



[12] 发明专利申请公开说明书

[21]申请号 94104665.6

[51]Int.Cl⁵

G10L 7/08

[43]公开日 1995年3月8日

[22]申请日 94.4.25

[30]优先权

[32]93.4.30 [33]US[31]056,044

[71]申请人 美国电话电报公司

地址 美国纽约

[72]发明人 迈克尔D·拉宾

马克斯S·肖艾弗勒

[74]专利代理机构 中国国际贸易促进委员会专利商

代理人 陆立英

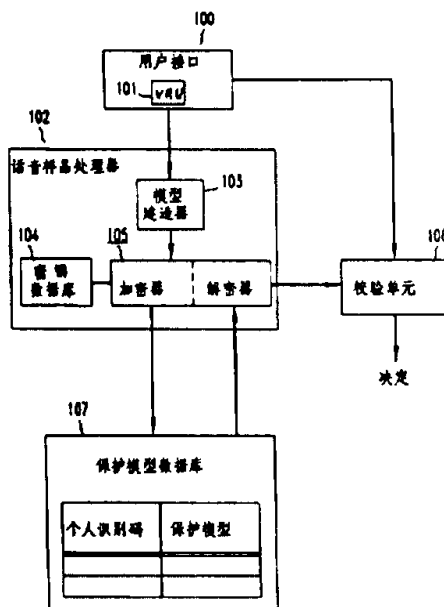
说明书页数:

附图页数:

[54]发明名称 讲话者校验系统和方法

[57]摘要

本系统包括用户接口在预约和接入过程分别提出个体和申报人提供其话语音样品和个人识别码; 话音处理器在预约过程处理个体话语音样品并由其内的话音模型建造器和加密单元利用密钥数据库产生相应的保护话语音模型, 在接入过程对保护话语音模型解密, 讲话者校验单元判定申报人的话语音样品与个体的保护话语音模型的解密形式是否匹配, 若匹配, 则验证了讲话者的身份; 以及保护模型数据库 (107) 用以存储保护模型和有关的个人识别码。



权 利 要 求 书

1. 一种用于确认一个个体申报的身份的方法，所述的个体具有一个个人识别码，所述的方法其特征在于包括以下步骤：

接收(301)该个体的话音的样品；

产生(303)上述话音的保护模型，所述的保护模型被设置得可使所述个体的话音的特征不能根据所述的保护话音模型来确定出来；

所述的保护模型与上述个人识别码结合在一起被存储(305)；

响应对身份确认的一个请求，接收(401)由上述申报人提供的申报人的话音样品和个人识别码；

检索(402)与上述申报人提供的上述个人识别码结合在一起的模型；

执行(403)比较，以确定所述的检索的模型是否与上述申报人的话音样品相对应；

只在上述比较表明肯定匹配时，确认(404)上述申报人的身份。

2. 根据权利要求1所述的方法，其特征在于，所述的产生步骤包括变换上述个体的话音样品的步骤；

所述的比较步骤包括变换上述申报人的话音样品的步骤，该

步骤与上面最后提到的步骤中使用的交换相同。

3. 根据权利要求 1 所述的方法,其特征在于,所述的产生步骤,包括产生具有所述个体的话音样品特性的模型和利用加密算法对上述的特性模型加密的步骤;

所述的比较步骤包括利用与上述加密算法互补的解密算法对上述特性模型解密的步骤。

4. 一种用以提供讲话者校验的方法,其中已存储的话音样品没有被误用的危险,其特征不在于包括以下步骤:

从每个用户取得一个话音样品;

处理上述话音样品,以产生一个保护的话音模型,该模型对于提供上述话音样品的用户是独特的,但是如果没附加的信息该模型不能再用于讲话者的校验;

上述保护的话音模型与识别该特定用户的密钥结合在一起来存储,所述保护的话音模型是从该特定用户取得的。

5. 一种利用话音校验来确认一个个体户申报的身份的系统,所述个体具有一个个人识别码,其特征不在于:

接收装置,用于接收上述个体的话音样品,上述个体的个人识别码、申报人的话音样品和申报人提供的个人识别码;

产生装置,用产生上述个体的话音的保护模型,上述保护模型被设置得可使上述个体话音的特征不能根据所述的保护模型来确定出来;

