



[12] 发明专利申请公开说明书

[21] 申请号 02118393.7

[43] 公开日 2003 年 11 月 5 日

[11] 公开号 CN 1453990A

[22] 申请日 2002.4.26 [21] 申请号 02118393.7

[71] 申请人 朗讯科技公司

地址 美国新泽西州

[72] 发明人 黄大威

[74] 专利代理机构 中国国际贸易促进委员会专利

商标事务所

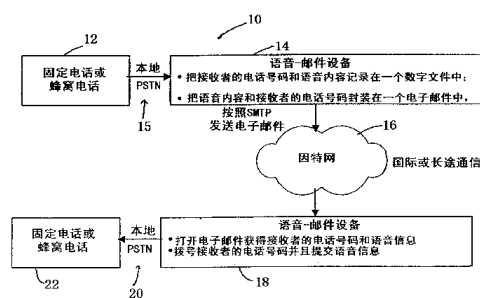
代理人 蒋世迅

权利要求书 3 页 说明书 17 页 附图 4 页

[54] 发明名称 语音信息系统和方法

[57] 摘要

一种通过网络传递发送者的语音信息给接收者的通信方法和系统。语音 - 邮件设备把接收者的电话号码和发送者的语音信息记录在一个数字文件中。该电话号码和语音信息被封装在一个电子邮件中，然后通过网络传递给作为网关的另一个语音 - 邮件设备。打开该电子邮件以获得接收者的电话号码和语音信息，并且经由一个电话网络拨打接收者的电话号码将语音信息传递给接收者。



1.一种通过网络传递发送者的语音信息给接收者的通信方法，包括以下步骤：

在一个语音-邮件设备中把接收者的电话号码和发送者的语音信息录制在一个数字文件中；

把该电话号码和语音信息封装在一个电子邮件中；

通过网络，将来自该语音-邮件设备的电子邮件传递给另一个语音-邮件设备；

打开该电子邮件以获得接收者的电话号码和语音信息；及

拨打接收者的电话号码并经由一个电话网络将语音信息传递给接收者。

2.根据权利要求1的方法，进一步包括步骤，由发送者使用公共交换电话网络或蜂窝电话网络呼叫一个语音-邮件设备。

3.根据权利要求1的方法，其中语音-邮件设备包括一个电话交换机。

4.根据权利要求1的方法，其中语音-邮件设备包括蜂窝基站或无线网络中心。

5.根据权利要1的方法，进一步包括步骤，通过公共交换电话网络或蜂窝电话网络将语音信息传递给接收者。

6.根据权利要求1的方法，进一步包括步骤，以数字语音方式将语音信息传递给接收者。

7.根据权利要求1的方法，进一步包括步骤，利用简单邮件传输协议通过网络传递电子邮件。

8.根据权利要求1的方法，进一步包括步骤，在语音-邮件设备中把语音信息记录为一个声波文件。

9.根据权利要求1的方法，进一步包括步骤，用一个邮件传送代理来传递电子邮件。

10.根据权利要求9的方法，进一步包括步骤，从该邮件传送代理

通过因特网来传递文文本件。

11.根据权利要求 1 的方法,进一步包括步骤,决定是否接收和/或发送任何单向语音信息,这一决定是建立在对未来通道容量的预测的基础上的,而这一预测是根据对无线网络系统所建立的排队论模型做出的。

12.根据权利要求 1 的方法,其中语音-邮件设备起以下作用:

a)播放提示音及接收电话拨号信号;b)记录及播放一个语音信息;c)找出电子邮件地址;及 d)发送及接收一个电子邮件。

13.根据权利要求 1 的方法,进一步包括步骤,在语音-邮件设备中用递归最小二乘算法以及游程长度编码来压缩语音信息。

14.一种通过网络将发送者的语音信息传递给接收者的通信系统,包括一个语音-邮件设备,作为网络中的网关,用于通过电话网接收发送者的语音信息和电话号码,所述语音-邮件设备起以下作用:

a)播放提示音及接收电话拨号信号;

b)记录及播放一个语音信息;

c)找出电子邮件地址;及

d)通过因特网和电话网对接收者发送和接收一个电子邮件。

15.根据权利要求 14 的系统,其中电话网包括公共交换电话网络。

16.根据权利要求 14 的系统,其中电话网包括蜂窝电话网络。

17.根据权利要求 14 的系统,其中语音-邮件设备包括电话交换机。

18.根据权利要求 14 的系统,其中语音-邮件设备包括蜂窝基站或无线网络中心。

19.根据权利要求 14 的系统,其中语音-邮件设备以声波文件形式记录和播放语音信息。

20.根据权利要求 14 的系统,其中语音-邮件设备基于对未来通道容量的预测发送任一单向语音信息,而这一预测是根据对无线网络系统所建立的排队论模型做出的。

21.根据权利要求 14 的系统,进一步包括一个邮件传送代理用于

传递电子邮件。

22. 根据权利要求 14 的系统，其中语音 - 邮件设备利用简单邮件传输协议通过网络传递电子邮件。

23. 根据权利要求 14 的系统，其中语音 - 邮件设备包括一个处理器用于文本到语音的转换。

24. 根据权利要求 14 的系统，其中语音 - 邮件设备包括一个处理器用于利用递归最小二乘算法来压缩语音信息。

25. 根据权利要求 1 的方法，进一步包括步骤，用递归最小二乘算法来压缩语音信息。

语音信息系统和方法

技术领域

本发明涉及通信系统，更具体地，涉及用电话和电子邮件通过网络传输语音信息的通信系统。

背景技术

对国际或长途通信，本领域的熟练技术人员已经发展了多种通信系统。其中一种是传统的、使用公用交换电话网络（PSTN）的电话通信系统。第二种系统是最近发展起来的，它使用互联网络传输语音（VoIP），来完成语音电话的通话。第三种系统是普通的电子邮件传输，它每天都被成千上万的人使用着。第四种系统，是被我们所知的“Fone2webTM”，也被称为“Phone2web”。

正如本领域熟练技术人员所知，“Fone2webTM”是通过因特网传输语音的通信系统，它使电话用户和个人计算机使用者互换语音信息。此系统在电话用户和因特网之间搭桥，以便于电话用户可以和电子邮件使用者进行语音通信。位于加利福尼亚硅谷的 Fone2web.com 公司发展了这个通信系统，中国电信使用这一系统的中国版。在这个系统中，电话对电话、电话对个人计算机、个人计算机对电话，都可以进行通信。

在这个系统的电话对电话的通信过程中，语音信息通过因特网传给其他的 Fone2webTM 用户。服务器接收到信息后，将其发送到离接收者最近的一个服务器，并把此语音信息存储在接收者的 Fone2webTM 语音信箱。在这个系统中，也可以用电话把语音信息发送给世界上任何一个有普通电子信箱的人。语音信息被传送到接收者的电子信箱内，接收者会看到这个语音信息的图标并且点击它，此时，将会出现“phone-to-web VE mailerTM”并播放语音信息。它也可以给世界上任何一个电子信箱发送语音信息，还可以向听众列表上的每个人发送

