

(12) 发明专利申请

(10) 申请公布号 CN 102969003 A

(43) 申请公布日 2013. 03. 13

(21) 申请号 201210461668. 7

(22) 申请日 2012. 11. 15

(71) 申请人 东莞宇龙通信科技有限公司

地址 523500 广东省东莞市松山湖科技产业
园区北部工业城C区

申请人 宇龙计算机通信科技(深圳)有限公
司

(72) 发明人 闫正航

(74) 专利代理机构 北京中博世达专利商标代理
有限公司 11274

代理人 申健

(51) Int. Cl.

G10L 21/0272(2013. 01)

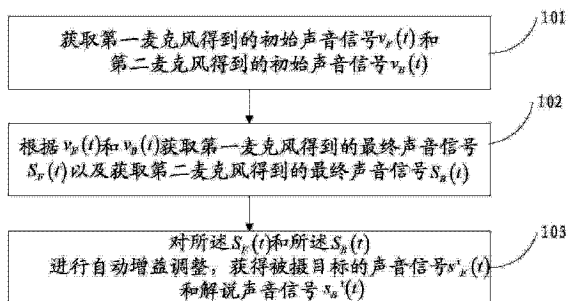
权利要求书 2 页 说明书 7 页 附图 4 页

(54) 发明名称

摄像声音提取方法及装置

(57) 摘要

本发明实施例公开了一种摄像声音提取方法及装置,所述方法包括:根据公式 $v_f(t) = s_f(t) + \beta_B s_B(t) + n(t)$ 获取第一麦克风得到的初始声音信号 $v_f(t)$,根据公式 $v_B(t) = s_B(t) + \beta_F s_F(t) + n(t)$ 获取第二麦克风得到的初始声音信号 $v_B(t)$,其中, $s_f(t)$ 为被摄目标的初始声音信号, $s_B(t)$ 为解说初始声音信号, β_B 为解说声音到第一麦克风的衰减系数, β_F 为被摄目标声音到第二麦克风的衰减系数, $n(t)$ 为噪声信号;根据公式 $S_f(t) = v_f(t) - \beta_B v_B(t)$ 获取第一麦克风得到的最终声音信号 $S_f(t)$,根据公式 $S_B(t) = v_B(t) - \beta_F v_f(t)$ 获取第二麦克风得到的最终声音信号 $S_B(t)$;对 $S_f(t)$ 和 $S_B(t)$ 进行自动增益调整,获得被摄目标的声音信号 $s'_f(t)$ 和解说声音信号 $s'_B(t)$ 。本发明适用于在摄像过程中对不同目标的声音进行区分和调整。



1. 一种摄像声音提取方法,其特征在于,包括:

根据公式 $v_F(t) = s_F(t) + \beta_B s_B(t) + n(t)$ 获取第一麦克风得到的初始声音信号 $v_F(t)$, 根据公式 $v_B(t) = s_B(t) + \beta_F s_F(t) + n(t)$ 获取第二麦克风得到的初始声音信号 $v_B(t)$, 其中, $s_F(t)$ 为被摄目标的初始声音信号, $s_B(t)$ 为解说初始声音信号, β_B 为解说声音到第一麦克风的衰减系数, β_F 为被摄目标声音到第二麦克风的衰减系数, $n(t)$ 为噪声信号;

根据公式 $S_F(t) = v_F(t) - \beta_B v_B(t)$ 获取所述第一麦克风得到的最终声音信号 $S_F(t)$, 根据公式 $S_B(t) = v_B(t) - \beta_F v_F(t)$ 获取所述第二麦克风得到的最终声音信号 $S_B(t)$;

对所述 $S_F(t)$ 和所述 $S_B(t)$ 进行自动增益调整, 获得被摄目标的声音信号 $s'_F(t)$ 和解说声音信号 $s'_B(t)$ 。

2. 根据权利要求 1 所述的方法, 其特征在于, 在所述对所述 $S_F(t)$ 和所述 $S_B(t)$ 进行自动增益调整之前, 还包括:

将所述 $v_F(t)$ 和 $v_B(t)$ 代入所述 $S_F(t)$, 得到 $S_F(t) = s_F(t) (1 - \beta_B \beta_F) + (1 - \beta_B) n(t)$; 将所述 $v_F(t)$ 和 $v_B(t)$ 代入所述 $S_B(t)$, 得到 $S_B(t) = s_B(t) (1 - \beta_B \beta_F) + (1 - \beta_F) n(t)$ 。

3. 根据权利要求 2 所述的方法, 其特征在于, 所述对所述 $S_F(t)$ 和所述 $S_B(t)$ 进行自动增益调整, 获得被摄目标的声音信号 $s'_F(t)$ 和解说声音信号 $s'_B(t)$ 包括:

调节所述 $S_F(t)$ 和所述 $S_B(t)$ 的信号增益值, 将调节后的 $S_F(t)$ 信号作为 $s'_F(t)$, 将调节后的 $S_B(t)$ 信号作为 $s'_B(t)$ 。

4. 根据权利要求 1 所述的方法, 其特征在于, 在所述获得被摄目标的声音信号 $s'_F(t)$ 和解说声音信号 $s'_B(t)$ 之后, 还包括:

根据用户需要分别对所述 $s'_F(t)$ 和所述 $s'_B(t)$ 的声音强度进行调节; 或者,

根据用户需要对所述 $s'_F(t)$ 和所述 $s'_B(t)$ 的声音强度进行联动调解, 当调高其中一个信号的声音强度时, 另一个信号的声音强度相应降低。

5. 根据权利要求 1-4 任一项所述的方法, 其特征在于, 所述第一麦克风与被摄目标的距离小于所述第二麦克风与所述被摄目标的距离; 所述第一麦克风与解说者的距离大于所述第二麦克风与所述解说者的距离。

6. 一种摄像声音提取装置, 其特征在于, 包括:

