



(12) 发明专利申请

(10) 申请公布号 CN 104462912 A

(43) 申请公布日 2015. 03. 25

(21) 申请号 201410438750. 7

(22) 申请日 2014. 08. 29

(30) 优先权数据

14/030, 919 2013. 09. 18 US

(71) 申请人 联想(新加坡)私人有限公司

地址 新加坡新加坡城

(72) 发明人 约翰·韦尔东·尼克尔森

斯蒂文·理查德·佩林

(74) 专利代理机构 北京集佳知识产权代理有限

公司 11227

代理人 朱胜 李春晖

(51) Int. Cl.

G06F 21/32(2013. 01)

G10L 17/24(2013. 01)

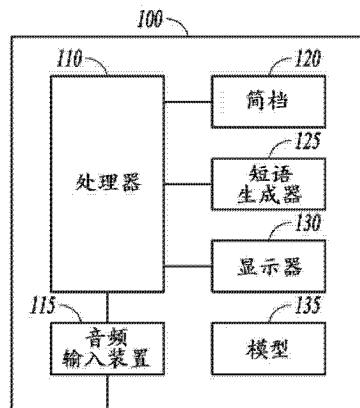
权利要求书2页 说明书5页 附图2页

(54) 发明名称

改进的生物密码安全

(57) 摘要

本申请涉及改进的生物密码安全，其中，一种系统包括：存储在计算机可读存储装置上的用户语音简档，该语音简档包含具有针对音素的用户识别特征的多个音素；以及语音处理器，该语音处理器被耦接以访问语音简档，从而基于针对该音素的用户识别特征与平均用户识别特征之间的差来生成包含用户区别音素的短语，以使得该短语具有与其他用户的可区分性。语音处理器还可以或者可替换地根据环境噪声选择短语。



1. 一种系统，包括：

存储在计算机可读存储装置上的用户语音简档，所述语音简档包含具有针对音素的用户识别特征的多个音素；以及

语音处理器，所述语音处理器被耦接以访问所述语音简档并且基于针对所述音素的用户识别特征与平均用户识别特征之间的差来生成包含用户区别音素的短语，以使得所述短语具有与其他用户的可区分性。

2. 根据权利要求 1 所述的系统，还包括音频输入装置，所述音频输入装置接收用户所说的生成的短语。

3. 根据权利要求 2 所述的系统，还包括显示器，所述显示器被耦接以接收所述生成的短语并将所述生成的短语显示给所述用户。

4. 根据权利要求 1 所述的系统，其中，所述语音处理器还经由音频输入装置接收用户所说的生成的短语并且验证所述短语是否由所述用户所说，以确认所述用户的身份。

5. 根据权利要求 1 所述的系统，还包括存储在计算机可读存储装置上的短语的库，并且其中，所述语音处理器通过搜索所述库中的、包含用户区别音素的短语来生成所述短语。

6. 根据权利要求 5 所述的系统，其中，生成的短语包含至少两个用户区别音素。

7. 根据权利要求 5 所述的系统，其中，生成的短语包含用户区别音素的至少两个实例。

8. 一种系统，包括：

音频输入装置，所述音频输入装置接收来自用户的语音，所述语音包括所述用户附近的环境噪声；

语音处理器，所述语音处理器被耦接至所述音频输入装置，以根据所述环境噪声生成包含用户区别音素的短语，从而使得所述短语具有与所述环境噪声的可区分性。

9. 根据权利要求 8 所述的系统，其中，所述语音处理器根据针对所述音素的用户识别特征与平均用户识别特征之间的差来选择所述短语，以使得所述短语具有增强的与其他用户的可区分性。

10. 根据权利要求 8 所述的系统，还包括：

音频输入装置，所述音频输入装置接收所述用户所说的生成的短语；以及

显示器，所述显示器被耦接以接收所述生成的短语并且将所述生成的短语显示给所述用户。

11. 根据权利要求 8 所述的系统，其中，所述语音处理器还经由所述音频输入装置接收所述用户所说的生成的短语，并且验证所述短语是否由所述用户所说，以确认所述用户的身份。

12. 根据权利要求 8 所述的系统，还包括存储在计算机可读存储装置上的短语的库，并且其中，所述语音处理器通过搜索所述库中的、包含用户区别音素的短语来生成所述短语。

