

# (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  
2008年4月10日 (10.04.2008)

(10) 国际公布号  
**WO 2008/040250 A1**

- (51) 国际专利分类号:  
*H04L 1/00* (2006.01)
- (21) 国际申请号: PCT/CN2007/070772
- (22) 国际申请日: 2007年9月25日 (25.09.2007)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:  
200610159697.2  
2006年10月1日 (01.10.2006) CN
- (71) 申请人 (对除美国外的所有指定国): 华为技术有限公司 (HUAWEI TECHNOLOGIES CO., LTD.) [CN/CN]; 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN).
- (72) 发明人; 及  
(75) 发明人/申请人 (仅对美国): 万华林 (WAN, Hualin) [CN/CN]; 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。王喆 (WANG, Zhe) [CN/CN]; 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。张军 (ZHANG, Jun) [CN/CN]; 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。
- (74) 代理人: 北京德琦知识产权代理有限公司 (DEQI INTELLECTUAL PROPERTY LAW CORPORATION); 中国北京市海淀区知春路1号学院国际大厦7层, Beijing 100083 (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, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE,

[见续页]

(54) Title: A METHOD, A DEVICE AND A SYSTEM FOR ERROR CONCEALMENT OF AN AUDIO STREAM

(54) 发明名称: 音频流错误隐藏的方法、装置和系统

对发送的音频帧按照内容进行分  
类, 得到该音频帧的类型信息 201

将音频帧的类型信息与该音频帧  
的编码结果封装并发送出去 202

当发生丢帧时, 确定丢失的音频  
帧的类型信息 203

当发生丢帧时, 对于丢失的音频  
帧, 确定其按照内容进行分类后  
得到的该音频帧的类型信息 204

201 CLASSIFYING THE AUDIO FRAME WHICH HAS BEEN TRANSMITTED BASED ON ITS CONTENTS AND GETTING THE TYPE INFORMATION OF THE AUDIO FRAME  
202 PACKAGING THE TYPE INFORMATION OF THE AUDIO FRAME WITH THE ENCODE RESULTS OF THE AUDIO FRAME AND TRANSMITTING THEM  
203 WHEN LOSING FRAME, DETERMINING THE TYPE INFORMATION OF THE LOST AUDIO FRAME  
204 WHEN LOSING FRAME, FOR THE LOST AUDIO FRAME, DETERMINING THE TYPE INFORMATION OF THE AUDIO FRAME WHICH HAS BEEN GOT BY CLASSIFYING IT BASED ON ITS CONTENTS

(57) Abstract: A method for concealing the error in the audio stream is provided which includes the following steps: classifying the audio frame which has been transmitted based on its contents and getting the type information of the audio frame; packaging the type information of the audio frame with the encode results of the audio frame and transmitting them; when losing frame, for the lost audio frame, determining the type information of the audio frame which has been got by classifying it based on its contents; according to the type information of the lost audio frame, reconstructing the audio frame by using the corresponding error recovery strategy. The said error concealing mode has better pertinency to the reconstruction of the lost frame and can reconstruct the lost audio frame self-adaptingly to achieve better compensating effect. A transmitting and receiving method for error concealment of an audio stream is provided. A transmitter, a receiver and a system for error concealment of an audio stream are also provided.

[见续页]



WO 2008/040250 A1



KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。

SD, SL, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), 欧洲 (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG)。

(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, LS, MW, MZ, NA,

本国际公布:

— 包括国际检索报告。

(57) 摘要:

本发明公开了一种音频流错误隐藏的方法, 该方法包括: 对发送的音频帧按照内容进行分类, 得到该音频帧的类型信息; 将音频帧的类型信息与该音频帧的编码结果封装打包发送出去; 当发生丢帧时, 对于丢失的音频帧, 确定其按照内容进行分类时得到的该音频帧的类型信息; 根据丢失的音频帧的类型信息, 采用相应的错误恢复策略进行音频帧重构。本发明所述的错误隐藏方式使得对丢失帧的重构具有更强的针对性, 能够自适应地重构音频帧, 以达到更好的补偿效果。本发明还公开了一种音频流错误隐藏的发送、接收方法。本发明又公开了音频流错误隐藏的发射机、接收机和系统。

## 音频流错误隐藏的方法、装置和系统

### 技术领域

本发明涉及实时音频通信技术，特别涉及音频流错误隐藏的方法、装置和系统。

### 5 发明背景

人们很早就开始从事音频分类研究，但对于不同的应用场景，音频分类的方法及分类的结果不尽相同。例如，1、在高级音响设备的噪音抑制中，常用调频分析或贝叶斯分类器将音频信号分类。2、为了更好地索引和检索因特网上音频资源，人们进行了基于内容的音频分类和检索研究。比较有代表性的基于内容的音频分类工作中详细分析了音频的区别性特征,包括响度(loudness)、基音(pitch)、谐度(harmonicity)等，并且设计了音频的分类器。3、音频分类另外一种应用是服务于音频——特别是语音编码器的语音激活检测器（VAD），目的是用来检测语音通信时是否有语音存在，对话音和非语音分别采用不同的编码方式，以达到在不降低通话质量的情况下节省话路资源的作用。

在实时音频传输系统（如 VoIP）中，网络传输造成音质下降的原因主要是时延、静电干扰和包丢失，其中，包丢失是音质下降的最主要原因。实际通讯网络总是存在着一定干扰，因此存在一定的包丢失概率，尽管协议低层有纠错编码，但只能解决包内误码，无法解决丢包问题。此外，由于实时音频业务有严格的延时限制，分组延迟超过一个极限也将被视为丢失。为了能够在一定丢包率情况下，仍然保证一定的通讯质量，很多编解码算法中都集成了错误隐藏技术，用来消除丢包的影响。

图 1 为音频流错误隐藏的框图。如图 1 所示，压缩音频信号经过 IP

网络或其他不可靠网络的传输后，通常将接收到的音频数据包存放在抖动缓存器，该抖动缓冲器用于解决迟包、早包的重排序等功能，然后进行丢包、错包检测。如果有丢包或错包发生，系统将启动错误隐藏进行丢包补偿，否则将正确接收音频包解码输出。

- 5 音频实时传输中的丢包恢复技术按照处理阶段可以划分为两大类：基于发送端的修复和基于接收端的修复。

● 基于发送端的错误隐藏

基于发送端的丢包恢复由发送端发起，并需要发送端和接收端协同进行。常见的方法有，增加冗余度、前向纠错、优先级设置和分类处理  
10 等。

1、增加冗余度：增加数据的冗余度可以提高系统的容错能力，但是同时增加了带宽。

2、前向纠错（FEC）：该策略也依赖于在传输流附加可修复丢失数据包的信息，利用块或代数码字生成能辅助纠错的额外传输包，同样需  
15 要增加带宽。

3、优先级设置方法：这种技术需要网络支持并按优先级传输分组，否则无法实现，并且只能改善网络拥塞造成的丢包概率。

4、分类处理方法：在语音编码中，为了在接收端更好地采用波形替换技术，发送端可以根据语音信号的特性分类处理，比如 3GPP2  
20 VMR-WB 和 ITU-T G.729.1 将语音帧进一步描述为 voiced, unvoiced, voiced transition, unvoiced transition, onset 等类型，而解码端接收到之后，利用其前一帧和后一帧的语音帧类型，可以推测出丢帧的类型，解码器得到丢帧类型后，可以较好的恢复丢失帧的信息。