第一信号获取单元, 用于根据公式 $v_F(t) = s_F(t) + \beta_B s_B(t) + n(t)$ 获取第一麦克风得到的初始声音信号 $v_F(t)$, 根据公式 $v_B(t) = s_B(t) + \beta_F s_F(t) + n(t)$ 获取第二麦克风得到的初始声音信号 $v_B(t)$, 其中, $s_F(t)$ 为被摄目标的初始声音信号, $s_B(t)$ 为解说初始声音信号, β_B 为解说声音到第一麦克风的衰减系数, β_F 为被摄目标声音到第二麦克风的衰减系数, $n(t)$ 为噪声信号;

第二信号获取单元, 用于根据公式 $S_F(t) = v_F(t) - \beta_B v_B(t)$ 获取所述第一麦克风得到的最终声音信号 $S_F(t)$, 根据公式 $S_B(t) = v_B(t) - \beta_F v_F(t)$ 获取所述第二麦克风得到的最终声音信号 $S_B(t)$;

调整单元, 用于对所述 $S_F(t)$ 和所述 $S_B(t)$ 进行自动增益调整, 获得被摄目标的声音信号 $s'_F(t)$ 和解说声音信号 $s'_B(t)$ 。

7. 根据权利要求 6 所述的装置, 其特征在于, 所述第二信号获取单元还包括:

第一计算子单元, 用于将所述 $v_F(t)$ 和 $v_B(t)$ 代入所述 $S_F(t)$, 得到 $S_F(t) = s_F(t) (1 - \beta_B \beta_F) + (1 - \beta_B) n(t)$;

第二计算子单元,用于将所述 $v_F(t)$ 和 $v_B(t)$ 代入所述 $S_B(t)$,得到 $S_B(t) = s_B(t) (1 - \beta_B \beta_F) + (1 - \beta_F)n(t)$ 。

8. 根据权利要求7所述的装置,其特征在于,所述调整单元还用于:

调节所述 $S_F(t)$ 和所述 $S_B(t)$ 的增益值,将调节后的 $S_F(t)$ 信号作为 $s'_F(t)$,将调节后的 $S_B(t)$ 信号作为 $s'_B(t)$ 。

9. 根据权利要求6所述的装置,其特征在于,还包括:

第一音量调节单元,用于根据用户需要分别对所述 $s'_F(t)$ 和所述 $s'_B(t)$ 的声音强度进行调节;或者,

第二音量调节单元,用于根据用户需要对所述 $s'_F(t)$ 和所述 $s'_B(t)$ 的声音强度进行联动调解,当调高其中一个信号的声音强度时,另一个信号的声音强度相应降低。

10. 根据权利要求6-9任一项所述的装置,其特征在于,所述第一麦克风与被摄目标的距离小于所述第二麦克风与所述被摄目标的距离;所述第一麦克风与解说者的距离大于所述第二麦克风与所述解说者的距离。

摄像声音提取方法及装置

技术领域

[0001] 本发明涉及通信技术领域,特别涉及一种摄像声音提取方法及装置。

背景技术

[0002] 随着音频技术的发展,人们对音乐品质的要求也越来越高。现有的音频技术主要包括单声道、双声道立体声和声道环绕声。其中,单声道音频在录制过程中使用一个麦克风,播放效果仅能够体现出声音的强弱;双声道立体声音频在录制过程中模拟人的双耳,在播放时使用两个喇叭分别播放左声道和右声道录制的信号,能够使听者既感知到声音的强弱,又感知到声音的水平方位变化;声道环绕声利用多个麦克风进行录制,并且使用多个喇叭进行播放,能够让听者感知声音的空间方位,具有立体的现场感觉。

[0003] 现有技术中至少存在如下问题:目前的单声道、双声道立体声以及声道环绕声只是将麦克风或麦克风阵列感知的信号还原为所录制的声音,但是对于所录制的声音并没有区分。例如,用户用智能手机为朋友所参加的某项表演比赛录像,并在录像的同时进行解说,当将其录制的声音还原时,既有解说音也有其朋友的声音,同时还有一定的干扰音源。由于用户距离手机要比其朋友近得多,所以解说声音要远大于其朋友的声音,再加上干扰音源的干扰,使得混合信号播放出来以后效果很差,无法对解说者的声音和被摄目标的声音进行清晰的分辨。

发明内容

[0004] 提供一种摄像声音提取方法及装置,能够解决录像过程中由于解说者和被摄目标距离麦克风远近不同而带来的声音强度差别很大的问题,将解说声音和被摄目标声音分别清晰的播放出来,并降低干扰音源的影响。

[0005] 第一方面,提供一种摄像声音提取方法,包括:

[0006] 根据公式 $v_F(t) = s_F(t) + \beta_B s_B(t) + n(t)$ 获取第一麦克风得到的初始声音信号 $v_F(t)$,根据公式 $v_B(t) = s_B(t) + \beta_F s_F(t) + n(t)$ 获取第二麦克风得到的初始声音信号 $v_B(t)$,其中, $s_F(t)$ 为被摄目标的初始声音信号, $s_B(t)$ 为解说初始声音信号, β_B 为解说声音到第一麦克风的衰减系数, β_F 为被摄目标声音到第二麦克风的衰减系数, $n(t)$ 为噪声信号;

[0007] 根据公式 $S_F(t) = v_F(t) - \beta_B v_B(t)$ 获取所述第一麦克风得到的最终声音信号 $S_F(t)$,根据公式 $S_B(t) = v_B(t) - \beta_F v_F(t)$ 获取所述第二麦克风得到的最终声音信号 $S_B(t)$;

[0008] 对所述 $S_F(t)$ 和所述 $S_B(t)$ 进行自动增益调整,获得被摄目标的声音信号 $s'_F(t)$ 和解说声音信号 $s'_B(t)$ 。

