(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
2010 年 1 月 7 日 (07.01.2010)

PCT

(10) 国际公布号
WO 2010/000161 A1

(51) 国际专利分类号:
G10L 13/00 (2006.01)

(21) 国际申请号: PCT/CN2009/071931

(22) 国际申请日: 2009 年 5 月 22 日 (22.05.2009)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
200810068262.6 2008 年 6 月 30 日 (30.06.2008) CN

(71) 申请人 (对除美国外的所有指定国): **腾讯科技（深圳）有限公司 (TENCENT TECHNOLOGY (SHENZHEN) COMPANY LIMITED)** [CN/CN]; 中国广东省深圳市福田区振兴路赛格科技园 2 栋东 4 楼, Guangdong 518044 (CN)。

(72) 发明人; 及

(75) 发明人/申请人 (仅对美国): **李大龙 (LI, Dalong)** [CN/CN]; 中国广东省深圳市福田区振兴路赛格科技园 2 栋东 4 楼, Guangdong 518044 (CN)。 **郑全战 (ZHENG, Quanzhan)** [CN/CN]; 中国广东省深圳市福田区振兴路赛格科技园 2 栋东 4 楼, Guangdong 518044 (CN)。 **盛馥钟 (SHENG, Fuzhong)** [CN/CN]; 中国广东省深圳市福田区振兴路赛格科技园 2 栋东 4 楼, Guangdong 518044 (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, AO, 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, 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, ST, SV, SY, 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, LS, MW, MZ, NA, 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, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO,

[见续页]

(54) Title: VOICE CONVERSATION METHOD AND APPARATUS BASED ON INSTANT COMMUNICATION SYSTEM

(54) 发明名称: 一种基于即时通讯系统的语音通话方法及装置

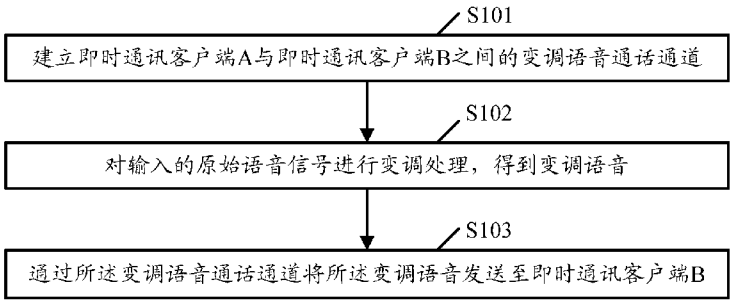


图 1 / Fig. 1

S101 ESTABLISHING A VOICE CHANGING CONVERSATION TUNNEL BETWEEN THE INSTANT COMMUNICATION CLIENT A AND THE INSTANT COMMUNICATION CLIENT B

S102 PERFORMING VOICE CHANGING PROCESS TO THE ORIGINAL VOICE SIGNAL INPUTTED, THUS GETTING THE CHANGED VOICE

S103 TRANSMITTING SAID CHANGED VOICE TO THE INSTANT COMMUNICATION CLIENT B THROUGH SAID VOICE CHANGING CONVERSATION TUNNEL

(57) **Abstract**: A voice conversation method and apparatus based on instant communication system are provided, wherein, said method includes the following steps: step a, establishing a voice changing conversation tunnel between the instant communication client A and the instant communication client B (S101); step b, performing voice changing process to the original voice information inputted, thus getting the changed voice (S102); transmitting said changed voice to the instant communication client B through said voice changing conversation tunnel (S103). The present invention performs the process of voice changing to the voice signal collected in the instant communication system at first, which realizes the voice changing conversation based on the instant communication system.

[见续页]

WO 2010/000161 A1

SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, 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))。

(57) 摘要:

一种基于即时通讯系统的语音通话方法及装置，其中，所述方法包括下述步骤：步骤 a，建立即时通讯客户端 A 与即时通讯客户端 B 之间的变调语音通话通道（S101）；步骤 b，对输入的原始语音信息进行变调处理，得到变调语音（S102）；通过所述变调语音通话通道将所述变调语音发送至即时通讯客户端 B（S103）。本发明将即时通讯系统中采集到的语音信号先进行语音变调处理，实现了基于即时通讯系统的变调语音通话。

一种基于即时通讯系统的语音通话方法及装置

技术领域

本发明属于通信领域，尤其涉及一种基于即时通讯系统的语音通话方法及装置。

发明背景

随着即时通讯技术的发展，即时通讯系统除具备基本的即时通讯功能外，还具有其他诸多附加功能，如语音通话功能。使用即时通讯系统进行语音通话已经成为一般社会大众所使用的沟通工具之一，但是现有的语音通话双方只能采用自己原来的声音进行通话，不能改变通话者的原始语音，功能比较单一，不能隐藏通话方的身份，缺乏新颖性、娱乐性，不能满足用户的个性化需求。

现在还没有基于即时通讯系统的变调语音通话方法。

发明内容

本发明实施例提供一种基于即时通讯系统的变调语音通话方法，旨在解决现在还没有基于即时通讯系统的变调语音通话方法的问题。

本发明实施例是这样实现的：

一种基于即时通讯系统的语音通话方法，所述方法包括下述步骤：

建立至少两个即时通讯客户端之间的变调语音通话通道；

对输入的原始语音信息进行变调处理，得到变调语音；通过所述变调语音通话通道将所述变调语音发送至即时通讯客户端 B。

本发明实施例的另一目的在于提供一种基于即时通讯系统的语音通话装置，所述装置包括：

请求发送单元，用于建立变调语音通话通道；

语音采集单元，用于采集输入的原始语音信息；

变调处理单元，用于将所述语音采集单元采集的原始语音信息进行变调处理，得到变调语音；

语音发送单元，用于通过所述请求发送单元建立的变调语音通话通道发送所述变调处理单元得到的变调语音。

在本发明实施例中，将即时通讯系统中采集到的语音信号先进行语音变调处理，实现了基于即时通讯系统的变调语音通话，为基于即时通讯场合下的语音通信带来极大的娱乐效果，给传统即时通讯服务带来新的增值服务增长点，增加用户对即时通讯产品的依赖性，从而提升产品竞争力，并且为语音通话用户提供崭新的业务体验。

