

实现复杂机载语音通信的混音方法^①

马永松, 冉进刚, 彭 健

(中电科航空电子有限公司, 成都 611731)

通讯作者: 马永松, E-mail: mays@cetcea.net.cn



摘 要: 随着通航飞机任务的越来越多元化、复杂化, 要求机载语音通信系统能够支持更多成员在任务中参与实时通话. 本文通过加入实时权重、动态箝位机制设计了改进的单输出通道混音算法, 并在此基础上进一步设计了可定制的多输出通道混音组织逻辑, 形成了一套适用于大型通用任务型飞机的“有优先级、用户可定制的机载会议系统”设计的实用混音方法.

关键词: 机载语音通信; 实时权重; 动态箝位; 混音组织逻辑; 可定制化; 混音

引用格式: 马永松, 冉进刚, 彭健. 实现复杂机载语音通信的混音方法. 计算机系统应用, 2019, 28(2): 101-106. <http://www.c-s-a.org.cn/1003-3254/6769.html>

Sound Mixing Method for Complex Airborne Voice Communication

MA Yong-Song, RAN Jin-Gang, PENG Jian

(China Electronics Technology Group Avionics Corporation (CETCA), Chendu 611731, China)

Abstract: With the increasing diversity and complexity of the mission of general-purpose plane, airborne voice communication system is required to involve more members in a real-time call in missions. This study introduced a mixing method with 2 critical algorithms: mixing algorithm with real-time weights and dynamic clamping mechanism for single-output, and mixing organization logic for the customizable meeting. The mixing method brings a new design method of “customizable airborne conferences system” on the large-scale general mission aircraft.

Key words: Airborne voice communication; real-time weights; dynamic clamping; mixing organization logic; customizable; sound mixing

机载语音通信系统一直是航空电子系统的最重要组成部分, 在通飞航电系统中机载语音通信功能包括飞机内部各岗位之间的通话和空地通话.

中小型通航飞机通信参与成员比较少, 大多采用对讲方式; 但是大型通航飞机 (括大型水陆两栖飞机、大型固定翼通用任务型飞机、民用重型直升机等) 通信参与成员变多、执行的任务也更加复杂, 在许多任务中要求不同任务岗位之间更紧密的协作^[1]. 随着国产大型通航飞机的研发, 需要机载语音通信系统成为支持飞行员、观察员、执行任务的各岗位、飞机状态提

示 (告警音、各种提示音) 以及地面指挥员之间进行的多边语音交互的平台. 这对机载语音通信系统提出了更高的期望, 如果能把所有的任务参与者都放在一个“会议”里, 所有信息将在第一时间传输到各岗位, 这对协作应对复杂任务和突发情况将极为有利.

本文提出了一套简单实用的混音方法, 并在此基础上设计出了“有优先级、用户可定制的机载会议系统”; 不但大幅提升机载语音系统的性能, 而且软件简单, 开发成本低、耗时短、安全性高, 对快速取得适航证、投入市场极为有利.

① 收稿时间: 2018-05-28; 修改时间: 2018-06-27, 2018-08-07, 2018-08-29; 采用时间: 2018-09-18; csa 在线出版时间: 2019-01-28

1 机载语音通信系统的需求变迁和挑战

传统的机载语音通信系统多采用对讲模式. 其优点包括:

(1) 采用“你一言”、“我一语”的沟通方式, 沟通清晰、准确;

(2) 一般主驾驶员具有最高的通话优先级, 随时可以抢占通话信道.

但其缺点也很突出的: 参与成员不能太多、只能满足相对简单的任务交互.

随着通航飞机的大型化、任务复杂化, 对机载语音通信系统的要求也越来越高. 这些飞机往往需要执行救援、灭火、空投等各种复杂任务. 要完成这些复杂任务, 需要包括驾舱(主驾驶、副驾驶、机械师、观察员)、后舱\任务舱(参与任务的各岗位)、飞机告警及各种提示音和地面指挥的有效协作. 这就给语音通信系统提出了更高的要求, 包括:

(1) 系统可以工作在“会议模式”, 容纳更多成员参与通信, 同时避免高优先级成员长时间占用信道, 导致延误; 通信参与成员在通信建立之后随时可以说话, 每个通信参与成员能听到所有成员(包括自己)的语音;

(2) 系统是“用户可实时定制”的; 在任务进行中, 由通信参与成员进行简单的设定就可以完成新的定制; 每个通信参与成员应该可以根据所在任务岗位的实际需要选择重点收听某路或者某几路音源, 突出重点避免次要音源的干扰.