13. 根据权利要求 12 所述的系统，其中，生成的短语包含至少两个用户区别音素。

14. 根据权利要求 12 所述的系统，其中，生成的短语包含用户区别音素的至少两个实例。

15. 一种方法，包括：

接收用户附近的环境噪声信号；以及

根据所述环境噪声生成包含用户区别音素的短语，以使得所述短语具有与所述环境噪

声的可区分性。

16. 根据权利要求 15 所述的方法,其中,还根据针对所述音素的用户识别特征与平均用户识别特征之间的差来选择生成的短语,以使得所述短语具有增强的与其他用户的可区分性。

17. 根据权利要求 15 所述的方法,还包括 :

将生成的短语显示给所述用户 ;

接收所述用户所说的生成的短语 ;以及

验证所述短语是否由所述用户所说以确认所述用户的身份。

18. 根据权利要求 15 所述的方法,其中,通过搜索库中的、包含用户区别音素的短语来执行生成所述短语。

19. 根据权利要求 18 所述的方法,其中,生成的短语包含至少两个用户区别音素。

20. 根据权利要求 18 所述的方法,其中,生成的短语包含用户区别音素的至少两个实例。

改进的生物密码安全

技术领域

[0001] 本申请涉及改进的生物密码安全。

背景技术

[0002] 语音生物密码被用作为一种形式的说话者验证，并且语音生物密码可以在一些系统上使用并且可能在近期逐渐变得无处不在。该密码基于对语音中的特征进行的识别，语音中的特征当被组合时以一定的概率唯一地识别个体。

[0003] 噪声降低算法可以用以降低背景噪声对以下的影响：采集样本以形成代表语音中的特征的模型以及使用模型以识别说话者二者。该噪声降低算法还会影响系统性能。

发明内容

[0004] 一种系统包括：存储在计算机可读存储装置上的用户语音简档，该语音简档包含具有针对音素的用户识别特征的多个音素；以及语音处理器，该语音处理器被耦接以访问语音简档，从而基于针对该音素的用户识别特征与平均用户识别特征之间的差来生成包含用户区别音素的短语，以使得该短语具有与其他用户的可区分性。

[0005] 在另一个实施方式中，系统包括：音频输入装置，该音频输入装置接收来自用户的语音，该语音包括用户附近的环境噪声；以及语音处理器，该语音处理器被耦接至音频输入装置，以根据环境噪声生成包含用户区别音素的短语，从而使得该短语具有与环境噪声、与其他用户或者与环境噪声和其他用户二者的可区分性。

[0006] 一种方法包括：接收用户附近的环境噪声信号，以及根据环境噪声生成包含用户区别音素的短语，以使得短语具有与环境噪声、与其他用户或者与环境噪声和其他用户二者的可区分性。

附图说明

[0007] 图 1 是根据示例实施方式的生成待由用户说以实现访问系统的短语的系统的框图。

[0008] 图 2 是示出了根据示例实施方式的音素和音素的分值的表。

[0009] 图 3 是示出了根据示例实施方式的短语的表，其中，在该短语中含有用于在生成短语以实现访问系统时使用的音素。

[0010] 图 4 是示出了根据示例实施方式的选择和使用短语以实现访问计算机系统的方法的流程图。

[0011] 图 5 是实现一个或更多个实施方式的示例计算机系统的框图。

具体实施方式

[0012] 在下面的描述中，参照了附图，附图构成本文的一部分并且在附图中通过说明的方式示出了可以被实行的特定的实施方式。足够详细地描述了这些实施方式以使本领域技

术人员能够实行本发明，并且将理解的是，在不偏离本发明的范围的情况下可以使用其他实施方式而且可以做出结构上的、逻辑上的以及电气方面的改变。因此，下面的示例实施方式的描述不是作为限制性的理解，并且本发明的范围由所附权利要求来限定。

[0013] 本文中描述的功能或算法在一个实施方式中可以以软件或者以软件与人类执行的步骤的结合来实现。软件可以包括存储在计算机可读介质如存储器或其他类型的存储装置上的计算机可执行指令。而且，该功能对应于模块，模块是软件、硬件、固件或其任意组合。可以根据需要将多个功能在一个或更多个模块中实行，并且所描述的实施方式仅是示例。软件可以在数字信号处理器、ASIC(专用集成电路)、微处理器或工作在计算机系统如个人计算机、服务器或其他计算机系统上的其他类型的处理器上执行。