语音广播。语音信息可以被存储。

普通的国际或长途电话给成千上万的日常用户提供了方便、高质量的双向通信。然而，这种通信系统仍然是昂贵的。因为它使用电路交换技术，并且要用到昂贵的数字交换机和本地环路载波。另一个系统 VoIP，使用的是较经济的分组交换通信方式，但是需要昂贵的设备，使得它在一些情况下不可避免价格昂贵。并且，同标准的国际或长途电话相比，它的话音质量比较差。

正如本领域的熟练技术人员所知，电子邮件已经成为因特网上最成功的应用之一。和普通邮件比较，它廉价、传送更快捷。一个电子邮件用户仅仅需要一台以某种方式与因特网相连的计算机。然而，它通常仍然需要使用计算机来进行电子邮件通信。

最近几年，许多研究致力于应用蜂窝电话或个人数字助理器（PDA）来发送/接收电子邮件信息，以便于用户获得更多移动性。但是，对于移动用户来说，这还不如发送语音信息方便。尤其是因为在蜂窝电话或是个人数字助理器（PDA）中键入中文字符非常困难。尽管最近 Fone2webTM 发展了通过因特网传送语音信息，但是这个系统仍然是昂贵的，而且被传送的语音信息需要存储在语音信箱里。接收者必须通过呼叫连接这个语音信箱，才能回复得到原信息。呼叫连接语音信箱对一些用户是不方便、不快捷的。

发明内容

因此，本发明的一个目的就是提供一种系统和方法，使用户克服现有技术的不足，可以不必直接使用计算机却能够通过因特网给接收者发送国际或长途的语音信息。

本发明的系统和方法给用户提供一个语音 - 邮件设备，使他们可以通过因特网给接收者发送国际或长途的语音信息却不用直接操作计算机。这个系统提供了费用便宜的语音通信，并且方便，可移动、话音质量高。和电子邮件以及 Fone2webTM 通信系统相似的是，本发明的系统和方法提供的是单向的通信系统。

本发明组合了电话 - 电子邮件 - 电话之间的通信。通过本地到语

音-邮件设备的连接,发送者和接收者都可以使用普通的电话呼叫。语音信息通过使用电子邮件在本发明的两个语音-邮件设备之间传送,这是本发明的一个方面。

本发明的一种方法是通过网络把发送者的语音信息传送给接收者。语音-邮件设备接收到来自发送者的电话呼叫,并且把接收者的电话号码和发送者的语音信息记录在一个数字文件中。该语音信息和电话号码被封装在一个电子邮件中,然后从这个语音-邮件设备通过网络发送给另一个语音-邮件设备。打开该电子邮件,获得接收者的电话号码和语音信息,拨叫接收者的电话号码后,语音信息就可以通过电话网传送给接收者。

本发明的另一个方面是,用户可以通过公用交换电话网为蜂窝电话网呼叫语音-邮件设备。语音-邮件设备可以包括一个电话交换机,蜂窝基站或无线网络中心。语音信息可以通过公用交换电话网或是蜂窝电话网传送给接收者。发送给接收者的语音信息可以用数字语音传递,使用简单的邮件传送协议即可通过网络,例如因特网,传送电子邮件。语音信息可以被语音-邮件设备记录为一个声波文件。

本发明的又一个方面是,可以根据无线网络系统中的排队论模型预测近期的信道容量,来决定是否发送任何单向语音信息。语音-邮件设备操作任务如下:(a)播放提示音和接收拨号音频信号;(b)录制和播放语音信息;(c)查找电子邮件地址;及(d)发送和接收电子邮件。语音-邮件设备利用递归最小二乘算法和游程长度编码方法压缩语音信息。

本发明还公开了一种通过网络将发送者的语音信息传送给接收者的系统,它包括一个语音-邮件设备作为网络中的网关,用于通过电话网接收发送者的语音信息和电话号码。语音-邮件设备操作任务如下:(a)提示和接收拨号音频;(b)录制和播放语音信息;(c)查找电子邮件地址;及(d)通过因特网和电话网向接收者发送和接收电子邮件。

附图说明

以下结合附图对本发明进行详细描述，其目的、特征、优点将会被明显地展现出来，附图中：

图 1 显示了在本发明第一实施例中固定电话或移动电话同语音 - 邮件设备之间的工作，以及通过因特网按照 SMTP 协议发送电子邮件；

图 2 是可以应用在本发明中的、对语音 - 邮件设备进行高层次描述的框图；

图 3 是应用了图 1 所示系统的本发明的方法的高层流程图；

图 4 描述了本发明另一个实施例，它和图 1 类似，但是它显示了在移动台和基站之间网关所起的作用；

图 5 是另一个框图，但是它描述的是一个和因特网及网关相连的邮件发送代理在本发明中怎样将文本向语音进行转换。

具体实施方式

下面参考附更详细地描述本发明，其中描述了本发明的优选实施例。然而，本发明可以具体表现为很多形式，不能仅限于这里展示的实施例。确切的说，利用这些实施例，表述会更完整更彻底，因而可以使本领域的技术人员将本发明最大程度地展现出来。图 1 至图 5 中相同的数字对应于本申请中的相同的部件。

本发明提供了低价格的长途和国际通信，并享有移动性和方便性，不用为进行电子邮件通信而键入字符。本发明也支持蜂窝电话不通过注册而进行漫游。比较传统的系统，包括普通国际或长途电话、因特网上传送语音、电子邮件、Fone2webTM 通信，它有更多的优点。

在本发明中，利用因特网传送语音信息。和更普遍使用在普通电话中的交换电路相比，它代价更低。和利用网络电话的语音通信相比，它能获得更高质量的话音，因为它使用国际标准 μ 律 ($m\mu$ 律) 编码。

如本领域的熟练技术人员所知， $m\mu$ 律编码是使用对数量化对语音信号进行编码的系统。信号分布在低电平比分布在高电平在统计上可能更多，这样，在低电平附近就比高电平有更多的量化点。在大多数 $m\mu$ 律系统中，14 - 16bits 的线性采样被混合为 8bits。这也经常应用在电话质量编译码中，例如 Spark 工作站语音编译码。

μ 律编码的命名来自感知曲线和听觉感知研究的应用。它是一个在对数域的非线性脉码调制，它加入的噪声和信号强度成比例。Sun Spark workstations .AU 设计的某些文件是关于 μ 律编码的一个最普遍的应用。例如，8bit μ 律编码会把 CD 语音的一个信道削减至 350Kbps。