(3) 由于空地通信的特殊性, 机上各岗位应该有权进行空地通信(但是其它机上乘员可以收听到双方的通话), 以便在某些任务的关键阶段, 地面可以直接指挥对应岗位, 而不是通过飞行员间接指挥.

(4) 赋予关键成员(主驾)管理者角色, 如有空地通话优先权、通话接入管理等.

当前航电各 LRU 之间的基本都实现了总线通讯; 随着 DSP、FPGA 以及数字化语音处理芯片的成熟和稳定, 机载语音通信系统也走入了数字化时代. 这给实现理想的“有优先级、用户可定制的机载会议系统”带来了机遇, 但同时也面临许多挑战:

(1) 在民用航空领域, 语音通信系统软件对飞行安全有重要作用, 属于 C 级软件. DO-178C 对 C 级软件的计划、需求、开发、集成验证、生命周期等有一系列规范, 遵照这些规范进行软件开发, 更容易满足适航要求、取得适航认证; 但是严苛的开发规范也导致了更

高的开发成本, 在国内, 一款 C 级民机机载软件从计划阶段到取得适航证往往要数年时间. 因此, 满足功能的情况下尽可能的精简设计, 对于保证安全、节约开发成本、缩短开发时间极为重要.

(2) 鉴于航电的特殊性, 如果能不改变机载语音通信系统硬件体系架构, 仅仅通过优化软件算法, 就可以实现“有优先级、用户可定制的机载会议系统”, 那么就可以最大限度的保证系统安全性、降低研发成本、控制研发周期, 这就给软件开发提出了很高的要求.

(3) 在混音环境中, 一般只能将混音后的音量整体调大或者调小. 而“有优先级、用户可定制的机载会议系统”要求各成员可以动态配置想要收听的信息以及各分量权重, 这要求软件算法有所突破.

(4) 机载语音通信系统需求对实时性、信噪比、系统稳定性等提出了明确的指标, 要求软件算法简单实用而且工作稳定.

2 有优先级、用户可定制的机载会议系统

2.1 方案背景

机载语音通信系统是 CN 系统的一部分, 涉及众多 LRU(如图 1)^[1], 且由于 CN 系统的重要性, 要将整个系统推翻重新设计其成本和耗费的时间是难以接受的. 因此, 本文正是在图 1 所示系统上通过优化 RIU 软件实现语音通信系统的品质提升.

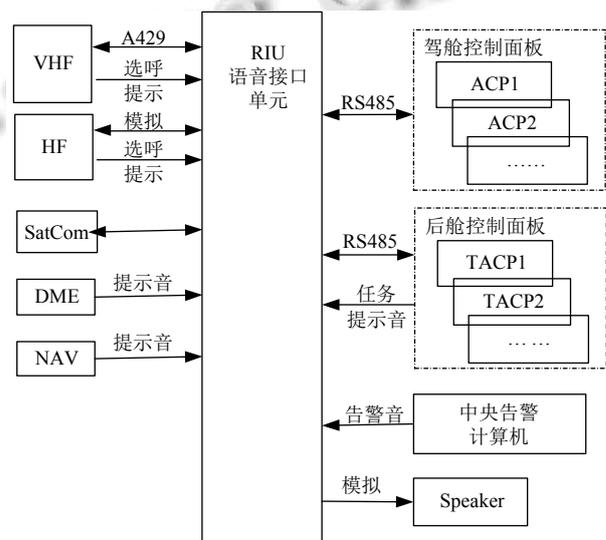


图 1 机载语音通信系统框图

由于 CN 系统中各 LRU 来自不同的厂家、接口各异, 系统中涉及的主要语音格式包括 A429 数字音频、

RS422/RS485 数字音频、模拟音频. RIU 作为音频接口和处理单元,是整个系统中枢,负责所有语音的信号、语音控制信号的收发和处理;也就是说 RIU 软件的质量直接影响整个机载语音系统的品质.