[0009] 在第一种可能的实现方式中,在所述对所述 $S_F(t)$ 和所述 $S_B(t)$ 进行自动增益调整之前,还包括:

[0010] 将所述 $v_F(t)$ 和 $v_B(t)$ 代入所述 $S_F(t)$,得到 $S_F(t) = s_F(t) (1 - \beta_B \beta_F) + (1 - \beta_B) n(t)$;将所述 $v_F(t)$ 和 $v_B(t)$ 代入所述 $S_B(t)$,得到 $S_B(t) = s_B(t) (1 - \beta_B \beta_F) + (1 - \beta_F) n(t)$ 。将所述 $v_F(t)$ 和 $v_B(t)$ 代入所述 $S_F(t)$,得到 $S_F(t) = s_F(t) (1 - \beta_B \beta_F) + (1 - \beta_B) n(t)$,将所述 $v_F(t)$

和 $v_B(t)$ 代入所述 $S_B(t)$, 得到 $S_B(t) = s_B(t)(1 - \beta_B \beta_F) + (1 - \beta_F)n(t)$ 。

[0011] 可选的, 所述对所述 $S_F(t)$ 和所述 $S_B(t)$ 进行自动增益调整, 获得被摄目标的声音信号 $s'_F(t)$ 和解说声音信号 $s'_B(t)$ 包括:

[0012] 调节所述 $S_F(t)$ 和所述 $S_B(t)$ 的增益值, 将调节后的 $S_F(t)$ 信号作为 $s'_F(t)$, 将调节后的 $S_B(t)$ 信号作为 $s'_B(t)$ 。

[0013] 进一步的, 在所述获得被摄目标的声音信号 $s'_F(t)$ 和解说声音信号 $s'_B(t)$ 之后, 还包括: 根据用户需要分别对所述 $s'_F(t)$ 和所述 $s'_B(t)$ 的声音强度进行调节; 或者,

[0014] 根据用户需要对所述 $s'_F(t)$ 和所述 $s'_B(t)$ 的声音强度进行联动调解, 当调高其中一个信号的声音强度时, 另一个信号的声音强度相应降低。

[0015] 其中, 所述第一麦克风与被摄目标的距离小于所述第二麦克风与所述被摄目标的距离; 所述第一麦克风与解说者的距离大于所述第二麦克风与所述解说者的距离。

[0016] 第二方面, 提供一种摄像声音提取装置, 包括:

[0017] 第一信号获取单元, 用于根据公式 $v_F(t) = s_F(t) + \beta_B s_B(t) + n(t)$ 获取第一麦克风得到的初始声音信号 $v_F(t)$, 根据公式 $v_B(t) = s_B(t) + \beta_F s_F(t) + n(t)$ 获取第二麦克风得到的初始声音信号 $v_B(t)$, 其中, $s_F(t)$ 为被摄目标的初始声音信号, $s_B(t)$ 为解说初始声音信号, β_B 为解说声音到第一麦克风的衰减系数, β_F 为被摄目标声音到第二麦克风的衰减系数, $n(t)$ 为噪声信号;

[0018] 第二信号获取单元, 用于根据公式 $S_F(t) = v_F(t) - \beta_B v_B(t)$ 获取所述第一麦克风得到的最终声音信号 $S_F(t)$, 根据公式 $S_B(t) = v_B(t) - \beta_F v_F(t)$ 获取所述第二麦克风得到的最终声音信号 $S_B(t)$;

[0019] 调整单元, 用于对所述 $S_F(t)$ 和所述 $S_B(t)$ 进行自动增益调整, 获得被摄目标的声音信号 $s'_F(t)$ 和解说声音信号 $s'_B(t)$ 。

[0020] 在第一种可能的实现方式中, 所述第二信号获取单元还包括:

[0021] 第一计算子单元, 用于将所述 $v_F(t)$ 和 $v_B(t)$ 代入所述 $S_F(t)$, 得到 $S_F(t) = s_F(t)(1 - \beta_B \beta_F) + (1 - \beta_B)n(t)$;

[0022] 第二计算子单元, 用于将所述 $v_F(t)$ 和 $v_B(t)$ 代入所述 $S_B(t)$, 得到 $S_B(t) = s_B(t)(1 - \beta_B \beta_F) + (1 - \beta_F)n(t)$ 。

[0023] 可选的, 所述调整单元还用于:

[0024] 调节所述 $S_F(t)$ 和所述 $S_B(t)$ 的增益值, 将调节后的 $S_F(t)$ 信号作为 $s'_F(t)$, 将调节后的 $S_B(t)$ 信号作为 $s'_B(t)$ 。

[0025] 进一步的, 所述装置还包括:

[0026] 第一音量调节单元, 用于根据用户需要分别对所述 $s'_F(t)$ 和所述 $s'_B(t)$ 的声音强度进行调节; 或者,

[0027] 第二音量调节单元, 用于根据用户需要对所述 $s'_F(t)$ 和所述 $s'_B(t)$ 的声音强度进行联动调解, 当调高其中一个信号的声音强度时, 另一个信号的声音强度相应降低。

[0028] 其中, 所述第一麦克风与被摄目标的距离小于所述第二麦克风与所述被摄目标的距离; 所述第一麦克风与解说者的距离大于所述第二麦克风与所述解说者的距离。

[0029] 与现有技术相比, 本发明实施例首先通过两个麦克风获取解说音初始信号和被摄目标初始信号; 然后利用新的算法消除解说音初始信号中被摄目标的干扰信号, 消除被摄

目标初始信号中解说音的干扰信号,通过公式变换获得二者最终的声音信号;最后通过自动增益调整,消除噪声信号的干扰,获得清晰的被摄目标声音和解说者的声音。能够获得清晰的解说者的声音和被摄目标的声音,可以在视频中播放解说者的声音和/或被摄目标的声音,消除了解说者和被摄目标由于距离麦克风的远近差异带来的声音强弱影响,并降低周围环境噪声的影响,提高用户体验。