附图简要说明

图 1 为本发明实施例提供的方法的基本流程图；

图 2 为本发明实施例提供的方法的详细流程图；

图 3 为本发明实施例提供的方法的另一详细流程图；

图 4 为本发明实施例提供的即时通讯客户端 B 接收到即时通讯客户端 A 发送的变调语音通话数据的处理流程图；

图 5 为本发明实施例提供的装置的基本结构图；

图 6 为本发明实施例提供的装置的详细结构图。

实施本发明的方式

为使本发明的目的、技术方案和优点更加清楚明白，以下举实施例，并参照附图，对本发明进一步详细说明。

在本发明实施例中，可以建立至少两个即时通讯客户端之间的变调

语音通话信道，比如，可以在即时通讯客户端 A、即时通信客户端 B 以及即时通讯客户端 C 之间建立变调语音通话信道，本发明实施例为便于描述，下述仅以在即时通讯客户端 A 和即时通信客户端 B 之间建立变调语音通话信道为例，其他情况类似，这里不再赘述。如此，即时通讯客户端 A 向即时通讯客户端 B 发送语音变调请求，并与即时通讯客户端 B 之间建立了变调语音通话通道后，对采集的原始语音进行变调处理，得到与原始语音对应的变调语音，并将该变调语音通过建立的变调语音通话通道发送至即时通讯客户端 B，从而实现即时通讯系统中各即时通讯客户端之间的变调语音通话。

参见图 1，图 1 为本发明实施例提供的基本流程图。如图 1 所示，该实施例以即时通信客户端 A 与即时通讯客户端 B 之间建立变调语音通话信道为例，该流程可包括以下步骤:

步骤 S101，建立即时通讯客户端 A 与即时通讯客户端 B 之间的变调语音通话通道。

步骤 S102，对输入的原始语音信号进行变调处理，得到变调语音。

步骤 S103，通过所述变调语音通话通道将所述变调语音发送至即时通讯客户端 B。

需要说明的是，上述即时通讯客户端 A 和即时通讯客户端 B 具体实现时可有多种实现形式，比如可以为 web 形式的终端，或者为无线形式的终端，本发明实施例并不具体限定。

还需要说明的是，上述步骤 S102 和步骤 S103 的操作可以由即时通讯客户端 A 执行，比如，即时通讯客户端 A 直接将输入的原始语音信息进行变调处理，得到变调语音；在变调语音通话通道上通过服务器中转或者 P2P 方式中转该变调语音至即时通讯客户端 B；或者可以由预设的变调处理设备比如服务器执行，比如，预设的服务器接收即时通讯客户

端 A 发送的原始语音信息，之后，对该原始语音信息进行变调处理，得到变调语音；通过所述变调语音通话通道将所述变调语音发送至即时通讯客户端 B。具体如何实现，本发明实施例并不具体限定。为便于描述，下述以在两个客户端之间执行语音通话操作为例。

如此，实现了本发明实施例提供的基于即时通讯系统的语音通话基本流程。

上述只是简单描述了本发明实施例提供的流程，下面结合具体实施例对本发明实施例提供的方法进行详细描述。

参见图 2，图 2 为本发明实施例提供的详细流程图，详述如下：

1、即时通讯客户端 A 向即时通讯客户端 B 发送变调语音通话请求。

2、即时通讯客户端 B 在接收到即时通讯客户端 A 发送的变调语音通话请求后，对该变调语音通话请求作出应答，并将应答信息返回至即时通讯客户端 A。当即时通讯客户端 A 接收到即时通讯客户端 B 返回的变调语音通话应答后，建立与即时通讯客户端 B 之间的变调语音通话通道。

为了成功建立通话通道，即时通讯客户端 A 和即时通讯客户端 B 在即时通讯服务器的协调下，建立变调语音通话通道。当然，即时通讯客户端 A 可以透明或者非透明地向即时通讯客户端 B 发送变调语音通话请求。如果即时通讯客户端 A 透明地向即时通讯客户端 B 发送变调语音通话请求，此过程即时通讯客户端 B 界面上不用显示。

3、即时通讯客户端 A 对采集的原始语音进行变调处理，得到与原始语音对应的变调语音。

在本发明实施例中，提供了多种语音变调方式，如改变语音的音调、性别变声（男声变女声、女声变男声）、年龄变声（青少年的声音变老年人的声音）、将用户原始话音变为某著名人物的声音、在用户声音上

加入背景声音（严格意义上，在用户声音上加入背景声音并不属于语音变调处理，而属于混音技术，但本发明所定义的变调语音通话包含此类应用）等。

其中对采集的原始语音进行变调处理，得到与原始语音对应的变调语音的具体过程如下：

A、采集用户输入的语音信息，并将采集的语音信息处理成计算机可以识别和处理的数字语音信息；

B、对该数字语音信息进行变调处理，得到与该数字语音信息对应的变调语音。

在本实施例中，语音变调处理可以使用线性预测（Linear Prediction，LP）分析合成语音模型将数字语音信号分解为谱包络部分（由线性预测编码（Linear predictive coding，LPC）系数表示）和激励部分（由 LPC 的残差表示）；然后在 LPC 系数上提取共振峰频率和频谱倾斜参数，再通过矢量量化码书法来实现语音转换。在转换函数的选择上，频率包络转换可以使用矢量量化法，对韵律（主要是基音周期）的转换可以使用时域基音同步叠加（time domain pitch synchronous overlap-add，TD-PSOLA）算法。

具体地，在本实施例中，语音变调处理操作中需要先确定当前要采用的语音变调方式。具体地，确定当前要采用的语音变调方式的操作可包括：确定当前变调信息；如此，当前要采用的语音变调方式可根据当前变调信息确定。其中，当前变调信息具体可包括用户选择信息，和/或授权变调信息。其中，用户选择信息为用户在提供的语音变调方式中所作出的选择；而授权变调信息为变调方用户在即时通讯系统中授权的变调信息。

优选地，为了给传统即时通讯服务带来新的增值服务增长点，即时

通讯服务提供商可以将语音变调或其中的部分语音变调方式作为增值服务项目，在本发明实施例中，提供的语音变调方式由变调方用户在即时通讯系统中的授权变调信息确定。在即时通讯客户端 A 用户选择语音变调方式前，用户通过即时通讯客户端 A 向服务器发送用户的授权变调方式查询信息，服务器根据该用户在即时通讯系统中的身份标识返回授权变调信息，即该用户可使用的语音变调方式。优选地，即时通讯客户端 A 还根据服务器返回的授权变调信息输入用户选择信息，从而根据用户选择信息和服务器返回的授权变调信息确定可使用的语音变调方式。当然，本实施例也可以根据用户选择信息、授权变调信息，用其他的业务选择逻辑确定变调方式；当用户只有一种可使用的语音变调方式时，可以只由授权变调信息确定变调方式。

由于语音变调处理是以用户的原始语音信号作为基础的，为了向用户提供更合适的语音变调方式以保证变调后的声音能被通信对方所辨识，作用本发明的另一个优选实施例，对原始语音进行变调处理的语音变调方式还要考虑用户个性特征信息，即主要是用户原始话音中的音段特征。此时，变调方式，根据用户选择信息及用户个性特征信息，或者用户选择信息、授权变调信息及用户个性特征信息，由业务选择逻辑确定。其中，该业务选择逻辑为即时通讯服务提供商所定义，以明确指出何等的授权变调信息、怎样的语音通信环境可以享受多少项变调语音服务（例如：“男声变女声”为一项变调语音服务）等，其用于确定语音变调方式。

