(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国 际 局

(43) 国际公布日
2018 年 6 月 28 日 (28.06.2018)

WIPO | PCT

(10) 国际公布号
WO 2018/113243 A1

(51) 国际专利分类号:
G10L 15/04 (2013.01)

(21) 国际申请号: PCT/CN2017/091310

(22) 国际申请日: 2017 年 6 月 30 日 (30.06.2017)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
2016111767919 2016年12月19日 (19.12.2016) CN

(71) 申请人: 平安科技（深圳）有限公司 (PING AN TECHNOLOGY (SHENZHEN) CO.,LTD.) [CN/CN]; 中国广东省深圳福田区八卦岭八卦三路平安大厦吴东勤, Guangdong 518000 (CN)。

(72) 发明人: 王健宗 (WANG, Jianzong); 中国广东省深圳福田区八卦岭八卦三路平安大厦吴东勤, Guangdong 518000 (CN)。 郭卉 (GUO, Hui); 中国广东省深圳福田区八卦岭八卦三路平安大厦吴东勤, Guangdong 518000 (CN)。 肖京(XIAO, Jing); 中国广东省深圳福田区八卦岭八卦三路平安大厦吴东勤, Guangdong 518000 (CN)。

(74) 代理人: 深圳市沃德知识产权代理事务所（普通合伙） (SHENZHEN WORLD INTELLECTUAL PROPERTY AGENCY (GENERAL PARTNERSHIP)); 中国广东省深圳福田区园岭街道八卦四路10号中浩大厦1528-1530室于志光, Guangdong 518000 (CN)。

(81) 指定国(除另有指明，要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK,

(54) **Title:** SPEECH SEGMENTATION METHOD, DEVICE AND APPARATUS, AND COMPUTER STORAGE MEDIUM

(54) **发明名称**: 语音分割的方法、装置、设备及计算机存储介质

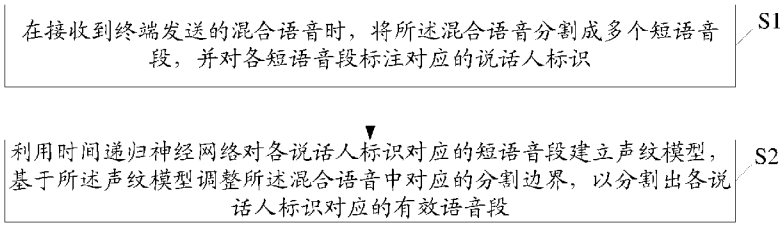


图 1

S1 When a mixed speech sent by a terminal is received, segmenting the mixed speech into a plurality of phrase segments, and labelling each phrase segment with a corresponding speaker identifier

S2 Using a recurrent neural network to build a voiceprint model for the phrase segment corresponding to each speaker identifier, and adjusting a corresponding segmentation boundary in the mixed speech based on the voiceprint model so as to segment out an effective speech segment corresponding to each speaker identifier

(57) **Abstract:** Disclosed are a speech segmentation method and device. The speech segmentation method comprises: when a mixed speech sent by a terminal is received, segmenting the mixed speech into a plurality of phrase segments, and labelling each phrase segment with a corresponding speaker identifier (S1); and using a recurrent neural network to build a voiceprint model for the phrase segment corresponding to each speaker identifier, and adjusting a corresponding segmentation boundary in the mixed speech based on the voiceprint model so as to segment out an effective speech segment corresponding to each speaker identifier (S2). The method can effectively improve the precision of speech segmentation and has a better speech segmentation effect especially on frequently alternate conversations and overlapping speeches.

(57) **摘要**: 一种语音分割的方法及装置，语音分割的方法包括：在接收到终端发送的混合语音时，将混合语音分割成多个短语音段，并对各短语音段标注对应的说话人标识（S1）；利用时间递归神经网络对各说话人标识对应的短语音段建立声纹模型，基于声纹模型调整混合语音中对应的分割边界，以分割出各说话人标识对应的有效语音段（S2）。该方法能够有效提高语音分割的精度，特别是对于对话交替频繁、以及有交叠的语音，语音分割的效果较好。

WO 2018/113243 A1

[见续页]

LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。

(84) 指定国(除另有指明，要求每一种可提供的地区保护)：ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

— 包括国际检索报告(条约第21条(3))。

语音分割的方法、装置、设备及计算机存储介质

优先权申明

本申请基于巴黎公约申明享有 2016 年 12 月 19 日递交的申请号为 CN201611176791.9、名称为“语音分割的方法及装置”的中国专利申请的优先权，该中国专利申请的整体内容以参考的方式结合在本申请中。

技术领域

本发明涉及语音处理技术领域，尤其涉及一种语音分割的方法、装置、设备及计算机存储介质。

背景技术

目前，呼叫中心接收到的语音很多都混杂有多人的语音，这时需要先对语音进行语音分割（speaker diarization），才能进一步对目标语音进行语音分析。语音分割是指：在语音处理领域，当多个说话人的语音被合并录在一个声道中时，把信号中每个说话人的语音分别进行提取。传统的语音分割技术是基于全局背景模型和高斯混合模型进行分割，由于技术的限制，这种语音分割的方法分割的精度并不高，特别是对于对话交替频繁、以及有交叠的对话分割效果差。

发明内容

本发明的目的在于提供一种语音分割的方法、装置、设备及计算机存储介质，旨在有效提高语音分割的精度。

为实现上述目的，本发明提供一种语音分割的方法，其特征在于，所述语音分割的方法包括:

S1，在接收到终端发送的混合语音时，将所述混合语音分割成多个短语音段，并对各短语音段标注对应的说话人标识;

S2，利用时间递归神经网络对各说话人标识对应的短语音段建立声纹模型，基于所述声纹模型调整所述混合语音中对应的分割边界，以分割出各说话人标识对应的有效语音段。

为实现上述目的，本发明还提供一种语音分割的装置，所述语音分割的装置包括:

分割模块，用于在接收到终端发送的混合语音时，将所述混合语音分割成多个短语音段，并对各短语音段标注对应的说话人标识;

调整模块，用于利用时间递归神经网络对各说话人标识对应的短语音段建立声纹模型，基于所述声纹模型调整所述混合语音中对应的分割边界，以分割出各说话人标识对应的有效语音段。

为实现上述目的，本发明还提供一种语音分割设备，所述语音分割设备包括存储器、处理器以及存储在所述存储器上并可在所述处理器上运行的语

音分割程序，所述语音分割程序被所述所述处理器执行时实现如下步骤:

S1，在接收到终端发送的混合语音时，将所述混合语音分割成多个短语音段，并对各短语音段标注对应的说话人标识;

S2，利用时间递归神经网络对各说话人标识对应的短语音段建立声纹模型，基于所述声纹模型调整所述混合语音中对应的分割边界，以分割出各说话人标识对应的有效语音段。

为实现上述目的，本发明还提供计算机可读存储介质，所述计算机可读存储介质上存储有语音分割程序，所述语音分割程序被处理器执行时实现如下步骤:

S1，在接收到终端发送的混合语音时，将所述混合语音分割成多个短语音段，并对各短语音段标注对应的说话人标识;

S2，利用时间递归神经网络对各说话人标识对应的短语音段建立声纹模型，基于所述声纹模型调整所述混合语音中对应的分割边界，以分割出各说话人标识对应的有效语音段。

本发明的有益效果是：本发明首先将混合语音进行分割，分割成多个短语音段，每一短语音段对应标识一个说话人，利用时间递归神经网络对各短语音段建立声纹模型，由于利用时间递归神经网络建立的声纹模型能够关联说话人跨时间点的声音信息，因此基于该声纹模型实现对短语音段的分割边界的调整，能够有效提高语音分割的精度，特别是对于对话交替频繁、以及有交叠的语音，语音分割的效果较好。

附图说明

图 1 为本发明语音分割的方法一实施例的流程示意图;

图 2 为图 1 所示步骤 S1 的细化流程示意图;

图 3 为图 1 所示步骤 S2 的细化流程示意图;

