**哈尔滨工业大学（深圳）**

**大一年度项目结题报告**

**项目名称：**基于MMSE语音降噪算法的音频自适应降噪处理

**项目负责人：** 黄靖翔  **学号：** 200110723

**联系电话：** 18154638575 **电子邮箱：** 356531063@qq.com

**院系及专业：** 计算机科学与技术学院 计算机类

**指导教师：** 丁宇新  **职称：** 副教授

**联系电话：** 86-755-26032193  **电子邮箱：** yxding@hit.edu.cn

**院系及专业：** 计算机科学与技术学院 计算机类

**填表日期： 2021年 10月 17日**

**一、项目团队成员**（包括项目负责人、按顺序）

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 姓名 | 性别 | 所在学院 | 学号 | 联系电话 | 本人签字 |
| 黄靖翔 | 男 | 计算机科学与技术学院 | 200110723 | 18154638575 | 未标题-1 |
| 任柏霖 | 男 | 计算机科学与技术学院 | 200111031 | 19981680377 | QQ图片20211018221348(已去底) |
| 魏晨希 | 男 | 计算机科学与技术学院 | 200111622 | 13678593216 | QQ图片20211018233516(已去底)G`8NPGN6WS7465$8TC01KT0 |
| 陈康宇 | 男 | 计算机科学与技术学院 | 200110219 | 13502743895 |  |
| 陈柏居 | 男 | 计算机科学与技术学院 | 200110705 | 15813726605 | QQ图片20211018220805(已去底) |

**二、指导教师意见**

|  |
| --- |
| 签 名：  年 月 日 |

**三、项目专家组意见**

|  |
| --- |
| 组长签名：  年 月 日 |

**四、项目成果**

|  |
| --- |
| Z%HDRZI9`9W1P{[$FA(E9DR  (75GBWCV6O(%~@JI)[Q}HZN@RH))TK_F5M[QOU_UAMN]6X 本项目开发出一个具有音频降噪、音频格式转换、音效处理的音频效果处理软件，其中核心功能为音频降噪。上图分别为软件界面、语音降噪效果图、音频效果处理时的数据拟合。在此处只做简要说明，在下文有详细阐述，同时具体成果展示会在答辩现场以视频的形式展现。 |

**五、项目研究结题报告**（字数应在3000字左右）

1.课题背景

当今社会，音频是信息交流不可或缺的一种媒介，但音频文件中的噪音会导致信息交流质量下降，有可能导致音频接受者在接受信号时出现误解或是难以接收到有效信号的问题，所以在音频降噪处理领域的探索具有现实意义。如今的音频处理软件存在算法的效率和质量较差、入门门槛比较高、收费较高等问题，无法满足大众的日常需求。故着力于开发出一款能够解决上述问题且用户使用体验比较好的音频处理软件具有重大意义。[1]

2.课题研究内容与方法

项目的研究内容包括软件的开发以及算法的研究运用两个方面，下面对于这两个方面的具体研究内容以及研究方法进行一一阐述。

2.1关于软件开发

我们的目的是开发出一款能够输入多种不同格式的音频（如mp3、flv、wav等），并对其进行降噪处理，最后输出既定格式的音频文件的桌面应用程序。

为实现以上需求，我们需要掌握python语言的运用以及开发应用时需具备的UI设计、交互设计等技能。同时，为实现课题背景中的“用户体验比较好”，我们小组相应地编写了用户手册。

2.2关于算法的研究应用

鉴于本项目的音频处理背景，涉及到的算法可划分为两大类：一是对音频文件的编解码算法，以实现处理音频前的准备工作；二是对音频文件的处理算法，主要体现在音频降噪算法。

在项目的前期准备中，我们进行了经典算法的研究学习，涉及到的经典算法包括但不局限于fft（快速傅里叶变换）、ifft（快速傅里叶变换的逆变换）、谱减法、MMSE自适应语音降噪算法。在项目中，我们以这些算法为基础编写程序。

1. 研究结果

可以呈现出的研究结果可以大致分为四个大点：一是音频的读入写出与格式转换；二是降噪功能的实现；三是应用程序的开发；四是额外功能的实现（即为本项目中的创新点之一）。下面依次简述研究结果。