在客户端 A 接收到用户选择信息后，对用户的原始语音信号进行分析，得到其个性特征信息，当其个性特征信息无法满足语音变调处理的要求时，需要对用户的语音变调方式请求进行修正，例如：某用户原始声音较粗且嘶哑，而其选择的语音变调方式为“低龄儿童”，此时的变

调效果会很差（不易为对方识别为“童音”），于是系统应建议用户重新选择语音变调方式。

为了提高通话接收方所听到的话音质量，向用户提供更合适的语音变调方式，作为本发明的又一个优选实施例，语音变调方式确认还要考虑对方语音环境信息。此时，变调方式，根据用户选择信息及对方语音环境信息，或者用户选择信息、授权变调信息及对方语音环境信息，由业务选择逻辑确定。即时通讯客户端 B 向即时通讯客户端 A 返回变调语音通话请求应答时，同时返回其自身的语音环境信息。该语音环境信息可以由即时通讯客户端 B 用户选择，或者由即时通讯客户端 B 根据麦克风采集到的声音信号分析得到等等。

在本发明实施例中，即时通讯客户端 A 的语音变调方式可以由授权变调信息、用户个性特征信息、对方语音环境信息等信息的一个或任意组合及用户选择信息，由业务选择逻辑确定。

在本发明实施例中，由于采集的语音信息中可能包含回音、噪声等不利于处理、传输和辨别的信号，因此，为了达到较好的变调语音通话效果，改善通话接收方所听到的话音质量，在对该数字语音信息进行语音变调处理之前，对数字语音信息进行去噪处理，即进行回音抵消、噪声抑制、信号增益调节等一种或者多种组合的处理。

4、即时通讯客户端 A 将得到的变调语音通过已经建立的变调语音通话通道发送至即时通讯客户端 B。

在本发明实施例中，为了便于传输变调语音，即时通讯客户端 A 在发送变调语音前，对变调语音进行分组、打包处理，得到变调语音数据包，并将变调语音数据包发送至即时通讯客户端 B。

在本发明实施例中，为了降低变调语音数据传输所需要的带宽，为实时变调语音通话创造条件，在对采集的原始语音进行变调处理后，按

照预设的编码规则，例如 G.729、G.729A、G.723.1 等，将得到的与原始语音对应的变调语音进行压缩编码。

在本发明实施例中，为了避免变调语音在网络传输过程中由于丢包、差错等原因造成的信号失真，在对变调语音进行压缩编码后，再利用信道编码技术，对压缩编码后得到的变调语音比特流进行冗余增强处理。

当即时通讯客户端 B 向即时通讯客户端 A 发送变调语音通话请求时，其实现流程同上，在此不再赘述。可以理解，即时通讯客户端 A 和即时通讯客户端 B 可以进行单向变调语音通话，也可以进行双向变调语音通话。上述语音通话基于有线互联网络或无线互联网络上的即时通讯系统。

如果通讯即时通讯客户端A或者即时通讯客户端B中的任意一方提出中断请求、网络发生故障等时，通话结束，释放变调语音通话通道。

图 3 为本发明实施例提供的另一详细流程图，该实施例为即时通讯客户端 A 与即时通讯客户端 B 之间建立语音通信通道，并在即时通讯客户端 A 与即时通讯客户端 B 之间进行语音通话方法的实现流程，详述如下：

1、即时通讯客户端 A 向即时通讯客户端 B 发送语音通话请求。

2、即时通讯客户端 B 在接收到即时通讯客户端 A 发送的语音通话请求后，对该语音通话请求作出应答，并将应答信息返回至即时通讯客户端 A。即时通讯客户端 A 接收到即时通讯客户端 B 返回的语音通话应答后，建立与即时通讯客户端 B 之间的语音通话通道。

语音通话通道后，即时通讯客户端 A 和即时通讯客户端 B 之间可以使用该语音通话通道进行语音通话。

3、即时通讯客户端 A 向即时通讯客户端 B 发送变调语音通话请求。

4、即时通讯客户端 B 在接收到即时通讯客户端 A 发送的变调语音通话请求后，对该变调语音通话请求作出应答，并将应答信息返回至即时通讯客户端 A。即时通讯客户端 A 接收到即时通讯客户端 B 返回的变调语音通话应答后，建立与即时通讯客户端 B 之间的变调语音通话通道。

变调语音通话通道建立后，可以释放先前建立的语音通话通道。当然，即时通讯客户端 A 可以透明或者非透明地向即时通讯客户端 B 发送变调语音通话请求。如果即时通讯客户端 A 透明地向即时通讯客户端 B 发送变调语音通话请求，此过程即时通讯客户端 B 界面上不用显示。

5、即时通讯客户端 A 对采集的原始语音进行变调处理，得到与原始语音对应的变调语音。

6、即时通讯客户端 A 将得到的变调语音通过已经建立的变调语音通话通道发送至即时通讯客户端 B。

需要说明的是，本实施例是以在即时通讯客户端 A 和即时通讯客户端 B 之间建立语音通话通道之后，再在即时通讯客户端 A 和即时通讯客户端 B 之间建立变调语音通话通道为例。为使本实施例更加简单，并且容易实现，本实施例中，即时通讯客户端 A 接收到即时通讯客户端 B 返回的变调语音通话应答后，也可不建立与即时通讯客户端 B 之间的变调语音通话通道，而是利用上述步骤 2 中建立的与即时通讯客户端 B 之间的语音通话通道将得到的变调语音发送至即时通讯客户端 B。这样，省略了步骤 4 中建立变调语音通话通道的操作。优选地，本实施例中可以当前的语音通话通道对应的带宽是否能够适合传输步骤 5 中得到的变调语音为原则来判断当前需不需要建立调语音通话通道。

7、通话结束，释放变调语音通话通道。

如果通讯即时通讯客户端 A 或者即时通讯客户端 B 中的任意一方提出中断请求、网络发生故障等时，通话结束，释放变调语音通话通道。

其中，即时通讯客户端 B 接收到即时通讯客户端 A 发送的变调语音通话数据后，对通话数据的处理流程与普通语音通话中的处理流程相同，处理流程如图 4 所示，详述如下：

在步骤 S401 中，通话数据接收、解包；

通过已经建立的变调语音通话通道接收分组通话数据，按照与即时通讯客户端 A 相同的网络传输协议进行数据包解包，并将分组数据组装，得到压缩码流。

在步骤 S402 中，将解包后的数据解码为语音信号；

将解包后的压缩码流采用即时通讯客户端 A 编码运算的逆运算进行解码，得到可被人耳识别的原始语音信号。

在步骤 S403 中，语音信号增强处理；

由于网络传输、语音信号压缩和语音变调处理等都可能造成信号的失真，需要对解码得到的原始语音信号进行信号增强处理。信号增强处理可以采用卡尔曼滤波法、短时谱幅度的最小均方误差估计法、或者自适应滤波法等。