如本领域的熟练技术人员所知，脉码调制的标准是 ITUUG.711，每 1/8,000 秒向每个样点分配一个电平。只有 8bits 送去编码每个样点，因而只有 256 个不同的电平被编码。这样产生的信道速率是 64Kbps。本领域的熟练技术人员知道，8-bit μ 律信号的质量等同于 12-bit 的 PCM 的质量。在低振幅对变化和表现语音比在高振幅更敏感，在低振幅使用的 bit 数较多，高振幅使用的较少。这种不规则量化可以由几种方式实现，如 μ 律和 a 律编码。

许多编码器使用时 - 频映射的方法，例如滤波器组或是快速傅立叶变换 (FFT)，把输入信号分解为子带信号。生理 - 声学模型通过考虑子带和原始信号，并使用生理 - 声学信息来决定遮盖门限。每个子带采样被量化、编码以保证量化噪声低于遮盖门限。这些量化过的采样被组合在要由解码器确定的帧中。解码时我们不需要生理 - 声学模型，这里的帧被打开并且子带样本被解码。然后频 - 时映射把它们转化成简单的语音输出信号。

在本发明中，声波文件被接收后，声音即被播放。这样，不会有延抖动发生。在拥挤时也很少出现错误。本发明建立在电子邮件系统上是很有利的，因为它不需要直接使用计算机。也不需要键入字符，键入字符对移动电话用户，尤其是中国的移动电话用户是很困难的。本发明建立在电话 - 网络通信系统上也是很有利的，因为本发明提供电话 - 电话的通信而不必从某个邮件信箱里“提取”信息。

本发明使用了一个语音 - 邮件设备，它在本发明中的作用是，即使国际距离范围，用户也可通过因特网给接收者发送信息并且不必直接使用计算机。在对本发明的整个说明书中，“语音 - 邮件设备”用来描述不同的处理器和各种部件，它们在本发明中有允许用户发送

语音邮件信息的功能。和电子邮件以及电话-网络通信系统类似，本发明的通信方法也是单向通信。它组合了电话-电子邮件-电话通信，其中发送者和接收者使用标准的电话呼叫用来本地连接语音-邮件设备。语音-邮件设备可以是某个电话交换机的一部分，也可以合并到公用交换电话网内。

根据本发明的第一方面，如图1所示，标记10表示该通信系统。通过电子邮件在两个语音-邮件之间发送信息，固定电话或蜂窝电话通过本地PSTN 15连接到对应于语音-邮件设备功能的语音邮件网关。接收者的电话号码以及语音信息都被录制在一个数字文件里，然后把语音内容和接收者的电话号码封装在电子邮件中。此电子邮件根据简单邮件传输协议(SMTP)通过因特网16传送给国际或长途距离的另一端，那里的语音邮件网关18中的语音-邮件设备接收到传来的电子邮件，它打开邮件在其中读出接收者的电话号码和语音信息。然后，它通过本地电话网20拨打接收者的电话号码传递给接收者的固定电话或是蜂窝电话22，提交该语音信息。

如本领域的熟练技术人员所知，简单邮件传输协议是建立在TCP/IP协议上的、在服务器之间发送电子邮件的协议。大多数电子邮件系统使用SMTP和POP3(电子邮局协议)服务器通过因特网发送电子邮件。收信端从POP3服务器检索信息。通常，当任何电子邮件客户应用软件配置时，都需要指定POP3和SMTP服务器。

如图2的本发明的一个非限定性的例子所示，某个语音-邮件设备14包括一个计算机处理器30，在这个计算机处理器里，包括相应的计算电路，还有四个主要的软件模块32、34、36、38，开发这些软件模块的技术众所周知。第一个软件模块32是对话软件，负责播放提示音和接收拨号音频。第二个软件34录制和播放应用本发明中的编解码方法压缩处理过的语音信息。第三个软件36含有一个可以根据接收者的国家和地区编码找到电子邮件地址的查询表36a。第四个软件模块38负责根据SMTP/POP3协议发送和接收电子邮件。

计算机里的语音卡或语音调制解调器40通过拨号音频检测呼叫

电话所在的国家地域。它把模拟语音转换成数字信号然后播放数字语音信号。用一个局域网网卡或是数据调制解调器 42 把语音 - 邮件设备和某个电子邮件用户代理连接起来。通过该语音卡或语音调制解调器, 该语音 - 邮件设备连接到一个标准电话交换机 44。给语音 - 邮件设备分配一个特别服务号码和电子邮件地址。

现在参考图 3, 描述本发明的方法的基本操作。在发送端, 对某个特别服务号码的呼叫被交换机 44 (框图 50) 转移给语音 - 邮件设备 14。用户 (发送者) 12 拨叫特服号码, 然后和语音 - 邮件设备进行对话 (框图 52)。语音 - 邮件设备里的第一软件模块 32 要求发送者键入接收者的电话号码包括国家、地域区号 (框图 54)。之后, 在提示音提示下, 开始录制发送者的语音信息 (框图 56)。该信息被语音卡/调制器 40 转换成数字形式, 处理后, 被第二软件模型 34 (框图 58) 存为一个声波文件。系统首先将声音滤波以增加压缩率 (框图 60)。根据国际标准 μ 律, 声音继续被压缩。下面的论述还会详细的解释基于递归最小二乘算法的无失真压缩。

第三软件模块 36 找到对应接收者 22 的电子邮件地址 (框图 62)。第四软件模块 38 在电子邮件中封装接收者的电话号码, 并把语音信息声波文件作为一个附件 (框图 64)。然后它把这封电子邮件根据 SMTP 协议发送出去。

在接收端, 语音 - 邮件设备 18 在局域网网卡或是数据调制解调器 42 (框图 66) 的作用下从互联网接收到那封电子邮件, 第四软件模块 38 传递出接收者的电话号码并把语音信息声波文件保存起来 (框图 68)。第二软件模块 34 将这个声波文件解压缩 (框图 70)。第一软件模块 32 控制语音卡/调制器 40 拨打接收者的电话号码 (框图 72)。得到接收者 32 的回应后, 语音 - 邮件设备通过语音卡/调制器 40 播放语音信息 (框图 74)。

本发明的使用语音 - 邮件设备 14, 18 的系统也可以包括以下功能: (a) 设置密码来保护语音信息的安全和版权; (b) 在所要求的时间呼叫接收者, 以解决发送者和接受者的时差问题; 及 (c) 多播语音信息给

多个接收者。

本发明用到的递归最小二乘算法在作者黄大威的文章“基于快速 RLS 算法的对 μ 律 (A 律) 及 IMA ADPCM 的无损失压缩”里有详细解释。本发明接下来含有的相关参考全部来自该文。