[0014] 在不同的说话者验证系统中使用了许多不同的技术，如频率估算、隐马尔科夫模型、高斯混合模型、模式匹配、神经网络以及其他技术。这些技术可以依靠音素的变化，该音素为语音。辅音是单个字母音，如“b”、“d”、“t”等。双子音字母是两个字母音，如“th”、“sh”、“ch”以及其他两个字母音。短元音对应于字母中的各种元音。其他类型的元音包括长元音、其他元音、双元音以及被“r”影响的元音。

[0015] 在各种实施方式中，生成人要说的短语，该短语将针对给定的人（如系统的用户）的可区分性最大化以提供给用户对系统的访问。用户会被要求说的短语包含已知为更容易与其他用户区分的音素。在另外的实施方式中，选择包含在存在环境噪声时最小可能受到不利影响的音素的短语。

[0016] 图1是用于生成待由人说的短语以帮助对使用系统的人的身份进行验证的系统100的框图。该短语可以由许多不同的系统使用以验证说密码或短语的用户，并且该短语可以提供对系统的访问。验证说话者的系统可以使用许多不同的语音验证技术中的一个或更多个语音验证技术，语音验证技术通常称为将用户简档与所说的短语进行匹配的模型。短语可以是单个单词或若干个连贯起来的单词。短语不需要在语法上是正确的。

[0017] 在一个实施方式中，系统100包括被耦接以接收音频输入装置115的处理器110。音频输入装置可以包括音频插口、或执行一个或更多个无线协议的无线电频率接收装置。在一个实施方式中，处理器可以包括存储器或者可以访问存储器以执行语音处理功能。处理器被耦接以接收对应于用户的简档120。该简档对用户可以以不同于其他用户的方式发音的所选择的音素以及每个音素与平均用户的差异的量的表示进行识别。可以认为简档120包含具有针对该音素的用户识别特征的多个音素。音素的发音越不同于大多数其他用户对该音素发音的方式，越可以将该音素认作为具有用户识别特征的音素。

[0018] 处理器110还被耦接至短语生成器125。短语生成器125可以是存储在计算机可读存储装置上的软件，以使得该软件当由处理器110执行时，基于针对该音素的用户识别特征与平均用户识别特征之间的差异来生成包含用户区别音素的短语，以使得该短语具有与其他用户的可区分性。可区分性可以基于差异的大小而被增强。

[0019] 在一个实施方式中，音频输入装置115被用以接收通过用户说生成的短语而接收到的音频信号。系统100可以包括显示器130，该显示器130被耦接以接收生成的短语并且将生成的短语显示给用户，以使得用户能大声将短语读出。系统100还可以包括模型135，该模型135对应于用户并且由处理器110使用以验证短语是否由用户说出，从而验证用户的身份。在一些实施方式中，模型可以是可以将生成的短语与对应于由用户所说的短语的

音频信号进行比较以确认用户的身份的任何种类的模型。

[0020] 在另一个实施方式中，音频输入装置 115 提供下述信号：该信号代表用户正试图访问系统 100 处附近的周边环境中的背景噪声。环境噪声是在用户试图提供音频短语或密码以实现访问系统处由麦克风收集的噪声。处理器 110 和短语生成器 125 进行工作以根据环境噪声生成包含用户区别音素的短语，以使得该短语具有增强的与环境噪声的可区分性。在一个实施方式中，表 200 可以替代地提供用户区别分值，可以具有代表比较不可能以类似于环境噪声的方式发音的音素的分值。例如，如果用户在乘火车时试图访问系统 100，火车会产生具有特定频率的附近环境噪声，该特定频率类似于对应于某些音素如“r”音素的频率。从而，包含“r”音素的短语不会那么容易来处理和使用以验证说这些短语的用户确实是正确的用户。

[0021] 在另一个实施方式中，表 200 中的分值可以对应于基于环境噪声和音素的用户区别发音二者的组合分值。在一个实施方式中，该分值可以是与用户差异相关的分值及与环境噪声相关的分值的简单平均值，或者可以对该分值更倾向于一个分值或另一分值进行加权。