附图说明

[0030] 为了更清楚地说明本发明实施例中的技术方案,下面将对实施例或现有技术描述中所需要使用的附图作简单地介绍,显而易见地,下面描述中的附图仅仅是本发明的一些实施例,对于本领域普通技术人员来讲,在不付出创造性劳动的前提下,还可以根据这些附图获得其它的附图。

[0031] 图 1 为本发明实施例一提供的方法流程图;

[0032] 图 2 为本发明实施例二提供的方法流程图;

[0033] 图 3 为本发明实施例二提供的解说者和被拍摄者的空间位置示意图;

[0034] 图 4 为本发明实施例二提供的音量调节示意图;

[0035] 图 5、图 6 为本发明实施例三提供的装置结构示意图。

具体实施方式

[0036] 下面将结合本发明实施例中的附图,对本发明实施例中的技术方案进行清楚、完整地描述,显然,所描述的实施例仅仅是本发明一部分实施例,而不是全部的实施例。基于本发明中的实施例,本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其它实施例,都属于本发明保护的范围。

[0037] 为使本发明技术方案的优点更加清楚,下面结合附图和实施例对本发明作详细说明。

[0038] 实施例一

[0039] 本实施例提供一种摄像声音提取方法,如图 1 所示,所述方法包括:

[0040] 101、获取第一麦克风得到的初始声音信号 $v_F(t)$ 和第二麦克风得到的初始声音信号 $v_B(t)$ 。

[0041] 例如,根据公式 $v_F(t) = s_F(t) + \beta_B s_B(t) + n(t)$ 得到 $v_F(t)$,根据公式 $v_B(t) = s_B(t) + \beta_F s_F(t) + n(t)$ 得到 $v_B(t)$,其中, $s_F(t)$ 为被摄目标的初始声音信号, $s_B(t)$ 为解说初始声音信号, β_B 为解说声音到第一麦克风的衰减系数, β_F 为被摄目标声音到第二麦克风的衰减系数, $n(t)$ 为噪声信号。

[0042] 102、根据 $v_F(t)$ 和 $v_B(t)$ 获取第一麦克风得到的最终声音信号 $S_F(t)$ 以及获取第二麦克风得到的最终声音信号 $S_B(t)$ 。

[0043] 例如,根据公式 $S_F(t) = v_F(t) - \beta_B v_B(t)$ 获取 $S_F(t)$,根据公式 $S_B(t) = v_B(t) - \beta_F v_F(t)$ 获取 $S_B(t)$ 。将所述 $v_F(t)$ 和 $v_B(t)$ 代入所述 $S_F(t)$,得到 $S_F(t) = s_F(t)(1 - \beta_B \beta_F) + (1 - \beta_B)n(t)$,将所述 $v_F(t)$ 和 $v_B(t)$ 代入所述 $S_B(t)$,得到 $S_B(t) = s_B(t)(1 - \beta_B \beta_F) + (1 - \beta_F)n(t)$ 。

[0044] 103、对所述 $S_F(t)$ 和所述 $S_B(t)$ 进行自动增益调整,获得被摄目标的声音信号 $s'_F(t)$ 和解说声音信号 $s'_B(t)$ 。

[0045] 例如,所述对所述 $S_F(t)$ 和所述 $S_B(t)$ 进行自动增益调整,获得被摄目标的声音信号 $s'_F(t)$ 和解说声音信号 $s'_B(t)$ 包括:调节所述 $S_F(t)$ 和所述 $S_B(t)$ 的信号增益值,将调节后的 $S_F(t)$ 信号作为 $s'_F(t)$,将调节后的 $S_B(t)$ 信号作为 $s'_B(t)$ 。

[0046] 可选的,在无解说音时,将 $S_F(t) = s_F(t)(1 - \beta_B \beta_F) + (1 - \beta_B)n(t)$ 式中的 β_B 调整为 1,将 β_F 调整为 0,获得被摄目标的声音信号 $S_F(t) = s_F(t)$ 。

[0047] 进一步的,在所述获得被摄目标的声音信号 $s'_F(t)$ 和解说声音信号 $s'_B(t)$ 之后,还包括:

[0048] 根据用户需要分别对所述 $s'_F(t)$ 和所述 $s'_B(t)$ 的声音强度进行调节;或者,

[0049] 根据用户需要对所述 $s'_F(t)$ 和所述 $s'_B(t)$ 的声音强度进行联动调解,当调高其中一个信号的声音强度时,另一个信号的声音强度相应降低。

[0050] 其中,所述第一麦克风与被摄目标的距离小于所述第二麦克风与所述被摄目标的距离;所述第一麦克风与解说者的距离大于所述第二麦克风与所述解说者的距离。

[0051] 与现有技术相比,本发明实施例首先通过两个麦克风获取解说音初始信号和被摄目标初始信号;然后利用新的算法消除解说音初始信号中被摄目标的干扰信号,消除被摄目标初始信号中解说音的干扰信号,通过公式变换获得二者最终的声音信号;最后通过自动增益调整,消除噪声信号的干扰,获得清晰的被摄目标声音和解说者的声音。能够获得清晰的解说者的声音和被摄目标的声音,可以在视频中播放解说者的声音和/或被摄目标的声音,消除了解说者和被摄目标由于距离麦克风的远近差异带来的声音强弱影响,并降低周围环境噪声的影响,提高用户体验。

[0052] 实施例二

[0053] 本实施例提供一种摄像声音提取方法,以通过智能终端进行摄像为例,如图 2 所示,所述方法包括:

[0054] 201、通过后置麦克风接收被拍摄者的初始声音信号 $s_F(t)$,同时通过前置麦克风接收解说者的初始声音信号 $s_B(t)$ 。

[0055] 例如,如图 3 所示,当用智能终端对目标进行摄像时,因为解说者(一般来说即为摄像者)位于终端的正面,距离前置的麦克风较近,而目标位于终端背面,距离后置的麦克风较近,所以为了提取更加清晰的声音,使用终端正面前置的麦克风接收解说者的声音信号,使用终端背面(与摄像头同侧)后置的麦克风接收目标的声音信号。

[0056] 当然,由于声音的传播特性,前置的麦克风在接收解说者声音信号的同时,也会接收到一定的被拍摄者的声音信号。同理,后置麦克风也接收到解说者的声音信号。再加上不可避免的噪声信号,所以需要综合考虑,利用步骤 202 得到一个总的声音信号

[0057] 202、获取第一麦克风得到的初始声音信号 $v_F(t)$ 和第二麦克风得到的初始声音信号 $v_B(t)$ 。

[0058] 例如,根据公式 $v_F(t) = s_F(t) + \beta_B s_B(t) + n(t)$ 获取第一麦克风得到的初始声音信号 $v_F(t)$,根据公式 $v_B(t) = s_B(t) + \beta_F s_F(t) + n(t)$ 获取第二麦克风得到的初始声音信号 $v_B(t)$ 。其中, β_B 为解说声音到第一麦克风的衰减系数, $\beta_B s_B(t)$ 即为后置麦克风接收到的经过衰减的解说者的声音信号, β_F 为被摄目标声音到第二麦克风的衰减系数, $\beta_F s_F(t)$ 即为前置麦克风接收到的经过衰减的被拍摄者的声音信号, $n(t)$ 为噪声信号。可选的,默认被拍摄者与后置麦克风之间的声音传播信道为 1,默认解说者与前置麦克风之间的声音传

播信道为 1。

[0059] 其中,经过此步骤得出的 $v_F(t)$ 和 $v_B(t)$ 为混合的声音信号,如果按照此信号进行声音的还原,那么用户听到的声音将会十分嘈杂,不一定能分辨出解说者和被拍摄者的声音,所以需要利用步骤 203 对这两个声音信号作进一步的处理。

[0060] 需要说明的是,两个衰减系数 β_B 和 β_F ,是根据时域和频域的匹配算法估计出来的,该匹配算法为本领域人员所熟知的现有技术,此处不再对其进行详细的说明。

[0061] 203、构造信号 $S_F(t)$ 和 $S_B(t)$ 。其中, $S_F(t)$ 为后置麦克风获得的被拍摄者的最终声音信号, $S_B(t)$ 为前置麦克风获得的解说者的最终声音信号。

[0062] 例如,使用公式 $S_F(t)=v_F(t)-\beta_B v_B(t)$ 得到信号 $S_F(t)$,利用此公式可以消除解说者给后置麦克风带来的声音干扰,将步骤 202 得到的 $v_F(t)$ 和 $v_B(t)$ 代入该公式,得到: $S_F(t)=s_F(t)-\beta_B \beta_F s_F(t)+(1-\beta_B)n(t)$,经过整理后得到: $S_F(t)=s_F(t)(1-\beta_B \beta_F)+(1-\beta_B)n(t)$,可以看出经过整理后的公式中并没有解说者的初始声音信号 $s_B(t)$,只包括被拍摄者的初始声音信号和噪声信号。

[0063] 同理,使用公式 $S_B(t)=v_B(t)-\beta_F v_F(t)$ 得到信号 $S_B(t)$,利用此公式可以消除被拍摄者给前置麦克风带来的声音干扰,将步骤 202 得到的 $v_F(t)$ 和 $v_B(t)$ 代入该公式,得到: $S_B(t)=s_B(t)-\beta_B \beta_F s_B(t)+(1-\beta_F)n(t)$,经过整理后得到: $S_B(t)=s_B(t)(1-\beta_B \beta_F)+(1-\beta_F)n(t)$,可以看出经过整理后的公式中并没有被拍摄者的初始声音信号 $s_F(t)$,只包括解说者的初始声音信号和噪声信号。

[0064] 进一步的,执行步骤 204,对整理后的声音信号进行微调,消除噪声的干扰,最大化的提取解说者和被拍摄者的声音信号。

[0065] 204、对信号 $S_F(t)$ 和 $S_B(t)$ 进行自动增益调整。

[0066] 例如,调整 $S_F(t)$ 和 $S_B(t)$ 的增益值,将最清晰的 $S_F(t)$ 信号作为 $s'_F(t)$,将最清晰的 $S_B(t)$ 信号作为 $s'_B(t)$ 。

[0067] 可选的,当仅有被拍摄者的声音,并没有解说时,可以令 β_B 为 1,令 β_F 为 0,获得被拍摄者的声音信号 $S_F(t)=s'_F(t)$,即仅存在清晰的被拍摄者的声音,噪声几乎为 0。在此种情形下,能够显著降低周围噪声对摄像的影响,提高录制效果。

[0068] 205、对被拍摄者的声音信号和解说者的声音信号进行还原。

[0069] 206、根据用户需要对被拍摄者的声音强度和解说者的声音强度进行调节。

[0070] 可选的,在实际播放时,可以根据用户喜好只播放被拍摄者的声音,或者只播放解说者的声音,或者同时播放二者的声音。例如,在进行调节时,如图 4 所示,调节方案 A 中,可以分别对解说音和目标音(被拍摄者的声音)进行调节,二者的调节互不干扰;调节方案 B 中,对解说音和目标音进行联动调节,即当调高解说音时目标音降低相应的音量,反之亦然;调节方案 C 中,使用旋钮对解说音和目标音进行联动调节,能够达到和方案 B 相同的效果。

[0071] 与现有技术相比,本发明实施例首先通过两个麦克风获取解说音初始信号和被摄目标初始信号;然后利用新的算法消除解说音初始信号中被摄目标的干扰信号,消除被摄目标初始信号中解说音的干扰信号,通过公式变换获得二者最终的声音信号;最后通过自动增益调整,消除噪声信号的干扰,获得清晰的被摄目标声音和解说者的声音。能够获得清晰的解说者的声音和被摄目标的声音,可以在视频中播放解说者的声音和/或被摄目标的

