

(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
2017 年 5 月 26 日 (26.05.2017)



(10) 国际公布号
WO 2017/084545 A1

- (51) 国际专利分类号:
H04M 11/06 (2006.01) **G10L 19/005** (2013.01)
- (21) 国际申请号: PCT/CN2016/105684
- (22) 国际申请日: 2016 年 11 月 14 日 (14.11.2016)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201510802586.8 2015 年 11 月 19 日 (19.11.2015) CN
- (71) 申请人: 电信科学技术研究院 (CHINA ACADEMY OF TELECOMMUNICATIONS TECHNOLOGY) [CN/CN]; 中国北京市海淀区学院路 40 号, Beijing 100191 (CN)。
- (72) 发明人: 邹莹 (ZOU, Ying); 中国北京市海淀区学院路 40 号, Beijing 100191 (CN)。
- (74) 代理人: 北京银龙知识产权代理有限公司 (DRAGON INTELLECTUAL PROPERTY LAW FIRM); 中国北京市海淀区西直门北大街 32 号院枫蓝国际中心 2 号楼 10 层, Beijing 100082 (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, JP, KE, KG, KN, KP, KR, KW, KZ, LA, LC, LK, 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))。

(54) Title: METHOD AND SYSTEM FOR VOICE PACKET LOSS CONCEALMENT

(54) 发明名称: 一种语音丢包补偿的方法及系统

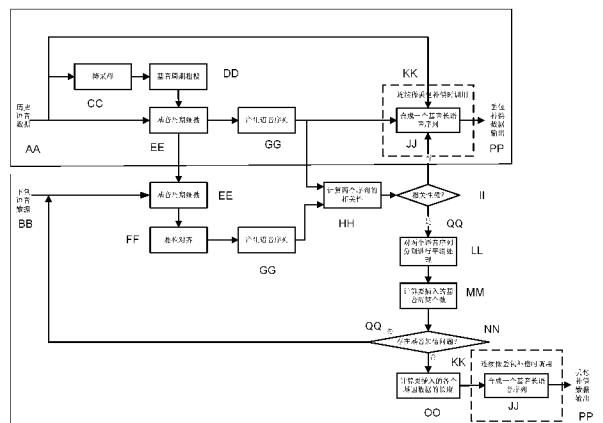


图 3

AA HISTORICAL VOICE DATA
BB VOICE DATA OF THE NEXT PACKET
CC DOWN-SAMPLING
DD ROUGH DETECTION OF THE PITCH PERIOD
EE FINE DETECTION OF THE PITCH PERIOD
FF PHASE ALIGNMENT
GG GENERATE A VOICE SEQUENCE
HH CALCULATE THE CORRELATION BETWEEN THE TWO SEQUENCES
II CORRELATION STRONG?
JJ SYNthesize TO OBTAIN A LONG VOICE SEQUENCE OF PITCHES

KK CALL THE LONG VOICE SEQUENCE WHEN CONTINUOUSLY PERFORMING PACKET LOSS CONCEALMENT
LL SMOOTH THE TWO VOICE SEQUENCES RESPECTIVELY
MM CALCULATE THE NUMBER OF PITCH PERIOD(S) TO BE INSERTED
NN PITCH DOUBLING PROBLEM EXISTS?
OO CALCULATE THE LENGTH OF EACH PITCH DATA TO BE INSERTED
PP OUTPUT THE PACKET LOSS CONCEALMENT DATA
QQ YES
RR NO

(57) Abstract: Provided are a method and a system for a voice packet loss concealment (PLC), said method comprising: when it is found for the first time that a voice packet is lost, generating, according to the historical voice data before the loss of the voice packet, a first voice sequence vec1; and if the voice data of the next packet after the loss of the voice packet can be acquired, generating, according to the voice data of the next packet after the loss of the voice packet, a second voice sequence vec2, and performing PLC according to the first voice sequence vec1 and the second voice sequence vec2.

(57) 摘要: 本公开提供了一种语音丢包补偿的方法及系统, 该方法包括: 当第一次发现丢失语音包时, 根据丢失语音包之前的历史语音数据, 生成第一语音序列 vec1; 若能够获取丢失语音包之后的下一包语音数据, 则根据丢失语音包之后的下一包语音数据, 生成第二语音序列 vec2, 根据第一语音序列 vec1 和所述第二语音序列 vec2 进行丢包补偿。

WO 2017/084545 A1

一种语音丢包补偿的方法及系统

相关申请的交叉引用

5 本申请主张在 2015 年 11 月 19 日在中国提交的中国专利申请号 No. 201510802586.8 的优先权，其全部内容通过引用包含于此。

技术领域

本公开涉及通信技术领域，尤其涉及一种语音丢包补偿的方法及系统。

背景技术

10 Voice-over-IP(VoIP, 一种由 IP 网络传送语音的技术服务)使用 IP 网络对打包后的语音数据进行传输。由于 IP 网络分布广泛, VoIP 电话费用低廉, 可实现多媒体通信、语音与数据的整合, 因此近些年来 VoIP 发展迅猛, 正有逐步替代传统电路交换网络 (如 Public Switched Telephone Network, PSTN-公共服务电话网络) 的趋势。鉴于目前 IP 网络主要用于传输数据业务, 其采用一种
15 尽力而为的无连接传输技术, 由此没有相应的服务质量保证。当网络出现拥塞时, 会导致在接收端有包丢失和延迟抖动的问题出现, 这将严重影响接收端的语音质量。因此在 VoIP 系统的接收端, 一般采用丢包补偿技术来重建丢失的语音包, 应用抗网络延迟抖动技术消除延迟抖动带来的不良影响。

 PLC (Packet Loss Concealment, 丢包补偿) 算法的目的是为了产生能够
20 替代丢失的语音的合成语音, 理想的合成语音应该具有与丢失语音相同的音色和频谱特性。因此最为常用的方法就是利用丢失语音数据前的历史数据产生合理的近似语音。如果丢失的数据长度不是太长, 并且丢失语音段不位于语音快速变化的区域, 经过 PLC 算法补偿后的语音听起来还是比较自然的。现有的一些标准中提出的基于接收端的 PLC 技术有 ANSI T1.521 (Annex B) ,
25 ITU-T Rec. G.711 (Appendix I), ITU-T Rec. G.722 (Appendix III) 等等。

 上面提及的标准中几种 PLC 算法的共性就是基音检测和合成语音两个部分 (如图 1 所示), 其中基音周期检测只是在发现丢包后的第一次丢包补偿时进行, 如果连续做丢包补偿时则用第一次检测出的基音周期。由于检测基音周期的运算复杂度较高, 因此一般都会分两步实现: 首先对降采样后的历史
30 数据 (原始信号或残差信号) 进行基音周期粗估, 再在粗估出的基音周期附

近细搜出更加精确的基音周期。之后按此长度合成出与原始语音近似的一帧语音，并且要在真实语音数据与合成的语音数据之间或合成语音与合成语音之间要做平滑过渡以免产生尖锐的噪声，同时在连续丢包的时候要通过能量递减的方式来减少后帧语音与前帧语音的相关性。

- 5 现有的 PLC 方法均仅利用丢失包前的历史数据来补偿丢失语音包数据，这样不能达到补偿效果的最优，尤其是在基音到基音的转折处，补偿的效果不太理想。

发明内容

10 鉴于上述技术问题，本公开提供一种语音丢包补偿的方法及系统，解决相关技术中语音丢包补偿的补偿效果不理想的问题。

依据本公开的一个方面，提供了一种语音丢包补偿的方法，所述方法包括：第一次发现丢失语音包时（即第一次做丢包补偿时），根据丢失语音包之前的历史语音数据，生成第一语音序列 $vec1$ ；若能够获取丢失语音包之后的下一包语音数据，则根据丢失语音包之后的下一包语音数据，生成第二语音序列 $vec2$ ；根据所述第一语音序列 $vec1$ 和所述第二语音序列 $vec2$ 进行丢包补偿。

可选地，若不能获取所述下一包语音数据，所述方法还包括：根据所述第一语音序列 $vec1$ 进行丢包补偿。

20 可选地，所述根据丢失语音包之前的历史语音数据，生成第一语音序列，具体包括：对所述历史语音数据进行基音粗估，得到基音粗估结果；根据所述基音粗估结果，进行基音细搜索，确定出第一基音周期 $Lag1$ ；根据所述第一基音周期 $Lag1$ 从所述历史语音数据生成所述第一语音序列 $vec1$ 。

25 可选地，所述根据丢失语音包之后的下一包语音数据，生成第二语音序列 $vec2$ ，具体为：从丢失语音包之后的下一包语音数据中找出一个与所述第一语音序列 $vec1$ 相位对齐的第二语音序列 $vec2$ 。

可选地，所述从丢失语音包之后的下一包语音数据中找出一个与所述第一语音序列 $vec1$ 相位对齐的第二语音序列 $vec2$ ，具体包括：根据所述第一基音周期 $Lag1$ ，在所述下一包语音数据中进行基音细搜索，确定出下一包语音数据的第二基音周期 $Lag2$ ；在所述下一包语音数据中找出与所述第一语音序

列 vec1 相关性最强的位置；

根据所述相关性最强的位置和第二基音周期 Lag2 ，从所述下一包语音数据中确定出所述第二语音序列 vec2 。

5 可选地，在根据所述第一语音序列 vec1 和所述第二语音序列 vec2 进行丢包补偿之前，所述方法还包括：计算出所述第一语音序列 vec1 和所述第二语音序列 vec2 的归一化互相关值；比较所述归一化互相关值与设定的门限值；若所述归一化互相关值大于所述设定的门限值，则进入根据所述第一语音序列 vec1 和所述第二语音序列 vec2 进行丢包补偿的步骤；若所述归一化互相关值小于或等于所述设定的门限值，则进入根据所述第一语音序列 vec1 进行丢包补偿的步骤。