存储装置,用于将所述保护模型与所述个体的个人识别的结合在一起存储,和响应身份确认的请求检索与所述的申报人提供的所述个人识别码在一起的保护模型;和

执行装置,用于执行比较,以确定上述检索的模型是否与所述申报人的话音样品相对应,只在上述比较表明肯定匹配时才确认所述申报人的身份。

6. 权利要求 5 所述的系统,其特征在于,所述产生装置包括变换装置,利用变换对上述个体的话音样品和上述申报人话音样品进行变换,该变换使上述样品的重新排列以致使一个授权的人也不能使用。

7. 根据权利要求 5 所述的系统,其特征在于,所述的产生装置包括:产生具有上述个体话音样品特性的模型的装置和利用加密算法的所述特性模型加密以和利用与上述加密算法互补的解密算法对所述特性模型的解密的装置。

8. 一种用于提供讲话者校验的装置,其中存储的话音样品不致被错用,其特征在于:

从用户取得话音样器的装置;

处理装置,用以处理上述话音样品,以产生一个保护的话音模型,该保护的话音模型对上述用户是独特的,但是在无附加信息时该模型不能进一步用于讲话者校验;和

存储装置,用以存储上述保护的话音模型,该保护的话音模型

以后用于与从申报自己所述用户的人得到的语音样品进行比较。

9. 根据权利要求 8 所述的装置, 其特征在于, 所述的处理装置包括一个加密装置。

10. 根据权利要求 8 所述的装置, 其特征在于所述处理装置包括不可逆的变换。

说 明 书

E9A0175

讲话者校验系统和方法

本发明涉及讲话者校验系统和方法，利用被存储的话音样品确认一个人个体 (*individual*) 的身份 (*identity*)，特别是涉及一种用于产生和存储话音样品和检索 (*retrieve*) 该话音样品的系统和方法，以便后来在该讲话者校验过程中使用时使这些样品不致被误用。

讲话者校验方法和系统用于确认一个个体的身份，以便控制或限制利用一种资源。从这个意义来说，一种资源可以是电话网络、银行帐簿或计算机。一般地讲，一个个体话音样品的一些特征称为“讲话人的发声 (*speaker utterance*)”，用以形成一种样板 (*template*) 并被存储起来。当一个人申报自己是寻求利用资源的个体 (下文称为“申报人”) 时，该申报人需要提供一个话音样品。将该申报人提供的话音样品的特征与存储的样板或模型比较，只在成功匹配时才确认其身份。

讲话者校验系统的使用者业已关注被存储的话音样品可能会被错用和/或误用而带来的可怕后果。这是因为话音样品所涉及的人就不能“排除”或/和以个人识别号 (*PIN*) 相同的方式不能改变他们

的个人话音特征提供一个新的样品的缘故。当一个 PIN 被涉及时，可给那个个体分配一个新的个人识别号，而旧号被删除或无效。可是，如果一个小偷窃取了一个特定人一个话音样品或话音样板或模型，则该小偷实际上可“变成”那个个体，至少用于向一个话音校验系统识别他自己的目的。

根据本发明，为了讲话者校验目的而得到的一个个体话音样品可用于产生该话音的“保护”模型。该保护模型被存储在与那个个体的个人识别码有关的一个数据库中，该保护模型可被设计得能使该人的声音特征不能从该保护的模型中被确认出来，而无需利用已存储在该系统中的一个密钥或其它信息。

当请求校验申报人身份的请求收到时，亦即一个人申报自己是一个特定的个人时，就要取得该申报人的话音样品，并检索该申报人提供的个人识别码有关的保护模型，利用已存储的密钥来处理该模型，或者利用已存储的信息来处理该话音样品，然后进行比较，以确定被检索的模型是否与申报人话音样品相对应。

在本发明的一个实施例中，一个个体的话音样品用于以一种常规的方式建立一个话音模型或样板，然后利用仅对该系统是已知的一个密钥对模型加密，以产生一个保护的模型，它与那个个体的个人识别码结合在一起地被存储。当一个人申报自己是那个个体寻求利用一种资源时，该申报人提供的话音样品与保护的模型的解密型式进行比较，只有在成功匹配时才验明其身份。由于话音模型是以