● 基于接收端的错误隐藏

25 不需要发送端参与的接收端错误隐藏技术，本质上是对接收到的数

据通过一系列的方法来估计丢失的数据，并根据人的生理特点进行优化，基本上是一种被动的修补，通常比较容易实现且不增加带宽需求。

基于接收端的错误隐藏方法可分为三类：

1、基于插入的策略：这类技术包括拼接（Splicing）、静音和噪声替代等方法。拼接技术会扰乱媒体流的时序，效果也不好。静音替代（用静音帧填充丢帧位置）的适用范围非常有限，当数据包的丢失频率很低（小于2%）且缺口宽度小于4 ms时，这种方法比较有效；当缺口宽度达40 ms时，其效果会让人无法接受。与静音替代相比，噪声替代（用噪声帧填充丢帧位置）可给人带来更好的主观听觉感受，同时能改善语音信号的可分辨性。当采用背景噪声而不是静音的时候，人脑能下意识地用正确的声音来修补语音信号中丢失的部分。插入方式与语音编码无关，也与分组的编码无关，只是对解码后丢失的语音进行处理。

2、基于插值的策略：与插入技术相比，插值技术使得处理得到的声音能给人带来相对更好的主观感受。

3、基于重新生成的策略：从丢失包周围的信息提取解码状态，并由此生成丢失包的替代包。这种方法的实现过程比较复杂，但会取得较好的结果。

一般来说，基于发送端错误隐藏会增加网络带宽和计算复杂性，效果比基于接收端的好，但是如果发送端错误隐藏独立于接收端，也就是说与媒体内容无关，那么它将不能根据所丢帧的特性采取相应的错误隐藏策略（例如，稳定语音帧与其前一帧非常相似，用帧复制策略就能取得很好的隐藏效果，过渡帧则需要考虑前后帧的状态才能确定隐藏策略）。接收端的技术简单，也能取得一定的隐藏效果，但如果隐藏策略与音频编码无关，也就是说，不分析当前丢失帧和周围音频帧的内容特性，从而采取有针对性的错误隐藏策略，可采用错误隐藏的策略将非常

有限。

目前在编码端和接收端都考虑错误隐藏需求的音频编码标准越来越多，有代表性的主要有两类方法：

(1) 编码器在正式编码前分析音频帧特性，对不同特性的音频帧采用不同的编码方法。例如，AMR-WB+根据音频帧内容对信号帧分别采用 ACELP 及 TCX 编码，形成 26 种超帧（每四帧组成一个超帧）编码模式。编码模式信息被用于错误隐藏，在某帧丢失的情况下，接收端根据超帧其余 3 帧的编码类型，推断或估计出超帧的编码模式，从而实现一定的错误隐藏功能。

(2) 根据语音帧的基音、频谱等特性将语音帧划分为 voiced, unvoiced, voiced transition, unvoiced transition, onset 等类型。例如，在最新的 ITU-T G.729.1 协议中，编码器根据帧内容及其特性，将语音帧分为 voiced, unvoiced, voiced transition, unvoiced transition, onset (VBR-WB 也将语音帧划分为这 5 类)，在层 2 用 2 比特标示其类型，为进一步增强错误隐藏的准确性，G.729.1 还计算帧的相位和能量，分别在其下一帧的层 3 和层 4 传送。解码器会试图从已知的类别标识中(包括前帧的类别标识)恢复出丢帧的模式标识，从而根据丢帧的类别模式，结合其相位和能量信息重构出音频波形。

对于第一类方法：虽然 AMR-WB+根据音频帧特性分别采用 ACELP、TCX256、TCX512 和 TCX1024 四种模式编码，分别用 2, 2, 4, 8 比特表示编码模式信息，并且在错误隐藏时利用编码模式信息，推断或估计出超帧（由 4 帧 1024 个采样点组成）的编码模式，从而实现一定的错误隐藏功能，但是标示出的仅仅是音频编码的编码模式，并不能根据音频帧的内容采用策略进行丢帧重构，因此不能实现高效的错误隐藏。

对于第二类方法：目前这类错误隐藏技术是针对语音帧设计的，在处理其他类型的音频帧时效果不好。对于音乐和自然声音等的分类检测、特别是它们在丢包情况下，如何重构丢包信息，使得音频通信也能够容忍较高的丢包率，目前尚没有有效的方法。

- 5 综上所述，目前的音频流错误隐藏技术尚不能实现对音频帧的高效错误隐藏，从而使得音频通信过程中对于丢包率的要求无法降低。

## 发明内容

有鉴于此，本发明实施例提供音频流错误隐藏的方法，能够实现对音频流的高效错误隐藏。

- 10 本发明实施例还提供音频流错误隐藏的装置和系统，应用该装置和系统，能够实现对音频流的高效错误隐藏。

为实现上述发明目的，采用如下的技术方案：

一种音频流错误隐藏的发送方法，包括：

对发送的音频帧按照内容进行分类，得到该音频帧的类型信息；

- 15 将音频帧的类型信息与该音频帧的编码结果封装打包发送出去。

一种音频流错误隐藏的接收方法，包括：

当发生丢帧时，对于丢失的音频帧，确定其按照内容进行分类时得到的该音频帧的类型信息；

- 20 根据丢失的音频帧的类型信息，采用相应的错误恢复策略进行音频帧重构。

一种音频流错误隐藏的方法，包括：

对发送的音频帧按照内容进行分类，得到该音频帧的类型信息；

将音频帧的类型信息与该音频帧的编码结果封装打包发送给接收端；

当发生丢帧时，对于丢失的音频帧，接收端确定其按照内容进行分类时得到的该音频帧的类型信息；

根据丢失的音频帧的类型信息，采用相应的错误恢复策略进行音频帧重构。

5 一种音频流错误隐藏的发射机，包括音频编码器模块、帧封装模块和音频帧分类器模块；

所述音频帧分类器模块，用于对发送的音频帧按照内容进行分类，得到该音频帧的类型信息，并将该类型信息发送给所述帧封装模块；

10 所述帧封装模块，用于接收所述音频帧分类器模块发送的音频帧的类型信息和所述音频编码器模块发送的音频帧的编码结果，将该音频帧的类型信息和音频帧的编码结果封装打包发送出去。

一种音频流错误隐藏的接收机，包括帧类型判别模块和错误隐藏模块，

15 所述帧类型判别模块，用于确定丢失的音频帧按照内容进行分类时得到的该音频帧的类型信息，并将该类型信息发送给所述错误隐藏模块；

所述错误隐藏模块，用于根据接收到的丢失的音频帧的类型信息，采用相应的错误恢复策略进行音频帧重构。

一种音频流错误隐藏系统，包括：发射机和接收机；

20 所述发射机，用于对发送的音频帧按照内容进行分类，得到该音频帧的类型信息，并将音频帧的类型信息与该音频帧的编码结果封装打包发送给所述接收机；

25 所述接收机，用于在发生丢帧时，确定丢失的音频帧按照内容进行分类时得到的类型信息，并根据该类型信息，采用相应的错误恢复策略进行音频帧重构。



由上述技术方案可见，本发明实施例在发送端，根据音频帧内容进行分类，并将音频帧的类型信息和音频帧的编码结果一同发送出去；在接收端，当发生丢包时，根据丢失的音频帧按照内容进行分类时得到的不同类型，采取相应的错误隐藏策略重构音频信号。可见，本发明实施例所述的错误隐藏方式使得对丢失帧的重构具有更强的针对性，能够自适应地重构音频帧，以达到更好的补偿效果，为收端用户带来更好的主观听觉感受，同时能改善音频帧信号的可分辨性，使得音频通信能够容忍更高的包丢失率。