在步骤 S404 中，输出增强处理后的语音信号。

通过输出设备，例如耳机、音箱、声卡等将增强处理后的语音信号输出。

为了得到可正确解码的语音比特流，在本发明实施例中，在数据接收、组包后，将解包后的数据进行逆冗余/容错处理，去除即时通讯客户端 A 加入压缩码流中的冗余信号，修改或者丢弃其中的错误数据。

上述对本发明实施例提供的方法进行了详述，下面对本发明实施例提供的装置进行描述。

参见图 5，图 5 为本发明实施例提供的装置的基本流程图。如图 5 所示，该装置可包括请求发送单元 501、语音采集单元 502、变调处理

单元 503 和语音发送单元 504。

其中，请求发送单元 501 用于建立变调语音通话通道。

语音采集单元 502 用于采集输入的原始语音信息。

变调处理单元 503 用于将语音采集单元 502 采集的原始语音信息进行变调处理，得到变调语音。

语音发送单元 504 通过请求发送单元 501 建立的变调语音通话通道发送变调处理单元 503 得到的变调语音。

如此，实现了本发明实施例提供的基于即时通讯系统的语音通话装置。

为使本发明实施例提供的基于即时通讯系统的语音通话装置更加清晰，下面详述本发明实施例提供的装置结构。

参见图 6，图 6 为本发明实施例提供的装置的详细结构图，为了便于说明，仅示出了与本发明实施例相关的部分。

该装置可以用于各种即时通讯客户端设备，例如计算机、笔记本电脑、个人数字助理（Personal Digital Assistant，PDA）、智能手机等，可以是运行于这些设备内的软件单元、硬件单元或者软硬件相结合的单元，也可以作为独立的挂件集成到这些设备中或者运行于这些设备的应用系统中，具体地，该装置可包括：请求发送单元 601、语音采集单元 602、变调处理单元 603 和语音发送单元 604。

其中，请求发送单元 601 用于建立变调语音通话通道。

语音采集单元 602 采集输入的原始语音信息。

变调处理单元 603 将语音采集单元 602 采集得到的原始语音信息进行变调处理，得到变调语音。

语音发送单元 604，通过请求发送单元 601 建立的变调语音通话通道发送变调处理单元 603 得到的变调语音。

需要说明的是，请求发送单元 601、语音采集单元 602、变调处理单元 603 和语音发送单元 604 可在同一个实体比如即时通讯客户端 A 上，也可不在同一实体上，比如，请求发送单元 601 和语音采集单元 602 在同一实体比如即时通讯客户端 A 上，而变调处理单元 603 和语音发送单元 604 在预设的变调处理设备比如服务器上。具体情况需要具体分析，本发明实施例并不具体限定。

具体地，请求发送单元 601 在接收到变调语音通话应答后，建立变调语音通话通道；其中，所述变调语音通话应答是针对请求发送单元 601 发送的变调语音通话请求所作出的应答。本实施例中，请求发送单元 601 还可用于接收用户输入的变调语音通话请求信息。

语音采集单元 603 还用于将采集到的语音信息转化为数字语音信息。其中，转化后的数字语音信息为计算机可以识别和处理的。

在本发明实施例中，变调处理单元 603 包括：变调信息确认模块 6031、业务逻辑模块 6032 和语音变调处理模块 6033。

其中，变调信息确认模块 6031 用于确定并输出当前变调信息，其中，当前变调信息包含用户选择信息，和/或授权变调信息。

业务逻辑模块 6032 用于生成业务选择逻辑，其中，该业务选择逻辑用于执行语音变调处理，并输出给语音变调处理模块 6033。这里，业务选择逻辑由即时通讯服务提供商所定义，以明确指出何等的授权变调信息、怎样的语音通信环境可以享受多少项变调语音服务（例如：“男声变女声”为一项变调语音服务）等。

语音变调处理模块 6033，用于根据接收的来自变调信息接收模块 6031 输出的变调信息和业务逻辑模块 6032 输出的业务选择逻辑来确定语音变调方式，将语音采集单元 602 得到的数字语音信息按照语音变调方式进行变调处理，得到与该数字语音信息对应的变调语音并输出。具

体地，语音变调处理模块 6033 根据变调信息中包含的用户选择信息，和/或授权变调信息，由业务选择逻辑确定语音变调方式。实现方式如上所述，不再赘述。

为了向用户提供更合适的语音变调方式以保证变调后的声音能被通信对方所辨识，作用本发明的一个优选实施例，变调处理单元 603 还包括：用户特征获取模块 6034。

其中，用户特征获取模块 6034 用于从语音采集单元 602 得到的该数字语音信息中提取个性特征信息，生成并输出该个性特征信息。

如此，语音变调处理模块 6033 解析出接收到的当前变调信息中包含的用户选择信息，和/或授权变调信息，结合接收到的用户个性特征信息，由业务选择逻辑确定语音变调方式。

为了提高通话接收方所听到的话音质量，向用户提供更合适的语音变调方式，作为本发明的另一个优选实施例，变调处理单元 603 还包括：对方环境获取模块 6035。

其中，对方环境获取模块 6035 用于获取请求发送单元 601 接收的变调语音通话应答所携带的对方语音环境信息。在本实施例中，通话对方返回的变调语音通话应答中包含有语音环境信息，请求发送单元 601 根据接收到语音环境信息生成对方语音环境信息，如此，具体地，对方环境获取模块 6035 获取请求发送单元 601 生成的对方语音环境信息。

这里，本装置并非必须包含用户特征获取模块 6034 和对方环境获取模块 6035；优选地，本实施例也可包含用户特征获取模块 6034 和对方环境获取模块 6035 中的一个或者多个。图 6 以变调处理单元 603 包含用户特征获取模块 6034 和对方环境获取模块 6035 为例。

如此，语音变调处理模块 6033 可以根据业务逻辑模块 6032 发送的业务选择逻辑、变调信息确认模块 6031 发送的当前变调信息和用户特

征获取模块 6034 发送的个性特征信息；或者根据业务逻辑模块 6032 发送的业务选择逻辑、变调信息确认模块 6031 发送的当前变调信息和对方环境获取模块 6035 发送的对方语音环境信息；或者根据业务逻辑模块 6032 发送的业务选择逻辑、变调信息确认模块 6031 发送的当前变调信息、用户特征获取模块 6034 发送的个性特征信息和对方环境获取模块 6035 发送的对方语音环境信息来确定语音变调方式。