加密形式存储的,而且需用密钥来解密,这样,误用小偷无益。

在本发明的另一个实施例中,一个个体的话音样品须在变换处理器中历经变换,该处理器的输出施加在模型发生器上,以产生一个保护的模型,该保护模型与那个个体的个人识别码结合在一起被存储。当一个申报人寻求身份验明时,该申报人提供的话音样品也使用相同的变换进行变换,并将变换结果与使用由该申报人提供的个人识别码检索的已存储的保护模型进行比较。只在成功匹配的情况下才提供验明。由于已存储的模型已历经变换,而且对申报人的话音样品也作了相同的变换,因此误用一个未经许可的个体不利。此外,即使保护模型被窃,其危害也可补救。该变换可以更改,而且那个个体可提供另一个话音样品,以产生另一个保护模型。然后使被窃的模型无效。该变换最好是“不可逆的”,以使一个个体的话音样品不可能根据被窃的保护模型再创生出来。

下面参照附图阅读对本发明的详细叙述就能对本发明更明了。

图 1 和 2 是根据本发明的原理设计的讲话者校验系统的两个实施例的方框图,该系统用于产生和存储被保护的话音模型,然后使用该模型控制利用资源;图 1 的实施例使用了加密/解密技术,而图 2 的实施例使用了变换技术。

图 3 和图 4 分别示出图 1 所示的本发明的第一实施例中执行的预约和接入过程的步骤的流程图;和

图 5 和 6 示出图 2 所示的本发明第二实施例执行的讲话者校验

方法的预约和接入过程的步骤的流程图。

首先参照图 1,该图示出按照本发明原理设计的讲话者校验系统的第一实施例的方框图,该系统用于产生和存储不同个体的保护话音模型,然后使用这些模型验明申报自己是那些个体的一个特定个人的身份。一般地讲,图 1 的系统执行两个主要过程,下文称为“预约”过程和“接入”过程。“预约”过程执行一次,例如在一个特定的个体预约或登记以便在此后能进行讲话者检验的时候执行。预约可在为了使一个人能得到购买信用卡的许可或者获得利用一个远端计算机或者始发长途电话呼叫的目的时发生。在预约过程期间,那个个体提供一个话音样品,并给该个体分配一个个人识别码。之后,在“接入”过程期间,申报自己是一个特定个体的人(申报人)的身份被验明。这个过程例如在每当一个人寻求利用一种资源如打电话或购买信用卡或使用远端计算机或简单地使另一个人确信他的/她的身份时执行。

图 1 的系统包括一个用户接口 100,在预约和接入过程期间都使用它。但是,在这里应该注意,也可以提供单独的用户接口以便在每一个过程期间使用。用户接口 100 包含一个交互式声音响应单元(VRU)101 例如可从 AT&T 公司购到的“*Conversant(R)Voice Response System*”,它具有以下能力:(a)播放通知,(b)收集由个人输入的信息,包括话音样品和例如经过按键盘输入的其它信息,以及(c)根据接收的响应执行预编程的逻辑操作。在预约过程期间,

VRU101 提示个体提供话音样品并分配给那个个体一个个人识别码以在后来的接入过程中使用。在接入过程期间,VRU101 提示申报人提供个人识别码和话音样品。

由用户接口 100 收集的信息施加到话音样品处理器 102 上,设置话音样品处理器 102 用以在预约过程期间处理从个体那里收集的话音样品并产生相应的保护的话音模型。在接入过程期间,处理器 102 还解密该保护话音模型,这将在下文更详细讨论。

对于建立一个保护话音模型而言,处理器 102 包括一个话音模型建造器 103,设置该建造器 103 用于接收和处理由 VRU101 收集的原始话音数据,以便得出一个独特的“声音复制品(voiceprint)或话音模型,它以足以把那个个人的话音与其他人的话音相区别的方式表现出那个人话音的特征。这个话音模型提供给加密单元 105 的加密部分,加密单元 105 还包括一个解密部分。在加密单元 105 中,可根据众所周知的数据加密标准(DES)技术或其它方法,利用存储在相关的密钥数据库 104 中的密钥进行加密和解密。不同的密钥例如根据分配给每个人的个人识别码可分配给不同的个人。

话音样品处理器 102 的输出施加在并且存储在数据库 107 中,该数据库可以是一个常规的数据库存储器,用于存储保护模型和有关的个人识别码。在接入过程期间利用个人识别码作为查阅密钥来询问数据库 107 时,相关的保护话音模型被检索出来。

