



(12) 发明专利

(10) 授权公告号 CN 101577062 B

(45) 授权公告日 2012. 07. 25

(21) 申请号 200810163752. 4

(22) 申请日 2008. 12. 30

(73) 专利权人 浙江工业大学

地址 310014 浙江省杭州市下城区朝晖六区

(72) 发明人 张宁宁 顾容 应文良 李晓

(74) 专利代理机构 杭州天正专利事务所有限公

司 33201

代理人 王兵 王利强

(51) Int. Cl.

G09B 21/00 (2006. 01)

G06F 3/01 (2006. 01)

G06F 3/048 (2006. 01)

G06T 13/00 (2011. 01)

(56) 对比文件

CN 101005574 A, 2007. 07. 25, 全文.

CN 1664807 A, 2005. 09. 07, 全文.

CN 1506871 A, 2004. 06. 23, 全文.

JP 特开平 5-241496 A, 1993. 09. 21, 全文.

KR 1020050032296 A, 2005. 04. 07, 全文.

CN 1369823 A, 2002. 09. 18, 全文.

审查员 杨雪

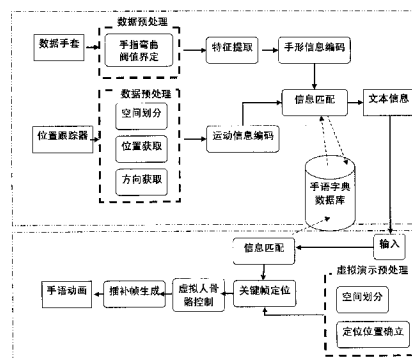
权利要求书 1 页 说明书 7 页 附图 4 页

(54) 发明名称

一种基于空间编码的手语运动信息与文本信息相互转换的实现方法

(57) 摘要

一种基于空间编码的手语运动信息与文本信息相互转换的实现方法, 首先建立文本信息、手语编码信息和动作动画建立对应序列的手语字典数据库, 通过检测佩戴者的手形变换的数据手套和佩戴者的手所在空间区域的位置跟踪器, 根据数据手套和位置跟踪器进行手语识别的智能识别, 所述智能识别包括手语运动信息转化为文本信息模块、文本信息转化为手语信息模块。本发明提供一种快速性好、实时性强、成本低的基于空间编码的手语运动信息与文本信息相互转换的实现方法。



1. 一种基于空间编码的手语运动信息转换为文本信息的实现方法,其特征在于:首先建立文本信息、手语编码信息和动作动画对应序列的手语字典数据库,所述的实现方法包括以下步骤:

(1)、信号数据采集:根据佩戴者身上的数据手套和位置跟踪器的波特率得到时间段的数据帧,得到输入数据的一系列向量数据;

(2)、数据预处理:根据数据手套佩戴者的手语习惯对弯曲度值域进行界定,同时使用位置跟踪器对佩戴者的个体特征的空间区域进行划分,定位嘴,左耳垂,右耳垂,左肩以及右肩的位置;

(3)、手语信息特征提取:根据数据手套输入的数据提取手势信息,从位置跟踪器输入的数据中提取手的方向、位置信息,构成输入样本的特征向量;

(4)、手形信息编码:根据获得的特征向量的手语信号进行编码,得到字符串,编码规则如下:

(4.1)、每个手指关节分为三个活动状态——伸直、半弯曲以及完全弯曲,采用二进制编码方式对指关节弯曲状态进行编码;

(4.2)、对手掌所在的位置空间进行二进制编码:将手掌可能出现的空间位置,按照嘴所在水平线、肩所在水平线为横向分界线,左耳和右耳所在竖向线为竖向分界线,横向分界线和竖向分界线相互交叉划分9个小空间,对每一个小空间及手掌的朝向进行二进制编码,从而判断手掌出现在哪一小空间进而快速识别手形,实时将手语转换为文本信息;

(5)、输出信息匹配:根据字符串的数值在手语字典数据库中进行查询,查询到的结果得到文本信息。

2. 如权利要求1所述的基于空间编码的手语运动信息转换为文本信息的实现方法,其特征在于:在所述步骤(4.1)中,将手指的三种状态分别编码为伸直00,半弯曲01,完全弯曲10,每个手指都按照上述的弯曲状态编码方式,用一个十位的字符串来表示五指的状态。