为了达到较好的变调语音通话效果，改善通话接收方所听到的话音质量，作为本发明的再一优选实施例，语音通话装置还包括：去噪单元 605。

其中，去噪单元 605 接收语音采集单元 602 得到的数字语音信息，进行去噪处理，得到去噪后的数字语音信息。

为了降低变调语音数据传输所需要的带宽，为实时变调语音通话创造条件，作为本发明的再一优选实施例，语音通话装置还包括：编码单元 606，和/或优化单元 607。图 6 以语音通话装置包括编码单元 606 和优化单元 607 为例。

编码单元 606 将变调处理单元 603 得到的变调语音进行压缩编码，得到的变调语音比特流。

优化单元 607，将编码单元 606 得到的变调语音比特流进行冗余增强处理，和/或分组、打包处理，并将处理后的变调语音数据输出给语音发送单元 604。这里，优化单元 607 主要目的是为了避免变调语音在网络传输过程中由于丢包、差错等原因造成的信号失真，或者便于传输变调语音。当然，若本装置不包括编码单元 606，则优化单元可将变调处理单元 603 得到的变调语音进行冗余增强处理，和/或分组、打包处理，并将处理后的变调语音数据输出给语音发送单元 604。

如图 6 所示，在本实施例中，优化单元 607 可包括：

冗余增强处理模块 6071，利用信道编码技术，对编码单元 606 得到的变调语音比特流或者变调处理单元 603 得到的变调语音进行冗余增强处理，输出处理后的变调语音比特流。

分组、打包模块 6072，将接收到的变调语音数据进行分组、打包处理，得到变调语音数据包。分组、打包模块 6072 可以接收变调处理单元 603、编码单元 606 或者冗余增强处理模块 6071 输出的变调语音、变调语音比特流。

需要说明的是，优化单元 607 可以只包括冗余增强处理模块 6071 或者分组、打包模块 6072。

如图 6 所示，为了使语音通话装置能够接收处理语音信息，作为本发明的再一优选实施例，语音通话装置还包括:

请求应答单元 608，接收请求发送单元 601 发送的变调语音通话请求，并返回变调语音通话应答，生成语音接收触发信息并输出给语音接收单元 609。

语音接收单元 609，接收到请求应答单元 608 输出的语音接收触发信息后，若当前接收的数据包为经过分组或者打包处理后的数据包，则按照与通话对方相同的网络传输协议进行数据包解包，并将分组数据组装，得到压缩码流并输出。

解码单元 610，将语音接收单元 609 得到的数据，即压缩码流，解码为语音信号。

语音信号增强处理单元 611，将解码单元 610 得到的数据解码，即将语音信号解码，得到原始语音信号并进行信号增强处理，得到增强处理后的语音信号。

语音输出单元 612，将得到的增强处理后的语音信号输出，可以为耳机、音箱、声卡等。

在本发明实施例中，若语音接收单元609当前接收的数据包为加入了压缩码流的冗余信号，则语音通话装置还包括：逆冗余/容错处理单元613。

其中，逆冗余/容错处理单元613，用于去除语音接收单元609得到的通话对方加入压缩码流中的冗余信号，并修改或丢弃其中的错误数据。如此，可大大改善话音质量。

优选地，请求应答单元608、语音接收单元609、解码单元610、语音信号增强处理单元611、语音输出单元612、逆冗余/容错处理单元613和去除语音接收单元614所在的实体与上述请求发送单元601、语音采集单元602、变调处理单元603、语音发送单元604、去噪单元605、编码单元606和优化单元607各个单元所在的实体不同，比如，若请求发送单元601、语音采集单元602、变调处理单元603、语音发送单元604、语音发送单元604、去噪单元605、编码单元606和优化单元607在即时通讯客户端A上，则请求应答单元608、语音接收单元609、解码单元610、语音信号增强处理单元611、语音输出单元612、逆冗余/容错处理单元613和去除语音接收单元614所在的实体可为即时通讯客户端A的对端（接收端），比如即时通讯客户端B。当然，若请求发送单元601、语音采集单元602在同一实体比如即时通讯客户端A上，而变调处理单元603和语音发送单元604在预设的变调处理设备比如服务器1上，则请求应答单元608、语音接收单元609、解码单元610、语音信号增强处理单元611、语音输出单元612、逆冗余/容错处理单元613和去除语音接收单元614所在的实体可为服务器1的对端（接收端），比如即时通讯客户端B。需要说明的是，上述只是一种举例，并非限定本发明实施例的应用。

在本发明实施例中，将即时通讯系统中采集到的语音信号先进行语

音变调处理，实现了基于即时通讯系统的变调语音通话，为基于即时通讯场合下的语音通信带来极大的娱乐效果，给传统即时通讯服务带来新的增值服务增长点，增加用户对即时通讯产品的依赖性，从而提升产品竞争力。并且为语音通话用户提供崭新的业务体验，例如：用变调语音通话来达到保护用户身份信息的目的。

以上所述仅为本发明的较佳实施例而已，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、等同替换和改进等，均应包含在本发明的保护范围之内。

权利要求书

1、一种基于即时通讯系统的语音通话方法，其特征在于，所述方法包括下述步骤:

步骤 a，建立至少两个即时通讯客户端之间的变调语音通话通道;

步骤 b，对输入的原始语音信息进行变调处理，得到变调语音；通过所述变调语音通话通道将所述变调语音发送至即时通讯客户端 B。

2、如权利要求 1 所述的方法，其特征在于，所述步骤 b 的操作是由即时通讯客户端 A 或者预设的变调处理设备执行的，其中，即时通讯客户端 A 为所述步骤 a 中已建立了所述变调语音通话通道的其中一个即时通讯客户端。

3、如权利要求 2 所述的方法，其特征在于，所述步骤 a 的操作是在即时通讯客户端 A 接收到所述即时通讯客户端 B 发送的变调语音通话应答后执行的，其中，变调语音通话应答是针对即时通讯客户端 A 发送的变调语音通话请求所作出的应答；或者，

所述步骤 a 中的变调语音通话通道为即时通讯客户端 A 接收到即时通讯客户端 B 返回的语音通话应答后，建立的与即时通讯客户端 B 之间的语音通话通道；其中，所述语音通话应答为针对所述即时通讯客户端 A 向即时通讯客户端 B 发送语音通话请求的应答。

4、如权利要求 1 所述的方法，其特征在于，所述步骤 b 中对输入的原始语音信号进行变调处理包括:

采集输入的原始语音信息，将该采集到的原始语音信息转化成数字语音信息；对所述数字语音信息进行变调处理。

5、如权利要求 1 至 4 任一所述的方法，其特征在于，所述变调处理包括:

确定语音变调方式;

根据确定出的语音变调方式执行变调处理。

6、如权利要求 5 所述的方法，其特征在于，在确定语音变调方式之前，进一步包括：确定当前变调信息和用于确定语音变调方式的业务选择逻辑;