在接入过程期间,如下文详细叙述的,利用讲话者校验来确定申

报自己是一个特定的个体的人是否实际上是他所说的那个个体。为此目的,讲话者校验单元 106 用以确定由申报人提供的语音样品与申报人申报自己是那个个体的保护语音模型的解密型式之间是否匹配。如果由校验单元 106 进行的比较结果表明匹配,则可产生一个信号以表示肯定的讲话者验明。这个信号可用于授权使用一种资源,例如电话网络或用于其它目的。

在本发明的第二实施例(示于图 2)中,用户接口 200、VRU201、数据库 207 和校验单元 206 都与图 1 中类似的元件相对应。在这个实施例中,语音样品处理器 202 包括变换单元 203,用以产生由 VRU201 得到的变换型式的话音样品。该变换可使用模拟的或数字的技术来实现,最好设计得,在语音样品已被变换之后使原始语音样品不可能根据该变换型式再创生出来。本领域的技术人员可得出实现这种变换的各种变换形式。在本发明的这个实施例中,设置检验单元 206 用以在来自申报人申报自己是一个特定的个体的、变换的话音样品与那个个体的已存储的变换的话音样品(保护模型)之间确定是否匹配。

参照图 3 和 4,该图分别示出图 1 所示的本发明的讲话者校验过程第一实施例所执行的“预约”和“接入”过程的步骤的流程图,用于产生和存储一个保护语音模型和在以后使用该模型来验明申报自己是一个特定个体的人的身份。在图 3 和图 4 示出图 1 执行特定步骤的元件在叙述该步骤的长方形框的右上角标示出其标号。

预约过程从步骤 301 开始,在该步骤中那个人与用户接口 100 或者经过电话线遥远地或者本地地例如使用连接到该系统的麦克风进行互相作用。在该步骤中,VRU101 提示那个个体提供一个或多个有代表性的话音样品,例如通过要求那个人说一个特定的词组就可得到一个样品。另一个可替代的方案是,对于利用询问与应答接入控制方式的更复杂的应用中,VRU101 可被设计得用于请求那个人以那个人的声音提供那个人双亲的中间名字、那个人的社会保险号码和他的(或她的)配偶、数字 0 至 9 的样品和其它类似信息。

接着,在步骤 302,分配一个个人识别码例如 PIN,这对该个人是独特的,或者对诸如一个家庭之类的一个个体小组共享的。

将步骤 301 得到的话音样品施加到话音样品处理器 102 上,以利用原始话音样品产生该话音的“保护模型”。在本发明的这个实施例中,保护模型是在两步骤(即步骤 303 和 304)中产生的。首先,在步骤 303,话音样品在模型建造器 103 中进行处理,以产生可被认为是一个独特的“声音复制品”的东西,用以代表用户话音的特征可从其他人的声音复制品中区分出来一个个人的话音样本(*pattern*)。业已提出了很多特征提取算法,并且可在这个步骤中使用,例如在 1985 年 11 月、卷 73、第 11 期的 *proceeding IEEE* 第 1651—1664 页 *Doddington, G. R.* 的讲话“*Speaker Recognition, Identifying People by their Voices*”和 1990 年“*International Conferencl of Acoustic Speech and Signal Processing*”第 269 页 *A. E. Lee, C.*,

Soong, F. K. 的 “Sub—word Unit Talker Verification Using Hidden Markov Models” 文章中所描述的那样。

在步骤 303 中产生的话音模型在步骤 304 中施加在加密单元 105 上, 以使该模型加密, 因而产生保护的话音模型, 这可防失窃, 而且在不使用存储在密钥数据库 104 中的密钥的情况下不能使用该保护模型。加密可根据数据加密标准 (DES) 技术或任何其它合适的技术来执行。在步骤 305 中, 保护的话音模型存储在数据库 107 中并与个人识别码 (在步骤 302 分配的) 结合在一起, 当在接入过程期间数据库 107 顺序地被查询时, 个人识别码用以作为查找密钥。