图 4 为本发明语音分割的装置一实施例的结构示意图;

图 5 为图 4 所示分割模块的结构示意图;

图 6 为图 4 所示调整模块的结构示意图;

图 7 为本发明实施例方案涉及的硬件运行环境的设备结构示意图。

具体实施方式

以下结合附图对本发明的原理和特征进行描述，所举实例只用于解释本发明，并非用于限定本发明的范围。

如图 1 所示，图 1 为本发明语音分割的方法一实施例的流程示意图，该语音分割的方法包括以下步骤:

步骤 S1，在接收到终端发送的混合语音时，将所述混合语音分割成多个短语音段，并对各短语音段标注对应的说话人标识;

本实施例提出的方法可应用于呼叫中心的自动应答系统中，例如保险呼叫中心的自动应答系统、各种客服呼叫中心的自动应答系统等等。自动应答

系统接收到终端发送的原始的混合语音，该混合语音中混合有多种不同的声源产生的声音，例如有多人说话混合的声音，多人说话的声音与其他噪声混合的声音等等。

本实施例可以利用预定的方法将混合语音分割成多个短语音段，例如可以利用高斯混合模型（Gaussian Mixture Model，GMM）将混合语音分割成多个短语音段，当然，也可以利用其他传统的方法将混合语音分割成多个短语音段。

其中，经本实施例的语音分割后，每一短语音段应只对应一说话人，不同的短语音段中可能有多个短语音段属于同一个说话人，将同一个说话人的不同短语音段进行相同的标识。

步骤 S2，利用时间递归神经网络对各说话人标识对应的短语音段建立声纹模型，基于所述声纹模型调整所述混合语音中对应的分割边界，以分割出各说话人标识对应的有效语音段。

本实施例中，时间递归神经网络模型（Long-Short Term Memory，LSTM）拥有递归神经网络在传统前向反馈神经网络中引入的定向循环，用以处理层间输入前后、层内输出前后的关联。用时间递归神经网络在语音序列上建模，可以得到跨越时间点的语音信号特征，可以用于对关联信息处于任何长度、任何位置的语音序列进行处理。时间递归神经网络模型通过神经网络层内设计多个交互层，可以记忆到更远时间节点上的信息，在时间递归神经网络模型中用“忘记门层”丢弃与识别任务不相关的信息，接着用“输入门层”决定需要更新的状态，最后确定需要输出的状态并处理输出。

本实施例对于各说话人标识对应的短语音段，利用时间递归神经网络建立声纹模型，通过该声纹模型可以得到说话人跨越时间点的声音信息，基于这些声音信息可以调整混合语音中对应的分割边界，以对每一说话人对应的所有短语音段调整其分割边界，最终分割出各说话人标识对应的有效语音段，该有效语音段可以看作对应的说话人的完整语音。

与现有技术相比，本实施例首先将混合语音进行分割，分割成多个短语音段，每一短语音段对应标识一个说话人，利用时间递归神经网络对各短语音段建立声纹模型，由于利用时间递归神经网络建立的声纹模型能够关联说话人跨时间点的声音信息，因此基于该声纹模型实现对短语音段的分割边界的调整，能够有效提高语音分割的精度，特别是对于对话交替频繁、以及有交叠的语音，语音分割的效果较好。

在一优选的实施例中，如图 2 所示，在上述图 1 的实施例的基础上，上述步骤 S1 包括:

步骤 S11，获取所述混合语音中的静音段，去除所述混合语音中的静音段，以根据所述静音段对所述混合语音进行分割，得到分割后的长语音段;

步骤 S12，对所述长语音段进行分帧，以提取每一长语音段的声学特征;

步骤 S13，对每一长语音段的声学特征进行 KL 距离分析，根据 KL 距

离分析结果对所述语音段进行切分，得到切分后的短语音段，

步骤 S14，利用高斯混合模型对各短语音段进行语音聚类，并对同一语音类的短语音段标注对应的说话人标识。

本实施例中，首先根据静音进行初步分割：确定混合语音中的静音段，将确定的静音段从混合语音中去除，以实现将混合语音根据静音段进行分割，静音段是通过对混合语音的短时语音能量和短时过零率的分析来确定的。

去除静音段后，首先假设在整个混合语音中，每人每次讲话时长为固定阈值 Tu，若某段语音大于该时长，则可能多人说话，若小于该时长，则更可能只有一个人说话，基于这种假设，可以对静音分割后的每个长语音段的时长大于固定阈值 Tu 的语音段的声学特征进行帧间 KL 距离分析。当然，也可以对所有的长语音段的声学特征进行帧间 KL 距离分析。具体地，对得到的长语音段进行分帧，以得到每一长语音段的语音帧，提取语音帧的声学特征，对所有长语音段的声学特征进行 KL 距离（也即相对熵）分析，其中，声学特征包括但不限定于线性预测系数、倒频谱系数 MFCC、平均过零率、短时频谱、共振峰频率及带宽。

其中，KL 距离分析的含义是对于两个离散型的声学特征概率分布集合 P={p1，p2，…，pn}和 Q={q1，q2，…，qn}，P 和 Q 间的 KL 距离：$KL(P\|Q)=\sum_{i=1}^{n} pi \lg\frac{pi}{qi}$，当 KL 距离越大时，PQ 两者差异越大，即 PQ 这两个集合来自两个不同人的语音。优选地，对时长大于预设时间阈值的长语音段在 KL 的最大值处进行切分，以提高语音分割的精度。

长语音段经过切分后得到短语音段，短语音段的数量大于长语音段的数量。然后进行短语音段聚类：对切分后的短语音段进行聚类，以将所有短语音段聚为多个语音类，并为各个短语音段标注对应的说话人标识，其中，属于同一语音类的短语音段标注相同的说话人标识，不属于同一语音类的短语音段标注不同的说话人标识。聚类方法是：采用 K 个成分的高斯混合模型拟合每段短语音段，以均值作为特征向量，使用 k-means 聚类方法把所有短语音段聚为多类。

在一优选的实施例中，如图 3 所示，在上述的实施例的基础上，上述步骤 S2 包括：

步骤 S21，利用所述时间递归神经网络对各说话人标识对应的短语音段建立声纹模型，基于所述声纹模型提取表征说话人身份特征的预设类型向量；

步骤 S22，基于所述预设类型向量计算每一语音帧属于对应的说话人的最大后验概率；

步骤 S23，基于所述最大后验概率并利用预定算法调整该说话人的混合高斯模型；

步骤 S24，基于调整后的混合高斯模型获取每一语音帧对应的概率最大的说话人，并根据概率最大的说话人与语音帧的概率关系调整所述混合语音中对应的分割边界;

步骤 S25，迭代更新所述声纹模型 n 次，每次更新所述声纹模型时迭代 m 次所述混合高斯模型，以得到各说话人对应的有效语音段，n 及 m 均为大于 1 的正整数。

本实施例中，利用时间递归神经网络对各说话人标识对应的短语音段建立声纹模型，基于所述声纹模型提取表征说话人身份特征的预设类型向量，优选地，该预设类型向量为 i-vector 向量，i-vector 向量是反映说话人声学差异的一个重要特征。

在整个混合语音中，根据预设类型向量计算每一语音帧属于某一说话人的最大后验概率，利用计算最大后验概率，在混合语音中通过预设算法重新调整说话人的混合高斯模型，例如，通过 Baum-Welch 算法重新调整说话人的混合高斯模型，该混合高斯模型为 k（一般为 3-5 个）个高斯模型的集合。利用重新调整后的混合高斯模型寻找每一语音帧概率最大的说话人。根据语音帧与寻找到的该说话人的概率关系调整混合语音的分割边界，例如将分割边界向前微调或者向后微调。最后，迭代更新上述声纹模型 n 次，每次更新声纹模型时迭代 m 次混合高斯模型，以得到各个说话人对应的有效语音段，n 及 m 均为大于 1 的正整数。