10 可选地，所述根据所述第一语音序列 vec1 和所述第二语音序列 vec2 进行丢包补偿，具体包括：根据所述第一语音序列 vec1 和所述第二语音序列 vec2 ，计算要填充的基音数据的个数；根据所述填充的基音数据的个数，计算要填充的各个基音数据的长度；对所述第一语音序列 vec1 和所述第二语音序列 vec2 分别进行平滑处理；根据所述各个基音数据的长度，对平滑后的第一语音序列 vec1 和第二语音序列 vec2 合成一个基音周期长度的语音序列。

20 可选地，所述第一语音序列 vec1 和所述第二语音序列 vec2 ，计算要填充的基音数据的个数，具体包括：根据丢失语音包的长度 d 、所述第一语音序列 vec1 和所述第二语音序列 vec2 相位对齐后确定的长度 rd 、所述第一基音周期 Lag1 和第二基音周期 Lag2 ，采用如下公式计算要填充的基音数据的个数 numPitch ：
$$\text{numPitch} = (rd + (\text{avgLag}/2)) / \text{avgLag}$$
，其中 avgLag 表示平均基音周期长度，
$$\text{avgLag} = (\text{Lag1} + \text{Lag2}) / 2$$
。

25 可选地，所述根据所述第一语音序列 vec1 和所述第二语音序列 vec2 进行丢包补偿，还包括：计算用 numPitch 个基音长度为 avgLag 的数据进行填充的长度与实际要填充长度的差值与平均基音周期长度 avgLag 的比值；根据所述比值判断是否存在基音加倍；若是，则根据校正的基音长度重新生成第二语音序列 vec2 ；若否，则根据所述填充的基音数据的个数，计算要填充的各个基音数据的长度。

可选地，所述根据所述各个基音数据的长度，利用平滑后的第一语音序

列 vec1 和第二语音序列 vec2 合成一个基音周期长度的语音序列,具体包括:
根据待合成语音的位置分别与所述第一语音序列 vec1 和所述第二语音序列 vec2 的距离关系,确定合成时所述第一语音序列 vec1 的权重 $w1$ 和所述第二语音序列 vec2 的权重 $w2$; 根据所述第一语音序列 vec1 和其权重 $w1$, 以及
5 所述第二语音序列 vec2 和其权重 $w2$, 利用平滑后的语音序列 vec1 和 vec2 合成得到一个基音周期长度的语音序列,若填充此合成的一个基音周期长度语音序列后,丢失的数据包还未到达,则只需要再次利用平滑后的第一语音序列 vec1 和第二语音序列 vec2 依次合成得其它的语音序列,每次合成一个基音周期长的数据后都会去检查丢失的数据包是否到达,若是则进行平滑衔接,
10 若否则继续用上述方法补偿。

依据本公开的另一个方面,还提供了一种语音丢包补偿的系统,所述系统包括:第一语音序列模块,用于当第一次发现丢失语音包时,根据丢失语音包之前的历史语音数据,生成第一语音序列 vec1 ; 第二语音序列模块,用于若能够获取丢失语音包之后的下一包语音数据,则根据丢失语音包之后的
15 下一包语音数据,生成第二语音序列 vec2 ; 补偿模块,用于根据所述第一语音序列 vec1 和所述第二语音序列 vec2 进行丢包补偿。

可选地,所述补偿模块,还用于若不能获取所述下一包语音数据,根据所述第一语音序列 vec1 进行丢包补偿。

可选地,所述第一语音序列模块具体用于对所述历史语音数据进行基音粗估,得到基音粗估结果; 根据所述基音粗估结果,进行基音细搜索,确定
20 出第一基音周期 Lag1 ; 根据所述第一基音周期 Lag1 从所述历史语音数据生成所述第一语音序列 vec1 。

可选地,所述第二语音序列模块具体用于:从丢失语音包之后的下一包语音数据中找出一个与所述第一语音序列 vec1 相位对齐的第二语音序列 vec2 。

25 可选地,所述第二语音序列模块具体包括:

第一单元,用于根据所述第一基音周期 Lag1 ,在所述下一个语音数据中进行基音细搜索,确定出所述下一个语音数据的第二基音周期 Lag2 ;

第二单元,用于在所述下一包语音数据中找出与所述第一语音序列 vec1 相关性最强的位置;

第三单元，用于根据所述相关性最强的位置和第二基音周期 Lag2，从所述下一包语音数据中确定出所述第二语音序列 vec2。

可选地，所述第二语音序列模块还包括：

5 第四单元，用于计算出所述第一语音序列 vec1 和所述第二语音序列 vec2 的归一化互相关值；

第五单元，用于比较所述归一化互相关值与设定的门限值；若所述归一化互相关值大于所述设定的门限值，则触发所述补偿模块根据所述第一语音序列 vec1 和所述第二语音序列 vec2 进行丢包补偿；若所述归一化互相关值小于或等于所述设定的门限值，触发所述补偿模块根据所述第一语音序列
10 vec1 进行丢包补偿。

可选地，所述补偿模块具体包括：

第六单元，用于根据所述第一语音序列 vec1 和所述第二语音序列 vec2，计算要填充的基音数据的个数；

15 第七单元，用于根据所述填充的基音数据的个数，计算要填充的各个基音数据的长度；

第八单元，用于对所述第一语音序列 vec1 和所述第二语音序列 vec2 分别进行平滑处理；

第八单元，用于根据所述各个基音数据的长度，利用平滑后的第一语音序列 vec1 和第二语音序列 vec2 合成一个基音周期长度的语音序列。

20 可选地，所述第六单元具体用于：根据丢失语音包的长度 d、所述第一语音序列 vec1 和所述第二语音序列 vec2 相位对齐后确定的长度 rd、所述第一基音周期 Lag1 和第二基音周期 Lag2，采用如下公式计算要填充的基音数据的个数 numPitch； $\text{numPitch} = (\text{rd} + (\text{avgLag}/2))/\text{avgLag}$ ，其中 avgLag 表示平均基音周期长度， $\text{avgLag} = (\text{Lag1} + \text{Lag2})/2$ 。

25 可选地，所述补偿模块还包括：

第十单元，用于计算用 numPitch 个基音周期长度为 avgLag 的数据进行填充后的长度与实际要填充长度的差值与平均基音周期长度 avgLag 的比值；根据所述比值判断是否存在基音加倍；若是，触发第二语音序列模块根据校正过的基音重新生成第二语音序列 vec2；若否，则触发所述第七单元根据所

述填充的基音数据的个数，计算要填充的各个基音数据的长度。

可选地，所述第九单元进一步用于：根据待合成语音的位置分别与所述第一语音序列 $vec1$ 和所述第二语音序列 $vec2$ 的距离关系，确定合成时所述第一语音序列 $vec1$ 的权重 $w1$ 和所述第二语音序列 $vec2$ 的权重 $w2$ ；根据所述
5 所述第一语音序列 $vec1$ 和其权重 $w1$ ，以及所述第二语音序列 $vec2$ 和其权重 $w2$ ，对平滑后的第一语音序列 $vec1$ 和第二语音序列 $vec2$ 合成得到一个基音周期长度的语音序列，其中若填充此合成的一个基音周期长度语音序列后，丢失的语音数据包还未到达，则只需要再次利用平滑后的第一语音序列 $vec1$ 和第二语音序列 $vec2$ 依次合成得其它的语音序列，每次合成一个基音周期长的
10 的数据后都会去检查丢失的数据包是否到达，若是则进行平滑衔接，若否则继续用上述方法补偿。

依据本公开的另一个方面，还提供了一种语音丢包补偿的系统，包括：存储器，以及处理器，用于读取存储器中的程序，执行下列过程：当第一次发现丢失语音包时，根据丢失语音包之前的历史语音数据，生成第一语音序
15 列 $vec1$ ；若能够获取丢失语音包之后的下一包语音数据，则根据丢失语音包之后的下一包语音数据，生成第二语音序列 $vec2$ ；根据所述第一语音序列 $vec1$ 和所述第二语音序列 $vec2$ 进行丢包补偿。

本公开的有益效果是：在本公开的实施例中在发现当前数据包丢失情况下，如果不能获取丢失包的下一个语音包时，则利用该丢失包前的历史数据进行丢包补偿；如果能获取丢失包的下一个语音包时，则利用丢失包前后（双向）的数据进行丢包补偿，以提高合成语音的质量。并且在丢包补偿的同时，能以较小的代价对基音加倍现象进行检测和纠正。该在发现丢包情况下，每次调用该方法时只补偿出一个基音周期长度的数据，这样精细的补偿长度，为在该过程中收到相应的延迟包时能利用此包数据更合理地做丢包补偿。

附图说明

图 1 为相关技术中丢包补偿算法流程简图；

图 2 为本公开的一些实施例中语音丢包补偿的方法流程图；

图 3 为本公开的一些实施例中语音丢包补偿的方法流程图；

图 4 为本公开的一些实施例中补偿方法示意图；

图 5 为本公开的一些实施例中对 $vec1$ 做平滑处理的示意图；

图 6 为本公开的一些实施例中余弦窗函数波形的示意图；

图 7 为本公开的一些实施例中语音合成示意图，其中左侧围伸展情况，右侧围压缩情况；

5 图 8 为丢失 20ms 数据没有补偿的波形的示意图；

图 9 为丢失 20ms 数据只利用历史数据补偿的波形的示意图；

图 10 为丢失 20ms 数据用双向数据补偿的波形的示意图；

图 11 为没有丢包的原始波形的示意图；

图 12 为本公开的语音丢包补偿的系统进行语音丢包补偿的流程图；

10 图 13 为本公开的一些实施例中语音丢包补偿的系统框图。

具体实施方式

下面将参照附图更详细地描述本公开的示例性实施例。虽然附图中显示了本公开的示例性实施例，然而应当理解，可以以各种形式实现本公开而不应被这里阐述的实施例所限制。相反，提供这些实施例是为了能够更透彻地理解本公开，并且能够将本公开的范围完整的传达给本领域的技术人员。