3. 如权利要求1或2所述的基于空间编码的手语运动信息转换为文本信息的实现方法,其特征在于:在所述步骤(4.2)中,对手掌所在的位置空间进行二进制编码过程为:将嘴部以上划分为三个空间,左耳以左的空间为“1001”,左耳到右耳之间的空间为“1000”,右耳以右的空间为“1010”;将嘴部以下、肩以上的空间划分为三个空间,左耳以左的空间为“0001”,左耳到右耳之间的空间为“0000”,右耳以右的空间为“0010”;将肩以下的空间也划分为三个空间,左肩以左的空间为“0101”,左肩到右肩的空间为“0100”,右肩以右的空间为“0110”;手掌方向根据手掌矢量对应的位置跟踪器的XYZ坐标,手掌朝上0010,手掌朝下0011,手掌朝左0101,手掌朝右0100,手掌朝前1001以及手掌朝身体1000。

一种基于空间编码的手语运动信息与文本信息相互转换的实现方法

技术领域

[0001] 本发明涉及一种聋哑人与正常人之间实现实时交流的方法,尤其是一种手语交流方法。

背景技术

[0002] 手语是聋哑人使用的语言。它是由手形动作辅之以表情姿势而构成的比较稳定的表达系统,是一种靠动作 / 视觉进行交际的特殊语言。我国有 2000 多万的听力障碍者,他们主要使用手语进行交流。由于手语不是社会的大多数人的常用语言,这在很大程度上制约了他们与社会的交流。手语交流系统的研制,可以一定程度上解决这个问题,为聋人创造无障碍环境方面发挥重要作用,并对推广规范的中国手语有很大作用。

[0003] 随着社会对聋哑人关爱的增长,越来越多的学者专家开始研究手语识别系统,以更好地实现正常人和聋哑人之间的交流。目前的手语识别系统主要分为基于数据手套的手语识别和基于视觉(图像)的手语识别系统。对于具有时空并发性的手语的识别,宜采用数据手套作为手形输入设备,并采用位置跟踪器采集手掌的运动。因为与摄像机相比,数据手套与位置跟踪器采集的数据简洁、准确,这两种采集设备易获取表明手语时空特性的特征,如手指关节运动信息,手掌运动信息等,数据手套采集的数据不受光照等环境变化的影响。世界上很多专家致力于手语识别方法的研究,实现了手语信号到文本、声音信息的转化,也有一些专家采用电子设备把文本信息转化为手语动画,从而实现人与机器终端单向的手语交流。

[0004] 然而,聋哑人绝大部分的交流障碍是在与正常人交流的过程中产生的。如果要实现实时的交流,就必须要求有响应迅速的输入输出设备,同时也要求高效的识别与转化方法,能在瞬间实现手语与文本信息的相互转化。为了进一步实现推广,所用的设备成本不能太昂贵。

发明内容

[0005] 为了克服已有的手语识别方法的快速性差、实时性差、成本高的不足,本发明提供一种快速性好、实时性强、成本低的基于空间编码的手语运动信息与文本信息相互转换的实现方法。

[0006] 本发明解决其技术问题所采用的技术方案是:

[0007] 一种基于空间编码的手语运动信息转换为文本信息的实现方法,首先建立文本信息、手语编码信息和动作动画建立对应序列的手语字典数据库,所述的实现方法包括以下步骤:

[0008] (1)、信号数据采集:根据佩戴者身上的数据手套和位置跟踪器的波特率得到时间段的数据帧,得到输入数据的一系列向量数据;

[0009] (2)、数据预处理:根据数据手套佩戴者的手语习惯对弯曲度值域进行界定,同时

使用位置跟踪器对佩戴者的个体特征对空间区域进行划分,定位嘴,左耳垂,右耳垂,左肩以及右肩的位置;

[0010] (3)、手语信息特征提取:根据数据手套输入的数据提取手势信息,从位置跟踪器输入的数据中提取手的方向、位置信息,构成输入样本的特征向量;

[0011] (4)、手形信息编码:根据获得的特征向量的手语信号进行编码,得到字符串,编码规则如下:

[0012] (4.1)、每个手指关节分为三个活动状态——伸直、半弯曲以及完全弯曲,采用二进制编码方式对指关节弯曲状态进行编码;

[0013] (4.2)、对手掌所在的位置空间进行二进制编码;

[0014] (5)、输出信息匹配:根据字符串的数值在手语字典数据库中进行查询,查询到的结果得到文本信息。

[0015] 作为优选的一种方案:在所述步骤(4.1)中,将手指的三种状态分别编码为伸直 00,半弯曲 01,完全弯曲 10,每个手指都按照上述的弯曲状态编码方式,用一个十位的字符串来表示五指的状态。