RIU 的外围设备主要包括空地通信电台、ACP/TACP(如图 2),其中 ACP/TACP 可以接入耳机和麦克风,也可以设置通信状态和音量.



图 2 常见的 ACP/TACP 操作面板

2.2 软件方案

软件方案简述如图 3^[2],首先将所有语音输入转换成相同的数据格式并进行降噪处理^[3],由于不同设备的采样率和采样精度的不同,可能对有些数据要进行裁剪或插值处理^[4].当所有数据对齐后即可按照既定的频率进行混音,本文将着重对混音算法的两个关键的环节进行阐释,首先是任意输出通道的混音算法,其次是多输出通道的混音组织逻辑.

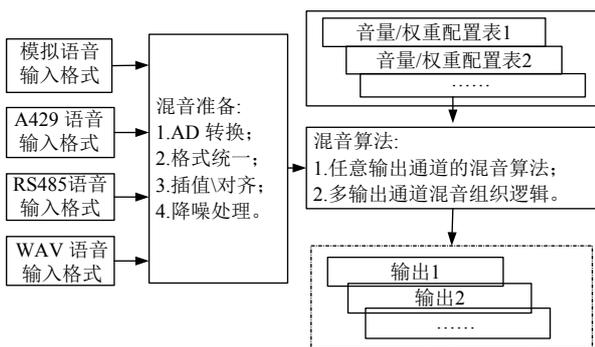


图 3 语音通信系统软件方案简图

2.2.1 任意输出通道的混音算法

传统的混音算法有很多,但对于机载语音通信系统来说都有一些缺点.

式 (1)^[5]描述了最简单的混音算法.直接将 t 时刻的语音输入线性叠加,这种算法在通信信道容量是最

大输入音量的 n 倍 (n 是通信参与成员的数量) 的时候显得简单实用;在复杂机载通信中,通信参与成员可能有近 10 路,信道容量要做到 10 倍余量是不可能的.

$$Voice_sum[t] = \left(\sum_{i=0}^{n-1} In[i, t] \right) \quad (1)$$

箝位算法 (如式 (3) 和式 (4))^[5,6]引入了箝位机制避免溢出,当混音参与数量少、极少出现溢出的时候是极为实用的,箝位算法是对式 (1) 的有益改进和补充.

当 $Voice_sum[t] < Voice_max$ 时,

$$Voice_sum[t] = Voice_sum[t] \quad (2)$$

当 $Voice_sum[t] \geq Voice_max$ 时,

$$Voice_sum[t] = Voice_max \quad (3)$$

平均值算法如式 (4)^[5,6].将语音输入的线性叠加值除以 n ,避免了溢出问题的产生;但是如果多达 10 路的语音同时输入,而每路音量又被缩减为原来的十分之一,可能会出现听不清楚任何一方讲话的情形.

$$Voice_sum[t] = \left(\sum_{i=0}^{n-1} In[i, t] \right) / n \quad (4)$$

针对机载环境的实际应用,本方案对混音算法进行了综合改进.首先加入实时权重,如式 (5) 和 (6) 所示^[6,7], $In[i, t]$ 是第 i 路输入的 t 时刻的声音强度, $Volume_i[j, t]$ 是输出到通道 i 的混音中来自 j 通道的输入分量的 t 时刻音量,来自通道 i 所对应的 ACP/TACP 面板上的用户配置.将权重因子 $f_i[t]$ 引入混音式 (6),即得到的第 i 通道混音输出 $Voice_sum_i[t]$.这种算法解决了两个问题:

(1) 混音权重由声音的输入强度和混音接收端主动设置的音量组合决定,即通过音量配置就可以影响混音输出中各个分量的权重.

