以看出，在安静的房间内，本人声音识别率为 96%，女音的误识别率为 12%，男音的误识别率为 38%。可以看出，本人的语音识别正确率较高，但对于他人语音的误识别率也较高，容易出现错误识别的现象。

在播放音乐的环境和嘈杂的大街环境中，本人识别正确率和有效距离受到影响，但是误识别率下降。

同时实验中发现，说话者的声音越大，识别正确率就越高，有效距离越远。对于救命宝而言，人们在呼救的时候，声音会明显比平时说话的声音更大，因此在实际呼救的场景中，识别正确率和有效距离均会增加。

2.下周工作计划

- 1.改进库函数中语音辨识的算法，降低误识别率和降低环境对识别准确率的影响。
- 2.学习 SPCE061A 单片机利用蓝牙与电脑上位机通讯的方法，完成代码的编写和测试。