[0016] 作为优选的再一种方案:在所述步骤(4.2)中,对手掌所在的位置空间进行二进制编码过程为:将嘴部以上划分为三个空间,左耳以左的空间为“1001”,左耳到右耳之间的空间为“1000”,右耳以右的空间为“1010”;将嘴部以下、肩以上的空间划分为三个空间,左耳以左的空间为“0001”,左耳到右耳之间的空间为“0000”,右耳以右的空间为“0010”;将肩以下的空间也划分为三个空间,左肩以左的空间为“0101”,左肩到右肩的空间为“0100”,右肩以右的空间为“0110”;手掌方向根据手掌矢量对应的位置跟踪器的 XYZ 坐标,手掌朝上 0010,手掌朝下 0011,手掌朝左 0101,手掌朝右 0100,手掌朝前 1001 以及手掌朝身体 1000。

[0017] 一种基于空间编码的文本信息转换为手语运动信息的实现方法,首先建立文本信息、手语编码信息和动作动画建立对应序列的手语字典数据库,所述的实现方法包括以下步骤:

[0018] (1)、根据虚拟人的骨骼参数设定虚拟人的演示空间,以空间划分规则界定虚拟人的演示空间;

[0019] (2)、记录下运动手掌在演示空间中各个区域的中心位置和相关的骨骼位置的矩阵;

[0020] (3)、以输入的信息作为关键字在手语字典数据库中进行检索,检索到相关信息则提取出手语编码信息;

[0021] (4)、根据手语编码信息设定虚拟人骨骼在手语意义涉及的各个区域中的定位位置,得到虚拟人动作关键帧;

[0022] (5)、由关键帧自动生成插补帧,得到虚拟人演示的手语动画,显示于屏幕终端。

[0023] 作为优选的一种方案:在所述步骤(3)、(4)中,手语编码信息的编码规则为:每个手指关节分为三个活动状态——伸直、半弯曲以及完全弯曲,采用二进制编码方式对指关节弯曲状态进行编码;对手掌所在的位置空间进行二进制编码。

[0024] 进一步,所述手语编码信息的编码规则中,将手指的三种状态分别编码为伸直 00,半弯曲 01,完全弯曲 10,每个手指都按照上述的弯曲状态编码方式,则用一个十位的字符

串来表示五指的状态。

[0025] 再进一步,所述手语编码信息的编码规则中,对手掌所在的位置空间进行二进制编码过程为:将嘴部以上划分为三个空间,左耳以左的空间为“1001”,左耳到右耳之间的空间为“1000”,右耳以右的空间为“1010”;将嘴部以下、肩以上的空间划分为三个空间,左耳以左的空间为“0001”,左耳到右耳之间的空间为“0000”,右耳以右的空间为“0010”;将肩以下的空间也划分为三个空间,左肩以左的空间为“0101”,左肩到右肩的空间为“0100”,右肩以右的空间为“0110”;手掌方向根据手掌矢量对应的位置跟踪器的 XYZ 坐标,手掌朝上 0010,手掌朝下 0011,手掌朝左 0101,手掌朝右 0100,手掌朝前 1001 以及手掌朝身体 1000。

[0026] 本发明的技术构思为:基于手语动作常用的统计规律,通过构造一个带有编码的手形和手势的划分空间,提出了专用于手语与文本之间实现实时交流的方法。根据手形与手势的编码组合以及手语字典数据库的引入,给出了一种快速编码方法,用于手语的识别;给出了一种快速的解码方法,用于手语的合成。该方法具有编码效率高和解码速度快的特点。

[0027] 本发明提出一种有效的手语字典数据库使用方法,用以存储手语编码信息并推动手语合成的应用。原有的将动作动画数据一一输入手语运动数据库方法,其数据库存储量大,系统运行效率低,形成的手语合成文件过大,不适合实时手语翻译。提出的数据库使用方法是一种更有效的方法。