递归最小二乘算法和游程长度编码方法可以实现在 μ 律 (A 律) 标准下的无损失压缩。无损压缩意味着给定一个原始输入, 产生的输出信号要和标准规定的完全一样, 同时却降低了压缩文件的比特率。为了保证信号的质量, 使用了与 μ 律 (A 律) 标准里相同的量化方法。使用预测和熵编码以降低比特率。预测是建立在快速递归最小二乘 (RLS) 算法上, 比现存的 RLS 算法需要的计算量更小。熵编码基于 Huffman 编码, 预测、量化和编码结合成一个适应的方案。对每个样点降低比特率都要保证相同的质量。比较每个样点 8 比特 μ 律或 A 律, 可以只用 3.24 比特每样点对 44,100Hz 采样频率的语音信号进行编码, 用 4.72 比特每样点对 11025Hz 采样频率的语音/音频信号编码。IMA ADPCM 标准也获得一些改进。

作为背景, μ 律和 A 律在 1972 年 ITU (CCITT) G.711 中被介绍到电话通信中。前面提到过, 根据一个对数公式, 编码标准把每个样点 16 比特的数字语音/音频信号转化成每个样点 8 比特。 μ 律 (A 律) 已经作为国际标准被用于国际或长途电话中。

本领域的熟练技术人员知道, 现在的压缩技术使用了以下三种方法: 预测、量化和熵编码。在 μ 律和 A 律里, 对数公式给出了量化步骤。对数律考虑了人类的感知特性, 并且它将原始信号中的 Laplace 分布 (此分布在零点附近有峰值出现) 转换成更平坦的分布。因而该统一编码方案 (对 256 个状态用 8 比特) 已经接近最优熵编码。本领域熟练技术人员对 RLS 算法的了解已经有一段时间了, 它优于一些简单的预测方法, 例如线性均方预测 (LMS)。但是它的应用不如 LMS 普遍, 因为它的计算量比较大。本发明极大地缩减了 RLS 算法的计算量。

递归最小二乘算法使用了结合预测、量化和熵编码的方案。此方

案使用和标准里一样的量化方法，用更复杂的预测和熵编码方法来降低比特率。和使用在传统的 PCM 编码中的无损压缩方法相反，此方案适应性更强。在现有技术中，有不同的方法出现在 256 样点模块的预测中。AR 模型的系数根据模块里的数据进行计算，它们对每个模块都是固定的。系数和残差都被编码用于存储和传输。

在本发明使用的递归最小二乘算法里，残差的计算由前面的预测器进行。预测器根据当时的观察数据进行更新，没必要对预测系数进行编码。只要过去的信号在解码端和编码端是相同的，系统就可以在编码和解码两端使未来信号和未来预测器做到同步。这种方法的优点就是系统可以在预测器中使用很多系数而不必增加码长。由于语音信号的非独立性，尤其是 44,100 采样频率的 CD 音质，长回归提高了预测的效率，降低了对残差编码时的比特率。

windows 平台的程序可以在这些算法基础上实现。可以在线对语音/音频信号进行编解码，并且在一台有 Pentium233kHz CPU 的笔记本电脑上对每秒 44,100 样点（CD 标准）进行实时编解码。

windows 平台的具有作为引擎 μ 律，A 律和 IMA ADPCM 压缩程序可以用 Visual C++ 语言写成。这些程序都可以下载、播放、编解码，也可以以声波格式被复制。下面的图表显示了不同采样的结果。这些

音频样本	Hz	大小(kBytes)	比特每 μ 律	样本ADPCM4
1	44,100	21,189	2.7	2.59
2	44,100	6,306	3.09	2.69
3	44,100	7,529	3.27	2.73
4	44,100	14,077	3.33	3.04
5	44,100	23,224	3.35	2.74
6	44,100	12,536	3.48	2.69
7	44,100	18,907	3.53	2.79
平均 语音/音频			3.24	3.75
1	11,025	6,482	4.36	3.08
2	11,025	12,942	4.63	3.19
3	11,025	12,920	4.7	3.22
4	11,025	12,920	4.74	3.15
5	11,025	38,007	4.82	3.17
平均			4.72	3.17

声音样本取自音乐光盘（CD）中的交响乐、歌曲等。这些语音/音频样本和可能的音乐背景，是从互联网下载的，被美国之音广播过。

从图表可以很明显的看出，软件比语音效果要好，因为音频样本是按 44, 100Hz 抽取的。对它们的预测只提前 1/44,100 秒，比语音样本的 11, 025Hz 实现要简单。另外，单纯的声音比有背景音乐的的声音更容易压缩。 μ 律压缩比比 IMA ADPCM 要高，可能是因为根据 μ 律再现的信号比 IMA ADPCM 再现的信号更容易使用 RLS 进行预测。

利用目前发展的信号处理技术，可以提高已经建立的、关于语音/声音压缩的国际标准中的压缩比，而不会在声音质量上有任何损失。

前面提到过，本发明应用语音 - 邮件设备 14, 18 和四个软件模型 32, 34, 36, 38 提供单向通信，给服务提供商一个高投资回报（ROI）。本发明扩展了互联网上最成功的应用电子邮件，更广泛的使用了电话终端，包括蜂窝电话和有线电话，允许更多的人不用计算机而使用本系统。本发明尤其适合中国和远东市场。

和长途电话、国际电话、互联网电话比较，本发明的系统不需要较大的投资。一封语音邮件的费用可以是互联网电话的 1/5 到 1/10，是中国电信市场标准的国际电话的 1/10 到 1/30。

下面的图表列出了目前中国长途、国际和互联网电话的通信收费标准。由下表可以很明显的看出本发明的价格优势所在。

中国大陆地区的长途电话	0.7 元人民币/分钟
VoIP* 中国大陆地区长途电话	0.3 元人民币/分钟
到香港、澳门、台湾的长途电话	2.0 元人民币/分钟
VoIP 中国大陆地区长途电话	1.5 元人民币/分钟
到美国、加拿大的国际电话	8.0 元人民币/分钟
VoIP 到美国、加拿大的国际电话	2.4 元人民币/分钟
到其他国家的国际电话	8.0 元人民币/分钟
VoIP 到其他国家的国际电话	3.2 元人民币/分钟
本地移动电话	0.4 元人民币/分钟
漫游移动电话	0.6 元人民币/分钟
长途/国际移动电话	上面费用加相应的固定电话费用

*VoIP 费用不包括本地呼叫费，即前 3 分钟 0.18 元人民币，3 分钟后每分钟 0.11 元人民币。

美金和元人民币兑换：1 美金 = 8.27 元人民币。

和文本的电子邮件相比，本发明的系统避免了字符的输入。这非常有利于中国的用户，因为中国文字的输入法很复杂。

本发明的另一个优点是，它便于实现漫游，对蜂窝电话漫游不要求注册本地代理和外地代理。对比 VoIP，它能提供高质量的话音。本发明中的系统传递语音信息时，不会出现停顿和时延。语音质量和国际标准 μ 律编码相同。