(2) 按照权重分配各输入分量在混音中的占比,直接避免了混音溢出问题.

$$f_i[t] = \frac{In[i, t] * Volume_i[i, t]}{\sum_{j=0}^{n-1} (In[j, t] * Volume_i[j, t])} \quad (5)$$

$$Voice_sum_i[t] = \sum_{i=0}^{n-1} (In[i, t] * f_i[t]) \quad (6)$$

式 (6) 解决了混音溢出的问题,但是牺牲了混音输出中各分量的音量.特别是在每路音量都比较小的时

候最后的混音结果中分配给各分量的音量将更小,甚至小到影响收听. 在加入混音的通道比较少的时候,用户可以通过调大某一路音量来适当的增加这一路输入分量混音后的音量,但是如果混音输入通道较多且要关注两路以上的输入的时候,这种弥补就显得不足了. 而另一方面不可能所有通道都以满负载输入,也不可能所有通道都选择满音量,所以 $Voice_sum_i[t]$ 的值可能远小于通道负荷. 所以,一方面混音后各分量音量大幅变小,另一方面通道又没有满负荷运行,造成了浪费;为解决这两个问题,对算法进一步进行了改进,如图 4.

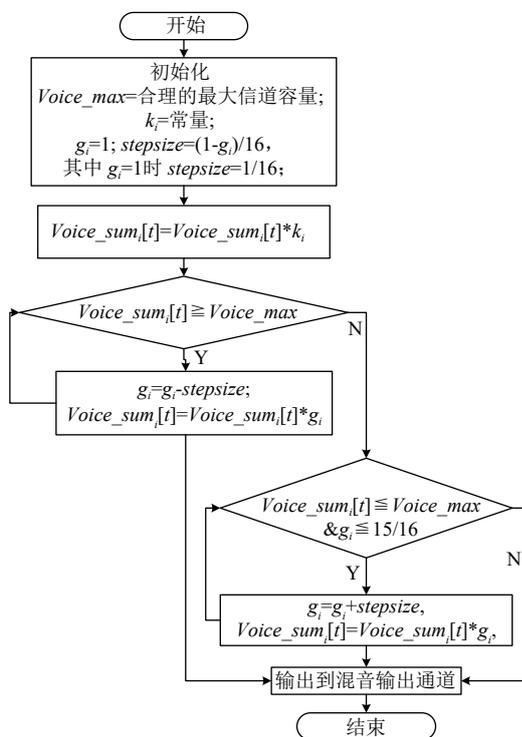


图 4 任意通道混音算法流程图

图 4 中, K_i 是放大系数、 $Voice_max$ 是软件限值, $Voice_sum_i[t]$ 和 $Voice_max$ 的比较结果会触发钳位因子 g_i 动态变化从而渐进的修正最终的 $Voice_sum_i[t]$ [8].

K_i 、 $Voice_max$ 是两个经验值, 需要综合音频系统信道的带宽和各通道混音的典型输入情况确定. 恰当的 K_i 值可以让最大混音输出值处于溢出临界状态, 最大限度利用信道宽度; 而 $Voice_max$ 和 g_i 的配合最大限度限制了在混音分量加入或者退出时音量的陡变. 总之, 按照这个算法得到的混音输出值达到了如下目的:

(1) 有新的分量加入/退出混音时, 不会导致混音输出值瞬间增大/减小; 无论混音参与分量如何变化, 最终

输出的音量都将循序增加/减小, 发送若干包后再趋于稳定;

(2) 通过实验, 选择恰当的 K_i 值, 可以保证混音输出基本不会溢出, 只有在声音陡变时可能会短时间工作在临界状态, 但是也能很快回到 $Voice_max$ 附近, 不会产生超出指标范围的噪音.

(3) 通过音量调节实现了混音权重分配. 当某个输入分量被调大时, 会挤占其它输入分量的信道资源, 从而获得重点关注.

(4) 混音算法没有造成明显的延时, 实时性良好.

2.2.2 多输出通道混音组织逻辑

算法的另外一个难点是系统中有多个通道需要同时输出混音, 并且每个通道的混音设定是其对应的 ACP/TACP 自行定制的. 为实现这一功能, 系统为每一个输出通道维护了一张音量/权重配置表, 配置表的内容根据对应输出通道的设置 (来自 ACP/TACP) 实时更新. CPU 调用任意输出通道混音算法依次轮询每张配置表, 即可得到各个通道定制的混音输出值, 流程如图 5.