所述语音变调方式由所述业务选择逻辑根据所述当前变调信息确定。

7、如权利要求 6 所述的方法，其特征在于，所述确定语音变调方式之前，进一步包括：获取输入的原始语音信息对应的个性特征信息;

所述语音变调方式由所述业务选择逻辑根据获取的个性特征信息和/或所述当前变调信息确定。

8、如权利要求 7 所述的方法，其特征在于，所述变调语音通话应答中携带了即时通讯客户端 B 的语音环境信息;

所述确定语音变调方式由所述业务选择逻辑根据所述语音环境信息、所述当前变调信息和所述个性特征信息中的一个或者任意一种组合确定。

9、如权利要求 4 所述的方法，其特征在于，在所述数字语音信息进行语音变调处理之前，进一步包括:

对所述数字语音信息进行去噪处理。

10、如权利要求 1 所述的方法，其特征在于，在所述通过变调语音通话通道将所述变调语音发送至即时通讯客户端 B 之前，进一步包括:

对所述变调语音进行压缩编码和/或冗余增强处理;

和/或，

对变调语音进行分组、打包处理。

11、一种基于即时通讯系统的语音通话装置，其特征在于，所述装

置包括:

请求发送单元，用于建立变调语音通话通道;

语音采集单元，用于采集输入的原始语音信息;

变调处理单元，用于将所述语音采集单元采集的原始语音信息进行变调处理，得到变调语音;

语音发送单元，用于通过所述请求发送单元建立的变调语音通话通道发送所述变调处理单元得到的变调语音。

12、如权利要求 11 所述的语音通话装置，其特征在于，所述语音采集单元还用于将采集到的语音信息转化为数字语音信息;

所述变调处理单元包括:

变调信息确认模块，用于确定并输出当前变调信息;

业务逻辑模块，用于生成并输出供语音变调处理模块执行语音变调处理的业务选择逻辑;

语音变调处理模块，用于根据接收的所述变调信息确认模块输出的当前变调信息和业务逻辑模块输出的业务选择逻辑来确定语音变调方式，将所述语音采集单元得到的数字语音信息按照所述语音变调方式进行变调处理，得到与所述数字语音信息对应的变调语音并输出。

13、如权利要求 12 所述的语音通话装置，其特征在于，所述变调处理单元还包括: 用户特征获取模块，和/或对方环境获取模块; 其中，

所述用户特征获取模块用于从所述语音采集单元得到的该数字语音信息中提取个性特征信息，生成并输出个性特征信息;

对方环境获取模块，用于获取并输出所述请求发送单元接收的变调语音通话应答所携带的对方语音环境信息;

所述语音变调处理模块根据所述业务选择逻辑、所述当前变调信息和所述个性特征信息; 或者所述业务选择逻辑、所述当前变调信息和所

述对方语音环境信息；或者所述业务选择逻辑、所述当前变调信息、所述个性特征信息和所述对方语音环境信息来确定语音变调方式。

14、如权利要求 12 所述的语音通话装置，其特征在于，该装置还包括:

去噪单元，接收所述语音采集单元得到的数字语音信息，进行去噪处理，得到去噪后的数字语音信息;

该装置还包括：编码单元和/或优化单元；其中，

所述编码单元用于将所述变调处理单元得到的变调语音进行压缩编码，得到变调语音比特流;

所述优化单元将所述变调处理单元得到的变调语音或者所述编码单元得到的变调语音比特流进行冗余增强处理，和/或分组、打包处理，并将处理后的变调语音数据输出给所述语音发送单元。

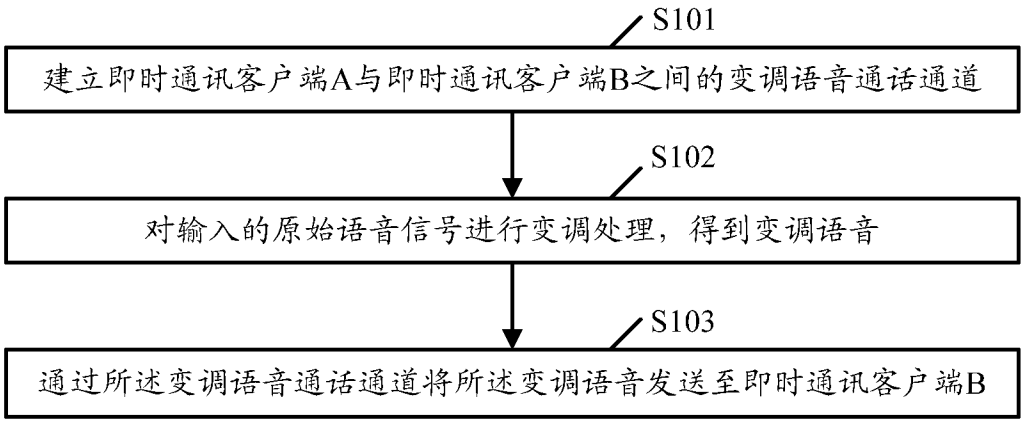

S101
建立即时通讯客户端A与即时通讯客户端B之间的变调语音通话通道
S102
对输入的原始语音信号进行变调处理，得到变调语音
S103
通过所述变调语音通话通道将所述变调语音发送至即时通讯客户端B

图 1

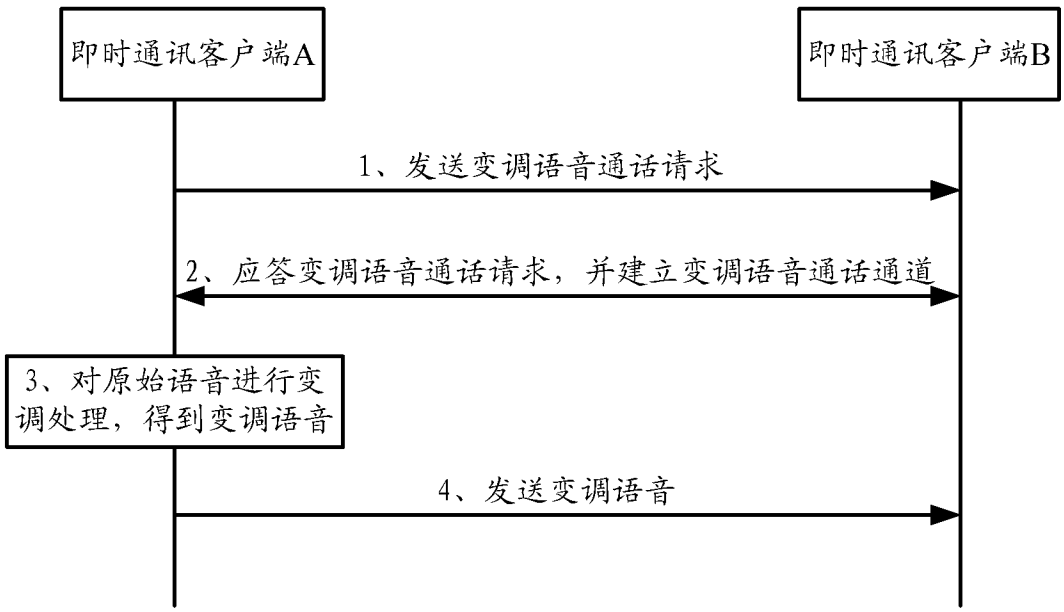

即时通讯客户端A
即时通讯客户端B
1、发送变调语音通话请求
2、应答变调语音通话请求，并建立变调语音通话通道
3、对原始语音进行变调处理，得到变调语音
4、发送变调语音