参见图 2，图中示出了一些实施例中语音丢包补偿的方法，具体步骤如下：

步骤 S201、当第一次发现丢失语音包时，根据丢失语音包之前的历史语音数据，生成第一语音序列 $vec1$ ，然后进入步骤 S203。

20 具体地，在步骤 S201 中，当第一次发现丢失语音包时（即第一次做丢包补偿时），对所述历史语音数据进行基音粗估，得到基音粗估结果；根据所述基音粗估结果，进行基音细搜索，确定出第一基音周期 $Lag1$ ；根据所述第一基音周期 $Lag1$ 从所述历史语音数据生成所述第一语音序列 $vec1$ 。

步骤 S203、若能够获取丢失语音包之后的下一包语音数据，则根据丢失语音包之后的下一包语音数据，生成第二语音序列 $vec2$ ，然后进入步骤 S205。

25 具体地，在步骤 S203 中，从丢失语音包之后的下一包语音数据中找出一个与所述第一语音序列 $vec1$ 相位对齐的第二语音序列 $vec2$ ，具体包括：根据所述第一基音周期 $Lag1$ ，在下一包语音数据中进行基音细搜索，确定出下一包语音数据的第二基音周期 $Lag2$ ；在所述下一包语音数据中找出与所述第一

语音序列 $vec1$ 相关性最强的位置；根据相关性最强的位置和第二基音周期 $Lag2$ ，从所述下一包语音数据中确定出所述第二语音序列 $vec2$ 。

可选地，若不能获取下一包语音数据，根据第一语音序列 $vec1$ 输出第二丢包补偿数据，并进行丢包补偿。

5 步骤 S205、根据第一语音序列 $vec1$ 和第二语音序列 $vec2$ 进行丢包补偿。

具体地，在步骤 S205 中，根据第一语音序列 $vec1$ 和第二语音序列 $vec2$ ，计算要填充的基音数据的个数；根据填充的基音数据的个数，计算要填充的各个基音数据的长度；对所述第一语音序列 $vec1$ 和所述第二语音序列 $vec2$ 分别进行平滑处理；根据各个基音数据的长度，利用平滑后的第一语音序列 $vec1$ 和第二语音序列 $vec2$ 合成一个基音周期长度的语音序列。可选的，利用平滑后的第一语音序列 $vec1$ 和第二语音序列 $vec2$ 采用波形内插法合成一个基音周期长度的语音序列。当然可以理解的是，在本实施例中并限定其具体插值算法。

需要说明的是，上述根据所述第一语音序列 $vec1$ 和所述第二语音序列 $vec2$ ，计算要填充的基音数据的个数，具体包括：根据丢失语音包的长度 d 、所述第一语音序列 $vec1$ 和所述第二语音序列 $vec2$ 相位对齐后确定的长度 rd 、所述第一基音周期 $Lag1$ 和第二基音周期 $Lag2$ ，采用如下公式计算要填充的基音数据的个数 $numPitch$ ； $numPitch = (rd + (avgLag/2)) / avgLag$ ，其中 $avgLag$ 表示平均基音周期长度， $avgLag = (Lag1 + Lag2) / 2$ 。

20 需要说明的是，上述根据各个基音数据的长度，合成一个基音周期长度的语音序列，具体包括：根据待合成语音的位置分别与所述第一语音序列 $vec1$ 和所述第二语音序列 $vec2$ 的距离关系，确定合成时所述第一语音序列 $vec1$ 的权重 $w1$ 和所述第二语音序列 $vec2$ 的权重 $w2$ ；根据所述第一语音序列 $vec1$ 和其权重 $w1$ ，以及所述第二语音序列 $vec2$ 和其权重 $w2$ ，对平滑后的第一语音序列 $vec1$ 和第二语音序列 $vec2$ （例如采用波形内插法）合成得到一个基音周期长度的语音序列，其中若填充此合成的一个基音周期长度语音序列后，丢失的语音数据包还未到达，则只需要再次利用平滑后的第一语音序列 $vec1$ 和第二语音序列 $vec2$ （例如采用波形内插法）依次合成得其它的语音序列，每次合成一个基音周期长的数据后都会去检查丢失的数据包是否到达，若是则

进行平滑衔接，若否则继续用上述方法补偿。

在本公开的实施例中在发现当前数据包丢失情况下，如果不能获取丢失包的下一个语音包时，则利用该丢失包前的历史数据进行丢包补偿；如果能获取丢失包的下一个语音包时，则利用丢失包前后（双向）的数据进行丢包补偿，以提高合成语音的质量。并且在丢包补偿的同时，能以较小的代价对基音加倍现象进行检测和纠正。该在发现丢包情况下，每次调用该方法时只补偿出一个基音周期长度的数据，这样精细的补偿长度，为在该过程中收到相应的延迟包时能利用此包数据更合理地做丢包补偿。

在一实施例中，参见图 3，图中示出了利用双向数据丢包补偿算法流程图，具体步骤如下：

第 1 步，当第一次进行丢包补偿时，先利用降采样后的历史数据进行基音粗估，降采样的目的是降低运算量，例如采用自相关法进行降采样处理，其中可以对数据降采样到 4K，自相关函数的计算方法如下：

$$\text{corr}(k) = \sum_{n=0}^{N-1-k} x(n)x(n-k) \quad 0 \leq k \leq p-1 \quad (4-1)$$

其中，N 为序列 $x(i)$ 的长度，p 为要计算的自相关值的个数。

从 p 个自相关值中找出三个最大值，它们分别对应一个待选基音周期。接着在三个待选的基因周期附近利用原始数据进行基因细搜索（细搜索的方法同样采用自相关法），并确定出最终的基音周期：Lag1。根据找到的基音周期从历史数据产生出一个语音序列（vec1，产生方法为从历史数据中的最新样点向前追溯 Lag1 个样点）用来合成最终输出的语音。

第 2 步，如果丢失包的下一包数据不存在，那么直接利用上一步生成的语音序列合成输出的语音序列；如果丢失包的下一包数据存在，则对其做基音周期细搜索。新的基音周期搜索的目标就锁定在 Lag1 附近，搜索以 Lag1 为中心，搜索范围为正负 Lag1/2，通过自相关法在此范围内找出最大自相关值来确定新的基音周期：Lag2。

第 3 步，从丢失包的下一包数据（下称：新数据）中找出一个与 vec1 相位对齐的语音序列 vec2。设历史数据存放在 speech buffer 中，那么 vec1 为 speech buffer 中最新数据向前追溯 Lag1 个样点，长度为 Lag1；可选的，用互相关法（互相关函数的计算方法见式 4-2）在新的数据（图 4 中的 Next packet data）

中找出与 vec1 相关性最强的位置 (corr(k)中最大值对应的位置 k), 并取出 Lag2 长数据, 记此序列为 vec2。根据相关性最强原则找出 vec2 的起始点的目的是为了两个序列相位对齐。

$$\text{corr}(k) = \sum_{n=0}^{N-1} \text{vec1}(n)\text{vec2}(n-k) \quad (4-2)$$

- 5 第 4 步, 计算出 vec1 与 vec2 的归一化互相关值 (归一化互相关函数的计算方法见式 4-3), 并与一个设定的门限做比较, 在 vec1 与 vec2 相关性较强 (大于门限) 的情况下才利用双向的数据进行丢包补偿, 否则只利用历史数据进行丢包补偿。

$$\text{corr}(k) = \frac{\sum_{n=0}^{N-1} \text{vec1}(n)\text{vec2}(n-k)}{\sqrt{\sum_{n=0}^{N-1} \text{vec1}^2(n) \sum_{n=0}^{N-1} \text{vec2}^2(n-k)}} \quad (4-3)$$

- 10 第 5 步, 根据丢失数据的长度 d (如图 3), 相位对齐后确定的长度 rd, Lag1 和 Lag2 计算出要填充的基音个数, 并根据填充个数判断前面确定的基音周期长度是否合理。方法如下:

1. 计算平均基音周期长度 $\text{avgLag} = (\text{Lag1} + \text{Lag2}) / 2$;
2. 计算要填充的基音个数 $\text{numPitch} = (\text{rd} + (\text{avgLag}/2)) / \text{avgLag}$;
- 15 3. 计算用 numPitch 个长为 avgLag 的数据进行填充后与实际要填充长度的差值与 avgLag 的比值: $\text{abs}(\text{numPitch} \times \text{avgLag} - \text{rd}) / \text{avgLag}$ 。

- 如果这个比值落在 0.4 到 0.6 之间, 那么证明前面确定的基音周期长度很可能是加倍了的基音周期, 为了确定这个可能性, 在第 1 步提到的三个待选的基音周期值中寻找是否有值落在 $[\text{avgLag}/2 - 5, \text{avgLag}/2 + 5]$ 之中, 如果有的话就用这个值更新 Lag1, 并且程序跳转到第 2 步重新执行; 如果不存在这个值, 那就说明基音周期搜索或相关性检测出现问题, 为了避免问题继续恶化, 只利用历史数据进行丢包补偿; 如果这个比值落在 $[0.2, 0.4]$ 或 $[0.6, 0.8]$ 之中, 说明基音周期搜索或相关性检测出现问题, 为了避免问题继续恶化, 只利用历史数据进行丢包补偿; 如果这个比值落在 $[0, 0.2]$ 或 $[0.8, 1]$ 之中, 说明基音周期搜索正确, 继续下面的步骤。
- 25