本实施例借助深度学习的时间递归神经网络建立声纹模型，用各说话人声纹对应的身份特征对应各语音帧以计算语音帧属于某一说话人的概率，基于该概率修正模型，最终调整语音分割的边界，可以有效提高说话人语音分割的精度，降低错误率，且可扩展性好。

在一优选的实施例中，在上述的实施例的基础上，该方法在上述步骤 S2 之后还包括：基于所述有效语音段获取对应的应答内容，并将所述应答内容反馈给所述终端。

本实施例中，自动应答系统关联对应的应答库，该应答库中存储有不同的问题对应的应答内容，自动应答系统在接收到终端发送的混合语音后，将其分割为说话人标识对应的有效语音段，从这些有效语音段中获取与该自动应答系统有关问题的一个有效语音段，针对该有效语音段在应答库中进行匹配，并将匹配得到的应答内容反馈给终端。

如图 4 所示，图 4 为本发明语音分割的装置一实施例的结构示意图，该语音分割的装置包括:

分割模块 101，用于在接收到终端发送的混合语音时，将所述混合语音分割成多个短语音段，并对各短语音段标注对应的说话人标识;

本实施例的语音分割的装置可以应用于自动应答系统，例如保险呼叫中心的自动应答系统、各种客服呼叫中心的自动应答系统等等。该语音分割的

装置的分割模块 101 接收到终端发送的原始的混合语音，该混合语音中混合有多种不同的声源产生的声音，例如有多人说话混合的声音，多人说话的声音与其他噪声混合的声音等等。

本实施例可以利用预定的方法将混合语音分割成多个短语音段，例如可以利用高斯混合模型（Gaussian Mixture Model，GMM）将混合语音分割成多个短语音段，当然，也可以利用其他传统的方法将混合语音分割成多个短语音段。

其中，经本实施例的语音分割后，每一短语音段应只对应一说话人，不同的短语音段中可能有多个短语音段属于同一个说话人，将同一个说话人的不同短语音段进行相同的标识。

调整模块 102，用于利用时间递归神经网络对各说话人标识对应的短语音段建立声纹模型，基于所述声纹模型调整所述混合语音中对应的分割边界，以分割出各说话人标识对应的有效语音段。

本实施例中，时间递归神经网络模型（Long-Short Term Memory，LSTM）拥有递归神经网络在传统前向反馈神经网络中引入的定向循环，用以处理层间输入前后、层内输出前后的关联。用时间递归神经网络在语音序列上建模，可以得到跨越时间点的语音信号特征，可以用于对关联信息处于任何长度、任何位置的语音序列进行处理。时间递归神经网络模型通过神经网络层内设计多个交互层，可以记忆到更远时间节点上的信息，在时间递归神经网络模型中用“忘记门层”丢弃与识别任务不相关的信息，接着用“输入门层”决定需要更新的状态，最后确定需要输出的状态并处理输出。

本实施例对于各说话人标识对应的短语音段，利用时间递归神经网络建立声纹模型，通过该声纹模型可以得到说话人跨越时间点的声音信息，基于这些声音信息可以调整混合语音中对应的分割边界，以对每一说话人对应的所有短语音段调整其分割边界，最终分割出各说话人标识对应的有效语音段，该有效语音段可以看作对应的说话人的完整语音。

在一优选的实施例中，如图 5 所示，在上述图 4 的实施例的基础上，上述分割模块 101 包括:

去除单元 1011，用于获取所述混合语音中的静音段，去除所述混合语音中的静音段，以根据所述静音段对所述混合语音进行分割，得到分割后的长语音段;

分帧单元 1012，用于对所述长语音段进行分帧，以提取每一长语音段的声学特征;

切分单元 1013，用于对每一长语音段的声学特征进行 KL 距离分析，根据 KL 距离分析结果对所述语音段进行切分，得到切分后的短语音段;

聚类单元 1014，用于利用高斯混合模型对各短语音段进行语音聚类，并对同一语音类的短语音段标注对应的说话人标识。

本实施例中，首先根据静音进行初步分割：确定混合语音中的静音段，

将确定的静音段从混合语音中去除，以实现将混合语音根据静音段进行分割，静音段是通过对混合语音的短时语音能量和短时过零率的分析来确定的。

去除静音段后，首先假设在整个混合语音中，每人每次讲话时长为固定阈值 Tu，若某段语音大于该时长，则可能多人说话，若小于该时长，则更可能只有一个人说话，基于这种假设，可以对静音分割后的每个长语音段的时长大于固定阈值 Tu 的语音段的声学特征进行帧间 KL 距离分析。当然，也可以对所有的长语音段的声学特征进行帧间 KL 距离分析。具体地，对得到的长语音段进行分帧，以得到每一长语音段的语音帧，提取语音帧的声学特征，对所有长语音段的声学特征进行 KL 距离（也即相对熵）分析，其中，声学特征包括但不限定于线性预测系数、倒频谱系数 MFCC、平均过零率、短时频谱、共振峰频率及带宽。

其中，KL 距离分析的含义是对于两个离散型的声学特征概率分布集合 P={p1，p2，…，pn}和 Q={q1，q2，…，qn}，P 和 Q 间的 KL 距离：$KL(P\|Q)=\sum_{i=1}^{n} pi \lg\frac{pi}{qi}$，当 KL 距离越大时，PQ 两者差异越大，即 PQ 这两个集合来自两个不同人的语音。优选地，对时长大于预设时间阈值的长语音段在 KL 的最大值处进行切分，以提高语音分割的精度。

长语音段经过切分后得到短语音段，短语音段的数量大于长语音段的数量。然后进行短语音段聚类：对切分后的短语音段进行聚类，以将所有短语音段聚为多个语音类，并为各个短语音段标注对应的说话人标识，其中，属于同一语音类的短语音段标注相同的说话人标识，不属于同一语音类的短语音段标注不同的说话人标识。聚类方法是：采用 K 个成分的高斯混合模型拟合每段短语音段，以均值作为特征向量，使用 k-means 聚类方法把所有短语音段聚为多类。

在一优选的实施例中，如图 6 所示，在上述实施例的基础上，上述调整模块 102 包括：

建模单元 1021，用于利用所述时间递归神经网络对各说话人标识对应的短语音段建立声纹模型，基于所述声纹模型提取表征说话人身份特征的预设类型向量；

计算单元 1022，用于基于所述预设类型向量计算每一语音帧属于对应的说话人的最大后验概率；

第一调整单元 1023，用于基于所述最大后验概率并利用预定算法调整该说话人的混合高斯模型；

第二调整单元 1024，用于基于调整后的混合高斯模型获取每一语音帧对应的概率最大的说话人，并根据概率最大的说话人与语音帧的概率关系调整所述混合语音中对应的分割边界；

迭代单元 1025，用于迭代更新所述声纹模型 n 次，每次更新所述声纹模

型时迭代 m 次所述混合高斯模型，以得到各说话人对应的有效语音段，n 及 m 均为大于 1 的正整数。

本实施例中，利用时间递归神经网络对各说话人标识对应的短语音段建立声纹模型，基于所述声纹模型提取表征说话人身份特征的预设类型向量，优选地，该预设类型向量为 i-vector 向量，i-vector 向量是反映说话人声学差异的一个重要特征。

在整个混合语音中，根据预设类型向量计算每一语音帧属于某一说话人的最大后验概率，利用计算最大后验概率，在混合语音中通过预设算法重新调整说话人的混合高斯模型，例如，通过 Baum-Welch 算法重新调整说话人的混合高斯模型，该混合高斯模型为 k（一般为 3-5 个）个高斯模型的集合。利用重新调整后的混合高斯模型寻找每一语音帧概率最大的说话人。根据语音帧与寻找到的该说话人的概率关系调整混合语音的分割边界，例如将分割边界向前微调或者向后微调。最后，迭代更新上述声纹模型 n 次，每次更新声纹模型时迭代 m 次混合高斯模型，以得到各个说话人对应的有效语音段，n 及 m 均为大于 1 的正整数。