图 2

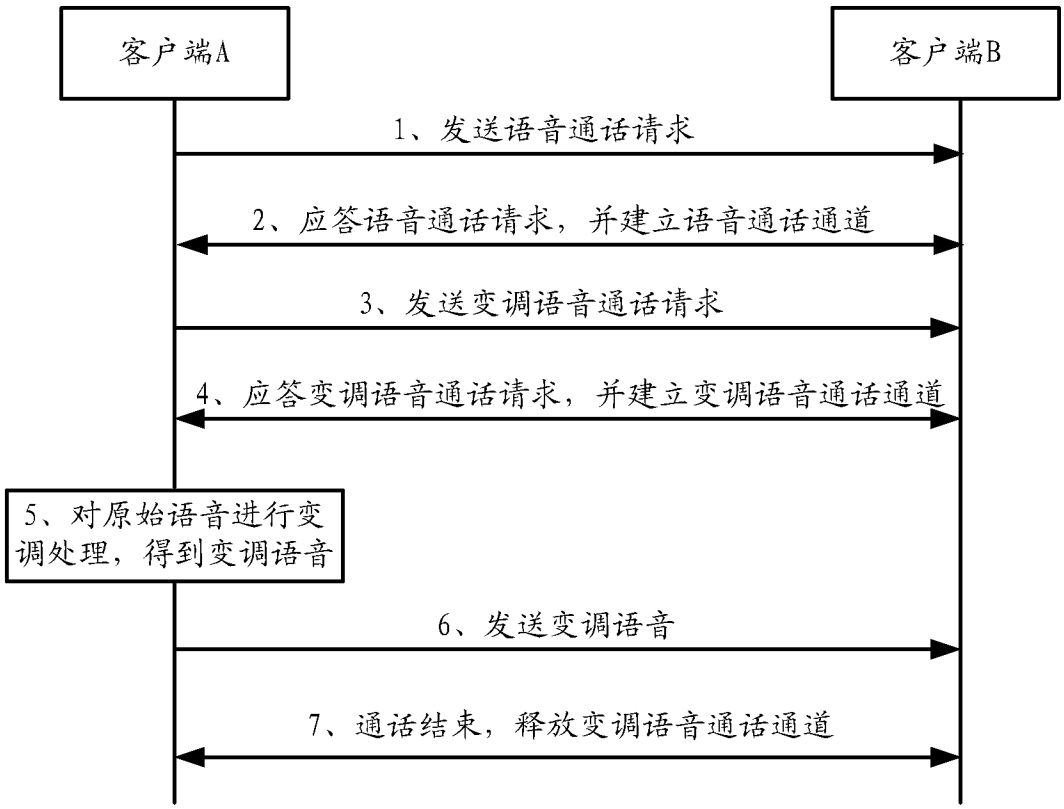

客户端A
客户端B
1、发送语音通话请求
2、应答语音通话请求，并建立语音通话通道
3、发送变调语音通话请求
4、应答变调语音通话请求，并建立变调语音通话通道
5、对原始语音进行变调处理，得到变调语音
6、发送变调语音
7、通话结束，释放变调语音通话通道