[0022] 图 2 是针对用户的包含音素 110 的表 100 的框图，每一个音素具有对应于与其他用户正常发音的差 115 的项。在一个实施方式中，该差包括 0 至 100 的分值，0 对应于正常发音。从而，分值越高，用户对该音素的发音越能与其他用户区别开来。在本示例中示出的分值从 10 至 90 变化，90 对应于“r”音。在这种情况下，也许用户对“r”音感到有困难，将其更像“w”来发音。从而，在其中具有“r”音的单词更可能引起该用户与其他用户更大的区分。类似地，该用户以显著地不同于其他用户的方式对“d”音和“sh”音进行发音。从而，与不包含这些音的短语相比，通过生成包含“r”音、“d”音和“sh”音的短语，更可能将该用户与其他用户区别出来。

[0023] 尽管在一个实施方式中选择了 1 至 100 的数值范围，但是在另外的实施方式中可以使用不同的数值范围，如基于方差的数值范围或音素发音差的其他统计度量。在另一些实施方式中，数值范围可以束缚于在执行用户验证时使用的特定的模型，从而导致针对对于所使用的特定的模型来说更易区别的短语来选择音素。

[0024] 图 3 是包含多个短语 310 的表 300，其中，对应的音素 320 包含在该短语中。在这个示例表中，仅示出了具有表 200 中较高的分值的音素。应该注意的是，虽然在短语中也具有其他音素，但是因为其他音素不能像列出的音素那样显著地将用户与其他用户区别开，所以并未示出其他音素以避免以不重要的信息将所示出的表变混乱。表 300 有效地表示存储在计算机可读存储装置上的短语的库。表 300 可以由处理器 110 和短语生成器 125 使用，以通过在库中搜索包含用户区别音素的短语来生成短语。在一个实施方式中，生成的短语应当包含至少两个用户区别音素。在另一些实施方式中，生成的短语包含用户区别音素的至少两个实例，这两个实例可以是同一音，如“r”或“ch”。

[0025] 在另一个实施方式中，表 300 包含具有关联的音素的各个单词。在重点放在可区分性的情况下，可以根据表中的单词生成具有增强的与其他用户的可区分性的短语。在一个实施方式中，可以根据单词随机生成短语。在再一个实施方式中，表 300 中的单词或短语可以来自一种或更多种不同的语言，如英语、法语、西班牙语、德语、日语、汉语等。在另一些实施方式中，还可以创造单词或短语。

[0026] 表 300 中一个示例短语是“short dogs run”。这个短语包含具有 70 的分值的“sh”音素、具有 80 的分值的“d”音素以及两个各自具有 90 的分值的“r”音素。从而该短语包含四个高度地区别的音素，并且可能是表 300 中的短语中的最具有区别的用户区别音素。

[0027] 在另一些实施方式中，可以从远程服务器（如基于云的服务器）选择用于短语的单词，并且可以对单词进行解析以识别音素。在再一些实施方式中，可以在每次用户试图实现访问系统时生成新的短语以将可能使用用户的记录以试图实现访问的风险最小化。

[0028] 图 4 是示出了根据示例实施方式的生成用户要说的短语以验证该用户访问系统的方法 400 的流程图。在 410 处，接收用户附近的环境噪声信号。还对用户进行识别。在 420 处，根据环境噪声生成包含用户区别音素的短语使得该短语具有增强的与环境噪声的可区分性。在一个实施方式中，还可以根据针对该音素的用户识别特征与平均用户识别特征之间的差来对生成的短语进行选择，使得短语具有增强的与其他用户的可区分性。在再一些实施方式中，环境噪声和用户差二者都用以生成短语。在再一个实施方式中，没有接收到环境噪声，而仅使用用户差来生成短语。

[0029] 当生成短语时，在 430 处将该短语显示给用户。用户然后可以对着麦克风大声读出该短语。在 440 处，接收用户所说的生成的短语，并且在 450 处，系统验证该短语是否由用户说出以验证用户的身份。

[0030] 在一个实施方式中，在 420 处生成短语是通过搜索库中的包含用户区别音素的短语而被执行的。生成的短语可以包含至少两个用户区别音素。在另一个实施方式中，生成的短语包含用户区别音素的至少两个实例。

[0031] 在另一个实施方式中，用户可以以其声音来训练系统并且设置将来待用以访问系统的短语。系统可以解析所设置的短语以确定短语是否提供了良好的与其他用户的可区分性，并且给用户提供对所提供的短语的反馈。该反馈可以指示短语没有提供适当的与其他用户的可区分性，并且可以以如上面所述的方式提示具有更高的可区分性的可替换的短语。该反馈还可以提示用户考虑包含更多被发现具有更高的与其他用户的可区分性的音素的不同的短语，如：“请提供包含更多 {x、y 和 z} 音（音素）的短语。”