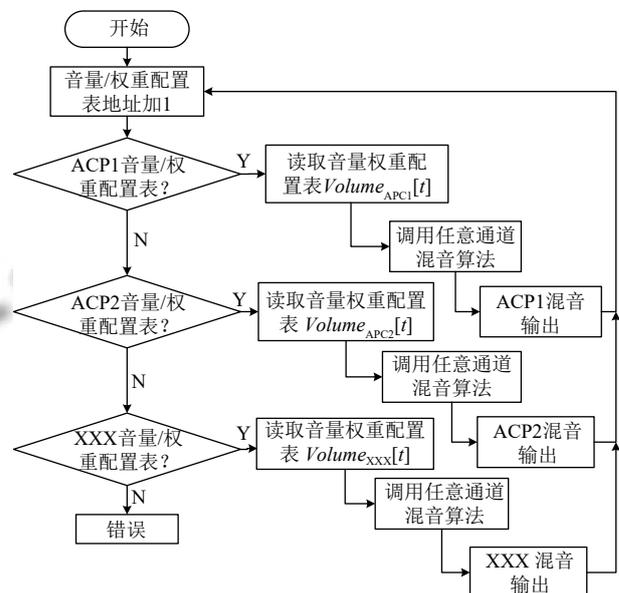


图 5 多通道混音组织逻辑流程图

系统采用了运算速率高达 2400MFLOPS 的 DSP, 而我们的混音周期是毫秒级的、混音输出通道一般不会超出 10 个, 所以技术上是可行的. 系统所需要提供的主要资源就是给每一个通道维护一个音量/权重配置表. 至此在新增加极少资源的情况下就可以实现“有优

“先级、用户可定制的机载会议系统”的预期目标。

本方案中多通道混音组织逻辑新增的代码量不超过 10%，整个混音方案总代码量不超过 2000 行，对于开发成本高昂的 C 级航空软件来说本方案是简洁有效的。

机载语音通信系统有一些高优先级的提示音和告警音需要突出展现，如图 6 所示。来自中央告警计算机的告警音具有最高优先级，应当被实时的清晰无疑的播报出来。当有中央告警音到来（离散信号有效）时，后舱呼叫和地面呼叫被同时抑制。告警音\提示音混音被分为两路，一路独占信道，以最大音量发送到机载广播；另一路作为混音输入加入到混音中，但是软件赋予这一路输入较高的权重（设整个信道权重为 1，这里赋予其 0.6 的权重），不能被调节。当告警、提示音有输入时，各个终端听到的结果就是其它音频分量音量瞬间降低，告警音\提示音混音被凸显出来，直到告警音\提示音混音值为 0（关断阈值）。

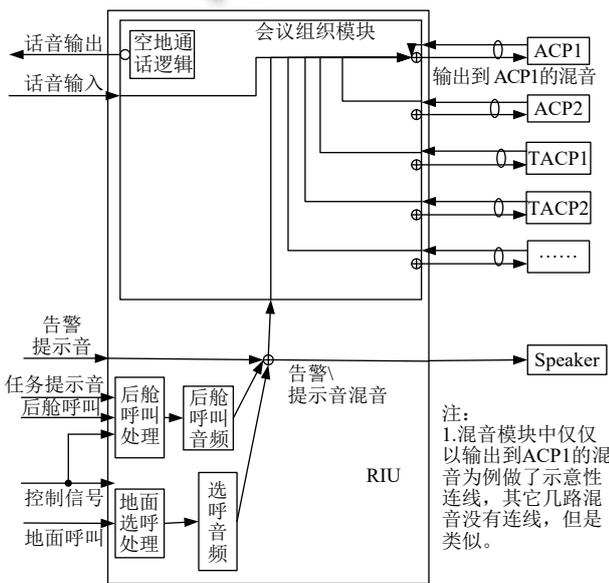


图 6 提示音和告警音混音

另外，本设计中主驾（设为 ACP1）负责会议的管理，处理接入请求和挂断请求。由于空地通话的特殊性，地面可能要求单独和某个终端通话、也可能要求收听机上会议，ACP1 有优先抢占和地面通话的权限。