图 3

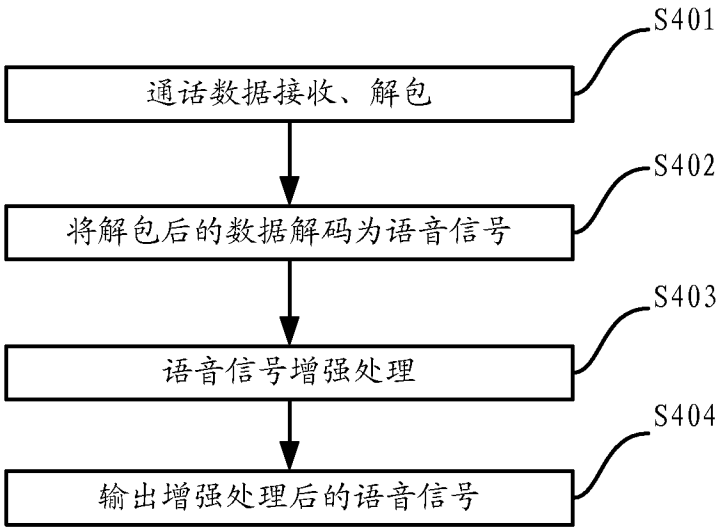

S401
通话数据接收、解包
S402
将解包后的数据解码为语音信号
S403
语音信号增强处理
S404
输出增强处理后的语音信号

图 4

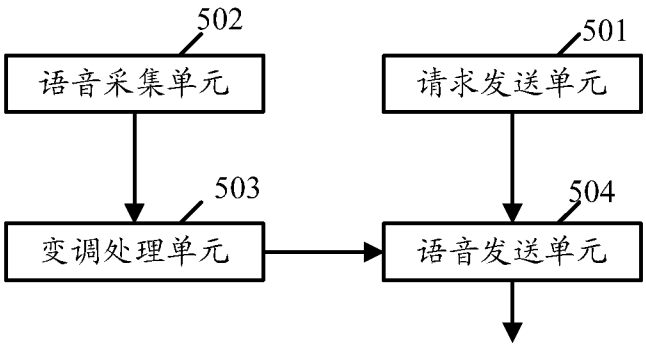

502
语音采集单元
501
请求发送单元
503
变调处理单元
504
语音发送单元

图 5

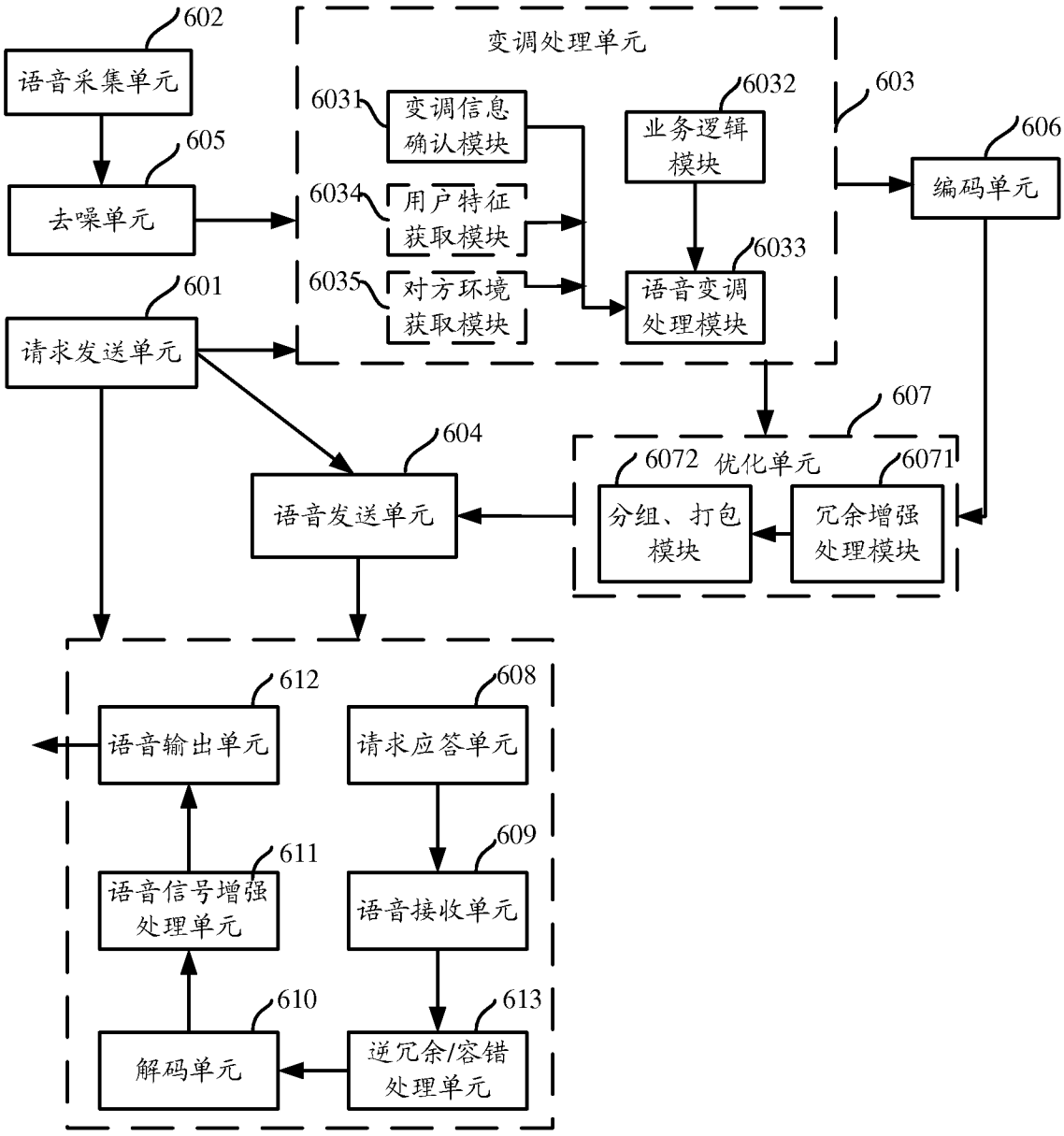

602
语音采集单元
605
去噪单元
601
请求发送单元
变调处理单元
6031
变调信息确认模块
6034
用户特征获取模块
6035
对方环境获取模块
6032
业务逻辑模块
6033
语音变调处理模块
603
606
编码单元
607
6072 优化单元
6071
分组、打包模块
冗余增强处理模块
604
语音发送单元
612
语音输出单元
611
语音信号增强处理单元
610
解码单元
608
请求应答单元
609
语音接收单元
613
逆冗余/容错处理单元

图 6

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2009/071931

A. CLASSIFICATION OF SUBJECT MATTER

G10L 13/00 (2006.01)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)

IPC: G10L, H04M, H04L

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)

CPRS, CNKI, WPI, EPODOC, PAJ: voice, sound, speech, audio, chang+, transform+, conversation, instant, real?time, network, internet

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	CN1805478A (HUAWEI TECHNOLOGIES CO LTD) 19 Jul. 2006 (19.07.2006) claims 1,2,11, fig. 2	1-5,9-11
A		6-8,12-14
Y	CN1719514A (INST AUTOMATION CAS) 11 Jan. 2006 (11.01.2006) description pages 1-2,8	1-5,9-11
A		6-8,12-14
A	US2005043951A1 (SCHURTER E T) 24 Feb. 2005 (24.02.2005) the whole document	1-14
P,X	CN101304391A (TENCENT SCI & TECHNOLOGY CO LTD) 12 Nov. 2008 (12.11.2008) the whole document	1-14

☐ 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
19 Aug. 2009 (19.08.2009)	**27 Aug. 2009 (27.08.2009)**
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 HAO Aixin Telephone No. (86-10)62411443

Form PCT/ISA/210 (second sheet) (April 2007)

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2009/071931

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN1805478A	19.07.2006	NONE	
CN1719514A	11.01.2006	NONE	
US2005043951A1	24.02.2005	NONE	
CN101304391A	12.11.2008	NONE	

国际检索报告

国际申请号
PCT/CN2009/071931

A. 主题的分类

G10L 13/00 (2006.01)i

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

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

IPC: G10L, H04M, H04L

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）)

CPRS, CNKI: 语音，声音，话音，音频，变调，变声，通话，即时通信，即时通讯，因特网，互联网

WPI, EPODOC, PAJ: voice, sound, speech, audio, chang+, transform+, instant, real?time, network, internet

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
Y	CN1805478A (华为技术有限公司) 19.7 月 2006 (19.07.2006) 权利要求 1、2、11，附图 2	1-5,9-11
A		6-8,12-14
Y	CN1719514A (中国科学院自动化研究所) 11.1 月 2006 (11.01.2006) 说明书第 1-2 页，第 8 页	1-5,9-11
A		6-8,12-14
A	US2005043951A1 (SCHURTER E T) 24.2 月 2005 (24.02.2005) 全文	1-14
P,X	CN101304391A (腾讯科技(深圳)有限公司) 12.11 月 2008 (12.11.2008) 全文	1-14

☐ 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

"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
19.8 月 2009 (19.08.2009)	**27.8 月 2009 (27.08.2009)**
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 郝爱昕 电话号码：(86-10) **62411443**

国际检索报告

关于同族专利的信息

国际申请号
PCT/CN2009/071931

检索报告中引用的专利文件	公布日期	同族专利	公布日期
CN1805478A	19.07.2006	无	
CN1719514A	11.01.2006	无	
US2005043951A1	24.02.2005	无	
CN101304391A	12.11.2008	无	