本发明的系统也可以增加无线信道的利用率。一个服务提供商在无线信道出现冗余时，可以分配冗余信道给语音电子邮件，这样就增加了信道利用率。进入到本地帐户的文本电子邮件也可以传递到蜂窝电话上。

如前所述，本发明利用普通的电话终端，提供低价格的国际和长途单向语音信息通信。根据本发明的第一个方面，本发明对现存的通信系统没有实质性的改变。软件是安装在一些网关上，例如，位于不同国家和城市的、带有语音调制解调器的个人计算机和笔记本电脑上。

根据本发明的第二方面，在移动电话的双向随机接入信道，当使用昂贵的无线通信资源出现信道冗余时，可以把单向的语音电子邮件分布在此信道中。通常，移动服务提供商只能使用 50% 的信道容量，因为现存的通信系统都是双向的。实时通信都被要求没有时延、随机接入。即将面世的第三代移动通信设备更加昂贵。单向的语音信息可以利用无线信道出现冗余时进行传输，即，传输优先级低于普通的无线通信。服务提供商利用此项低价格却非常方便的服务，可以吸引大量本地电话用户和国际或长途电话用户。

在 IP 上传递语音信息的基本思想是创造一种单向通信以代替分组交换系统中的复杂的和/或不可控的双向通信。这种思想可以扩展到其他方面。它在即将到来的多媒体信息服务 (MMS) (例如，www.nokia.com) 中会发挥重要的作用。在现在的和下一代的移动电

话系统中增加语音信息传递功能以充分利用信道容量；转寄电子邮件的内容发送到办公室计算机上或是家庭电话和移动电话上这些都是其中的例子。

图 4 显示的框图用来说明本发明的第二方面。图中，“MS”对应于移动台，例如蜂窝电话或其它无线电话（无线 PDA）。“BS”对应 GSM 网络中的基站，也可以是专业人士建议的其他网络的基站。“RNC”对应于 GPRS/UMTS 中的无线网络中心。“RTS”表示请求发送，“CTS”表示同意发送。M/M/B 是一个马氏排队模型，其中到达时间和服务时间都服从指数分布，而模型中最大信道容量是 B。这样，在用无线网络系统中的排队模型对未来信道容量预测的基础上，系统决定是否接收和/或发送任何语音信息，通常是单向语音信息。

如图示，发送者也即移动台 80 有一封压缩的或是编码的语音电子邮件。发出请求发送信号，接收到同意发送信号后，数字语音被传送到发送方的网关 82，这个网关可以是具有语音 - 邮件设备功能的某个基站或是某个无线网络中心。网关 82 通过对 M/M/B 模型的判断，及是否有剩余的信道容量来决定是否同意请求、是否发送信号。信息通过因特网 84 发送到接收方的网关 86，如专业人士所知，这个网关也可以是具有语音 - 邮件设备功能的某个基站或是某个无线网络中心。它通过对 M/M/B 模型的判断，及是否有剩余的信道容量来决定是否传输信息。该数字语音以解压解码后的语音电子邮件格式被传送到接收者的移动台或移动电话 88。

如本领域的熟练技术人员所知，进入某个排队等待序列可以是马氏序列，可以假定是平稳状态，可以忽略等待序列的非先进先出标准。其中某些进入序列可以按照 Poisson 准则、delta 速率处理，不同传输的传输时间按照设置的均值进行指数分配。

图 5 显示了本发明的第三方面，即提供以下服务给用户：从用户的本地帐户上获取电子邮件，然后使用文本 - 语音转换装置转换后收听，或者使用蜂窝电话短消息（SMS）收看。虽然一些 2.5 代、3 代移动电话可以从用户的本地帐户上接收电子邮件，但是这些费用是很昂

贵的。而且 GSM/IS95 系统也会在将来相当一段时间内工作。并且，固定电话仍然是价格低而被广泛使用的。本发明可以使用任何一种通用话机包括蜂窝电话和固定电话，通过使用贝尔实验室语音模型小组发展的文本转换语音（TTS）技术接收电子邮件。TTS 技术也可以广泛应用在例如有源移动网络等更多服务上，使用它，也可以做到把那些不太长的电子邮件用蜂窝电话短消息发送。

如图 5 所示，信件传送代理 90 通过因特网 92 传送文本电子邮件，在此有语音 - 邮件设备功能的网关 94 根据 (a) 信道容量和 (b) M/M/B 模型来决定是否传送信息。网关内的处理器使用 TTS 技术将文本转换成语音。然后数字语音以解压解码后的语音电子邮件或 SMS 格式传送给移动台/移动电话 96。

文本转换成语音系统的基本算法可以从美国专利 Nos.5,751,907 to Moebius et al.;5,790,978 to Olive et al.;6,272,464 to Kiraz et al. 中找到。以上所有属于朗讯科技。接下来文本中含有的相关参考，全部都来自以上专利。

众所周知，电子邮件是跟文本转换语音综合系统竞争激烈的文本格式之一，因为普通的文本经常参杂其他的格式例如表格、逐条列表或者用键盘字符做出的美术图案，很少有正规的设备可以把这些区域的边界区分开来。电子邮件文本也可以包含各种嵌入信息，传送材料和报表，这些都应该被检测出并被发送出去，帮助收听者接收附件。另外还有一些商标和电子格式的地址不能被 TTS 系统正确地处理。

朗讯科技发展的这个系统中，字符文本转换成语音（TTS）可以分成三个主要任务：语义分析，韵律模型，以及语音合成。语音合成把给定的语音表述，例如，一串带有语法、语调、强调信息的语音符号，利用一种合适的综合性的方法，转换成人工的、机器合成的声音。文本分析模块从字符文本中通过计算得到语言表述。

TTS 系统结构有多种语言合成器的功能。目前有效的是英语、法语、西班牙语、意大利语、德语。俄语、罗马尼亚语、中文、日语系统。这个多语言系统，是因为它语义分析、语音合成的基本软件对所

有语言是相同的，英语除外。一些特殊语义的信息是必要的；每种语言唯一对应一种声学目录，还有一些语义分析的特殊规则。但是，这些数据，存放在外部表格和参数文件中，当被使用时，TTS引擎才会把它们载入。这样，在例如对话或者电子邮件阅读等的应用中，才可以在运行时，把声音和语言按照原意转换。

该系统的多语言特性可以和一个文本处理程序相对照，文本处理程序提供特殊语言字体，允许用户几乎可以用任何一种语言编辑文件。考虑文本格式化和输出，相同的基本原则和选项被使用，而不管现在语言是怎样被处理的。多语言的文本转换声音合成器中的统一的软件结构有利于推动新语言的扩展，它的模块结构有利于增加现存系统中完善的组件之间的集成。