现在参看图 4, 当申报自己是先前预约个体的一个人请求使用图 1 的系统来做讲话者验明时, 在步骤 401 中该申报人向用户接口 100 提供他 (或她) 的个人识别码和话音样品。在步骤 402 中, 利用个人识别码从数据库 107 中寻找和检索该申报人申报自己的是的那个个体的保护话音模型。然后, 在步骤 403, 将该保护话音模型施加在加密单元 105 上以便解密。根据个人识别码从密钥数据库 104 中获得合适的密钥, 并在解密过程中使用它。在步骤 404, 加密单元 105 的输出与该申报人提供的话音样品一起提供给检验单元 106, 该加密单元 105 的输出代表从先前提供话音样品的那个人得到的原始话音模型。检验单元 106 将该话音样品与解密后的模型钥相比较, 以确定是否匹配。如果比较结果是肯定的, 这表示非常匹配, 则申报人可被确定是先前进行预约过程的那个个体。另一种情况是, 如果比较结

果是否定的,则拒绝验明。

在图 2 所示的本发明的实施例在预约和接入过程中执行的步骤分别示于图 5 和图 6 中。在图 5 所示的接入过程中,在步骤 501,以类似于上文提到的步骤 301 叙述的方式获得该个体的代表的话音样品,并在步骤 502 分配一个个人识别码。接着,在步骤 503,该话音样品提供给模拟的或数字的变换单元(图 2 中用标号 203 表示),以产生该话音样品的一种型式,这种型式不能被一个未授权的人士用来产生该个体的话音样品。例如,可采用一个简单的时间分隔的端点声频信号(*a simple time slice of the endpointed acoustic signal*),而且该时间分隔按照预定的安全的序列重新安排。该变换最好是不可逆的,亦即,一旦该话音样品已被处理,该转换后的型式就不能再被处理以获取该话音样品。根据那个个体的个人识别码使一种特定的变换与一个特定的个体相关联的信息存储在变换 ID 数据库 204 中。

在步骤 504 中,将变换后的话音样品提供给模型建造器 205,以提取重要特征,用以建立该个体的话音模型,以在接入过程中使用。在步骤 505,将保护的模型与在步骤 502 中分配的个人识别码一起存储在数据库 207 中,个人识别码在接入过程期间在查询数据库时用以作为查找密钥。请注意,在一个未授权的人士从数据库 107 已获取存储的信息时,他不能用该信息来模仿提供该样品的那个个体,为该话音样品先前已被变换而且不像那个个体户的实际话音了。

在图 6 的接入过程期间,在步骤 601,申报自己是一个已授权的个体的那个人提供一个话音样品和个人识别码。在步骤 602,该话音样品提供给变换单元 203,并将该单元 203 的输出提供给校验单元 206。在步骤 603,使用申报人提供的个人识别码从数据库 207 中提取一个保护的话音模型,并将这个保护的话音模型也提供给校验单元 206。然后,在步骤 604,将申报人的变换后的话音样品与已检索的保护模型相比较,以确定是否匹配。如果比较结果是肯定的,这表示非常匹配,则可确定请求讲话者验明的人士是先前进行预约过程的那个个体。另一种情况是,如果比较结果是否定的,则申报人的身份不被确认。

对本发明可作出各种变型和修改。例如,可对该系统提供分开的用户接口,以在接入过程和预约过程期间单独使用。还要注意,本领域的技术人员可设计适应的学习技术,其中的话音模型可在预约过程期间构成和存储,在接入过程期间更新,以适应变化的声音特性。此外,尽管本文这里叙述的本发明实施例参照个人帐户识别码,用以在该系统和/或方法的多个用户中进行区分,但是本领域技术人员应该理解:本发明可适用于一个用户系统,诸如用于验明持有“智能卡”(smart card)的人的身份而设置的系统,该卡本身含有从该人取得的话音样品,在这样的系统中,智能卡本身就是存储保护模型的数据库,这就不需要单独的数据库(107 或 207),也不需要个人识别码了。