[0028] 本发明的基于空间编码的特点为:(1)、常用手语中手的空间位置主要集中在头部周围,而少量分布于上半身身体周围,所以头部空间是手语运动频率最高的空间。竖直方向分别以嘴部,肩为分界点,将空间划分为 3 个横向空间,水平方向分别以左耳、右耳、左肩、右肩分界点将空间划分为 6 个纵向空间。这种划分方法有效地对头部周围的空间进行细致的划分,而对身体周围的空间则进行大致的划分,从而提高空间的识别效率。(2)、根据手语设备佩戴者个体大小的不同事先对手语运动空间进行预定义,使运动空间具有自适应性。(3)、编码效率高。利用空间的划分,手形的编码和手势的编码过程中充分采用了二进制编码方式,没有使用长码字和特殊码字的出现,可以使手语字典数据库的存储量更小,读写效率更高。利用本方法对手语运动信息编码,根据句子的长度数据量可以控制在 0.75K 以内。(4)、解码速度快。利用编码序列映射到虚拟人骨骼矩阵的特点,解码过程中可以有效避免图片或动画的编辑工作,因此可以很容易实现手语合成。(5)、方法简单,实现方便。整个算法只采用二进制匹配运算,避免了复杂的运算,是一种简单且有效的编码方法,可以方便的在多个不同版本的开发平台上移植。

[0029] 本发明的有益效果主要表现在:(1)、借助于手语运动空间的划分与编码,实现了手语交流的实时功能,减少了无意义的等待与响应时间;(2)、带有手指弯曲传感器的数据手套就能实现手形特征的提取,降低系统配套设施的成本;(3)、编码是高效的,佩戴者可以通过手语的输入完善手语字典数据库,保证了手语词汇的完备性;(4)、所用的快速合成手语方法能保证手语合成具有较快的速度;(5)、手语实时交流系统是具有培训性的系统,可以为不懂手语的用户提供手语教学训练。

附图说明

- [0030] 图 1 是实时手语交流系统框架示意图。
- [0031] 图 2 是手形编码说明示意图。
- [0032] 图 3 是手语空间划分的示意图。
- [0033] 图 4 是手掌方向测量参考坐标示意图。
- [0034] 图 5 是汉语手指字母图。
- [0035] 图 6 是“你好”手语的示意图。
- [0036] 图 7 是“很高兴认识你”手语的示意图。

具体实施方式

[0037] 下面结合附图对本发明作进一步描述。

[0038] 实施例 1

[0039] 参照图 1～图 7, 一种基于空间编码的手语运动信息转换为文本信息的实现方法, : 首先建立文本信息、手语编码信息和动作动画建立对应序列的手语字典数据库, 所述的实现方法包括以下步骤:

[0040] (1)、信号数据采集: 根据佩戴者身上的数据手套和位置跟踪器的波特率得到时间段的数据帧, 得到输入数据的一系列向量数据;

[0041] (2)、数据预处理: 根据数据手套佩戴者的手语习惯对弯曲度值域进行界定, 同时使用位置跟踪器对佩戴者的个体特征对空间区域进行划分, 定位嘴, 左耳垂, 右耳垂, 左肩以及右肩的位置;

[0042] (3)、手语信息特征提取: 根据数据手套输入的数据提取手势信息, 从位置跟踪器输入的数据中提取手的方向、位置信息, 构成输入样本的特征向量;

[0043] (4)、手形信息编码: 根据获得的特征向量的手语信号进行编码, 得到字符串, 编码规则如下:

[0044] (4.1)、每个手指关节分为三个活动状态——伸直、半弯曲以及完全弯曲, 采用二进制编码方式对指关节弯曲状态进行编码;

[0045] (4.2)、对手掌所在的位置空间进行二进制编码;

[0046] (5)、输出信息匹配: 根据字符串的数值在手语字典数据库中进行查询, 查询到的结果得到文本信息。

[0047] 参照图 2、图 3, 根据手势的运动信息进行编码: 通过手语词汇实验结果, 在充分考虑手语交流的实时性, 本文给出的基于空间编码的手语识别方法是一种性能稳定, 识别迅速的方法。该方法在中国手语词汇上进行应用说明。

[0048] 当使用者佩戴上数据手套时, 根据使用者手的大小及使用者对弯曲程度的不同定义阈值。如数据手套原始数据范围为 0 ~ 4095, 对于食指而言, 相应的数据为小于 1862 的数表示手指处于伸直状态, 显示值为 1862-2268 之间的数值表示手指处于半弯曲状态, 显示值为 2268-4095 之间的数值则表示手指处于完全弯曲状态。由于每个手指的灵活程度不同, 各个手指的阈值也不尽相同。

[0049] 位置跟踪器固定于使用者的手腕上, 绝对坐标原点位于接收器位置。位置跟踪器获得的角度信息可以确定手掌方向, 将使用者以竖直方向分别以嘴部, 肩为分界点, 将空间划分为 3 个横向空间, 水平方向分别以左耳、右耳、左肩、右肩分界点将空间划分为 6 个纵向