### 附图简要说明

- 10 图 1 为音频流错误隐藏框图；  
图 2 为本发明实施例中音频流错误隐藏的方法总体流程图；  
图 3 为本发明实施例中音频流错误隐藏的系统总体结构图；  
图 4 为本发明实施例中音频流错误隐藏的发射机总体结构图；  
图 5 为本发明实施例中音频流错误隐藏的接收机总体结构图；  
15 图 6 为本发明实施例中音频流错误隐藏的发送方法具体流程图；  
图 7 为本发明实施例中对音频帧分类的示意图；  
图 8 为本发明实施例中音频流错误隐藏的接收方法具体流程图；  
图 9 为本发明实施例中音频流错误隐藏的发射机具体结构图；  
图 10 为本发明实施例中音频流错误隐藏的接收机具体结构图。

### 20 实施本发明的方式

为使本发明的目的、技术手段和优点更加清楚明白，以下结合附图并举实施例，说明本发明的具体实施方式。

图 2 为本发明实施例中音频流错误隐藏的方法总体流程图。如图 2

所示，该方法包括：

步骤 201，对发送的音频帧按照内容进行分类，得到该音频帧的类型信息。

5 步骤 202，将音频帧的类型信息与该音频帧的编码结果封装打包发送出去。

步骤 203，当发生丢帧时，对于丢失的音频帧，确定其按照内容进行分类时得到的该音频帧的类型信息；

步骤 204，根据丢失的音频帧的类型信息，采用相应的错误恢复策略进行音频帧重构。

10 其中，步骤 201~202 构成了音频流错误隐藏的发送方法总体流程；步骤 203~204 构成了音频流错误隐藏接收方法总体流程。

图 3 为本发明实施例中音频流错误隐藏的系统总体结构图。如图 3 所示，该系统包括发射机 301 和接收机 302。在该系统中，发射机 301，用于对发送的音频帧按照内容进行分类，得到该音频帧的类型信息，并将音频帧的类型信息与该音频帧的编码结果封装打包发送给接收机 15 302；接收机 302，用于在发生丢帧时，确定丢失的音频帧按照内容进行分类时得到的类型信息，并根据该类型信息，采用相应的错误恢复策略进行音频帧重构。该系统中的发射机和接收机可以分别采用下面图 4 和图 5 所示的发射机 400 和接收机 500 的具体结构。

20 图 4 为本发明中音频流错误隐藏的发射机总体结构图。如图 4 所示，该发射机 400 包括音频编码器模块 410、音频帧分类器模块 420 和帧封装模块 430。

在该发射机 400 中，音频编码器模块 410，用于将发送的音频帧进行编码，并将编码结果发送给帧封装模块 430。音频帧分类器，用于对 25 发送的音频帧按照内容进行分类，得到该音频帧的类型信息，并将该类

型信息发送给帧封装模块 430。帧封装模块 430，用于接收音频编码器模块 410 发送的音频帧的编码结果和音频帧分类器模块 420 发送的音频帧的类型信息，将该音频帧的类型信息和音频帧的编码结果封装打包发送出去。在音频编码器 410 中对音频帧进行编码时，可以根据音频帧分类器发送的该音频帧的类型信息，对音频帧采用不同的编码方式，或者直接对所有的编码帧采用相同的编码方式。

与发射机相应地，图 5 为本发明中音频流错误隐藏的接收机总体结构图。该接收机包括：帧类型判别模块 510 和错误隐藏模块 520。

在该接收机中，帧类型判别模块 510，用于丢失的音频帧按照内容进行分类时得到的该音频帧的类型信息，并将该类型信息发送给错误隐藏模块 520。错误隐藏模块 520，用于根据接收到的丢失的音频帧的类型信息，采用相应的错误恢复策略进行音频帧重构。

由上述可见，本发明在发送端对音频帧按内容进行分类得到音频帧的类型信息发送给接收端，接收端在发生丢帧时，根据丢失的音频帧的类型信息，采用不同的错误恢复策略进行音频帧重构，以高效进行错误隐藏。

图 6 为本发明实施例中音频流错误隐藏的方法具体流程图。如图 6 所示，该方法包括：

步骤 601，将音频信号分成等间隔的音频帧。

本步骤中，音频帧的帧长根据编码协议而定。

步骤 602，对音频帧的内容及特性进行分析，得到音频帧的类型信息。

本步骤中，将音频帧分为语音信号帧、噪音信号帧、静音信号帧、乐音信号帧等类型，然后还可以再对每种类型进一步细分，比如，语音信号帧还可进一步分为 voiced, unvoiced, voiced transition, unvoiced

transition, onset 等类型, 乐音信号帧可以简单地根据信号的稳定特性分为稳定乐音帧 (steady)、过渡乐音帧 (transition) 等类型。

步骤 603, 对发送的音频帧进行编码压缩。

本步骤中, 可以对整个音频信号采用同样的编码方法, 或者也可以  
5 根据音频帧的类型不同, 采用不同的编码方法。

步骤 604, 将音频帧的类型和编码压缩的结果封装打包, 并发送出去。

本步骤中, 在封装打包时, 音频帧的类型信息可以在本帧或下一帧的帧头标识出来。

10 至此, 本实施例中音频流错误隐藏的发送流程结束。

在上述流程中, 步骤 602 中进行音频帧分类时, 可以采用图 7 所示的方法进行。参见图 7, 首先利用 VAD 检测该音频帧是否为噪声信号帧。若是噪声信号帧, 则对该音频帧进行频谱能量分析, 若是非噪声信号帧, 则对该音频帧进行频谱稳定性分析。

15 根据频谱能量分析的结果, 将该音频帧划分为静音信号帧和噪音信号帧, 然后可以进一步对该静音信号帧或噪音信号帧做细化的分类, 得到该音频帧的类型信息。

根据频谱稳定性分析的结果, 将该音频帧划分为语音信号帧和乐音信号帧, 然后可以进一步对该语音信号帧或乐音信号帧做细化的分类,  
20 如对于语音信号帧可以细化为 voiced, unvoiced, voiced transition 等, 对于乐音信号帧可以细化为稳定乐音帧、过渡乐音帧等。

与上述发送方法中对音频帧进行的分类, 在接收端, 本实施例中采用图 8 所示的方法进行接收。如图 8 所示, 该接收方法包括:

步骤 801, 对音频信号进行丢帧检测, 若发生丢帧, 则执行步骤 804  
25 及其后续步骤, 否则执行步骤 802 及其后续步骤。

本步骤中，根据音频帧中携带的帧序号判定是否出现音频帧的丢失。

步骤 802，对音频帧的类型进行检测并记录。

当发生丢帧时，可以利用本步骤中记录的音频帧类型确定丢失的音频帧的类型信息。

步骤 803，对音频帧进行解码，并输出解码结果，结束本流程。

本步骤中，根据发送端对音频帧的编码方法，采用相应的解码方法进行解码。

步骤 804，确定丢失的音频帧按照内容进行分类时得到的类型信息。

本步骤中，若音频帧的类型信息是携带在本音频帧中传送，则接收端提取历史数据，根据正确接收帧的类型信息推断当前丢失帧的类型；若音频帧的类型信息是携带在其他正确接收音频帧中传送的，则接收端直接在相应正确接收音频帧中提取当前丢失帧的类型信息即可。