[0032] 图 5 是执行一个或更多个实施方式的计算机系统 500 的示意性框图。在一个实施方式中，在分布式网络中使用多个这样的计算机系统以实现多个部件。可以通过一个计算机系统来执行短语生成并且通过一个或更多个其他计算机系统来使用该短语生成以验证用户的身份。更进一步，可以使用一个或更多个计算机系统以针对用户正试图访问的其他计算机系统验证用户的身份。在一个实施方式中，计算机系统可以是具有内置麦克风的智能电话。可以使用面向对象的、面向服务的或其他架构来实现该功能并且在多个系统和部件之间进行通信。为计算机 500 形式的一个示例计算装置可以包括处理单元 502、存储器 503、可移除存储装置 510 以及不可移除存储装置 512。存储器 503 可以包括易失性存储器 514 和非易失性存储器 508。计算机 500 可以包括多种计算机可读介质如易失性存储器 514 及非易失性存储器 508、可移除存储装置 510 及不可移除存储装置 512，或可以访问包括多种计算机可读介质如易失性存储器 514 及非易失性存储器 508、可移除存储装置 510 及不可移除存储装置 512 的计算环境。计算机存储装置包括随机存取存储器 (RAM)、只读存储器 (ROM)、可擦除可编程只读存储器 (EPROM) & 电可擦除可编程只读存储器 (EEPROM)、闪存或

其他存储技术、光盘只读存储器 (CD ROM)、数字通用磁盘 (DVD) 或其他光盘存储器、磁盒、磁带、磁盘存储器或其他磁学存储装置、或能够存储计算机可读指令的任何其他介质。计算机 500 可以包括或访问包括输入装置 506、输出装置 504 以及通信连接 516 的计算环境。计算机可以使用通信连接在网络环境中工作以连接到一个或更多个远程计算机 (如数据库服务器)。远程计算机可以包括个人计算机 (PC)、服务器、路由器、网络 PC、对等装置或其他普通的网络节点等。通信连接可以包括局域网 (LAN)、广域网 (WAN) 或其他网络。

[0033] 存储在计算机可读介质上的计算机可读指令可由计算机 500 的处理单元 502 执行。硬盘驱动器、CD-ROM 以及 RAM 是包括非暂态计算机可读介质的制品的一些示例。例如，能够提供类属技术以针对数据存取和 / 或在基于部件对象模型 (COM) 的系统中的各个服务器之一上进行操作来执行存取控制检查的计算机程序 518 可以包括在 CD-ROM 上并且从 CD-ROM 加载至硬盘驱动器。计算机可读指令使得计算机 500 能够在具有多个用户和服务器的基于 COM 的计算机网络系统中提供类属存取控制。

[0034] 尽管以上已详细描述了一些实施方式，但是也可以进行其他修改。例如，在附图中描绘的逻辑流不需要所示出的特定的顺序或连续的顺序来取得理想的结果。可以设置其他步骤，或者可以将步骤从所描述的流程中删去，以及可以将其他部件添加到所描述的系统或从所描述的系统移除。其他实施方式可以在所附权利要求的范围内。

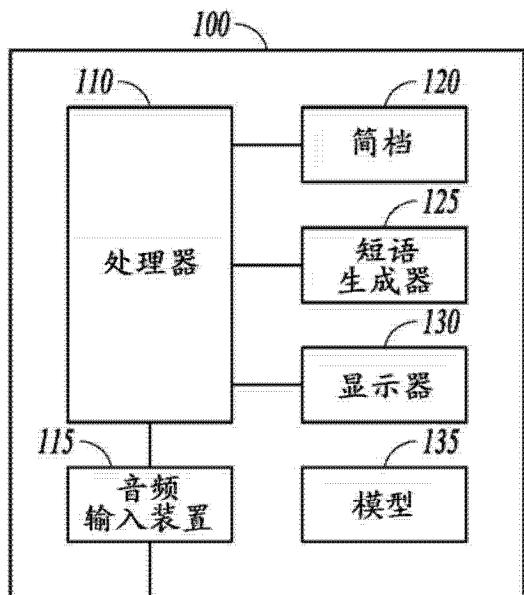


图 1

音素	Δ
“b”	50
“d”	80
“t”	30
⋮	⋮
“th”	10
“sh”	70
“ch”	60
“r”	90
⋮	⋮

图 2

短语	音素
THE CHURCH	“th”, “ch”, “ch”, “r”
BIG SHIRT TOY	“b”, “sh”, “t”, “r”
SHORT DOGS RUN	“sh”, “d”, “r”, “r”
⋮	⋮