第 6 步, 计算要插入的各个基音的长度。因为基音周期长度是渐变的, 所以丢包前后的基音周期长度很可能不同, 为了与实际中的周期长度的变化更接近, 我们对填充的基音周期长度也要做变化, 并且变化后的长度之和与

rd 相等。具体做法如下：

● 将 Lag1 作为 numPitch 个要填充的基音的初始值长度，统计一下按目前这种填充方式填充后的总长度 len，如果 len 与 rd 相等，说明目前的填充方式最优，继续进行第 7 步。否则进行下面的步骤。

5 ● 如果 len 大于等于 rd，要分两种情况处理：

 ◇ 情况 1，Lag2 大于等于 Lag1。从补偿的第一个基音开始逐个对其长度减 1，如果减到最后一个基音还不能让 len 等于 rd，那就再从第一个基音开始重复上面的过程直到 len 等于 rd。

 ◇ 情况 2，Lag2 小于 Lag1。从补偿的最后一个基音开始倒着向前逐个对其长度减 1，如果减到第一个基音还不能让 len 等于 rd，那就再从最后一个基音开始重复上面的过程直到 len 等于 rd。

 如果 len 小于 rd，要分两种情况处理

 ◇ 情况 1，Lag2 大于等于 Lag1。从补偿的最后一个基音开始倒着向前逐个对其长度加 1，如果加到第一个基音还不能让 len 等于 rd，那就再从最后一个基音开始重复上面的过程直到 len 等于 rd。

 ◇ 情况 2，Lag2 小于 Lag1。从补偿的第一个基音开始逐个对其长度加 1，如果加到最后一个基音还不能让 len 等于 rd，那就再从第一个基音开始重复上面的过程直到 len 等于 rd。

 第 7 步，对 vec1，vec2 做平滑处理。因为之后要插入的基音数据都是用 vec1 和 vec2 插值得到的，如果 vec1 或 vec2 各自首尾之间不做下平滑的话那么插值出的基音数据之间就会出现不连续现象，严重的话会有尖锐的噪声出现。下面以 vec1 为例，说明一下如何做平滑处理。

 如图 5 所示，取 vec1 前面一段数据，称为 o1，取 vec1 最新的一段数据，称为 o2，o2 与 o1 长度相等。对 o1 施加三角窗的左半部，对 o2 施加三角窗的右半部，然后将两段信号重叠相加，结果更新 o2 区域。

 第 8 步，产生出基音周期长的数据。每次调用 PLC 算法时，根据第 6 步计算得到的各个基音的长度依次合成新的数据。合成的方法可选采用波形内插法，具体如下：

 ● 根据目标长度 N（要合成的基音的长度）调整 vec1 和 vec2 的长度。

调整的方法如下：

设原始长度为 L , 目标长度为 N 。当 L 等于 N 时不用调整, 不相等的时候不能通过变采样的方式进行调整, 否则会破坏原始语音的音色, 要通过加余弦窗的方式做伸展或压缩处理。参见图 6, 余弦窗函数如下：

$$w(n) = 0.5 + 0.5 * \cos\left(\frac{2\pi n}{N}\right) \quad n = 0, 1, \dots, N-1$$

✧ 伸展处理 (当 L 小于 N 时): 生成一个长为 $2 \times L$ 的余弦窗, 左/右半余弦窗左/右边界与原序列的左/右边界对齐, 加到原序列上, 得到数据 $p1/p2$ 。合成时把 $p1$ 的左边界与目标序列的左边界对齐 (图 7 左 a), $p2$ 的右边界与目标序列的右边界对齐 (图 7 左 b), 然后将 $p1$ 与 $p2$ 对应点相加 (图 7 左 c) 得到最终合成的语音。

✧ 压缩处理 (当 L 大于 N 时): 生成一个长为 $2 \times N$ 的余弦窗, 左/右半余弦窗的左/右边界与原序列的左/右边界对齐, 加到原序列上, 得到数据 $p1/p2$ 。合成时把 $p1$ 的左边界与目标序列的左边界对齐 (图 7 右 a), $p2$ 的右边界与目标序列的右边界对齐 (图 7 右 b), 然后将 $p1$ 与 $p2$ 对应点相加 (图 7 右 c) 得到最终合成的语音。

● 按上述方法将 $vec1$ 和 $vec2$ 按目标长度调整之后, 要计算各自的权重 $w1$ 和 $w2$ 。权重的计算根据待合成语音的位置与 $vec1$ 和 $vec2$ 的距离关系。原理是合成语音离哪个序列近其波形应该更接近哪个序列。步骤 6 中得到要插入的基音个数为 $numPitch$, 要合成的第 i ($i=1, 2, \dots, numPitch$) 个基音的数据时, 相应权重为 $w2 = i/(numPitch+1), w1=1-w2$ 。

● 合成一个基音周期长度的语音。 $vec1$ 乘以权重 $w1$, $vec2$ 乘以权重 $w2$, 二者相加就得到了最终的合成语音。在连续做丢包补偿时, 只要重复执行第 8 步, 依次生成 $numPitch$ 段合成语音即可。

参见图 8~图 11, 图中示出了丢失包前后数据相关性较强情况下的丢包补偿的实验数据, 实验条件为包长 20ms, 5%的丢包率。如图 8 所示, 中间部分为丢失一包的情况, 当算法发现有丢包情况发生时, 便按如下步骤进行丢包补偿:

(1) 先从缓存 BUFFER 中取出 30 毫秒数据, 将数据降采样到 4K, 对降采样后的数据用公式 4-1 求出 50 个自相关值, 其中 N 取 60 个样点 (60 个样点

在 4K 采样率下对应最大基音周期长度), 并找出其中最大的三个, 相应的 k 值对应粗估出的基音周期对应的样点数。再把三个 k 值映射到原采样率下的对应位置 k' , 在 k' 附近 (左右 0.5ms 范围内) 时, 可选的再用公式 4-1 计算自相关值, 此范围内算出的最大值对应的位置设为 k'' , 3 个待选的 k'' 中的最大值为最终的基音周期, 设为 $Lag1$ 。并从缓存 BUFFER 中最新的样本向前追溯 $Lag1$ 个样点, 取出语音序列 $vec1$ 。

(2) 在能获取丢失包下一个包的情况下, 我们对拿到的数据包中的数据按步骤 2 所述方法找出新的基音周期 $Lag2$, 同样可选的用公式 4-1 求自相关值, N 取 $Lag1$ 。

(3) 按上述第 3 步对 $vec1$ 和 $vec2$ 进行相位对齐, $vec1$ 在新的数据中逐点移动 $Lag2$ 次, 每移动一次都计算 $vec1$ 和与之对应的 $vec2$ 的互相关值, 按公式 4-2 计算 $Lag2$ 个互相关值, 其中 N 取 $\text{MIN}(Lag1, \text{数据包的长度 } LEN - Lag2)$, 最后从 $Lag2$ 个互相关值中找出最大的一个, 从此最大值对应位置处取出 $Lag2$ 个样点数据作为最终的序列 $vec2$ 。

(4) 按上述第 4 步计算 $vec1$ 和 $vec2$ 的归一化互相关值 $bestCorr$ 。其中 N 取 $\text{MIN}(Lag1, \text{数据包的长度 } LEN - Lag2)$, k 取 0。根据测试经验将比较门限取为 0.6, 当 $bestCorr$ 大于 0.6 时认为两个序列相关性强, 反之相关性弱。

(5) 按上述第 5 步计算要填充的基音个数, 并判断出前面确定的基音周期长度是否合理。

(6) 按上述第 6 步计算要插入的各个基音的长度。

(7) 按步骤 7 所述方法对序列进行平滑, 平滑区域长度: $5 * \text{samplerate} / 8000$ 个样点, 其中 samplerate 为数据的采样率。

(8) 在每次需要合成数据的时候, 按上述第 8 步进行基音周期长的数据的合成。按如上步骤合成出的数据如图 10 所示, 可以看出用本算法补偿出的波形相对于图 9 更接近于原始波形。

在有一些实施例中, 参见图 13, 图中示出了第三实施例中语音丢包补偿的系统框图, 该系统 13 包括:

第一语音序列模块 131, 用于当第一次发现丢失语音包时, 根据丢失语音包之前的历史语音数据, 生成第一语音序列 $vec1$;

第二语音序列模块 133, 用于若能够获取丢失语音包之后的下一包语音数据, 则根据丢失语音包之后的下一包语音数据, 生成第二语音序列 $vec2$;

补偿模块 135, 用于根据所述第一语音序列 $vec1$ 和所述第二语音序列 $vec2$ 进行丢包补偿。

5 需要说明的是, 补偿模块 135, 还用于若不能获取所述下一包语音数据, 根据所述第一语音序列 $vec1$ 进行丢包补偿。

需要说明的是, 第一语音序列模块 131 具体用于对所述历史语音数据进行基音粗估, 得到基音粗估结果; 根据所述基音粗估结果, 进行基音细搜索, 确定出第一基音周期 $Lag1$; 根据所述第一基音周期 $Lag1$ 从所述历史语音数
10 据生成所述第一语音序列 $vec1$ 。