熟悉本领域的人员皆知，某些语言如中文不标记工作区的边界，系统必须重新构建这些边界。另外一些语言，例如俄语，有过于复杂的词法，它们必须被妥当处理，因为在俄语中词法会影响语调的表达位置。另外，不同的语言通常使用不同的书写系统。在中文中，系统必须处理中文字符文本，而在俄语中，必须处理成串的斯拉夫语字符。然而，用一种更抽象的观点也是可能的：每个语义分析问题可以被看作是一个从一串字符（例如拼音字符）到另一串字符（例如注释的语义分析）的转换。

字符串转换可以利用有限状态转换器（FSTs）被模拟计算。FST按其特征可以被描述为一个抽象设备，它包括有限个数的状态。每个状态向其它状态的转换都要依据一张表格，它决定哪个输入符号将被处理。表格也决定符号的输出。语言的描述，例如发音规则，由专家指定，它可以自动地被编辑成FSTs。在加权的FSTs中（WFSTs），权值（或者是价值）被加入到代选名单中，使我们可以做一有序的替换分析。最好的分析值会从替换值中被拣选出来。

我们也可以使用所谓的韵律模型，在这里，区间持续时间模块对每一个音素片断决定一个持续时间。在片断串的已被综合起来的基础上，每一个片断都被一个特征向量来标记，这个向量里包含了一系列

因素，例如片断的标识、音节的重音、口音的状态、片断内容或者是在一个短语中位置。一个重要要求是我们能够从文本中计算这些因素。构造持续时间模型可以分成两个阶段：对语音主体的相关和统计分析以及参数拟合。这个系统使用一个量化持续时间模型，这个模型是作为一个“和积”模型的特例，其中模型的参数符合一个分段的语音数据库。这种方案使用统计技术来处理混淆不清的因素以及这些因素有多少水平，还有数据不足的问题。这些分析对每一个特定的说话人产生一个复杂的有关持续时间特征的图案。

这个语调模块通过加入三种类型的与时间有关的曲线来计算一个基音频率等高线 (F0)。 这些时间相关曲线是：相位曲线，依赖于相位类型，例如，是陈述词对疑问词的曲线；口音曲线，每一种口音（经常是取决于重读音节和非重读音节的位置）都有自己的口音曲线；混乱曲线，在后辅音的元音中捕获阻塞音对音调的效果程度。这一方式与所谓的叠加语调模型使用同一些观念。这一系统能建立口音曲线是如何依赖于口音种类的组成和持续时间的详细的模型。因为收听者对音节序列中基音的微小变化都很敏感，所以这很重要。先前在时间和基音等高线高度方面的结果将整合到新的模型中。类似于构造持续时间模块，对这些依赖关系的建模涉及到对一个语音主体的参数配置。

该系统能凭藉将自然语音片段串联起来以产生合成语音。声学库中多数单位是双音素，例如，包含在两个毗连的语言片段之间的转变单位，它开始于第一个片段的稳定时期而终止于第二个片段的稳定时期。基于包括谱差异和能源度量的各种不同的标准，我们来选择语音库中储存的单位。

前后关系上的或压在一起的效果可能要求储存上下文敏感的“全音素”单位或三音素单位。对于声学库的构造，该系统用一个自动的最佳元素选择算法来决定取舍音素。对一个给定的元音，这一方式选择音素使得在共同元素之间的谱差异被那个元音减到最少，而所要求的元素涵盖达到极大。我们提供了一个工具组来帮助减少人工参与选择声学库中的元素。选择到声学库的元素先是被提取（“剪切”），然

后规范化它的振幅,再在总计中编入索引而且作为声学库元素储存。

单位选择和连锁模块选择并且连接声学库元素。这些模块检索所需的单位,决定新的持续时间,基音等高线和振幅分布,然后将这些参以矢量形式传递到合成模块。我们的语音合成系统使用矢量量化的线性预测编码(LPC)和参数化的音带声波来进行合成。

该系统能将电子邮件信息分解为下列概念上不同的阶段来进行分析:(i)分析和分解文本区域(打标记),(ii)文本的规范化(独立于装置的翻译),和(iii)文本合成(声音翻译)。

打标记阶段从尝试识别文本的关键区域开始。对于文件头来说,由于这里有相当可靠的线索,例如以“从哪里(From)”或“主题(Subject)”开始的行,问题是相对地简单。其他的情形则复杂得多,例如从纯文本或美国信息交换标准代码(ASCII)图形中区分表格。对于较复杂的情形,系统使用主体训练的统计模型来对每一行文本产生有加权的各种不同的可能分类,这些模型包括不同的符号类(照字母次序的符号,数字,和预先定义了的非字母符号)。我们对每一行给分来说明它与一个表格,一个ASCII图标,一个签字,或一个简单的文本的规范行相似的程度。一个文本块被定义为(概略地)由一条或多条空白行所界定的任何区域。对于每个文本块,系统进一步地限制所有的行必须属于相同的类。然后系统选择得分最高的类做为这一文本块的类。在检测到区域之后,它们被列入一个分阶层的文档结构,这里遵从标准广泛结构语言(SGML),分层中的每个节点被注明分类标签和属性。

装置-独立的翻译尝试使非标准文本材料的各种不同的类型规范化,包括电子邮件地址,统一资源定义符(URL),以及混合商标如WinNT。一个例子是电子邮件地址 brsnyder@netcom.com。“@”应该被读当做“在于”(不是,例如,“at符号”),而那个“.”应该被读如“点”(不应被跳越)。此外,称号brsnyder应该断开如b r snyder。这个规范化由一个知道电子邮件地址结构的有限状态转换器来实现。在这个转换器中嵌入了一个经由很好形成地英文单字主体训练的有限状

态模型。这个模型将会发现 `brsnyder` 是不太可能是拼字正确的字而且建议把它分解为 `b r snyder`。

在最后的阶段，声音翻译从有记号及规范化文本处开始，而且插入适当的标签告知 TTS 系统将会如何发音。除了其他的事物之外，这一翻译阶段还将会决定使用哪一组被预先定义的声音于例如一个被引述的区域，使当一个被引述的区域被发现的时候，TTS 将会转变到这种比较能帮助收听者了解文件的结构的聲音。

显然，本发明是先进的并且提供了一种系统和方法使得使用者得以经过因特网传递长途和国际语音信息给接收者，它克服了现有技术的一些不足之处。本发明中的语音-邮件设备允许使用者经过因特网传递这些长途和国际语音信息给接收者而不用直接使用计算机。该系统既便宜又方便，它支持移动通信并且提供高质量音频信号的语音通信。它提供了一个单向通信系统。发送者和接收者都可以用与当地的语音-邮件设备相连接的标准电话进行通信。作为本发明的一个方面，信息在本发明的两个语音-邮件设备之间以电子邮件的方式传送。