本实施例借助深度学习的时间递归神经网络建立声纹模型，用各说话人声纹对应的身份特征对应各语音帧以计算语音帧属于某一说话人的概率，基于该概率修正模型，最终调整语音分割的边界，可以有效提高说话人语音分割的精度，降低错误率，且可扩展性好。

在一优选的实施例中，在上述的实施例的基础上，所述语音分割的装置还包括：反馈模块，用于基于所述有效语音段获取对应的应答内容，并将所述应答内容反馈给所述终端。

本实施例中，自动应答系统关联对应的应答库，该应答库中存储有不同的问题对应的应答内容，自动应答系统在接收到终端发送的混合语音后，分割模块 101 将其分割为说话人标识对应的有效语音段，从这些有效语音段中获取与该自动应答系统有关问题的一个有效语音段，针对该有效语音段在应答库中进行匹配，反馈模块将匹配得到的应答内容反馈给终端。

如图 7 所示，图 7 为本发明实施例方案涉及的硬件运行环境的设备结构示意图。

本发明实施例语音分割设备可以是 PC（Personal Computer，个人电脑），也可以是智能手机、平板电脑、电子书阅读器、便携计算机等终端设备。

参照图 7 所示，该设备包括：存储器 11、处理器 12，通信总线 13，以及网络接口 14。其中，存储器 11 至少包括一种类型的可读存储介质，所述可读存储介质包括闪存、硬盘、多媒体卡、卡型存储器（例如，SD 或 DX 存储器等）、随机访问存储器（RAM）、静态随机访问存储器（SRAM）、只读存储器（ROM）、电可擦除可编程只读存储器（EEPROM）、可编程只读存储器（PROM）、磁性存储器、磁盘、光盘等。

存储器 11 在一些实施例中可以是语音分割设备的内部存储单元，例如该语音分割设备的硬盘或内存。存储器 11 在另一些实施例中也可以是语音分割设备的外部存储设备，例如语音分割设备上配备的插接式硬盘，智能存储卡（Smart Media Card, SMC），安全数字（Secure Digital, SD）卡，闪存卡（Flash Card）等。进一步地，存储器 11 还可以既包括语音分割设备的内部存储单元也包括外部存储设备。存储器 11 不仅可以用于存储安装于语音分割设备的应用软件及各类数据，例如语音分割程序的代码等，还可以用于暂时地存储已经输出或者将要输出的数据。

处理器 12 在一些实施例中可以是一中央处理器（Central Processing Unit, CPU）、控制器、微控制器、微处理器或其他数据处理芯片，用于运行存储器 11 中存储的程序代码或处理数据，例如执行语音分割程序等。

通信总线 13 用于实现这些组件之间的连接通信。

网络接口 14 可选的可以包括标准的有线接口、无线接口（如 WI-FI 接口），通常用于在该设备与其他电子设备之间建立通信连接。

图 7 仅示出了具有组件 11-14 以及语音分割程序的语音分割设备，但是应理解的是，并不要求实施所有示出的组件，可以替代的实施更多或者更少的组件。

可选地，该设备还可以包括用户接口，用户接口可以包括显示器（Display）、输入单元比如键盘（Keyboard），可选的用户接口还可以包括标准的有线接口、无线接口。可选地，在一些实施例中，显示器可以是 LED 显示器、液晶显示器、触控式液晶显示器以及 OLED（Organic Light-Emitting Diode，有机发光二极管）触摸器等。其中，显示器也可以适当的称为显示屏或显示单元，用于显示在语音分割设备中处理的信息以及用于显示可视化的用户界面。

在图 7 所示的装置实施例中，作为一种计算机存储介质的存储器 11 中可以包括操作系统、以及语音分割程序；其中，网络接口 14 用于与终端进行通信连接，以接收终端发送的混合语音，处理器 12 执行存储器 11 中存储的语音分割程序时实现如下步骤：

S1，在接收到终端发送的混合语音时，将所述混合语音分割成多个短语音段，并对各短语音段标注对应的说话人标识；

S2，利用时间递归神经网络对各说话人标识对应的短语音段建立声纹模型，基于所述声纹模型调整所述混合语音中对应的分割边界，以分割出各说话人标识对应的有效语音段。

进一步地，步骤 S13 包括：

对每一长语音段的声学特征进行 KL 距离分析，对时长大于预设时间阈值的长语音段在 KL 距离的最大值处进行切分，得到切分后的短语音段。

进一步地，步骤 S1 包括：

S11，获取所述混合语音中的静音段，去除所述混合语音中的静音段，以根据所述静音段对所述混合语音进行分割，得到分割后的长语音段；

S12，对所述长语音段进行分帧，以提取每一长语音段的声学特征;

S13，对每一长语音段的声学特征进行 KL 距离分析，根据 KL 距离分析结果对所述语音段进行切分，得到切分后的短语音段;

S14，利用高斯混合模型对各短语音段进行语音聚类，并对同一语音类的短语音段标注对应的说话人标识。

进一步地，步骤 S2 之后还包括:

基于所述有效语音段获取对应的应答内容，并将所述应答内容反馈给所述终端。

本发明还提供一种计算机可读存储介质，该计算机可读存储介质上存储有语音分割程序，所述语音分割程序被处理器执行时实现如下步骤:

S1，在接收到终端发送的混合语音时，将所述混合语音分割成多个短语音段，并对各短语音段标注对应的说话人标识;

S2，利用时间递归神经网络对各说话人标识对应的短语音段建立声纹模型，基于所述声纹模型调整所述混合语音中对应的分割边界，以分割出各说话人标识对应的有效语音段。

进一步地，步骤 S1 包括:

S11，获取所述混合语音中的静音段，去除所述混合语音中的静音段，以根据所述静音段对所述混合语音进行分割，得到分割后的长语音段;

S12，对所述长语音段进行分帧，以提取每一长语音段的声学特征;

S13，对每一长语音段的声学特征进行 KL 距离分析，根据 KL 距离分析结果对所述语音段进行切分，得到切分后的短语音段;

S14，利用高斯混合模型对各短语音段进行语音聚类，并对同一语音类的短语音段标注对应的说话人标识。

进一步地，步骤 S13 包括:

对每一长语音段的声学特征进行 KL 距离分析，对时长大于预设时间阈值的长语音段在 KL 距离的最大值处进行切分，得到切分后的短语音段。

进一步地，步骤 S2 包括:

S21，利用所述时间递归神经网络对各说话人标识对应的短语音段建立声纹模型，基于所述声纹模型提取表征说话人身份特征的预设类型向量;

S22，基于所述预设类型向量计算每一语音帧属于对应的说话人的最大后验概率;

S23，基于所述最大后验概率并利用预定算法调整该说话人的混合高斯模型;

S24，基于调整后的混合高斯模型获取每一语音帧对应的概率最大的说话人，并根据概率最大的说话人与语音帧的概率关系调整所述混合语音中对应的分割边界;

S25，迭代更新所述声纹模型 n 次，每次更新所述声纹模型时迭代 m 次所述混合高斯模型，以得到各说话人对应的有效语音段，n 及 m 均为大于 1

的正整数。

本发明计算机可读存储介质的具体实施例与上述语音分割的方法和装置各实施例基本相同，在此不作赘述。

通过以上的实施方式的描述，本领域的技术人员可以清楚地了解到上述实施例方法可借助软件加必需的通用硬件平台的方式来实现，当然也可以通过硬件，但很多情况下前者是更佳的实施方式。基于这样的理解，本发明的技术方案本质上或者说对现有技术做出贡献的部分可以以软件产品的形式体现出来，该计算机软件产品存储在一个存储介质（如ROM/RAM、磁碟、光盘）中，包括若干指令用以使得一台终端设备（可以是手机，计算机，服务器，或者网络设备等）执行本发明各个实施例所述的方法。

以上所述仅为本发明的较佳实施例，并不用以限制本发明，凡在本发明的精神和原则之内，所作的任何修改、等同替换、改进等，均应包含在本发明的保护范围之内。

权 利 要 求 书

1、一种语音分割的方法，其特征在于，所述语音分割的方法包括:

S1，在接收到终端发送的混合语音时，将所述混合语音分割成多个短语音段，并对各短语音段标注对应的说话人标识;

S2，利用时间递归神经网络对各说话人标识对应的短语音段建立声纹模型，基于所述声纹模型调整所述混合语音中对应的分割边界，以分割出各说话人标识对应的有效语音段。

2、根据权利要求 1 所述的语音分割的方法，其特征在于，所述步骤 S1 包括:

S11，获取所述混合语音中的静音段，去除所述混合语音中的静音段，以根据所述静音段对所述混合语音进行分割，得到分割后的长语音段;

S12，对所述长语音段进行分帧，以提取每一长语音段的声学特征;

S13，对每一长语音段的声学特征进行 KL 距离分析，根据 KL 距离分析结果对所述语音段进行切分，得到切分后的短语音段;

S14，利用高斯混合模型对各短语音段进行语音聚类，并对同一语音类的短语音段标注对应的说话人标识。

3、根据权利要求 2 所述的语音分割的方法，其特征在于，所述步骤 S13 包括:

对每一长语音段的声学特征进行 KL 距离分析，对时长大于预设时间阈值的长语音段在 KL 距离的最大值处进行切分，得到切分后的短语音段。

4、根据权利要求 2 所述的语音分割的方法，其特征在于，所述步骤 S2 包括:

S21，利用所述时间递归神经网络对各说话人标识对应的短语音段建立声纹模型，基于所述声纹模型提取表征说话人身份特征的预设类型向量;

S22，基于所述预设类型向量计算每一语音帧属于对应的说话人的最大后验概率;

S23，基于所述最大后验概率并利用预定算法调整该说话人的混合高斯模型;

S24，基于调整后的混合高斯模型获取每一语音帧对应的概率最大的说话人，并根据概率最大的说话人与语音帧的概率关系调整所述混合语音中对应的分割边界;

S25，迭代更新所述声纹模型 n 次，每次更新所述声纹模型时迭代 m 次所述混合高斯模型，以得到各说话人对应的有效语音段，n 及 m 均为大于 1 的正整数。

5、根据权利要求 3 所述的语音分割的方法，其特征在于，所述步骤 S2 包括:

S21，利用所述时间递归神经网络对各说话人标识对应的短语音段建立声纹模型，基于所述声纹模型提取表征说话人身份特征的预设类型向量;

S22，基于所述预设类型向量计算每一语音帧属于对应的说话人的最大后验概率;

S23，基于所述最大后验概率并利用预定算法调整该说话人的混合高斯模型;

S24，基于调整后的混合高斯模型获取每一语音帧对应的概率最大的说话人，并根据概率最大的说话人与语音帧的概率关系调整所述混合语音中对应的分割边界;

S25，迭代更新所述声纹模型 n 次，每次更新所述声纹模型时迭代 m 次所述混合高斯模型，以得到各说话人对应的有效语音段，n 及 m 均为大于 1 的正整数。

6、根据权利要求 1 所述的语音分割的方法，其特征在于，所述步骤 S2 之后还包括:

基于所述有效语音段获取对应的应答内容，并将所述应答内容反馈给所述终端。

7、一种语音分割的装置，其特征在于，所述语音分割的装置包括:

分割模块，用于在接收到终端发送的混合语音时，将所述混合语音分割成多个短语音段，并对各短语音段标注对应的说话人标识;

调整模块，用于利用时间递归神经网络对各说话人标识对应的短语音段建立声纹模型，基于所述声纹模型调整所述混合语音中对应的分割边界，以分割出各说话人标识对应的有效语音段。

8、根据权利要求 7 所述的语音分割的装置，其特征在于，所述分割模块包括:

去除单元，用于获取所述混合语音中的静音段，去除所述混合语音中的静音段，以根据所述静音段对所述混合语音进行分割，得到分割后的长语音段;

分帧单元，用于对所述长语音段进行分帧，以提取每一长语音段的声学特征;

切分单元，用于对每一长语音段的声学特征进行 KL 距离分析，根据 KL 距离分析结果对所述语音段进行切分，得到切分后的短语音段;

聚类单元，用于利用高斯混合模型对各短语音段进行语音聚类，并对同一语音类的短语音段标注对应的说话人标识。

9、根据权利要求 8 所述的语音分割的装置，其特征在于，所述切分单元具体用于对每一长语音段的声学特征进行 KL 距离分析，对时长大于预设时间阈值的长语音段在 KL 距离的最大值处进行切分，得到切分后的短语音段。

10、根据权利要求 8 所述的语音分割的装置，其特征在于，所述调整模块包括:

建模单元，用于利用所述时间递归神经网络对各说话人标识对应的短语音段建立声纹模型，基于所述声纹模型提取表征说话人身份特征的预设类型向量;

计算单元，用于基于所述预设类型向量计算每一语音帧属于对应的说话人的最大后验概率;

第一调整单元，用于基于所述最大后验概率并利用预定算法调整该说话人的混合高斯模型;

第二调整单元，用于基于调整后的混合高斯模型获取每一语音帧对应的概率最大的说话人，并根据概率最大的说话人与语音帧的概率关系调整所述混合语音中对应的分割边界;

迭代单元，用于迭代更新所述声纹模型 n 次，每次更新所述声纹模型时迭代 m 次所述混合高斯模型，以得到各说话人对应的有效语音段，n 及 m 均为大于 1 的正整数。

11、根据权利要求 9 所述的语音分割的装置，其特征在于，所述调整模块包括:

建模单元，用于利用所述时间递归神经网络对各说话人标识对应的短语音段建立声纹模型，基于所述声纹模型提取表征说话人身份特征的预设类型向量;

计算单元，用于基于所述预设类型向量计算每一语音帧属于对应的说话人的最大后验概率;

第一调整单元，用于基于所述最大后验概率并利用预定算法调整该说话人的混合高斯模型;

第二调整单元，用于基于调整后的混合高斯模型获取每一语音帧对应的概率最大的说话人，并根据概率最大的说话人与语音帧的概率关系调整所述混合语音中对应的分割边界;

迭代单元，用于迭代更新所述声纹模型 n 次，每次更新所述声纹模型时迭代 m 次所述混合高斯模型，以得到各说话人对应的有效语音段，n 及 m 均为大于 1 的正整数。

12、根据权利要求 7 所述的语音分割的装置，其特征在于，所述语音分割的装置还包括: 反馈模块，用于基于所述有效语音段获取对应的应答内容，

并将所述应答内容反馈给所述终端。

13、一种语音分割设备，其特征在于，所述语音分割设备包括存储器、处理器以及存储在所述存储器上并可在所述处理器上运行的语音分割程序，所述语音分割程序被所述所述处理器执行时实现如下步骤：

S1，在接收到终端发送的混合语音时，将所述混合语音分割成多个短语音段，并对各短语音段标注对应的说话人标识；

S2，利用时间递归神经网络对各说话人标识对应的短语音段建立声纹模型，基于所述声纹模型调整所述混合语音中对应的分割边界，以分割出各说话人标识对应的有效语音段。

14、根据权利要求 13 所述的语音分割设备，其特征在于，所述步骤 S13 包括：

对每一长语音段的声学特征进行 KL 距离分析，对时长大于预设时间阈值的长语音段在 KL 距离的最大值处进行切分，得到切分后的短语音段。

15、根据权利要求 14 所述的语音分割设备，其特征在于，所述步骤 S1 包括：

S11，获取所述混合语音中的静音段，去除所述混合语音中的静音段，以根据所述静音段对所述混合语音进行分割，得到分割后的长语音段；

S12，对所述长语音段进行分帧，以提取每一长语音段的声学特征；

S13，对每一长语音段的声学特征进行 KL 距离分析，根据 KL 距离分析结果对所述语音段进行切分，得到切分后的短语音段；

S14，利用高斯混合模型对各短语音段进行语音聚类，并对同一语音类的短语音段标注对应的说话人标识。

16、根据权利要求 14 所述的语音分割设备，其特征在于，所述步骤 S2 之后还包括：

基于所述有效语音段获取对应的应答内容，并将所述应答内容反馈给所述终端。

17、一种计算机可读存储介质，其特征在于，所述计算机可读存储介质上存储有语音分割程序，所述语音分割程序被处理器执行时实现如下步骤：

S1，在接收到终端发送的混合语音时，将所述混合语音分割成多个短语音段，并对各短语音段标注对应的说话人标识；

S2，利用时间递归神经网络对各说话人标识对应的短语音段建立声纹模型，基于所述声纹模型调整所述混合语音中对应的分割边界，以分割出各说话人标识对应的有效语音段。

18、根据权利要求 17 所述的计算机可读存储介质，其特征在于，所述步骤 S1 包括:

S11，获取所述混合语音中的静音段，去除所述混合语音中的静音段，以根据所述静音段对所述混合语音进行分割，得到分割后的长语音段;

S12，对所述长语音段进行分帧，以提取每一长语音段的声学特征;

S13，对每一长语音段的声学特征进行 KL 距离分析，根据 KL 距离分析结果对所述语音段进行切分，得到切分后的短语音段;

S14，利用高斯混合模型对各短语音段进行语音聚类，并对同一语音类的短语音段标注对应的说话人标识。

19、根据权利要求 18 所述的计算机可读存储介质，其特征在于，所述步骤 S13 包括:

对每一长语音段的声学特征进行 KL 距离分析，对时长大于预设时间阈值的长语音段在 KL 距离的最大值处进行切分，得到切分后的短语音段。

20、根据权利要求 18 所述的计算机可读存储介质，其特征在于，所述步骤 S2 包括:

S21，利用所述时间递归神经网络对各说话人标识对应的短语音段建立声纹模型，基于所述声纹模型提取表征说话人身份特征的预设类型向量;

S22，基于所述预设类型向量计算每一语音帧属于对应的说话人的最大后验概率;

S23，基于所述最大后验概率并利用预定算法调整该说话人的混合高斯模型;

S24，基于调整后的混合高斯模型获取每一语音帧对应的概率最大的说话人，并根据概率最大的说话人与语音帧的概率关系调整所述混合语音中对应的分割边界;

S25，迭代更新所述声纹模型 n 次，每次更新所述声纹模型时迭代 m 次所述混合高斯模型，以得到各说话人对应的有效语音段，n 及 m 均为大于 1 的正整数。

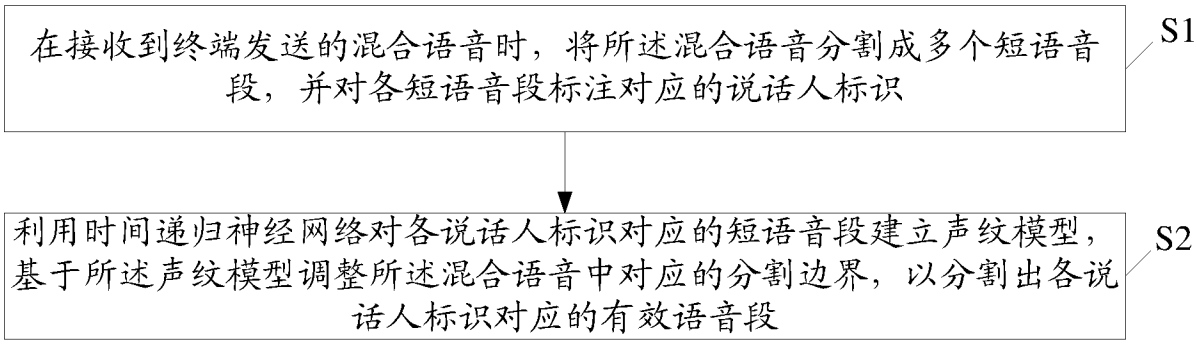

在接收到终端发送的混合语音时，将所述混合语音分割成多个短语音段，并对各短语音段标注对应的说话人标识
S1
利用时间递归神经网络对各说话人标识对应的短语音段建立声纹模型，基于所述声纹模型调整所述混合语音中对应的分割边界，以分割出各说话人标识对应的有效语音段
S2

图 1

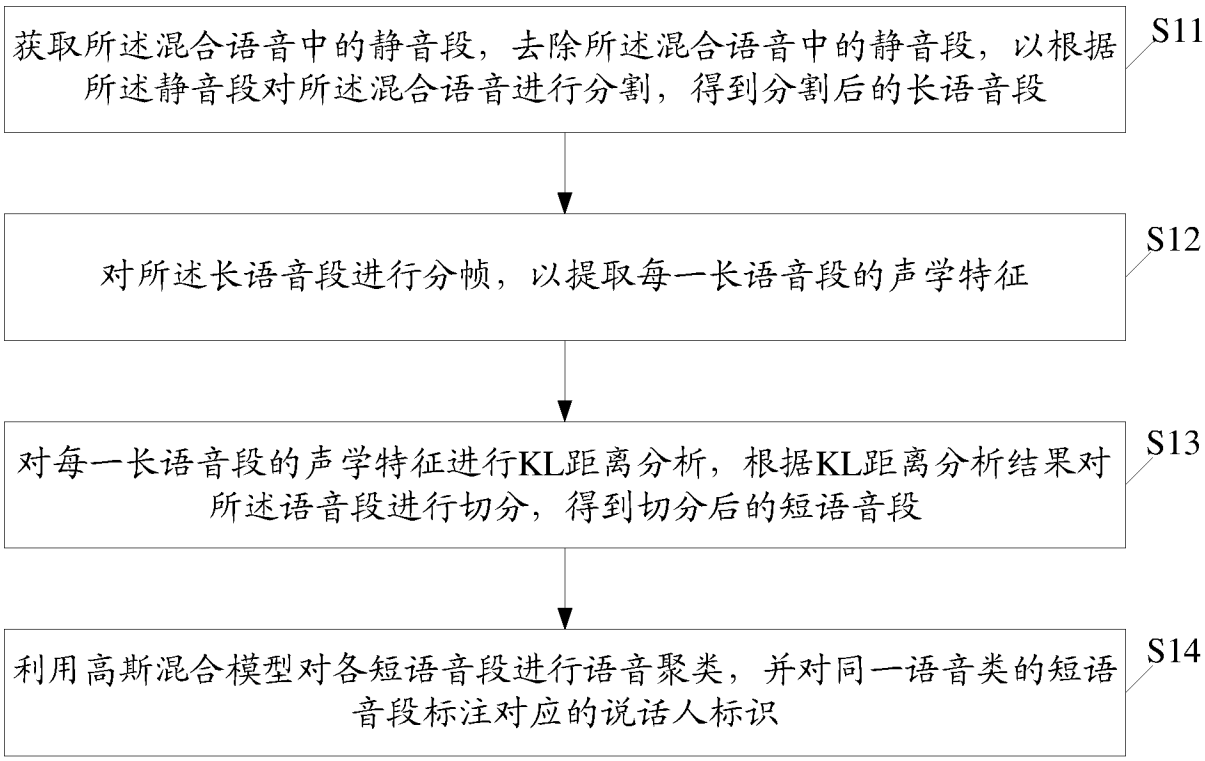

获取所述混合语音中的静音段，去除所述混合语音中的静音段，以根据所述静音段对所述混合语音进行分割，得到分割后的长语音段
S11
对所述长语音段进行分帧，以提取每一长语音段的声学特征
S12
对每一长语音段的声学特征进行KL距离分析，根据KL距离分析结果对所述语音段进行切分，得到切分后的短语音段
S13
利用高斯混合模型对各短语音段进行语音聚类，并对同一语音类的短语音段标注对应的说话人标识
S14