声音,消除了解说者和被摄目标由于距离麦克风的远近差异带来的声音强弱影响,并降低周围环境噪声的影响,提高用户体验。

[0072] 实施例三

[0073] 本实施例提供一种摄像声音提取装置 30,如图 5 所示,所述装置 30 包括:

[0074] 第一信号获取单元 31,用于根据公式 $v_F(t) = s_F(t) + \beta_B s_B(t) + n(t)$ 获取第一麦克风得到的初始声音信号 $v_F(t)$,根据公式 $v_B(t) = s_B(t) + \beta_F s_F(t) + n(t)$ 获取第二麦克风得到的初始声音信号 $v_B(t)$,其中, $s_F(t)$ 为被摄目标的初始声音信号, $s_B(t)$ 为解说初始声音信号, β_B 为解说声音到第一麦克风的衰减系数, β_F 为被摄目标声音到第二麦克风的衰减系数, $n(t)$ 为噪声信号;

[0075] 第二信号获取单元 32,用于根据公式 $S_F(t) = v_F(t) - \beta_B v_B(t)$ 获取所述第一麦克风得到的最终声音信号 $S_F(t)$,根据公式 $S_B(t) = v_B(t) - \beta_F v_F(t)$ 获取所述第二麦克风得到的最终声音信号 $S_B(t)$;

[0076] 调整单元 33,用于对所述 $S_F(t)$ 和所述 $S_B(t)$ 进行自动增益调整,获得被摄目标的声音信号 $s'_F(t)$ 和解说声音信号 $s'_B(t)$ 。

[0077] 进一步的,如图 6 所示,所述第二信号获取单元 32 还可以包括:

[0078] 第一计算子单元 321,用于将所述 $v_F(t)$ 和 $v_B(t)$ 代入所述 $S_F(t)$,得到 $S_F(t) = s_F(t) (1 - \beta_B \beta_F) + (1 - \beta_B) n(t)$;

[0079] 第二计算子单元 322,用于将所述 $v_F(t)$ 和 $v_B(t)$ 代入所述 $S_B(t)$,得到 $S_B(t) = s_B(t) (1 - \beta_B \beta_F) + (1 - \beta_F) n(t)$ 。

[0080] 进一步的,如图 6 所示,所述调整单元 33 还用于:

[0081] 调整所述 $S_F(t)$ 和所述 $S_B(t)$ 的增益值,将调整后的 $S_F(t)$ 信号作为 $s'_F(t)$,将调整后的 $S_B(t)$ 信号作为 $s'_B(t)$ 。

[0082] 进一步的,如图 6 所示,所述装置 30 还可以包括:

[0083] 第一音量调节单元 34,用于根据用户需要分别对所述 $s'_F(t)$ 和所述 $s'_B(t)$ 的声音强度进行调节;或者,

[0084] 第二音量调节单元 35,用于根据用户需要对所述 $s'_F(t)$ 和所述 $s'_B(t)$ 的声音强度进行联动调解,当调高其中一个信号的声音强度时,另一个信号的声音强度相应降低。

[0085] 其中,所述第一麦克风与被摄目标的距离小于所述第二麦克风与所述被摄目标的距离;所述第一麦克风与解说者的距离大于所述第二麦克风与所述解说者的距离。

[0086] 与现有技术相比,本发明实施例首先通过两个麦克风获取解说音初始信号和被摄目标初始信号;然后利用新的算法消除解说音初始信号中被摄目标的干扰信号,消除被摄目标初始信号中解说音的干扰信号,通过公式变换获得二者最终的声音信号;最后通过自动增益调整,消除噪声信号的干扰,获得清晰的被摄目标声音和解说者的声音。能够获得清晰的解说者的声音和被摄目标的声音,可以在视频中播放解说者的声音和/或被摄目标的声音,消除了解说者和被摄目标由于距离麦克风的远近差异带来的声音强弱影响,并降低周围环境噪声的影响,提高用户体验。

[0087] 本发明实施例提供的摄像声音提取装置可以实现上述提供的方法实施例,具体功能实现请参见方法实施例中的说明,在此不再赘述。本发明实施例提供的摄像声音提取方法及装置可以适用于在摄像过程中对不同目标的声音进行区分和调整,但不仅限于此。

[0088] 本领域普通技术人员可以理解实现上述实施例方法中的全部或部分流程,是可以
通过计算机程序来指令相关的硬件来完成,所述的程序可存储于一计算机可读取存储介质
中,该程序在执行时,可包括如上述各方法的实施例的流程。其中,所述的存储介质可为磁
碟、光盘、只读存储记忆体(Read-Only Memory, ROM)或随机存储记忆体(Random Access
Memory, RAM)等。

[0089] 以上所述,仅为本发明的具体实施方式,但本发明的保护范围并不局限于此,任何
熟悉本技术领域的技术人员在本发明揭露的技术范围内,可轻易想到的变化或替换,都应
涵盖在本发明的保护范围之内。因此,本发明的保护范围应该以权利要求的保护范围为准。