需要说明的是, 第二语音序列模块 133 具体用于: 从丢失语音包之后的下一包语音数据中找出一个与所述第一语音序列 $vec1$ 相位对齐的第二语音序列 $vec2$ 。

需要说明的是, 第二语音序列模块 133 具体包括: 第一单元, 用于根据
15 所述第一基音周期 $Lag1$, 在所述下一个语音数据中进行基音细搜索, 确定出所述下一个语音数据的第二基音周期 $Lag2$; 第二单元, 用于在所述下一包语音数据中找出与所述第一语音序列 $vec1$ 相关性最强的位置; 第三单元, 用于根据所述相关性最强的位置和第二基音周期 $Lag2$, 从所述下一包语音数据中确定出所述第二语音序列 $vec2$ 。

20 需要说明的是, 第二语音序列模块还包括: 第四单元, 用于计算出所述第一语音序列 $vec1$ 和所述第二语音序列 $vec2$ 的归一化互相关值; 第五单元, 用于比较所述归一化互相关值与设定的门限值; 若所述归一化互相关值大于所述设定的门限值, 则触发所述补偿模块根据所述第一语音序列 $vec1$ 和所述第二语音序列 $vec2$ 进行丢包补偿; 若所述归一化互相关值小于或等于所述设
25 定的门限值, 触发所述补偿模块根据所述第一语音序列 $vec1$ 进行丢包补偿。

需要说明的是, 补偿模块 135 具体包括: 第六单元, 用于根据所述第一语音序列 $vec1$ 和所述第二语音序列 $vec2$, 计算要填充的基音数据的个数; 第七单元, 用于根据所述填充的基音数据的个数, 计算要填充的各个基音数据的长度; 第八单元, 用于对所述第一语音序列 $vec1$ 和所述第二语音序列 $vec2$

分别进行平滑处理；第九单元，用于根据所述各个基音数据的长度，对平滑过的第一语音序列 $vec1$ 和第二语音序列 $vec2$ （例如采用波形内插法）分别合成一个基音周期长度的语音序列。需要说明的是，所述第七单元具体用于：根据丢失语音包的长度 d 、所述第一语音序列 $vec1$ 和所述第二语音序列 $vec2$ 相位对齐后确定的长度 rd 、所述第一基音周期 $Lag1$ 和第二基音周期 $Lag2$ ，采用如下公式计算要填充的基音数据的个数 $numPitch$ ； $numPitch = (rd + (avgLag/2))/ avgLag$ ，其中 $avgLag$ 表示平均基音周期长度， $avgLag = (Lag1 + Lag2) / 2$ 。

需要说明的是，补偿模块 135 还包括：第十单元，用于计算用 $numPitch$ 个基音周期长度为 $avgLag$ 的数据进行填充后的长度与实际要填充长度的差值与平均基音周期长度 $avgLag$ 的比值；根据所述比值判断是否存在基音加倍；若是，触发第二语音序列模块重新生成第二语音序列 $vec2$ ；若否，则触发所述第七单元根据所述填充的基音数据的个数，计算要填充的各个基音数据的长度。

需要说明的是，第九单元进一步用于：根据待合成语音的位置分别与所述第一语音序列 $vec1$ 和所述第二语音序列 $vec2$ 的距离关系，确定合成时所述第一语音序列 $vec1$ 的权重 $w1$ 和所述第二语音序列 $vec2$ 的权重 $w2$ ；根据所述第一语音序列 $vec1$ 和其权重 $w1$ ，以及所述第二语音序列 $vec2$ 和其权重 $w2$ ，合成得到一个基音周期长度的语音序列。具体地，对平滑后的第一语音序列 $vec1$ 和第二语音序列 $vec2$ （例如采用波形内插法）合成得到一个基音周期长度的语音序列，其中若填充此合成的一个基音周期长度语音序列后，丢失的语音数据包还未到达，则只需要再次利用平滑后的第一语音序列 $vec1$ 和第二语音序列 $vec2$ （例如采用波形内插法）依次合成得其它的语音序列，每次合成一个基音周期长的数据后都会去检查丢失的数据包是否到达，若是则进行平滑衔接，若否则继续用上述方法补偿。

结合图 12，本丢包补偿系统应用在播放线程中，播放线程每隔 10 毫秒进行 10 毫秒的语音数据播放。本系统中有一个语音包缓冲区用来缓存未被播放的接收到的语音包。通过时间标的关系来判断当前包是否丢失， $WantedTS$ 表示当前播放需要的数据包的时间标， $CurrTS$ 表示语音包缓存区中的存在的

有效的最小的时间标。当 WantedTS 与 CurrTS 相等时，表示当前需要的语音包存在，则对此包数据进行解码并更新 WantedTS，这种情况下还要判断上一次的操作是否为丢包补偿，如果是，还要对新的数据与历史数据（包含补偿出的数据）进行相关最强的位置确定，然后在相关性最强的位置处做平滑衔接保证边界处的平滑过渡，再将处理后的数据拷贝到播放缓冲进行播放。如果上一次操作不是丢包补偿，则将解码后的数据拷贝到播放缓冲区中进行播放就可以了；当 WantedTS 小于 CurrTS 时，表示当前需要的数据包不存在于语音包缓冲区中，这包可能是真的丢失了，也可能是由于延迟在指定的时间还没到达，不管何种原因，由于播放的实时性都要调用丢包补偿算法进行补包操作，一次补偿出一个基因周期长的数据，补偿完之后判断一下语音缓冲区中的数据是不是够 10 毫秒，够了则进行播放，否则程序再跳到开始处根据时间标检测需要的语音包到达否，如果到了则按上述方法利用延迟包进行平滑过渡，如果没到则继续进行丢包补偿，直到语音缓冲区中的数据够 10 毫秒播放为止。

在本申请所提供的几个实施例中，应该理解到，所揭露方法和装置，可以通过其它的方式实现。例如，以上所描述的装置实施例仅仅是示意性的，例如，所述单元的划分，仅仅为一种逻辑功能划分，实际实现时可以有另外的划分方式，例如多个单元或组件可以结合或者可以集成到另一个系统，或一些特征可以忽略，或不执行。另一点，所显示或讨论的相互之间的耦合或直接耦合或通信连接可以是通过一些接口，装置或单元的间接耦合或通信连接，可以是电性，机械或其它的形式。

另外，在本公开各个实施例中的各功能单元可以集成在一个处理单元中，也可以是各个单元单独物理包括，也可以两个或两个以上单元集成在一个单元中。上述集成的单元既可以采用硬件的形式实现，也可以采用硬件加软件功能单元的形式实现。

上述以软件功能单元的形式实现的集成的单元，可以存储在一个计算机可读取存储介质中。上述软件功能单元存储在一个存储介质中，包括若干指令用以使得一台计算机设备（可以是个人计算机，服务器，或者网络设备等）执行本公开各个实施例所述方法的部分步骤。而前述的存储介质包括：U 盘、

移动硬盘、只读存储器（Read-Only Memory，简称ROM）、随机存取存储器（Random Access Memory，简称RAM）、磁碟或者光盘等各种可以存储程序代码的介质。

- 5 以上所述的是本公开的可选实施方式，应当指出对于本技术领域的普通人员来说，在不脱离本公开所述的原理前提下还可以做出若干改进和润饰，这些改进和润饰也在本公开的保护范围内。

权 利 要 求 书

1. 一种语音丢包补偿的方法，包括：

当第一次发现丢失语音包时，根据丢失语音包之前的历史语音数据，生
5 成第一语音序列 $vec1$ ；

若能够获取丢失语音包之后的下一包语音数据，则根据丢失语音包之后
的下一包语音数据，生成第二语音序列 $vec2$ ；根据所述第一语音序列 $vec1$ 和
所述第二语音序列 $vec2$ 进行丢包补偿。

2. 根据权利要求 1 所述的方法，其中，若不能获取所述下一包语音数据，
10 所述方法还包括：

根据所述第一语音序列 $vec1$ 进行丢包补偿。

3. 根据权利要求 1 所述的方法，其中，所述根据丢失语音包之前的历史
语音数据，生成第一语音序列，具体包括：

对所述历史语音数据进行基音粗估，得到基音粗估结果；

15 根据所述基音粗估结果，进行基音细搜索，确定出第一基音周期 $Lag1$ ；
根据所述第一基音周期 $Lag1$ 从所述历史语音数据生成所述第一语音序
列 $vec1$ 。

4. 根据权利要求 1 所述的方法，其中，所述根据丢失语音包之后的下一
包语音数据，生成第二语音序列 $vec2$ ，具体为：

20 从丢失语音包之后的下一包语音数据中找出一个与所述第一语音序列
 $vec1$ 相位对齐的第二语音序列 $vec2$ 。

5. 根据权利要求 4 所述的方法，其中，所述从丢失语音包之后的下一包
语音数据中找出一个与所述第一语音序列 $vec1$ 相位对齐的第二语音序列 $vec2$ ，
具体包括：

25 根据所述第一基音周期 $Lag1$ ，在所述下一包语音数据中进行基音细搜索，
确定出所述下一包语音数据的第二基音周期 $Lag2$ ；

在所述下一包语音数据中找出与所述第一语音序列 $vec1$ 相关性最强的位
置；

根据所述相关性最强的位置和第二基音周期 $Lag2$ ，从所述下一包语音数

据中确定出所述第二语音序列 vec2 。

6. 根据权利要求 5 所述的方法, 其中, 在根据所述第一语音序列 vec1 和所述第二语音序列 vec2 进行丢包补偿之前, 所述方法还包括:

计算出所述第一语音序列 vec1 和所述第二语音序列 vec2 的归一化互相关值;