图 2

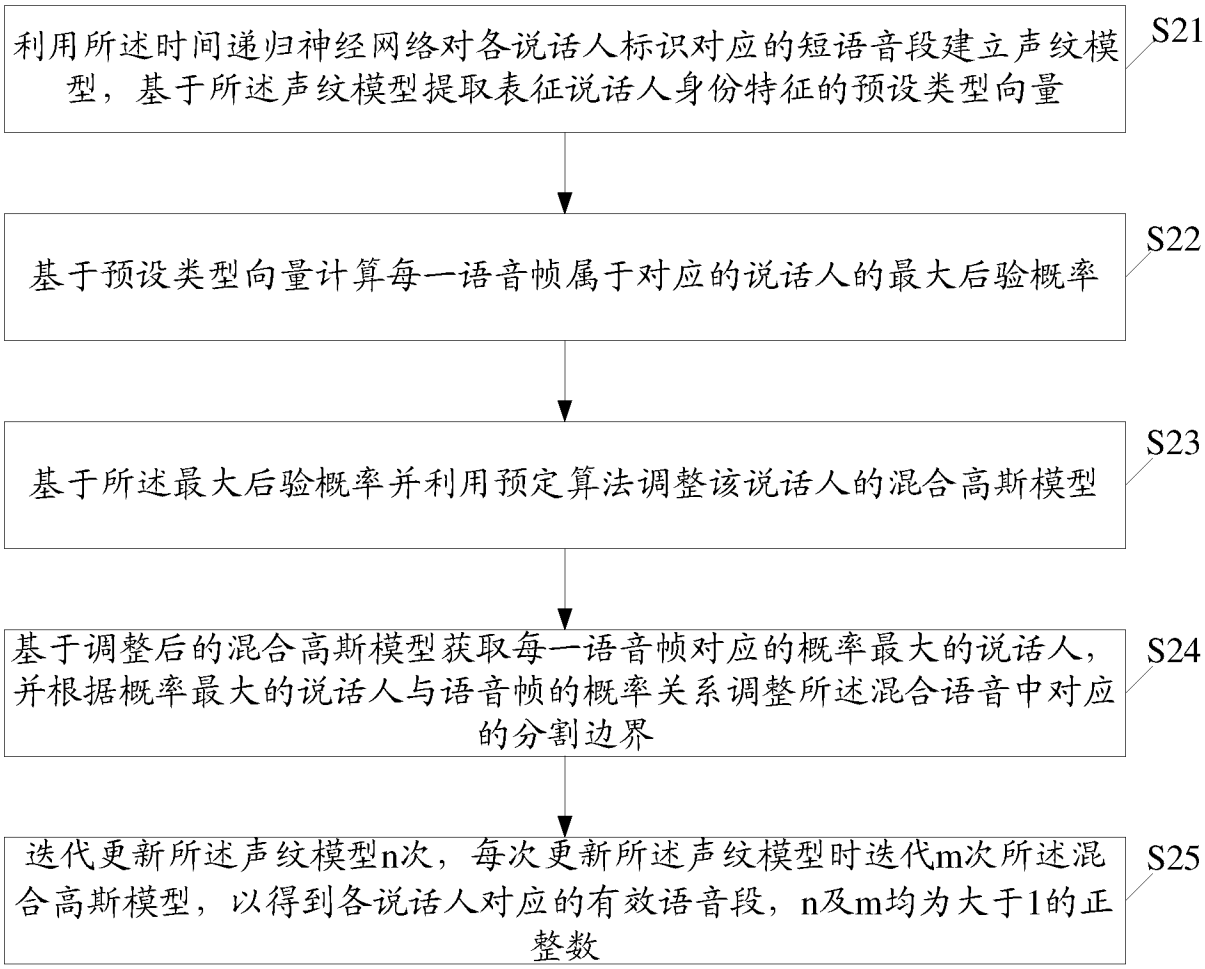

利用所述时间递归神经网络对各说话人标识对应的短语音段建立声纹模型，基于所述声纹模型提取表征说话人身份特征的预设类型向量
S21
基于预设类型向量计算每一语音帧属于对应的说话人的最大后验概率
S22
基于所述最大后验概率并利用预定算法调整该说话人的混合高斯模型
S23
基于调整后的混合高斯模型获取每一语音帧对应的概率最大的说话人，并根据概率最大的说话人与语音帧的概率关系调整所述混合语音中对应的分割边界
S24
迭代更新所述声纹模型n次，每次更新所述声纹模型时迭代m次所述混合高斯模型，以得到各说话人对应的有效语音段，n及m均为大于1的正整数
S25