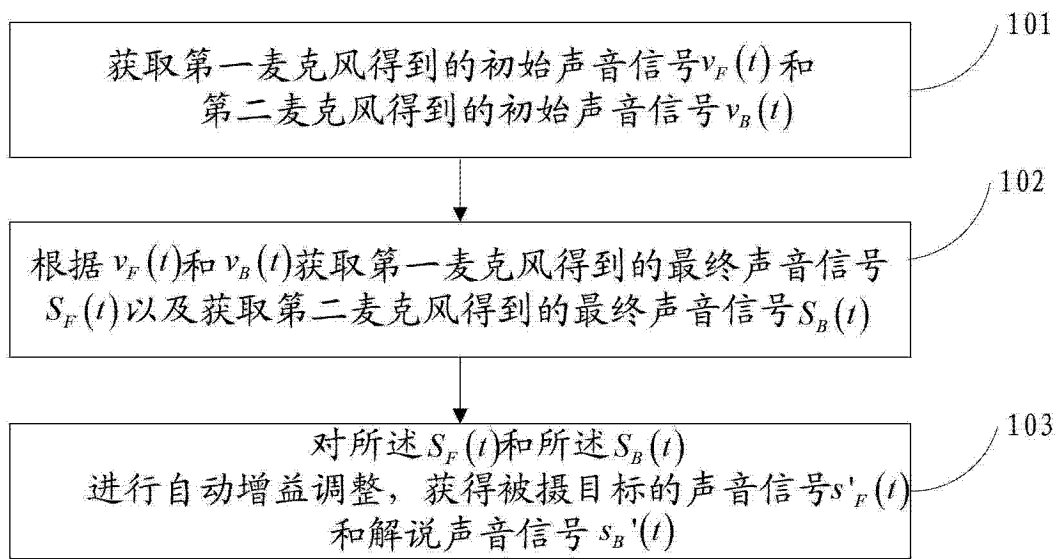


图 1

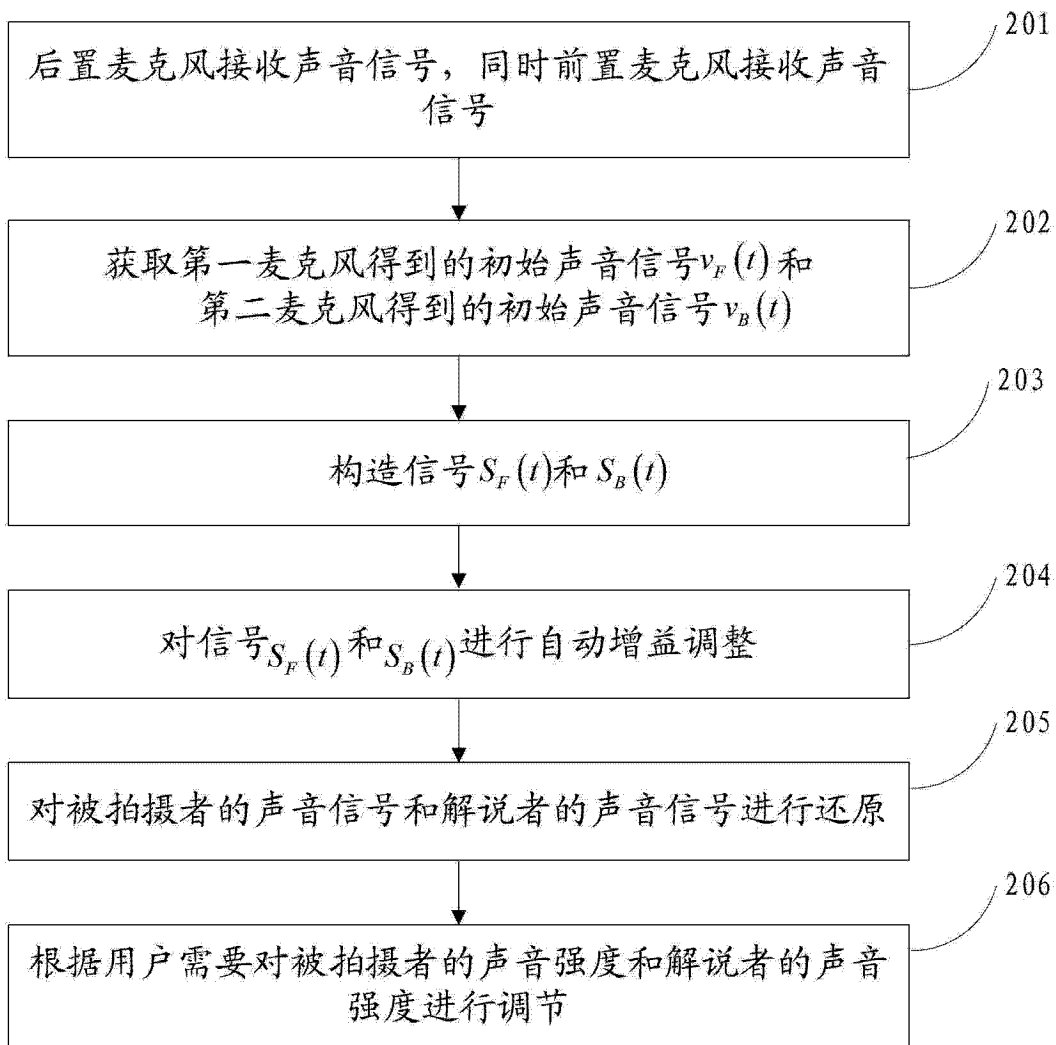


图 2

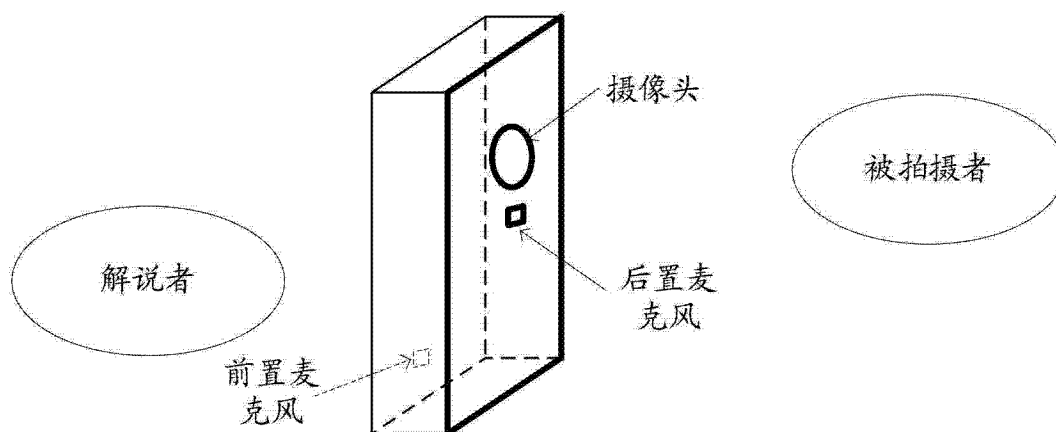


图 3