图1

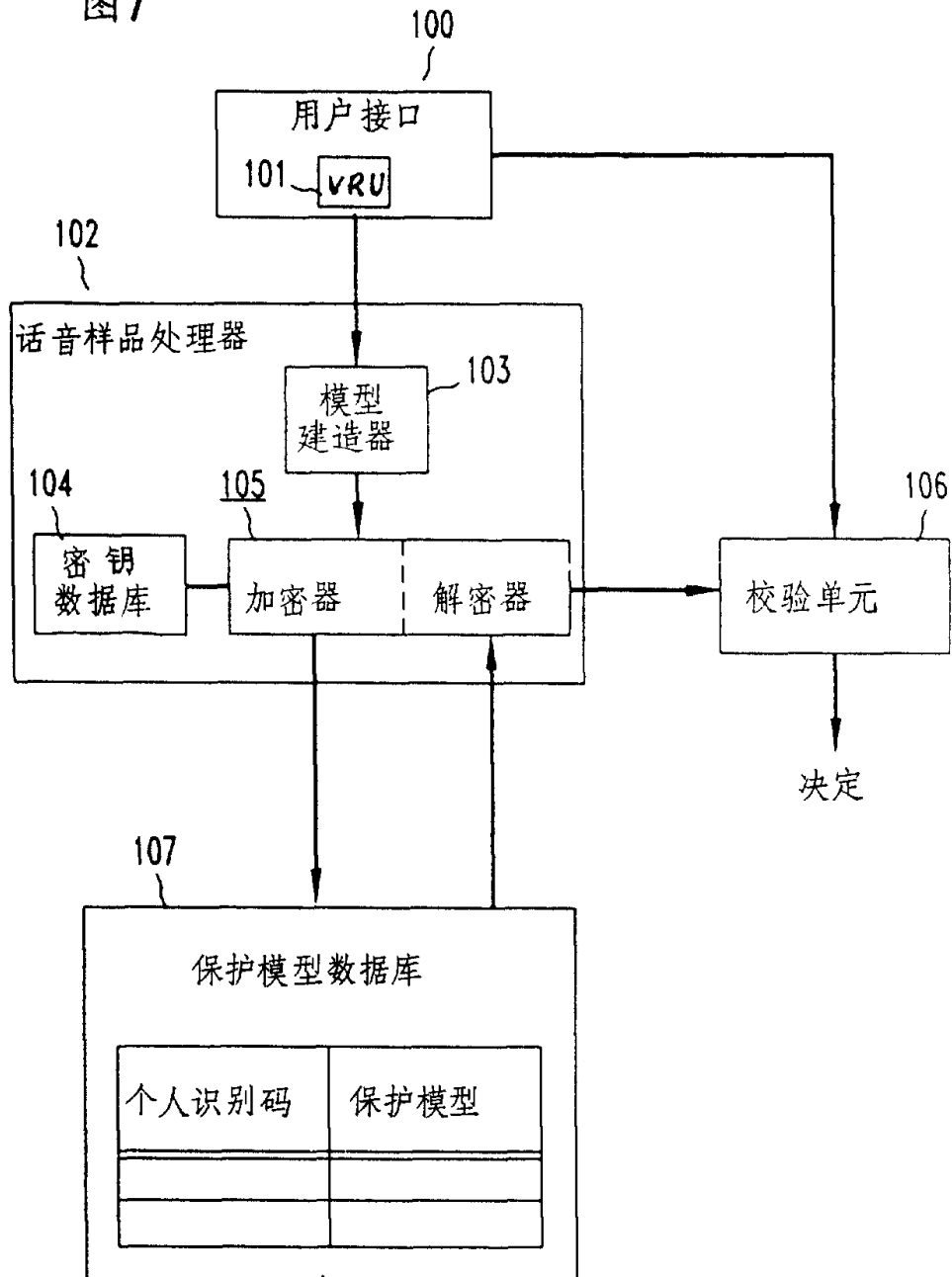


图 2

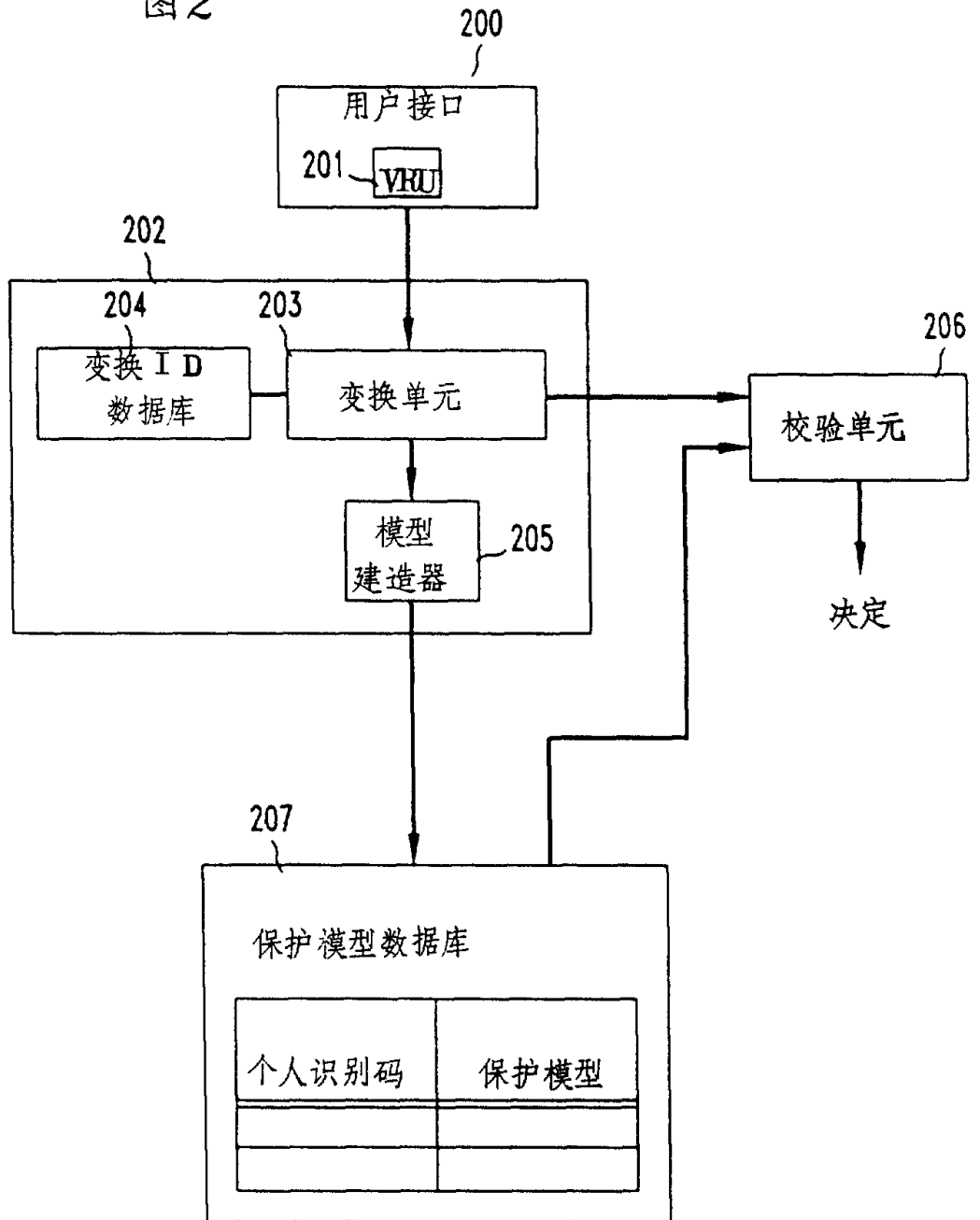