步骤 805，根据丢失的音频帧的类型，自适应地采用相应的错误恢复策略重构音频帧，并输出重构的结果，结束本流程。

本步骤中，可以根据丢失的音频帧的类型，选用针对该类型最合适的错误恢复策略重构音频帧。如，稳定语音帧与其前一帧非常相似，用帧复制策略就能取得很好的隐藏效果，过渡帧则需要考虑前后帧的状态来确定隐藏策略等。

至此，音频流错误隐藏的接收方法结束。

在本实施例的发送方法中，对音频帧进行分类时采用的是图 7 所示的方式进行分类，当然也可以采用其他基于内容的音频帧分类方式，只要能够达到对音频帧根据内容分类的目的即可。

由上述可见，发送端和接收端相配合，就能够利用对音频帧的分类，来高效实现错误隐藏，大大提高实时音频通信对于丢包率的容忍度。

上述为本实施例中提供的音频流错误隐藏的发送和接收方法的具体实施方式。由该两种实施方式相互配合，即可以构成本发明中音频流错误隐藏的方法具体实施方式。另外，本实施例还提供了相应的音频流错误隐藏的发射机和接收机的具体实施方式。

5 图 9 为本发明实施例中音频流错误隐藏的发射机具体结构图。如图 9 所示，该发射机 900 包括：音频编码器模块 910、音频帧分类器模块 920、帧封装模块 930 和音频帧划分模块 940。

在该发射机 900 中，音频帧划分模块 940，用于根据不同的编码协议，将音频信号划分为等间隔的音频帧，并将音频帧发送给音频编码器  
10 模块 910 和音频帧分类器模块 920。

音频编码器模块 910，用于对音频帧进行编码，并将编码结果发送给帧封装模块 930。音频帧分类器模块 920，用于对音频帧按照内容进行分类，其具体分类方式可以采用图 7 所示的方式，并将音频帧的类型信息发送给帧封装模块 930。

15 帧封装模块 930，用于接收音频编码器模块 910 发送的音频帧编码结果和音频帧分类器模块 920 发送的音频帧类型信息，并将类型信息和音频帧的编码结果封装打包，并发送出去。在进行封装打包时，可以将音频帧的类型信息封装在本音频帧或下一音频帧中，具体可以位于帧头的部分。

20 在音频编码器 910 中对音频帧进行编码时，可以根据音频帧分类器发送的该音频帧的类型信息，对音频帧采用不同的编码方式，或者直接对所有的编码帧采用相同的编码方式。

图 10 为本发明实施例中音频流错误隐藏的接收机具体结构图。如图 10 所示，该接收机 1000 包括帧类型判别模块 1010、错误隐藏模块  
25 1020、差错检测模块 1030 和音频解码器模块 1040。其中，帧类型判别

模块 1010 包括判别子模块 1011 和存储子模块 1012; 错误隐藏模块 1020 包括策略判决子模块 1021 和错误隐藏子模块 1022。

在该接收机 1000 中, 差错检测模块 1030, 用于从信道上接收音频帧, 将接收到的音频帧发送给帧类型判别模块 1010 中的判别子模块 1011, 并检测是否出现丢帧, 若出现丢帧, 则通知帧类型判别模块 1010 中的判别子模块 1011。

在帧类型判别模块 1010 中, 在确定音频帧按照内容进行分类得到的类型时, 若音频帧的类型信息在正确接收到的音频帧中携带, 则直接将该类型信息提取出来存储到存储子模块 1012 中; 若音频帧的类型信息在丢失的音频帧中携带, 则根据前后帧的类型推断该丢失的音频帧按照内容进行分类时得到的类型信息。

在错误隐藏模块 1020 中, 策略判决子模块 1021, 用于接收判别子模块 1011 发送的丢失帧的类型信息, 并根据该类型信息, 判定采用的错误恢复策略, 并将结果发送给错误隐藏子模块 1022。错误隐藏子模块 1022, 用于根据策略判决子模块 1021 发送的错误恢复策略判决结果, 对丢失的音频帧进行重构。

音频帧解码器模块 1040, 用于对接收到的音频帧进行解码, 并输出解码结果。

在本实施例的发射机 900 中, 音频帧分类器模块 920 采用的是图 7 的方式对音频帧进行分类, 当然也可以采用其他基于内容的音频帧分类方式, 这里就不再赘述。在接收机 1000 中, 将帧类型判决模块 1010 细化为判决子模块 1011 和存储子模块 1012, 分别进行帧类型的判决和存储; 错误隐藏模块 1020 细化为策略判决子模块 1021 和错误隐藏子模块 1022, 分别进行策略判决和错误隐藏。其中, 错误隐藏子模块 1022 还可以进一步划分为多种不同类型的错误隐藏单元, 如噪音错误隐藏单

元、话音错误隐藏单元等，用于处理不同类型的音频帧的错误隐藏。

本发明音频流错误隐藏系统的实施方式可以为：利用上述图 9 和图 10 所示的发射机 900 和接收机 1000 作为音频流错误隐藏系统中发射机和接收机的具体实施方式，并且，将发射机 900 中帧封装模块 930 输出的音频帧发送给接收机 1000 中的差错检测模块 1030。这样便可以构成本发明的音频流错误隐藏系统的一种实施方式。

由上述本发明方法、装置和系统的具体实施方式可见，采用本发明的技术方案，使得对丢失帧的重构具有更强的针对性，能够自适应地重构音频帧，以达到更好的补偿效果，为收端用户带来更好的主观听觉感受，同时能改善音频帧信号的可分辨性，使得音频通信能够容忍更高的包丢失率。

以上仅为本发明的较佳实施例而已，并非用于限定本发明的保护范围。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、一种音频流错误隐藏的发送方法，其特征在于，该方法包括：  
对发送的音频帧按照内容进行分类，得到该音频帧的类型信息；  
将音频帧的类型信息与该音频帧的编码结果封装打包发送出去。
- 5 2、根据权利要求1所述的方法，其特征在于，所述对发送的音频帧按照内容进行分类包括：  
利用语音激活检测器 VAD 检测音频帧是否为噪声信号帧，若是，则对噪声信号帧进行频谱能量分析，确定该音频帧为静音信号帧或噪音信号帧；否则，对非噪声信号帧进行频谱稳定性分析，确定该音频帧为语  
10 音信号帧或乐音信号帧。
- 3、根据权利要求2所述的方法，其特征在于，对所述静音信号帧、噪音信号帧、语音信号帧或乐音信号帧作进一步细致划分。
- 4、根据权利要求1所述的方法，其特征在于，所述将音频帧的类型信息与该音频帧的编码结果封装打包为：将音频帧的类型信息打包到音  
15 频帧的帧头中。
- 5、根据权利要求4所述的方法，其特征在于，所述音频帧的帧头为：音频帧的类型信息所代表的音频帧的帧头或音频帧的类型信息所代表的音频帧的下一帧的帧头。
- 6、根据权利要求1所述的方法，其特征在于，所述音频帧的编码结果  
20 为：所有音频帧采用相同的编码方法得到的编码结果，或者，根据音频帧的类型不同，采用不同的编码方法得到的编码结果。
- 7、一种音频流错误隐藏接收方法，其特征在于，该方法包括：  
当发生丢帧时，对于丢失的音频帧，确定其按照内容进行分类时得到的该音频帧的类型信息；  
25 根据丢失的音频帧的类型信息，采用相应的错误恢复策略进行音频