3 测试分析

3.1 定量测试

所有输入通道均输入 1 kHz 的等相位的单音、且

所有通道音量开到最大时，测试结果如表 1 和图 7 所示。结果显示：

- (1) 混音输入值幅度的变化对混音输出幅度（稳定后的值）基本不产生影响，混音输出幅度由对应输出通道的音量设置决定；
- (2) 混音输出失真率小于 3%，并且随输入音量的增大而减小，满足需求。

表 1 混音输出幅度和失真率测试统计表

序号	各通道输入幅度 (mV)	混音输出幅值 (V)	混音输出失真率 (%)
1	100	7.540	2.881 970
2	250	7.713	1.313 989
3	450	7.816	1.015 194



图 7 混音输出幅度和失真率测试结果截图示例

3.2 系统通话测试

在图 8 所示的语音通信系统硬件平台上测试了本文描述的混音方法设计的机载会议系统。测试结果显示，系统功能工作正常，性能良好。达到以下预期效果：

- (1) ACP/TACP 面板上旋钮可以实时调整各个语音输入通道的音量/权重占比，达到清晰关注某一个或者几个输入分量的目的，不存在混音路数太多，导致每一路都听不清的情形，达到预定目的；
- (2) 在混音分量加入\退出或者用户快速调整音量时，混音输出中各个分量没有明显的音量突变，音量调整渐进柔和；
- (3) 在有告警到来时，混音中其它分量陡然变小、告警音被凸显，提示作用十分明显；
- (4) 所有音量均开到最大时执行混音分量加入或者退出没有引入明显噪音，没有削峰现象；
- (5) 从音频信号输入到混音输出信号延迟小于 5 ms，不影响实时交流。

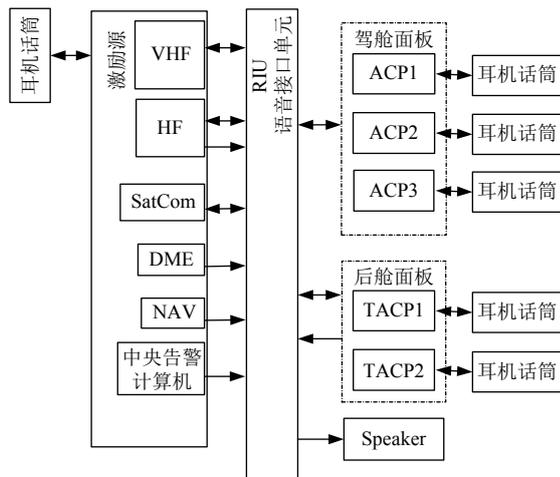


图8 通信导航系统地面测试平台

4 结论与展望

本文采用的算法简洁,非常符合高安全性要求的航空软件的设计思路.这种简单明了的设计缩短了产品开发周期、对快速取得适航证有相当积极意义,有利于产品快速进入市场,有利于加速追赶国际先进航电软件.

本文描述的算法是基于国内主流CN系统平台实

现的,既能代表未来机载语音通信系统的发展方向,同时又能兼顾成本、向后兼容,有利于快速类似CN系统上移植.

参考文献

- 1 王洪涛. 民用机载音频系统模拟器的研究及实现[硕士学位论文]. 成都: 电子科技大学, 2011. 3-4.
- 2 González AJ, Abdel-Wahab H. Audio mixing for interactive multimedia communications. Proceedings of JCIS'98. NC, USA. 1998. 217-220.
- 3 王秋菊. 机载噪声环境下语音增强研究[硕士学位论文]. 武汉: 华中师范大学, 2017.
- 4 黄梅, 洪攻, 卢扬, 等. 数字混音模型的设计与实现. 计算机工程与设计, 2010, 31(16): 3625-3627, 3631.
- 5 胡斌. 一种多路音频混音算法的实现与应用. 探索与观察, 2017, (15): 13-14.
- 6 李红光, 郭英, 苏令华, 等. 多路数字混音方案设计. 计算机仿真, 2014, 31(12): 169-173.
- 7 徐保名, 王秀玲. 一个改进的混音算法. 电子与信息学报, 2003, 25(12): 1709-1713.
- 8 Radenkovic M, Greenhalgh C, Benford S. A scalable audio service to support many simultaneous speakers. www.crg.cs.nott.ac.uk/research/projects/nait/, 2002.