3.1音频的读入写出与格式转换

想要进行音频文件的处理，首先是要对音频文件进行采样。人耳听觉的频率上限在20kHz左右，高保真度音频信号的上限频率取20kHz，传输或普通音频信号的上限频率取15kHz。为了防止采样频谱混叠，采样前需采用前置低通滤波器把高频分量及杂波滤除。考虑到低通滤波器的截止特性，通常按进行取样（其中表示采样频率，表示信号最高频率）。[2]因此，本程序的音频采样频率为48kHz。

其次要进行傅里叶变换，将处理空间从时域转变至频域上，这样才能对不同频率的信号进行后续处理。同样的，想要输出音频文件，就需要进行逆傅里叶变换，将频域信息转换为时域信息。故想顺利处理音频，fft和ifft是基础。

为实现上述功能，我们通过调用了python中的numpy库完成了fft与ifft的过程，并用plt库实现了时域图和频谱图的展示（同时通过现存的音频分析软件Adobe Audition验证出fft与ifft的结果是正确的）。

|  |
| --- |
| $82S%A[JGHEFVZV_5W2ZDWW |
| 上图为两组音频的时域图（左）与频谱图（右） |

除此之外，我们通过阅读相关知识以及博客的学习，我们理解了音频文件的编码方式以及量化方式，并将之运用到自己程序当中。通过调用python中的pydub库，现可以实现对16、24、32三中量化位数的mp3、wav、flv、ogg音频文件进行识别并提取其中信息，后进行下一步的处理。至于音频的输出，同样是可以选择mp3、wav、flv、ogg这四种格式当中的一种。

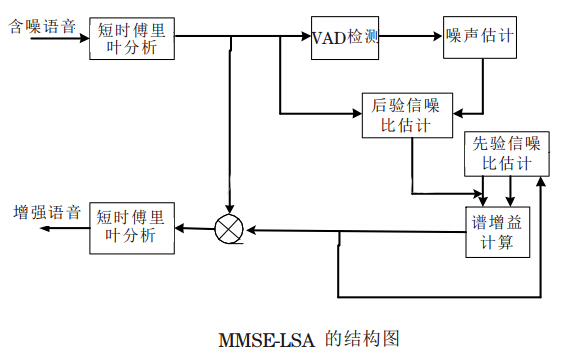
|  |
| --- |
| Z%HDRZI9`9W1P{[$FA(E9DR |
| 上图为应用程序的UI界面，涵盖了四种输出文件格式的选择 |

3.2降噪功能的实现

通过自适应的语音降噪算法的运用，我们实现了对规律性噪音的去噪消除。下面先介绍本程序应用的语音降噪算法。

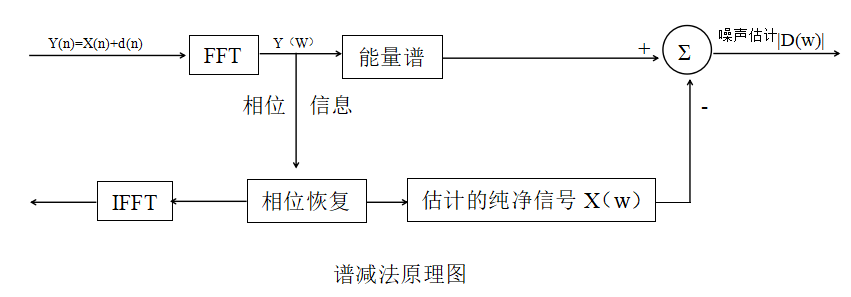
3.2.1MMSE语音降噪算法

1. 对带噪信号分帧，加窗（这里是汉明窗，能改善矩形窗造成的频谱泄露的情况）；
2. 对每帧带噪信号计算FFT；
3. 估计后验信噪比：，这里是在非语音片段（语音开始之前几帧或语音间隙）估计的噪声能量谱，然后用判决引导法估计先验信噪比；
4. 用最优的MMSE-LSA估计器的公式估计增强信号幅度；
5. 重建增强信号谱，然后计算增强信号谱的IFFT，得到对应输入语音帧的增强的时域信号x(n)。