帧重构。

8、根据权利要求 7 所述的接收方法，其特征在于，所述确定其按照内容进行分类时得到的该音频帧的类型信息为：

当丢失的音频帧的类型信息位于正确接收的音频帧中时，直接从该正确接收的音频帧中提取丢失的音频帧的类型信息；

当丢失的音频帧的类型信息位于该丢失的音频帧中时，根据正确接收帧的类型信息推断丢失的音频帧的类型信息。

9、根据权利要求 7 或 8 所述的接收方法，其特征在于，在对丢失的音频帧执行所述确定其按照内容进行分类时得到的该音频帧的类型信息的操作前，该方法进一步包括：对接收到的音频信号检测是否有丢帧，若是，则执行所述确定其按照内容进行分类时得到的该音频帧的类型信息的操作，否则，对正确接收的音频帧进行解码，提取该帧中携带的音频帧的类型信息并存储。

10、一种音频流错误隐藏的方法，其特征在于，该方法包括：  
对发送的音频帧按照内容进行分类，得到该音频帧的类型信息；  
将音频帧的类型信息与该音频帧的编码结果封装打包发送给接收端；

当发生丢帧时，对于丢失的音频帧，接收端确定其按照内容进行分类时得到的该音频帧的类型信息；

根据丢失的音频帧的类型信息，采用相应的错误恢复策略进行音频帧重构。

11、根据权利要求 10 所述的方法，其特征在于，所述对发送的音频帧按照内容进行分类包括：

利用语音激活检测器 VAD 检测音频帧是否为噪声信号帧，若是，则对噪声信号帧进行频谱能量分析，确定该音频帧为静音信号帧或噪音信

号帧；否则，对非噪声信号帧进行频谱稳定性分析，确定该音频帧为语音信号帧或乐音信号帧。

12、根据权利要求 10 所述的方法，其特征在于，所述将音频帧的类型信息与该音频帧的编码结果封装打包为：将音频帧的类型信息打包到  
5 音频帧的帧头中。

13、根据权利要求 12 所述的方法，其特征在于，所述音频帧的帧头为：音频帧的类型信息所代表的音频帧的帧头或音频帧的类型信息所代表的音频帧的下一帧的帧头。

14、根据权利要求 10 所述的方法，其特征在于，所述确定其按照内  
10 容进行分类时得到的该音频帧的类型信息为：

当丢失的音频帧的类型信息位于正确接收的音频帧中时，直接从该正确接收的音频帧中提取丢失的音频帧的类型信息；

当丢失的音频帧的类型信息位于该丢失的音频帧中时，根据正确接收帧的类型信息推断丢失的音频帧的类型信息。

15 15、根据权利要求 10 到 14 中任意一项所述的方法，其特征在于，在执行所述将音频帧的类型信息与该音频帧的编码结果封装打包发送给接收端的操作后、接收端对丢失的音频帧执行所述确定其按照内容进行分类时得到的该音频帧的类型信息的操作前，该方法进一步包括：对接收到的音频信号检测是否有丢帧，若是，则执行所述确定其按照内容  
20 进行分类时得到的该音频帧的类型信息的操作，否则，对正确接收的音频帧进行解码，提取该帧中携带的音频帧的类型信息并存储。

16、一种音频流错误隐藏的发射机，包括音频编码器模块和帧封装模块，其特征在于，该发射机还包括音频帧分类器模块；

所述音频帧分类器模块，用于对发送的音频帧按照内容进行分类，  
25 得到该音频帧的类型信息，并将该类型信息发送给所述帧封装模块；

所述帧封装模块，用于接收所述音频帧分类器模块发送的音频帧的类型信息和所述音频编码器模块发送的音频帧的编码结果，将该音频帧的类型信息和音频帧的编码结果封装打包发送出去。

17、一种音频流错误隐藏的接收机，其特征在于，该接收机包括帧  
5 类型判别模块和错误隐藏模块，

所述帧类型判别模块，用于确定丢失的音频帧按照内容进行分类时得到的该音频帧的类型信息，并将该类型信息发送给所述错误隐藏模块；

所述错误隐藏模块，用于根据接收到的丢失的音频帧的类型信息，  
10 采用相应的错误恢复策略进行音频帧重构。

18、根据权利要求 17 所述的接收机，其特征在于，该接收机进一步包括差错检测模块和音频解码器模块，

所述差错检测模块，用于从信道上接收音频帧，将接收到的音频帧发送给所述帧类型判别模块，并检测是否出现丢帧，若出现丢帧，则通  
15 知所述帧类型判别模块；

所述帧类型判别模块，进一步用于将音频帧转发给所述音频帧解码器模块；

所述音频帧解码器模块，用于对音频帧进行解码。

19、根据权利要求 18 所述的接收机，其特征在于，所述帧类型判别  
20 模块包括判别子模块和存储子模块，

所述判别子模块，用于确定丢失的音频帧按照内容进行分类时得到的该音频帧的类型信息，并将该类型信息发送给所述存储子模块，还用于在接收到所述差错检测模块发送的丢帧通知后，将该丢失帧的类型信息发送给所述错误隐藏模块，并将接收到的音频帧转发给所述音频帧解  
25 码器模块；

所述存储子模块，用于保存音频帧的类型信息。

20、根据权利要求 18 或 19 所述的接收机，其特征在于，所述错误隐藏模块包括策略判决子模块和错误隐藏子模块，

所述策略判决子模块，用于接收所述帧类型判别模块发送的丢失帧  
5 的类型信息，并根据该类型信息，判定采用的错误恢复策略，并将结果发送给所述错误隐藏子模块；

所述错误隐藏子模块，用于根据所述策略判决子模块发送的错误恢复策略判决结果，对丢失的音频帧进行重构。

21、一种音频流错误隐藏系统，其特征在于，该系统包括：发射机  
10 和接收机；

所述发射机，用于对发送的音频帧按照内容进行分类，得到该音频帧的类型信息，并将音频帧的类型信息与该音频帧的编码结果封装打包发送给所述接收机；

所述接收机，用于在发生丢帧时，确定丢失的音频帧按照内容进行分类时得到的类型信息，并根据该类型信息，采用相应的错误恢复策略  
15 进行音频帧重构。

22、根据权利要求 21 所述的系统，其特征在于，所述发射机包括音频编码器模块、帧封装模块和音频帧分类器模块，所述接收机包括帧类型判别模块和错误隐藏模块；

所述音频帧分类器模块，用于对发送的音频帧按照内容进行分类，  
20 得到该音频帧的类型信息，并将该类型信息发送给所述帧封装模块；

所述帧封装模块，用于接收所述音频帧分类器模块发送的音频帧的类型信息和所述音频编码器模块发送的音频帧的编码结果，将该音频帧的类型信息和音频帧的编码结果封装打包发送出去；