图 3

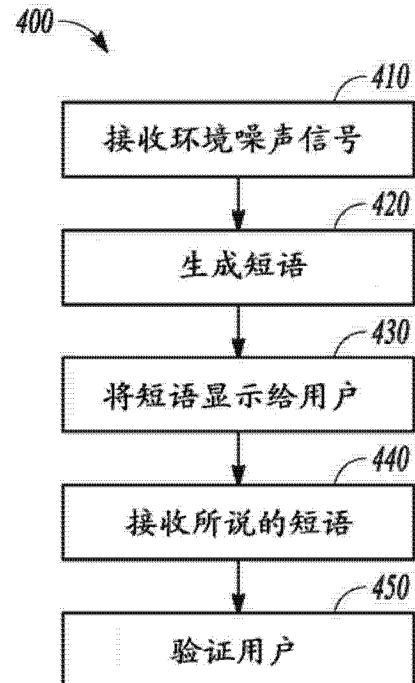


图 4

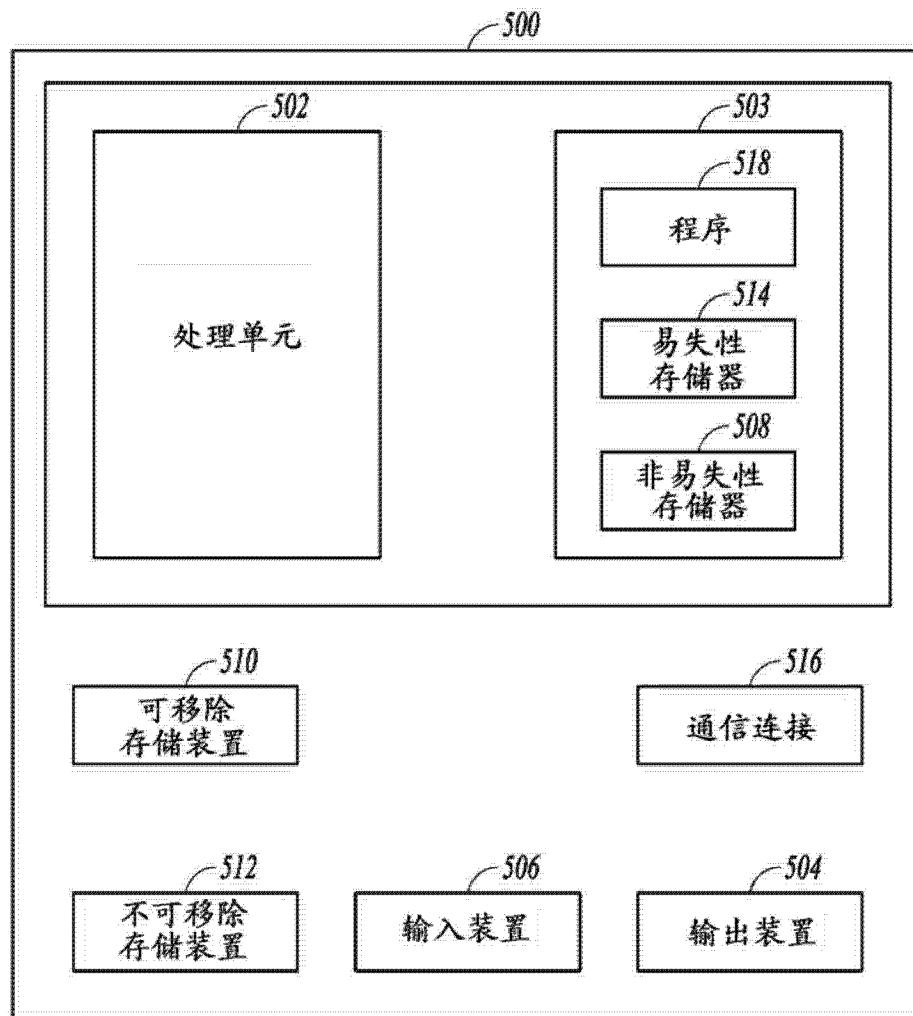


图 5