一个熟习本领域的人员将会从前述和相关图表中联想到与本发明有关的许多改变和其他具体方式。因此，我们理解本发明是不受限制于特定的描述和具体实施例的。权利要求书的范围包括了各种改变和实施例。

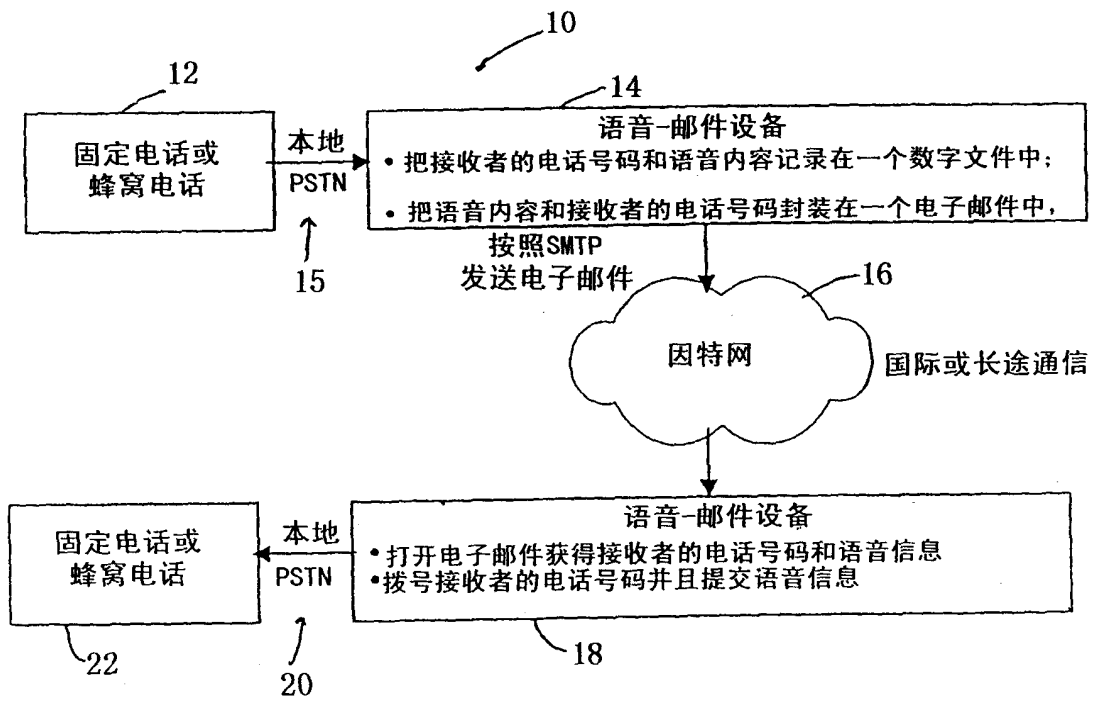


图1

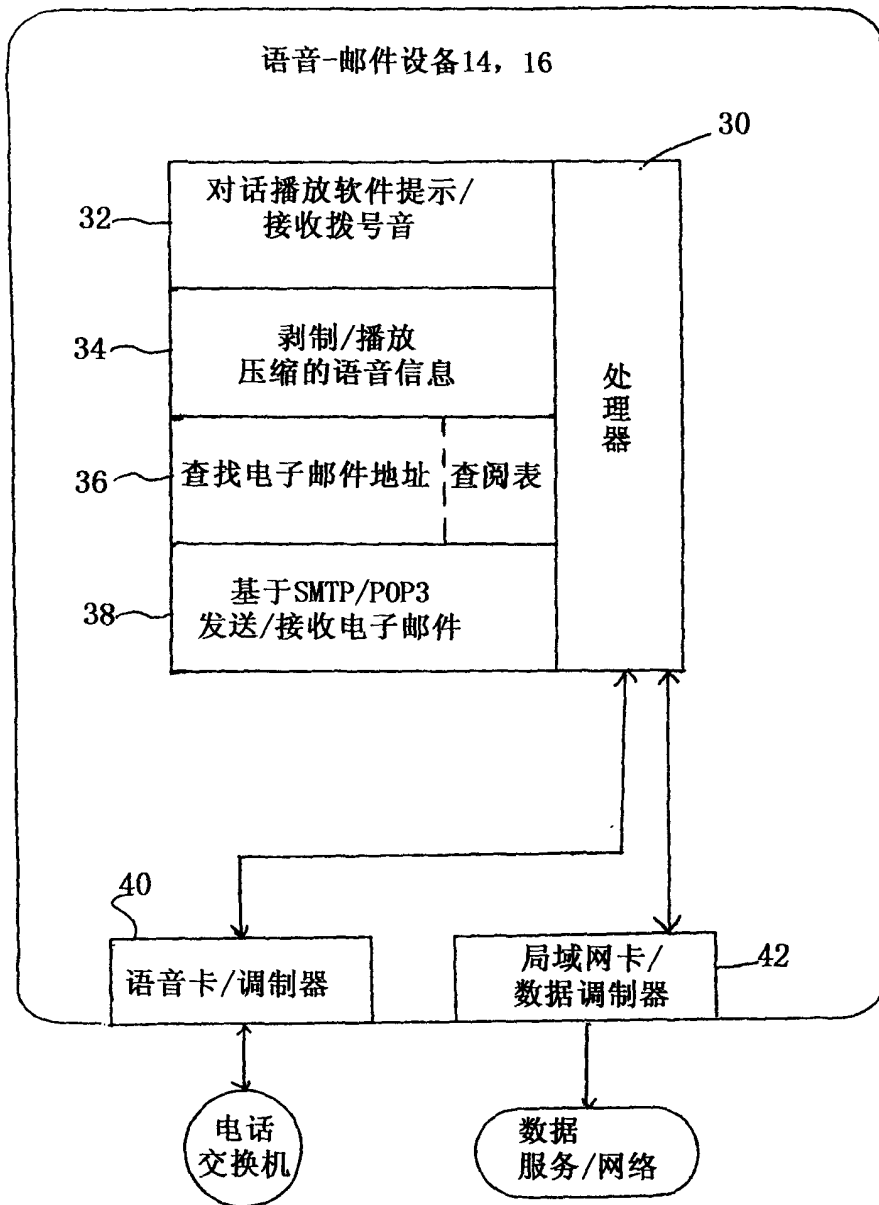