比较所述归一化互相关值与设定的门限值;

若所述归一化互相关值大于所述设定的门限值, 则进入根据所述第一语音序列 vec1 和所述第二语音序列 vec2 进行丢包补偿的步骤;

若所述归一化互相关值小于或等于所述设定的门限值, 则进入根据所述第一语音序列 vec1 进行丢包补偿的步骤。

7. 根据权利要求 6 所述的方法, 其中, 所述根据所述第一语音序列 vec1 和所述第二语音序列 vec2 进行丢包补偿, 具体包括:

根据所述第一语音序列 vec1 和所述第二语音序列 vec2 , 计算要填充的基音数据的个数;

根据所述填充的基音数据的个数, 计算要填充的各个基音数据的长度;

对所述第一语音序列 vec1 和所述第二语音序列 vec2 分别进行平滑处理;

根据所述各个基音数据的长度, 利用平滑后的第一语音序列 vec1 和第二语音序列 vec2 合成一个基音周期长度的语音序列。

8. 根据权利要求 7 所述的方法, 其中, 所述根据所述第一语音序列 vec1 和所述第二语音序列 vec2 , 计算要填充的基音数据的个数, 具体包括:

根据丢失语音包的长度 d 、在所述第一语音序列 vec1 和所述第二语音序列 vec2 相位对齐后确定的长度 rd 、所述第一基音周期 Lag1 和第二基音周期 Lag2 , 采用如下公式计算要填充的基音数据的个数 numPitch ;

$$\text{numPitch} = (rd + (\text{avgLag}/2)) / \text{avgLag}$$
 其中 avgLag 表示平均基音周期长度, $\text{avgLag} = (\text{Lag1} + \text{Lag2}) / 2$ 。

9. 根据权利要求 8 所述的方法, 其中, 所述根据所述第一语音序列 vec1 和所述第二语音序列 vec2 进行丢包补偿, 还包括:

计算用 numPitch 个基音周期长度为 avgLag 的数据进行填充后的长度与实际要填充长度的差值与平均基音周期长度 avgLag 的比值;

根据所述比值判断是否存在基音加倍；

若是，则要根据校正过的基音重新生成第二语音序列 $vec2$ ；

若否，则根据所述填充的基音数据的个数，计算要填充的各个基音数据的长度。

- 5 10. 根据权利要求 8 所述的方法，其中，所述根据所述各个基音数据的长度，利用平滑后的第一语音序列 $vec1$ 和第二语音序列 $vec2$ 合成一个基音周期长度的语音序列，具体包括：

根据待合成语音的位置分别与所述第一语音序列 $vec1$ 和所述第二语音序列 $vec2$ 的距离关系，确定合成时所述第一语音序列 $vec1$ 的权重 $w1$ 和所述第二语音序列 $vec2$ 的权重 $w2$ ；

10

根据所述第一语音序列 $vec1$ 和其权重 $w1$ ，以及所述第二语音序列 $vec2$ 和其权重 $w2$ ，对平滑后的第一语音序列 $vec1$ 和第二语音序列 $vec2$ 合成得到一个基音周期长度的语音序列，其中若填充此合成的一个基音周期长度语音序列后，丢失的语音数据包还未到达，则只需要再次利用平滑后的第一语音序列 $vec1$ 和第二语音序列 $vec2$ 依次合成得其它的语音序列，每次合成一个基音周期长的数据后都会去检查丢失的数据包是否到达，若是则进行平滑衔接，若否则继续用上述方法补偿。

15

11. 一种语音丢包补偿的系统，包括：

第一语音序列模块，用于当第一次发现丢失语音包时，根据丢失语音包之前的历史语音数据，生成第一语音序列 $vec1$ ；

20

第二语音序列模块，用于若能够获取丢失语音包之后的下一包语音数据，则根据丢失语音包之后的下一包语音数据，生成第二语音序列 $vec2$ ；

补偿模块，用于根据所述第一语音序列 $vec1$ 和所述第二语音序列 $vec2$ 进行丢包补偿。

- 25 12. 根据权利要求 11 所述的系统，其中，所述补偿模块，还用于若不能获取所述下一包语音数据，根据所述第一语音序列 $vec1$ 进行丢包补偿。

13. 根据权利要求 11 所述的系统，其中，所述第一语音序列模块具体用于对所述历史语音数据进行基音粗估，得到基音粗估结果；根据所述基音粗估结果，进行基音细搜索，确定出第一基音周期 $Lag1$ ；根据所述第一基音周

期 Lag1 从所述历史语音数据生成所述第一语音序列 vec1。

14. 根据权利要求 11 所述的系统, 其中, 所述第二语音序列模块具体用于: 从丢失语音包之后的下一包语音数据中找出一个与所述第一语音序列 vec1 相位对齐的第二语音序列 vec2。

5 15. 根据权利要求 14 所述的系统, 其中, 所述第二语音序列模块具体包括:

第一单元, 用于根据所述第一基音周期 Lag1, 在所述下一个语音数据中进行基音细搜索, 确定出所述下一个语音数据的第二基音周期 Lag2;

10 第二单元, 用于在所述下一包语音数据中找出与所述第一语音序列 vec1 相关性最强的位置;

第三单元, 用于根据所述相关性最强的位置和第二基音周期 Lag2, 从所述下一包语音数据中确定出所述第二语音序列 vec2。

16. 根据权利要求 15 所述的系统, 其中, 所述第二语音序列模块还包括:

15 第四单元, 用于计算出所述第一语音序列 vec1 和所述第二语音序列 vec2 的归一化互相关值;

第五单元, 用于比较所述归一化互相关值与设定的门限值; 若所述归一化互相关值大于所述设定的门限值, 则触发所述补偿模块根据所述第一语音序列 vec1 和所述第二语音序列 vec2 进行丢包补偿; 若所述归一化互相关值小于或等于所述设定的门限值, 触发所述补偿模块根据所述第一语音序列
20 vec1 进行丢包补偿。

17. 根据权利要求 16 所述的系统, 其中, 所述补偿模块具体包括:

第六单元, 用于根据所述第一语音序列 vec1 和所述第二语音序列 vec2, 计算要填充的基音数据的个数;

25 第七单元, 用于根据所述填充的基音数据的个数, 计算要填充的各个基音数据的长度;

第八单元, 用于对所述第一语音序列 vec1 和所述第二语音序列 vec2 分别进行平滑处理;

第九单元, 用于根据所述各个基音数据的长度, 利用平滑后的第一语音序列 vec1 和第二语音序列 vec2 合成一个基音周期长度的语音序列。

18. 根据权利要求 17 所述的系统，其中，所述第六单元具体用于：根据丢失语音包的长度 d 、所述第一语音序列 $vec1$ 和所述第二语音序列 $vec2$ 相位对齐后确定的长度 rd 、所述第一基音周期 $Lag1$ 和第二基音周期 $Lag2$ ，采用如下公式计算要填充的基音数据的个数 $numPitch$ ； $numPitch = (rd + (avgLag/2))/avgLag$ ，其中 $avgLag$ 表示平均基音周期长度， $avgLag = (Lag1 + Lag2)/2$ 。

19. 根据权利要求 18 所述的系统，其中，所述补偿模块还包括：

第十单元，用于计算用 $numPitch$ 个基音周期长度为 $avgLag$ 的数据进行填充后的长度与实际要填充长度的差值与平均基音周期长度 $avgLag$ 的比值；根据所述比值判断是否存在基音加倍；若是，触发第二语音序列模块根据校正过的基音重新生成第二语音序列 $vec2$ ；若否，则触发所述第七单元根据所述填充的基音数据的个数，计算要填充的各个基音数据的长度。

20. 根据权利要求 18 所述的系统，其中，所述第九单元进一步用于：根据待合成语音的位置分别与所述第一语音序列 $vec1$ 和所述第二语音序列 $vec2$ 的距离关系，确定合成时所述第一语音序列 $vec1$ 的权重 $w1$ 和所述第二语音序列 $vec2$ 的权重 $w2$ ；根据所述第一语音序列 $vec1$ 和其权重 $w1$ ，以及所述第二语音序列 $vec2$ 和其权重 $w2$ ，对平滑后的第一语音序列 $vec1$ 和第二语音序列 $vec2$ 合成得到一个基音周期长度的语音序列，其中若填充此合成的一个基音周期长度语音序列后，丢失的语音数据包还未到达，则只需要再次利用平滑后的第一语音序列 $vec1$ 和第二语音序列 $vec2$ 依次合成得其它的语音序列，每次合成一个基音周期长的数据后都会去检查丢失的数据包是否到达，若是则进行平滑衔接，若否则继续用上述方法补偿。