图 3

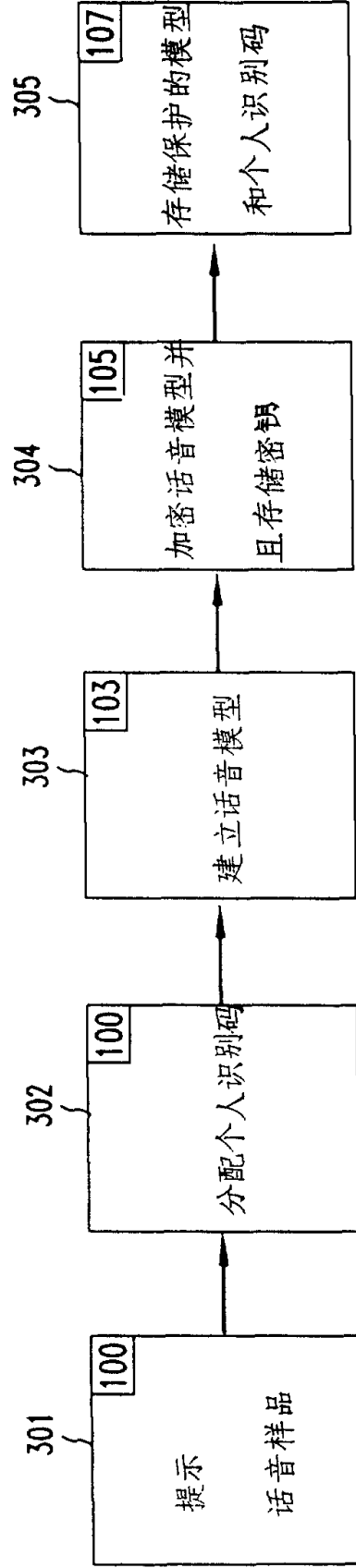


图4