所述帧类型判别模块，用于确定丢失的音频帧按照内容进行分类时  
25

得到的该音频帧的类型信息，并将该类型信息发送给所述错误隐藏模块；

所述错误隐藏模块，用于根据接收到的丢失的音频帧的类型信息，采用相应的错误恢复策略进行音频帧重构。

- 5 23、根据权利要求 22 所述的系统，其特征在于，所述接收机进一步包括差错检测模块和音频解码器模块，

所述差错检测模块，用于从信道上接收音频帧，将接收到的音频帧发送给所述帧类型判别模块，并检测是否出现丢帧，若出现丢帧，则通知所述帧类型判别模块；

- 10 所述帧类型判别模块，进一步用于将音频帧转发给所述音频帧解码器模块；

所述音频帧解码器模块，用于对音频帧进行解码；

所述帧封装模块，用于将该音频帧的类型信息和音频帧的编码结果封装打包发送给所述差错检测模块。

- 15 24、根据权利要求 22 或 23 所述的接收机，其特征在于，所述错误隐藏模块包括策略判决子模块和错误隐藏子模块，

所述策略判决子模块，用于接收所述帧类型判别模块发送的丢失帧的类型信息，并根据该类型信息，判定采用的错误恢复策略，并将结果发送给所述错误隐藏子模块；

- 20 所述错误隐藏子模块，用于根据所述策略判决子模块发送的错误恢复策略判决结果，对丢失的音频帧进行重构。

1/5

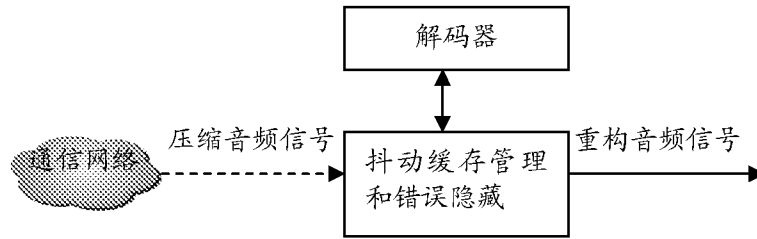


图 1

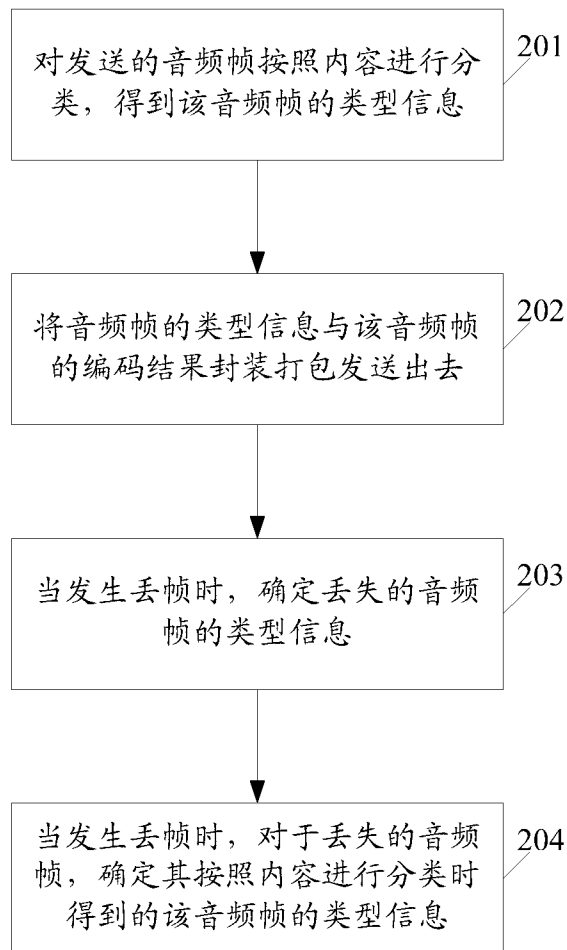


图 2

2/5

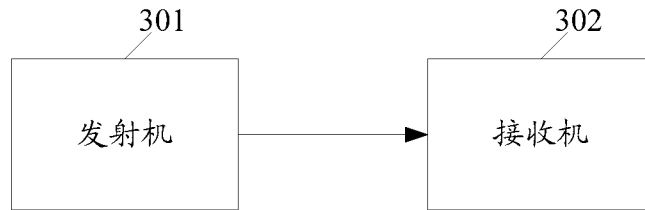


图 3

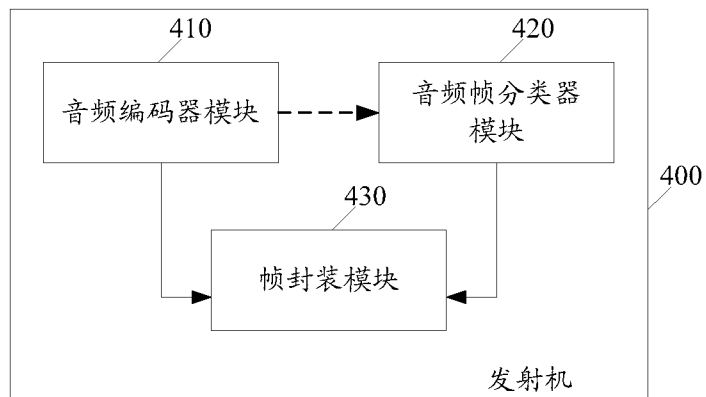


图 4

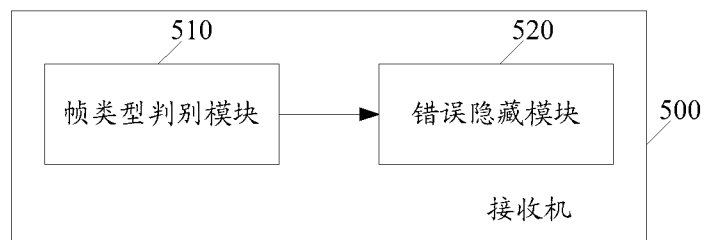


图 5



3/5

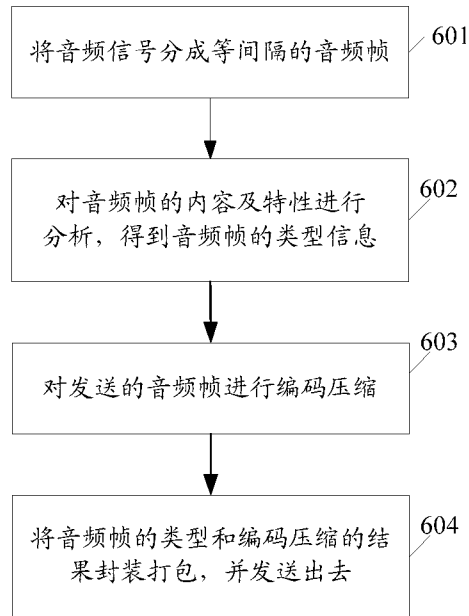


图 6

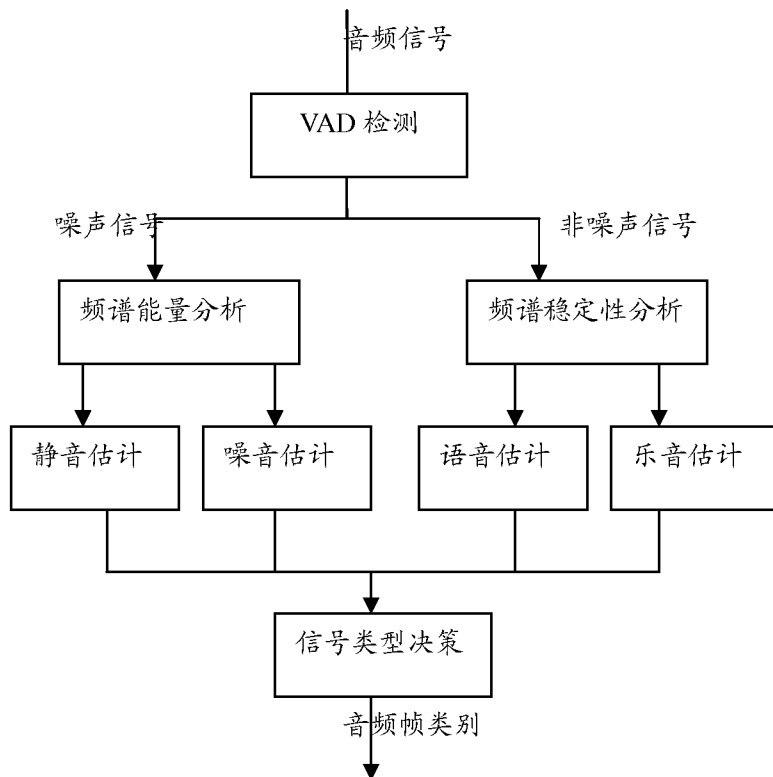


图 7

4/5

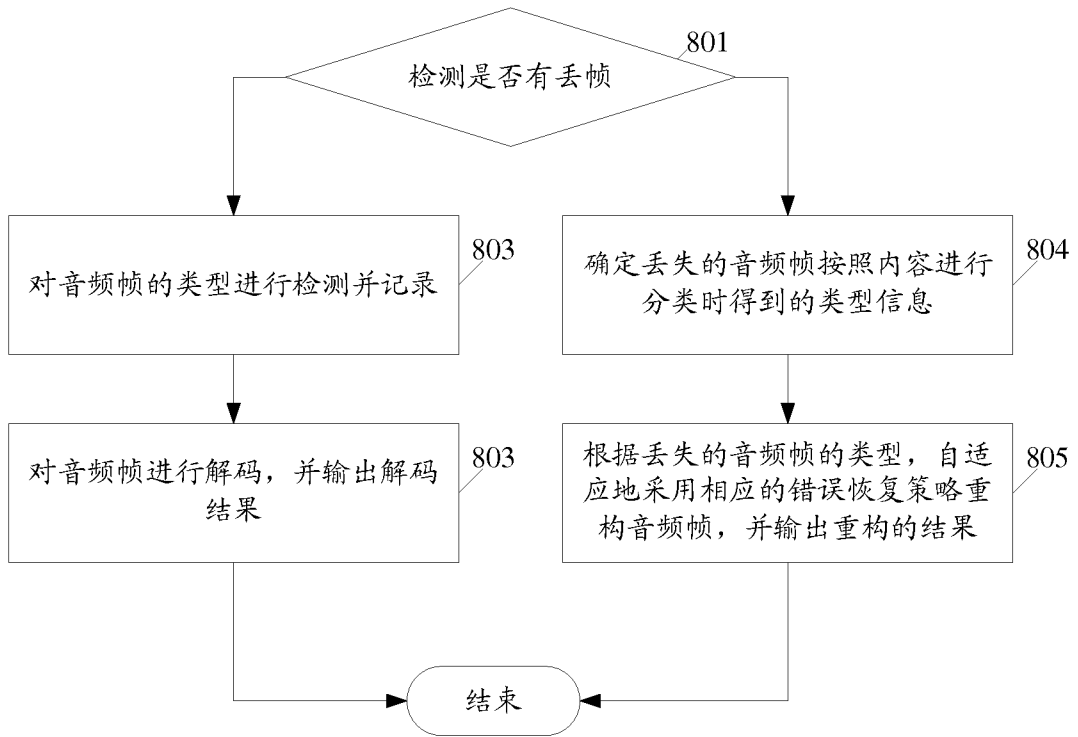


图 8

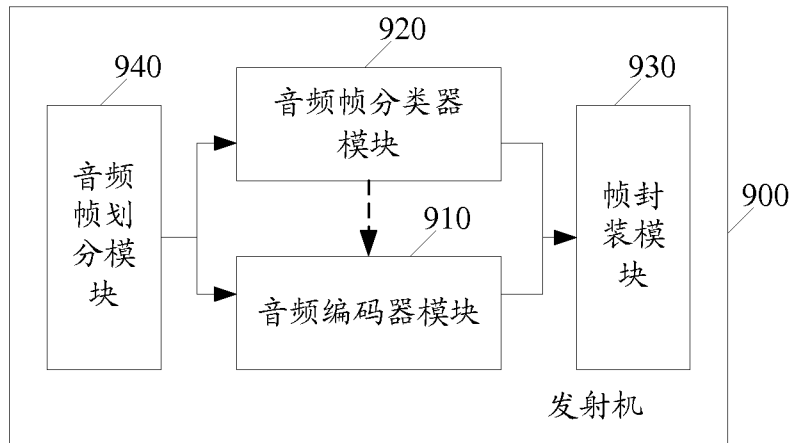


图 9

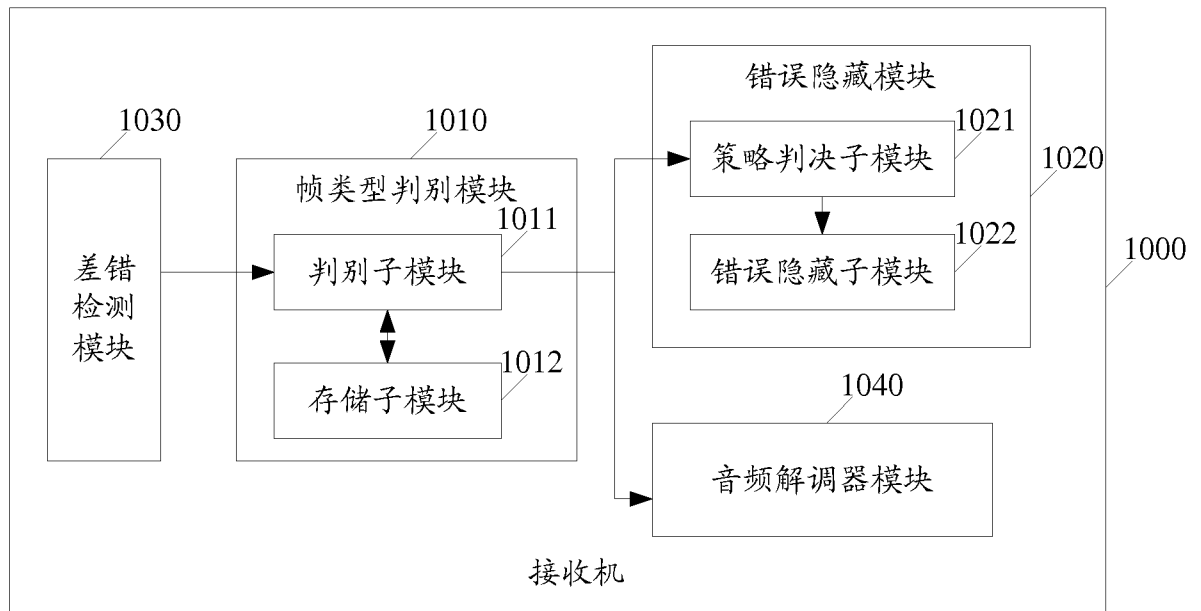


图 10

# INTERNATIONAL SEARCH REPORT

International application No.  
PCT/CN2007/070772

<b>A. CLASSIFICATION OF SUBJECT MATTER</b>  <p style="text-align: center;">H04L1/00 (2006.01) i</p> <p>According to International Patent Classification (IPC) or to both national classification and IPC</p>		
<b>B. FIELDS SEARCHED</b>  <p>Minimum documentation searched (classification system followed by classification symbols)</p> <p style="text-align: center;">IPC: H04L1; G10L19</p> <p>Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched</p> <p>Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)</p> <p style="text-align: center;">CPRS; CNKI ;WPI; EPODOC; PAJ: voice,error,content,type,noise,encod+,eras+,conceal+,encapsulat+,VAD</p>		
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US2005/0154584A1 (Jelinek et al.) 14 July 2005(14.07.2005)  paragraphs [0031]-[0234] of the description, Figs. 1-7	1-24
A	CN1589550A (MATSUSHITA ELECTRIC IND CO LTD) 02 Mar. 2005(02.03.2005)  the whole document	1-24
A	US6754188B1 (MeshNetworks, Inc.) 22 June 2004(22.06.2004)  the whole document	1-24
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> 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  "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	