3.2.2谱减法

在语音降噪方面,最普遍的方法就是谱减法及其改进算法[3]。谱减法的原理较为简单，可表示为：将噪声假定为加性噪声，那么可以在估计噪声谱的基础上，通过减法将带噪语音谱中的噪声谱减去，剩下的部分即为纯净的信号谱。下面给出谱减法的算法流程图。



相对于MMSE语音降噪算法，谱减法的适用范围比较广，故本程序将在无法估计噪声能量谱的情况下使用谱减法进行音频降噪。

3.2.3降噪结果的分析

下面给出降噪前后音频的时域图进行对比。（具体实现过程由视频展示）

|  |
| --- |
| (75GBWCV6O(%~@JI)[Q}HZN |
| 上图为降噪前（下）与降噪后（上）的时域图 |

测试音频的捕捉使用的是录音室级别的录音设备（雅马哈ur12＋铁三角AT2035），设备底噪远小于环境噪音，故测试数据具备说服力。

通过对数字信号的相关知识的学习，我们知道要精确地对一个实测噪声值与该噪声的技术要求进行比较是不可能的。[4]因此，需要引入能够客观的检测噪音（或是降噪效果）的指标，故为了检测语音降噪效果，我们引入一个指标“信噪比”。信噪比是指一个电子设备或者电子系统中信号与噪声的比例。记为。（Ps表示音频功率、Pn表示噪音功率）语音降噪需要有效的提高语音信号的信噪比。[5]下面就用此公式计算用以测试的音频文件的前后信噪比。

|  |  |  |
| --- | --- | --- |
|  | 处理前信噪比 | 处理后信噪比 |
| 测试数据1 | 11.24 | 21.92 |
| 测试数据2 | 3.01 | 12.01 |
| 测试数据3 | -2.42 | 4.97 |

通过超过一百组的噪音测试，我们可以简要得出以下结论：对规律性噪音，在不影响人声的前提下基本可以消除；对于非规律性噪音或是偶发性噪音，处理效果差强人意；对音频实现降噪效果基本不会改变原本的音调，但会稍许改变音色。若以信噪比作为参数进行下结论，即可得出“噪音强度越小，降噪效果越好”的结论。

3.3应用程序的开发

在数据分析、交互式计算以及数据可视化方面，Python将不可避免地与其他开源和商业的领域特定编程语言/⼯具进⾏对⽐，如R、MATLAB、SAS、Stata等。近年来，由于Python的库（例如pandas和scikit-learn）不断改良，使其成为数据分析任务的⼀个优选方案。结合其在通用编程方面的强⼤实⼒，我们完全可以只使⽤Python这⼀种语言构建以数据为中⼼的应用。[6]

漂亮的界面是一个GUI程序必不可少的一部分，wxPython可以做到这一点，加之Python强大的功能和简洁的语法，使得它在Python的gui中成为一种主流。[7]故我们主要是运用了wxPython来编写软件的交互界面。软件具备调节音效、格式转换、音频降噪的功能，用户可以在交互界面实现上述操作。为了方便用户进行使用，我们在菜单栏中附上了用户手册。

迄今为止，本软件进行了三个版本的更新，现行版本为v3.0。v1.0版本仅能实现对音频的格式转换；v2.0版本实现了对音频格式的转换及音频降噪功能；v3.0版本实现了格式转换、音频降噪及不同频段的截取（现行版本）。

|  |
| --- |
| 6G00MQH@FCD7ZW{N9(Z%GDLY{WK6FW9]P2F[_3K$ZU(JC8 |
| 上图为用户手册 |

本软件能够在Windows、Linux系统中运行，运行环境为：python3环境，包括库：numpy、wavio、wxpython、pydub、ffmpeg、ctypes。其中，UI界面的实现运用了python中wxPython界面开发os识别文件路径并管理处理音频产生的临时文件库，ctypes实现对不同系统、不同电脑分辨率的适配。目前仍有开发v4.0版本的计划，目标是实现开发出可执行文件，脱离python3的运行环境。