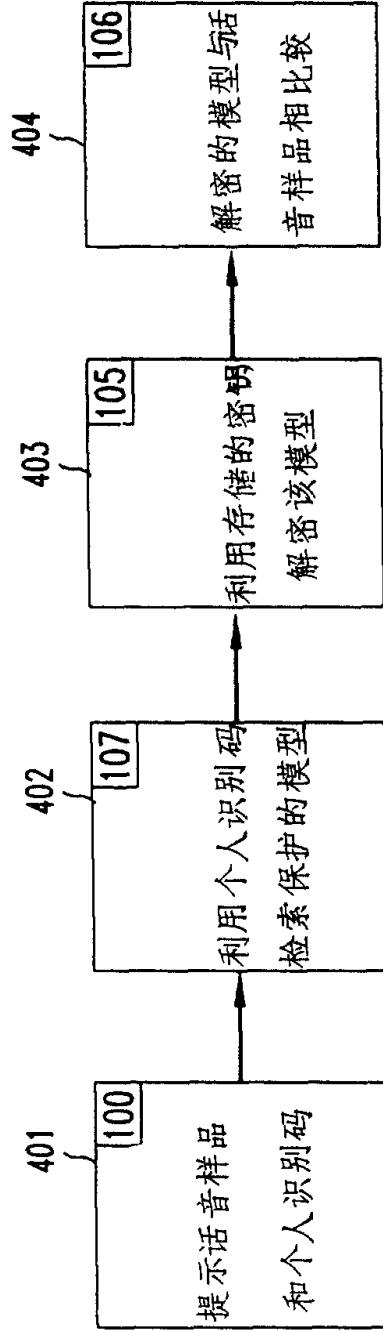


图 5

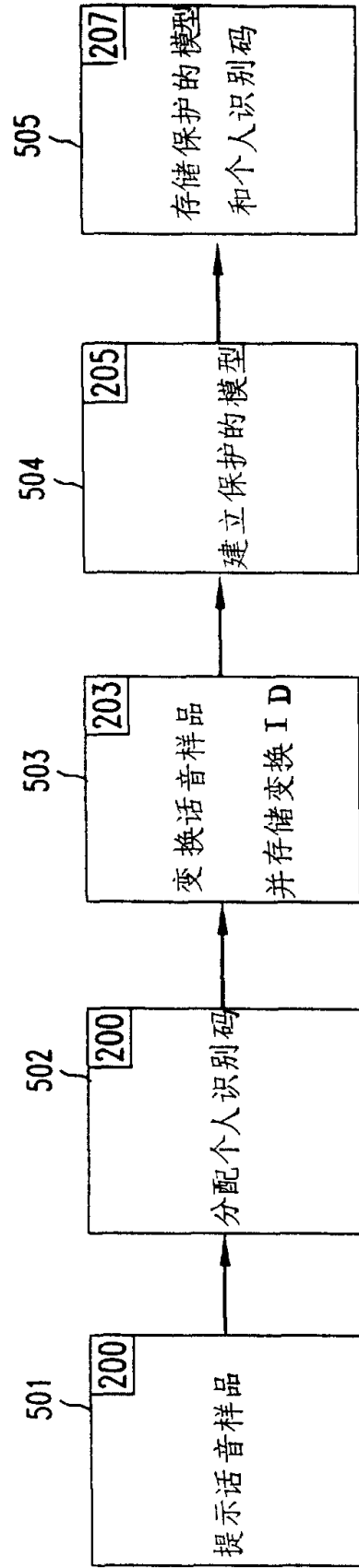


图6