图2

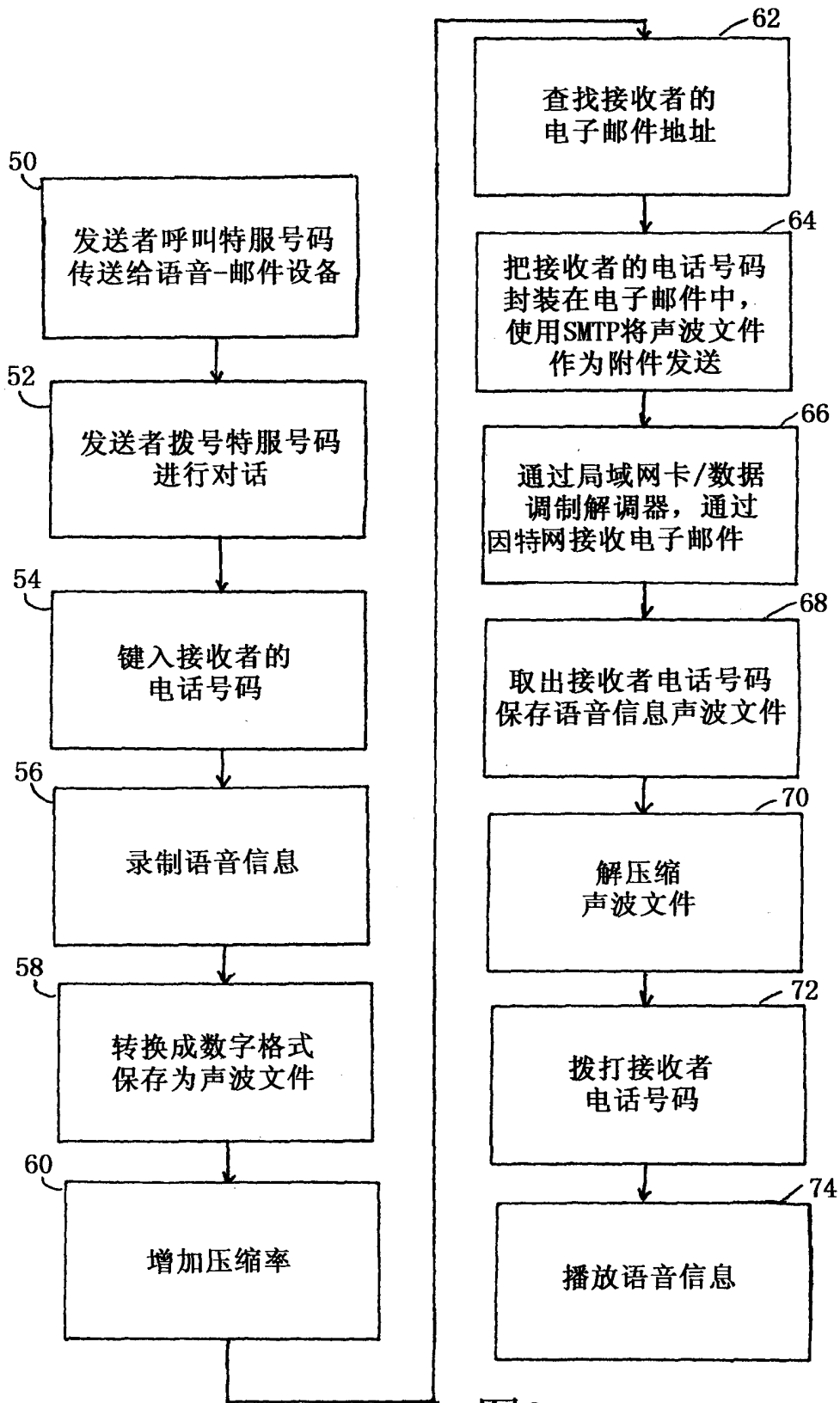


图3

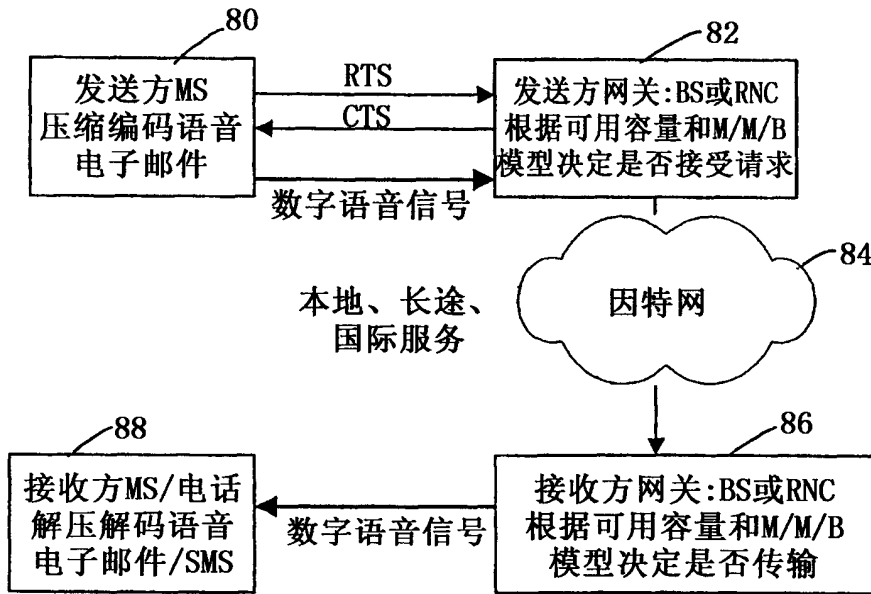


图4

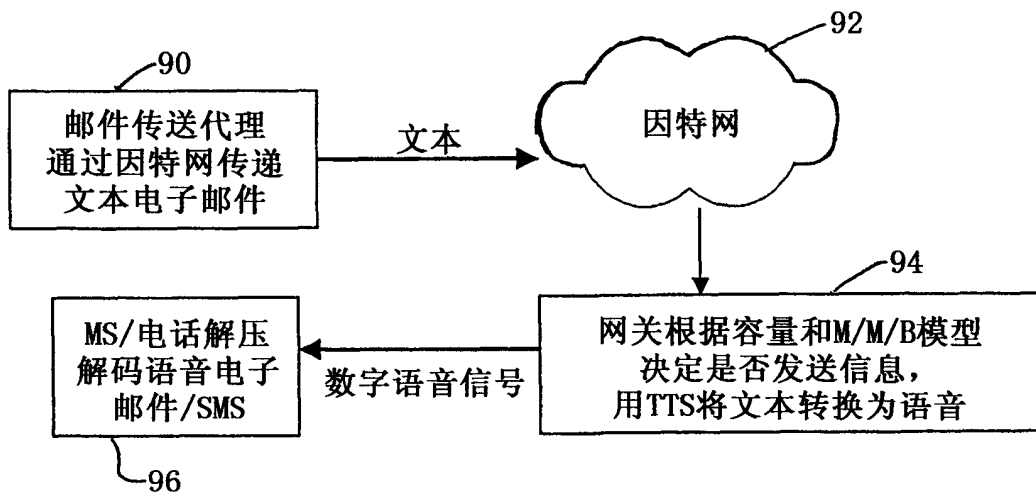


图5