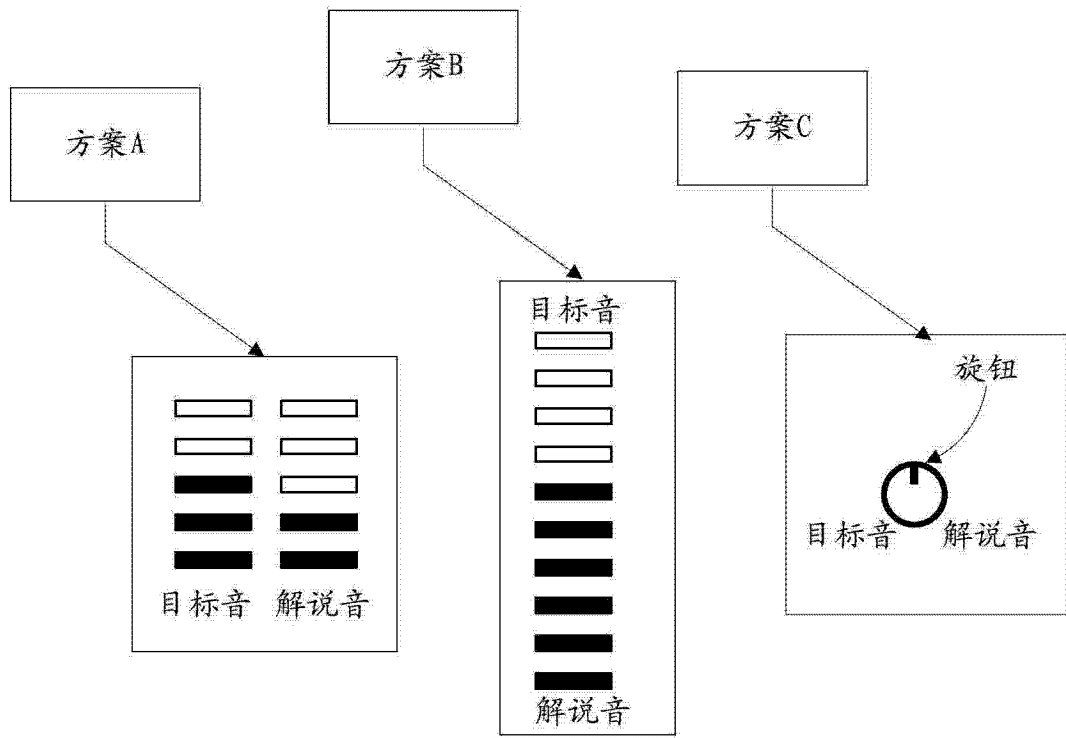


图 4

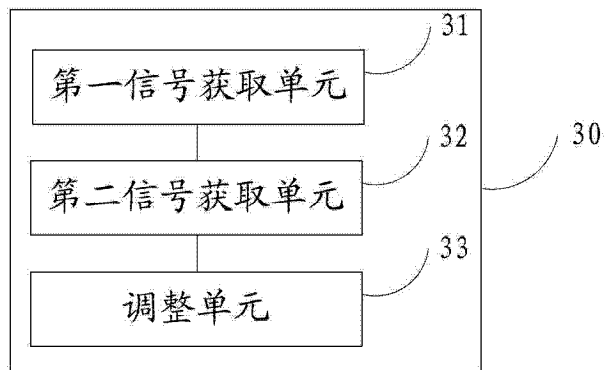


图 5

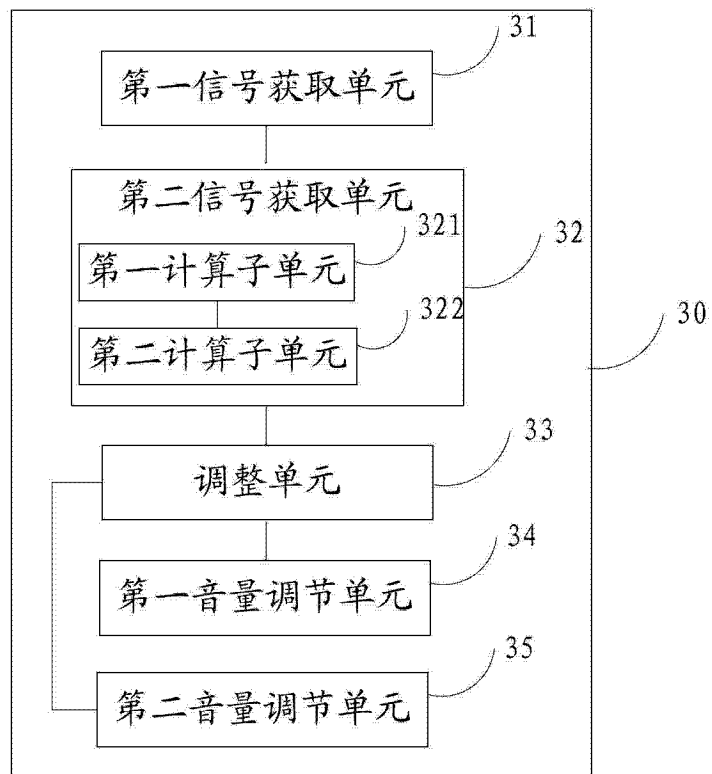


图 6