图 3

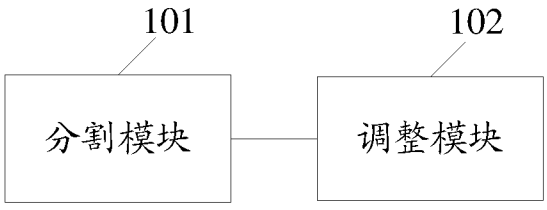

101
102
分割模块
调整模块

图 4

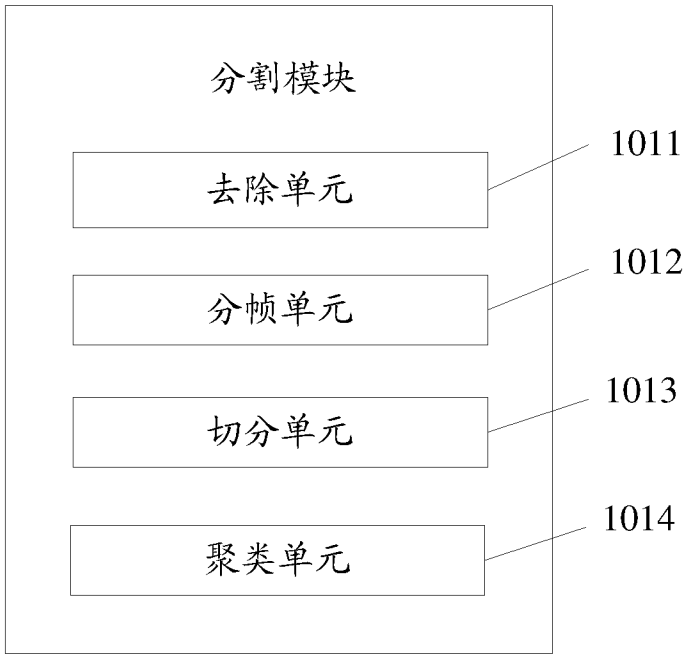

分割模块
去除单元
1011
分帧单元
1012
切分单元
1013
聚类单元
1014

图 5

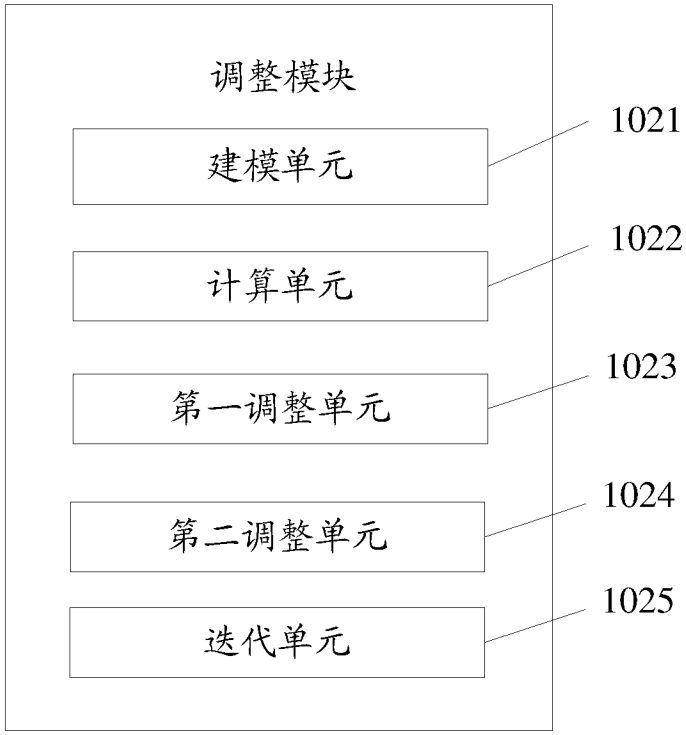

调整模块
建模单元
1021
计算单元
1022
第一调整单元
1023
第二调整单元
1024
迭代单元
1025

图 6

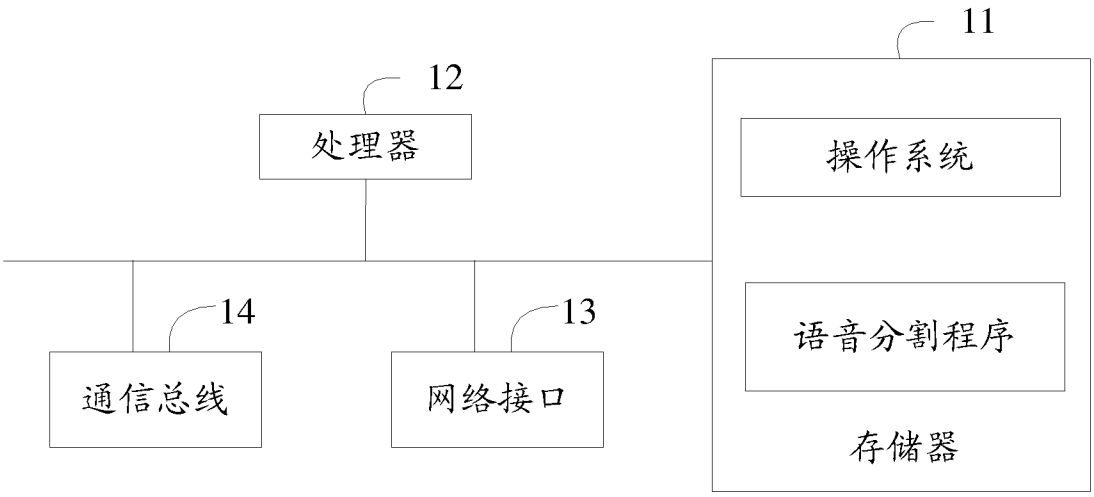


图 7

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2017/091310

A. CLASSIFICATION OF SUBJECT MATTER

G10L 15/04 (2013.01) i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G10L 15/-

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNABS, CNTXT, CNKI, WPI, EPODOC: 平安科技, 语音, 说话人, 分割, 模型, 边界, speech, speaker?, segment+, model+, boundary

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 105161093 A (IFLYTEK CO., LTD.), 16 December 2015 (16.12.2015), description, paragraphs [0078]-[0097], and figure 2	1-20
PX	CN 106782507 A (PING AN TECHNOLOGY (SHENZHEN) CO., LTD.), 31 May 2017 (31.05.2017), claims 1-10	1-20
A	CN 102760434 A (HUAWEI DEVICE CO., LTD.), 31 October 2012 (31.10.2012), entire document	1-20
A	CN 105913849 A (INSTITUTE OF THE CHINESE PEOPLE'S LIBERATION ARMY GENERAL STAFF ARMY AVIATION), 31 August 2016 (31.08.2016), entire document	1-20
A	CN 1716380 A (ZHEJIANG UNIVERSITY), 04 January 2006 (04.01.2006), entire document	1-20
A	CN 102543063 A (SOUTH CHINA UNIVERSITY OF TECHNOLOGY), 04 July 2012 (04.07.2012), entire document	1-20
A	US 7930179 B1 (AT&T INTELLECTUAL PROPERTY II, L. P.), 19 April 2011 (19.04.2011), entire document	1-20

☐ Further documents are listed in the continuation of Box C. ☒ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
06 September 2017	28 September 2017
Name and mailing address of the ISA State Intellectual Property Office of the P. R. China No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088, China Facsimile No. (86-10) 62019451	Authorized officer WANG, Xinning Telephone No. (86-10) 61648159

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2017/091310

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 105161093 A	16 December 2015	None	
CN 106782507 A	31 May 2017	None	
CN 102760434 A	31 October 2012	JP 5996783 B2	21 September 2016
		JP 2015516091 A	04 June 2015
		WO 2014008843 A1	16 January 2014
		US 9685161 B2	20 June 2017
		AU 2013289660 A1	30 October 2014
		KR 20140144233 A	18 December 2014
		US 2015112680 A1	23 April 2015
		EP 2821992 A1	07 January 2015
		AU 2013289660 B2	30 June 2016
CN 105913849 A	31 August 2016	None	
CN 1716380 A	04 January 2006	CN 100505040 C	24 June 2009
CN 102543063 A	04 July 2012	CN 102543063 B	24 July 2013
US 7930179 B1	19 April 2011	US 7295970 B1	13 November 2007

国际检索报告

国际申请号 PCT/CN2017/091310

A. 主题的分类

G10L 15/04(2013.01)i

按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

G10L 15/-

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称，和使用的检索词（如使用））

CNABS，CNTXT，CNKI，WPI，EPODOC：平安科技，语音，说话人，分割，模型，边界，speech，speaker?，segment+，model+，boundary

C. 相关文件

类 型*	引用文件，必要时，指明相关段落	相关的权利要求
X	CN 105161093 A (科大讯飞股份有限公司) 2015年 12月 16日 (2015 - 12 - 16) 说明书[0078]-[0097]段、附图2	1-20
PX	CN 106782507 A (平安科技深圳有限公司) 2017年 5月 31日 (2017 - 05 - 31) 权利要求1-10	1-20
A	CN 102760434 A (华为终端有限公司) 2012年 10月 31日 (2012 - 10 - 31) 全文	1-20
A	CN 105913849 A (中国人民解放军总参谋部陆航研究所) 2016年 8月 31日 (2016 - 08 - 31) 全文	1-20
A	CN 1716380 A (浙江大学) 2006年 1月 4日 (2006 - 01 - 04) 全文	1-20
A	CN 102543063 A (华南理工大学) 2012年 7月 4日 (2012 - 07 - 04) 全文	1-20
A	US 7930179 B1 (AT&T INTELLECTUAL PROPERTY II， L.P.) 2011年 4月 19日 (2011 - 04 - 19) 全文	1-20

☐ 其余文件在C栏的续页中列出。

☑ 见同族专利附件。

* 引用文件的具体类型：

“A” 认为不特别相关的表示了现有技术一般状态的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“L” 可能对优先权要求构成怀疑的文件，或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件（如具体说明的）

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布，与申请不相抵触，但为了理解发明之理论或原理的在后文件

“X” 特别相关的文件，单独考虑该文件，认定要求保护的发明不是新颖的或不具有创造性

“Y” 特别相关的文件，当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时，要求保护的发明不具有创造性

“&” 同族专利的文件

国际检索实际完成的日期	国际检索报告邮寄日期
2017年 9月 6日	2017年 9月 28日

ISA/CN的名称和邮寄地址	受权官员
中华人民共和国国家知识产权局(ISA/CN) 中国北京市海淀区蓟门桥西土城路6号 100088	王馨宁
传真号 (86-10)62019451	电话号码 (86-10)61648159

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2017/091310

检索报告引用的专利文件			公布日（年/月/日）	同族专利			公布日（年/月/日）
CN	105161093	A	2015年 12月 16日	无			
CN	106782507	A	2017年 5月 31日	无			
CN	102760434	A	2012年 10月 31日	JP	5996783	B2	2016年 9月 21日
				JP	2015516091	A	2015年 6月 4日
				WO	2014008843	A1	2014年 1月 16日
				US	9685161	B2	2017年 6月 20日
				AU	2013289660	A1	2014年 10月 30日
				KR	20140144233	A	2014年 12月 18日
				US	2015112680	A1	2015年 4月 23日
				EP	2821992	A1	2015年 1月 7日
				AU	2013289660	B2	2016年 6月 30日
CN	105913849	A	2016年 8月 31日	无			
CN	1716380	A	2006年 1月 4日	CN	100505040	C	2009年 6月 24日
CN	102543063	A	2012年 7月 4日	CN	102543063	B	2013年 7月 24日
US	7930179	B1	2011年 4月 19日	US	7295970	B1	2007年 11月 13日