3.4额外功能的实现

在项目的后期进行数据测试的时候，我们通过matlab实现数据拟合，自建了若干滤波器并应用到了程序当中，最终实现出的效果是不同频段的截断。（具体实现过程由视频展示）

|  |
| --- |
| @RH))TK_F5M[QOU_UAMN]6X |
| Matlab进行数据拟合界面 |

1. 创新点

在运用自适应降噪算法前，我们进行了适配化处理，即对处理前后的数据进行了标准化处理、反标准化还原，其目的是去除少数非法数据，让最后的信息更准确。同时，我们也对算法中的某些变量进行了适配化处理，实现了处理效果与处理速度之间的相对平衡。

在实现原有的目标“降噪”之后，我们增添了软件的功能“音效”，可以视为项目的创新点。下面简要阐述实现该创新点的研究过程。

项目的后期阶段，我们录制了大量的音频文件进行数据测试，以测试降噪算法对于不同情况的噪音的消除效果。在测试数据的过程中，我们通过plt库绘制出不同音频文件的频谱图，并对音频文件的不同频段的信息进行分析，分析出“低音”、“中音”、“高音”所占的频段，最终使用Matlab进行数据拟合，自建出不同频段谱的滤波函数，实现了对一段音频截取低音、中音、高音的功能。

5、结束语

从去年11月至今，项目进行了也差不多一年。这一年的学习研究过程中，我们经历过理论知识学习、将理论知识的实际应用、捕捉数据进行测试……不同的工作的侧重点不同，为了将这些工作完成得更好，我们小组的五个人都尽己所能去完成目标。如：擅长编程的同学就主要负责软件的开发、数学能力强的同学就主要负责原理的学习、数据的处理、对项目整体框架比较明朗的同学就在组会上进行统筹规划，分析项目的现状以及下一步的方向。可以说，没有我们五个人这一年的齐心协力，是不会有如今的项目成果的。

除此之外，我们还要感谢在项目进行的过程中曾经帮助过我们的人。首先是我们小组的导师丁宇新老师，在大一项目的选题上，丁老师鼓励我们“兴趣驱动”，不限制我们的选题，我们才有了后来推进项目的动力；在每次有疑惑，向丁老师寻求帮助时，他总能用自己丰富的经验来给我们指明方向；最重要的是在丁老师的带领下，我们了解了科研项目的流程，让我们能在大一到时候就能接触到研究与创新。第二是在项目的中期答辩时遇到的评委罗文坚老师，我们在中期答辩时曾提出“降噪算法难度较大，学习原理与实际应用之间存在着壁垒”这一困难，他在答辩的点评中给出了“你们可以去看看微软的相关开源项目，说不定能在里边学到一点东西”的建议。正是这一建议，给我们项目的中后期工作指明了方向。

经过这一次大一项目，我们收获良多。这次项目所代表的不仅仅是一次共同完成科研项目的经历，还有在实现项目成果的过程中所培养的自主学习、探索未知领域知识的能力、团队配合协作的能力……希望我们能把大一项目作为起点，由此开始我们的科研探索吧。

6、参考文献

[1]卢官明，宗昉. 数字音频原理及应用（第3版）[M]. 北京：机械工业出版社，2017.4：31.

[2]王一海.一种改进型谱减法的电子商务语音噪声去噪方法研究[J].信息化研究，2020，46（2）：21-25.

[3]王方杰，金赟. 基于会议电话中的实时语音降噪算法研究[J]. 电子器件，2019，42（1）：1-2.

[4]卢官明，宗昉. 数字音频原理及应用（第3版）[M]. 北京：机械工业出版社，2017.4：329.

[5]王涛. 语音降噪处理技术的研究[D]. 兰州：兰州交通大学，2018.2.

[6]Wes McKinney(O'Reilly) .Python for Data Analysis[M]. 唐学韬等译. 北京，机械工业出版社，北京华章图文信息有限公司，2013. 19.

[7]HarriPasanen and RobinDunn.wxPython in Action[M].滴水译.Manning Publications.2006.1.