21. 一种语音丢包补偿的系统，包括：

存储器，以及

处理器，用于读取存储器中的程序，执行下列过程：

当第一次发现丢失语音包时，根据丢失语音包之前的历史语音数据，生成第一语音序列 $vec1$ ；

若能够获取丢失语音包之后的下一包语音数据，则根据丢失语音包之后的下一包语音数据，生成第二语音序列 $vec2$ ；

根据所述第一语音序列 vec1 和所述第二语音序列 vec2 进行丢包补偿。

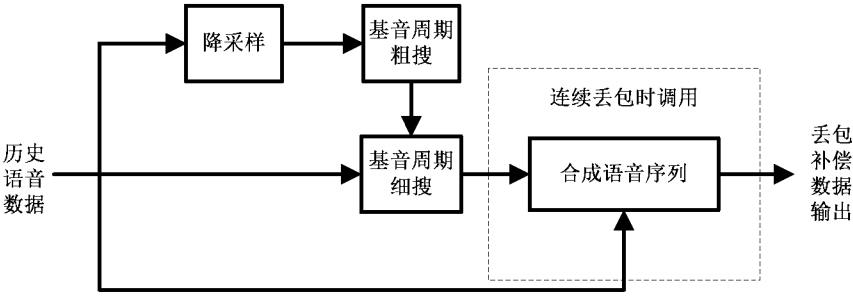


图 1

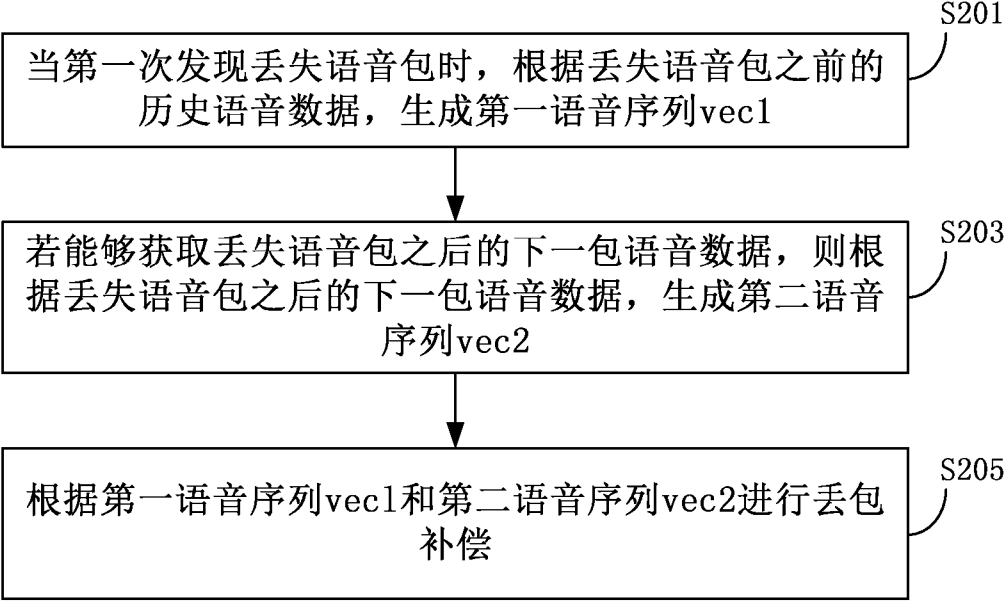


图 2

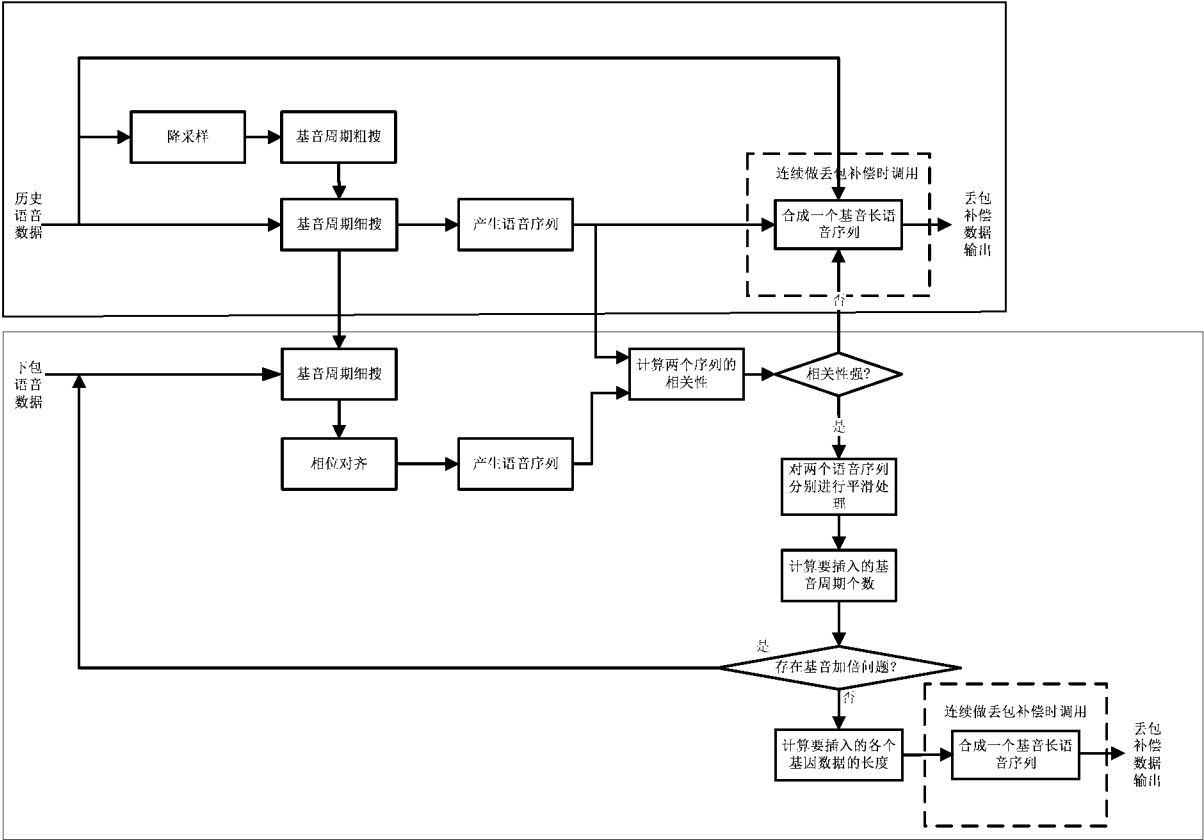


图 3

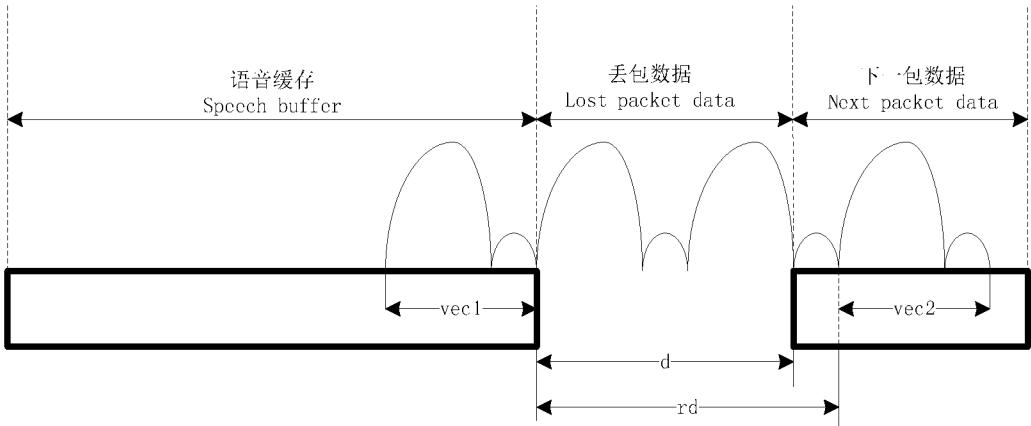


图 4

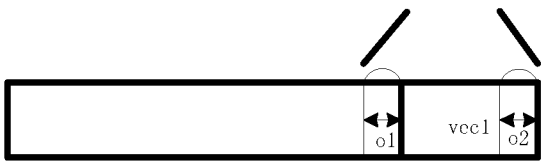


图 5

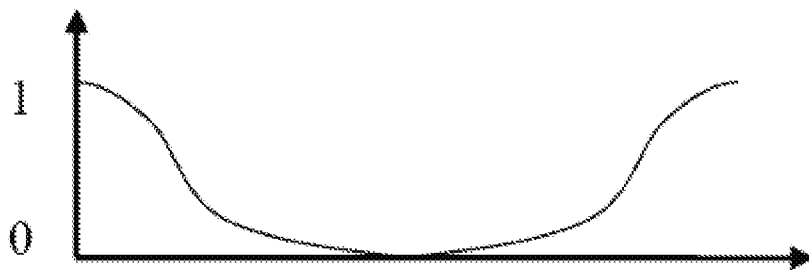


图 6

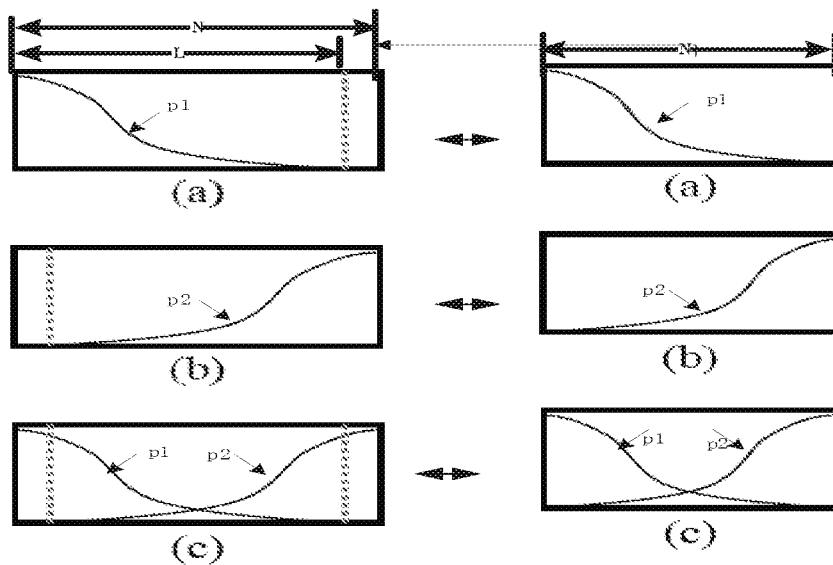


图 7

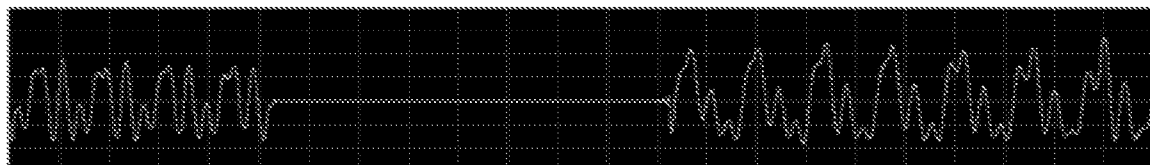


图 8



图 9



图 10

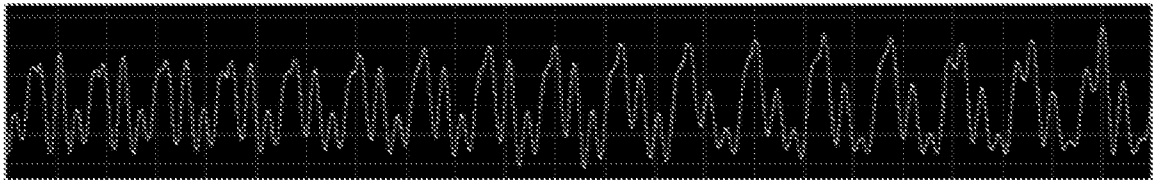


图 11

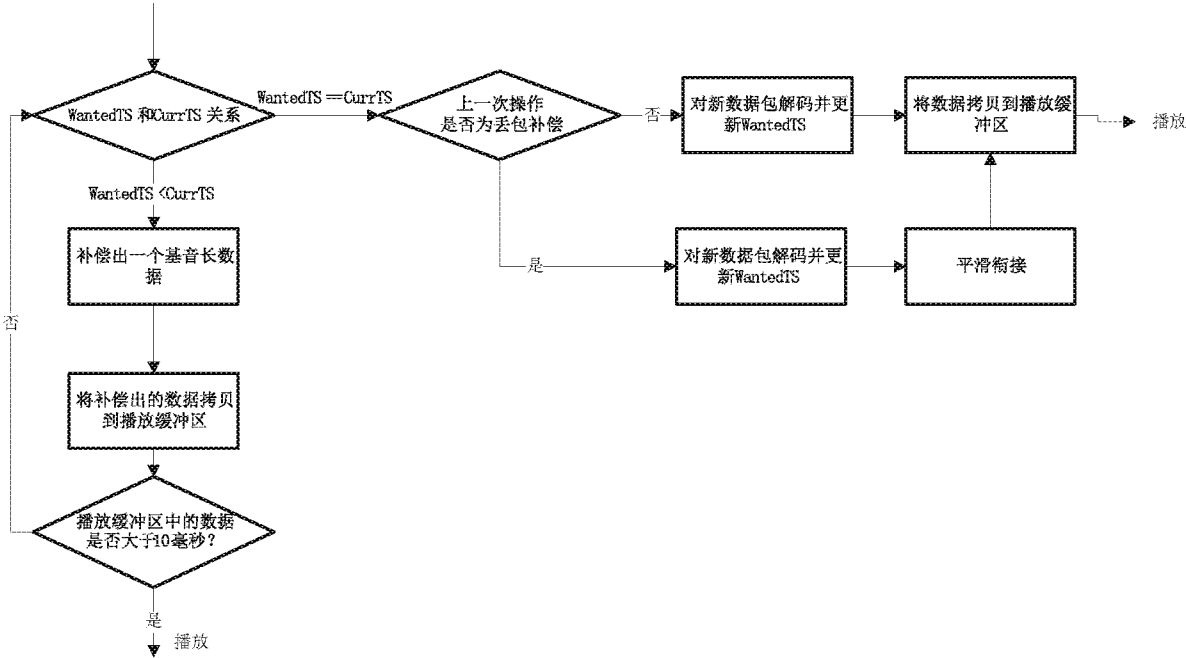


图 12

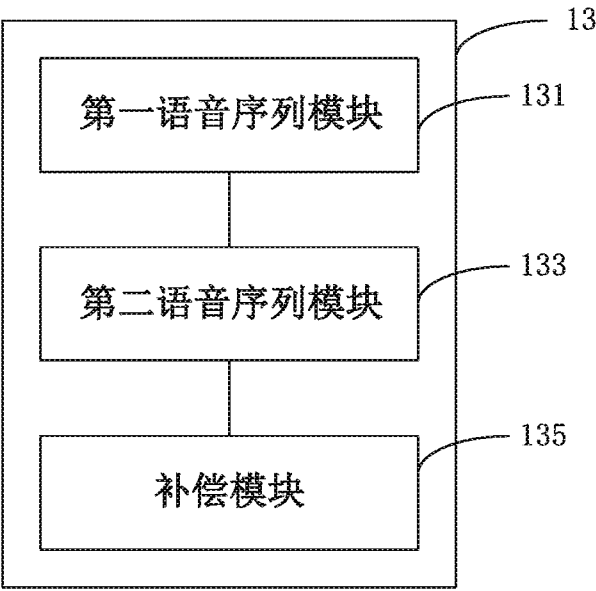


图 13

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2016/105684

A. CLASSIFICATION OF SUBJECT MATTER

H04M 11/06 (2006.01) i; G10L 19/005 (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)

H04M; G10L

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)

CNPAT; CNKI; WPI; EPODOC; ITU: packet loss, compensate, after, VoIP, PLC, voice, packet, loss, conceal+, pre+, next, frame, sequence

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 1984203 A (HUAWEI TECHNOLOGIES CO., LTD.), 20 June 2007 (20.06.2007), description, page 1, paragraph 2 to page 2, paragraph 4 and page 5, last paragraph to page 8, paragraph 5	1-21
A	CN 104978966 A (TENCENT TECHNOLOGY SHENZHEN CO., LTD.), 14 October 2015 (14.10.2015), the whole document	1-21
A	CN 103489448 A (GUANGZHOU ROPENTE TECHNOLOGY CO., LTD.), 01 January 2014 (01.01.2014), the whole document	1-21

☐ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents:	“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
“A” document defining the general state of the art which is not considered to be of particular relevance	“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
“E” earlier application or patent but published on or after the international filing date	“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
“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)	“&” document member of the same patent family
“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	

Date of the actual completion of the international search 10 January 2017 (10.01.2017)	Date of mailing of the international search report 06 February 2017 (06.02.2017)