"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		
Date of the actual completion of the international search 03 Dec. 2007(03.12.2007)	Date of mailing of the international search report <b>20 Dec. 2007 (20.12.2007)</b>	
Name and mailing address of the ISA/CN The State Intellectual Property Office, the P.R.China 6 Xitucheng Rd., Jimen Bridge, Haidian District, Beijing, China 100088 Facsimile No. 86-10-62019451	Authorized officer  <b>GAO Haiyan</b>  Telephone No. (86-10)62086089	

**INTERNATIONAL SEARCH REPORT**  
**Information on patent family members**

International application No.

PCT/CN2007/070772

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
US2005/0154584A1	14.07.2005	CA2388439A1	30.11.2003
		CA2483791A	11.12.2003
		WO03102921A1	11.12.2003
		AU2003233724A1	19.12.2003
		NO20045578A	22.02.2005
		EP1509903A1	02.03.2005
		BRPI0311523A	08.03.2005
		MXPA04011751A	08.06.2005
		RU2004138286A	10.06.2005
		CN1659625A	24.08.2005
		JP2005534950T	17.11.2005
		ZA200409643A	28.06.2006
		NZ536238A	30.06.2006
		KR20050005517A	13.01.2005
		INKOLNP200401665E	02.06.2006
CN1589550A	02.03.2005	WO03043277A1	22.05.2003
		JP2003218932A	31.07.2003
		EP1458145A1	15.09.2004
		US2005044471A1	24.02.2005
US6754188B1	22.06.2004	WO03030472A1	10.04.2003
		AU2002336612A1	14.04.2003

国际检索报告

国际申请号  
PCT/CN2007/070772

<p><b>A. 主题的分类</b></p> <p style="text-align: center;">H04L1/00 (2006.01) i</p> <p>按照国际专利分类表(IPC)或者同时按照国家分类和 IPC 两种分类</p>														
<p><b>B. 检索领域</b></p> <p>检索的最低限度文献(标明分类系统和分类号)</p> <p style="text-align: center;">IPC: H04L1; G10L19</p> <p>包含在检索领域中的除最低限度文献以外的检索文献</p> <p>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))</p> <p style="text-align: center;">CPRS; CNKI; WPI; EPODOC; PAJ: 声音, 错误, 内容, 类型, 噪音, 编码, 消除, 隐藏, 封装, VAD; voice, error, content, type, noise, encod+, eras+, conceal+, encapsulat+, VAD</p>														
<p><b>C. 相关文件</b></p> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="width: 10%;">类 型*</th> <th style="width: 60%;">引用文件, 必要时, 指明相关段落</th> <th style="width: 30%;">相关的权利要求</th> </tr> </thead> <tbody> <tr> <td style="text-align: center;">X</td> <td>US2005/0154584A1 (Jelinek et al.) 14.7 月 2005(14.07.2005) 说明书第[0031]-[0234]栏, 图 1-7</td> <td style="text-align: center;">1-24</td> </tr> <tr> <td style="text-align: center;">A</td> <td>CN1589550A (松下电器产业株式会社) 