Name and mailing address of the ISA/CN: 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 YAN, Yue Telephone No.: (86-10) 010-62413472

International application No.
PCT/CN2016/105684

Form PCT/ISA/210 (patent family annex) (July 2009)

国际检索报告

国际申请号

PCT/CN2016/105684

A. 主题的分类 H04M 11/06(2006.01)i; G10L 19/005(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) H04M; G10L 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)) CNPAT;CNKI;WPI;EPDOC;ITU:语音, 丢包, 补偿, 前, 后, 帧, 序列, VoIP, PLC, voice, packet, loss, conceal+, pre+, next, frame, sequence														
C. 相关文件 <table border="1"> <thead> <tr> <th>类 型*</th> <th>引用文件, 必要时, 指明相关段落</th> <th>相关的权利要求</th> </tr> </thead> <tbody> <tr> <td>X</td> <td>CN 1984203 A (华为技术有限公司) 2007年 6月 20日 (2007 - 06 - 20) 说明书第1页第2段-第2页第4段, 第5页最后1段-第8页第5段</td> <td>1-21</td> </tr> <tr> <td>A</td> <td>CN 104978966 A (腾讯科技深圳有限公司) 2015年 10月 14日 (2015 - 10 - 14) 全文</td> <td>1-21</td> </tr> <tr> <td>A</td> <td>CN 103489448 A (广州日滨科技发展有限公司) 2014年 1月 1日 (2014 - 01 - 01) 全文</td> <td>1-21</td> </tr> </tbody> </table>			类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求	X	CN 1984203 A (华为技术有限公司) 2007年 6月 20日 (2007 - 06 - 20) 说明书第1页第2段-第2页第4段, 第5页最后1段-第8页第5段	1-21	A	CN 104978966 A (腾讯科技深圳有限公司) 2015年 10月 14日 (2015 - 10 - 14) 全文	1-21	A	CN 103489448 A (广州日滨科技发展有限公司) 2014年 1月 1日 (2014 - 01 - 01) 全文	1-21
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求												
X	CN 1984203 A (华为技术有限公司) 2007年 6月 20日 (2007 - 06 - 20) 说明书第1页第2段-第2页第4段, 第5页最后1段-第8页第5段	1-21												
A	CN 104978966 A (腾讯科技深圳有限公司) 2015年 10月 14日 (2015 - 10 - 14) 全文	1-21												
A	CN 103489448 A (广州日滨科技发展有限公司) 2014年 1月 1日 (2014 - 01 - 01) 全文	1-21												
☐ 其余文件在C栏的续页中列出。 ☒ 见同族专利附件。 <table border="0"> <tr> <td> * 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 </td> <td> “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 </td> </tr> </table>			* 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										
* 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 2017年 1月 10日		国际检索报告邮寄日期 2017年 2月 6日												
ISA/CN的名称和邮寄地址 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 传真号 (86-10)62019451		受权官员 颜悦 电话号码 (86-10)010-62413472												

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2016/105684

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	1984203	A	2007年 6月 20日	无	
CN	104978966	A	2015年 10月 14日	无	
CN	103489448	A	2014年 1月 1日	无	