02.3 月 2005(02.03.2005) 全文</td> <td style="text-align: center;">1-24</td> </tr> <tr> <td style="text-align: center;">A</td> <td>US6754188B1 (MeshNetworks, Inc.) 22.6 月 2004(22.06.2004) 全文</td> <td style="text-align: center;">1-24</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	US2005/0154584A1 (Jelinek et al.) 14.7 月 2005(14.07.2005) 说明书第[0031]-[0234]栏, 图 1-7	1-24	A	CN1589550A (松下电器产业株式会社) 02.3 月 2005(02.03.2005) 全文	1-24	A	US6754188B1 (MeshNetworks, Inc.) 22.6 月 2004(22.06.2004) 全文	1-24
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求												
X	US2005/0154584A1 (Jelinek et al.) 14.7 月 2005(14.07.2005) 说明书第[0031]-[0234]栏, 图 1-7	1-24												
A	CN1589550A (松下电器产业株式会社) 02.3 月 2005(02.03.2005) 全文	1-24												
A	US6754188B1 (MeshNetworks, Inc.) 22.6 月 2004(22.06.2004) 全文	1-24												
<p><input type="checkbox"/> 其余文件在 C 栏的续页中列出。      <input checked="" type="checkbox"/> 见同族专利附件。</p>														
<p>* 引用文件的具体类型:</p> <p>“A” 认为不特别相关的表示了现有技术一般状态的文件</p> <p>“E” 在国际申请日的当天或之后公布的在先申请或专利</p> <p>“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</p> <p>“O” 涉及口头公开、使用、展览或其他方式公开的文件</p> <p>“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</p>		<p>“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</p> <p>“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</p> <p>“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</p> <p>“&amp;” 同族专利的文件</p>												
<p>国际检索实际完成的日期 03.12 月 2007(03.12.2007)</p>		<p>国际检索报告邮寄日期 20.12 月 2007 (20.12.2007)</p>												
<p>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</p>		<p>授权官员  高海燕 电话号码: (86-10) 62086089</p>												

国际检索报告

关于同族专利的信息

国际申请号

PCT/CN2007/070772

检索报告中引用的 专利文件	公布日期	同族专利	公布日期
US2005/0154584A1	14.07.2005	CA2388439A1	30.11.2003
		CA2483791A	11.12.2003
		WO03102921A1	11.12.2003
		AU2003233724A1	19.12.2003
		NO20045578A	22.02.2005
		EP1509903A1	02.03.2005
		BRPI0311523A	08.03.2005
		MXPA04011751A	08.06.2005
		RU2004138286A	10.06.2005
		CN1659625A	24.08.2005
		JP2005534950T	17.11.2005
		ZA200409643A	28.06.2006
		NZ536238A	30.06.2006
		KR20050005517A	13.01.2005
INKOLNP200401665E	02.06.2006		
CN1589550A	02.03.2005	WO03043277A1	22.05.2003
		JP2003218932A	31.07.2003
		EP1458145A1	15.09.2004
US6754188B1	22.06.2004	US2005044471A1	24.02.2005
		WO03030472A1	10.04.2003
		AU2002336612A1	14.04.2003