空间。

[0050] 本发明的编码方法为：根据对于所述终端上获得的手形和手势信息，依次设定每个手形和手势信息的编码值，得到码流数据，该码表对应动态链表存储时间序列。按照时间与手形的变换情况可分为静态手势，静态复合手势，和动态手势。

[0051] (1)、静态手势：如表 1 所示对应的手语手形，编码规则如表 1 所示。表 1 为汉字手指字母表对应编码：

[0052]

字母	手掌方向	手形编码
A	0100	0010101010
B	1001	1000000000
C	0101	0101010101
D	1001	1010101010
E	1000	1010000000
F	1000	1000000101
G	1000	1000101010
H	1001	1000001010
I	1001	1000101010
J	0101	1001101010
K	0101	0000001010
L	0101	0000101010
M	1001	1001010110
N	1001	1001011010
O	0101	1001010101
P	0101	1010000000
Q	1001	0001011010
R	1000	0000101010
S	1001	0010101010

T	1001	1000101000
U	1001	0000000000
V	1001	0100001010
W	1001	0100000010
X	1001	0101001010
Y	1001	0010101000
Z	1000	0100101000
ZH	1000	0100001000
CH	0011	0001010101
SH	1001	0001011010
NG	1000	1010101000

[0053] 表 1

[0054] (2)、静态复合手势：如图 6 所示“你好”手语，按照时间顺序记录各个时间序列的手形，如表 2 所示：

[0055]

时间序列	手掌方向	手掌位置	手形编码
1	0011	0100	1000101010
2	0011	0100	1001101010
3	0011	0100	0010101010

[0056] 表 2

[0057] (3)、动态手势：如图 7 所示“很高兴认识你”手语，按照时间顺序记录各个时间序列的手形，如表 3 所示，左手手形同理可得。

[0058]

时间序列	手掌方向	手掌位置	手形编码
1	1000	0100	0000000000
2	1000	0100	1000101010
3	0010	0100	1000101010

4	1000	0100	0000000000
5	0011	1000	1000001010
6	0101	0001	1000101010
7	0101	0100	1000101010

[0059] 表 3

[0060] 上述编码信息经由手语字典数据库进行信息匹配,若对应各个时间序列上的各个编码均相同,则输出相对应文字;若对应各个时间序列上的编码不同,则输出“未知手语信息”,用户可以重新演示手语以确认手语信息或者作为新手语信息输入手语字典数据库。

[0061] 实施例 2

[0062] 参照图 1-图 7,一种基于空间编码的文本信息转换为手语运动信息的实现方法,首先建立文本信息、手语编码信息和动作动画建立对应序列的手语字典数据库,所述的实现方法包括以下步骤:

[0063] (1)、根据虚拟人的骨骼参数设定虚拟人的演示空间,以空间划分规则界定虚拟人的演示空间;

[0064] (2)、记录下运动手掌在演示空间中各个区域的中心位置和相关的骨骼位置的矩阵;

[0065] (3)、以输入的信息作为关键字在手语字典数据库中进行检索,检索到相关信息则提取出手语编码信息;

[0066] (4)、根据手语编码信息设定虚拟人骨骼在手语意义涉及的各个区域中的定位位置,得到虚拟人动作关键帧;

[0067] (5)、由关键帧自动生成插补帧,得到虚拟人演示的手语动画,显示于屏幕终端。

[0068] 在所述步骤 (3)、(4) 中,手语编码信息的编码规则为:每个手指关节分为三个活动状态——伸直、半弯曲以及完全弯曲,采用二进制编码方式对指关节弯曲状态进行编码;对手掌所在的位置空间进行二进制编码。

[0069] 本实施例的手语编码信息的编码规则与实施例 1 相同。

[0070] 当不懂手语的用户看到文字信息并以文字信息作为回应时,参照图 3 设定虚拟人的演示空间。控制虚拟人骨骼,使手掌位置经过演示空间中各个区域的中心位置,并记录骨骼关节矩阵以及相关的变换信息。以输入的文字信息作为关键字在手语数据词典中进行检索,检索到相关信息则提取出手语动作的编码,若检索不到则提醒用户重新输入文本信息。根据手语动作的时间序列设置动画时间与各个关键时刻点。提取手掌编码设定虚拟人骨骼在图 3 所示的各个区域中各个关键时刻点定位位置,提取手掌方向编码设定虚拟人手骨骼的方向,提取手形编码设定虚拟人手指骨骼的动作,从而得到各个时间序列中的关键帧;由关键帧生成插补帧,得到虚拟人演示的手语关键帧动画,显示于屏幕终端。

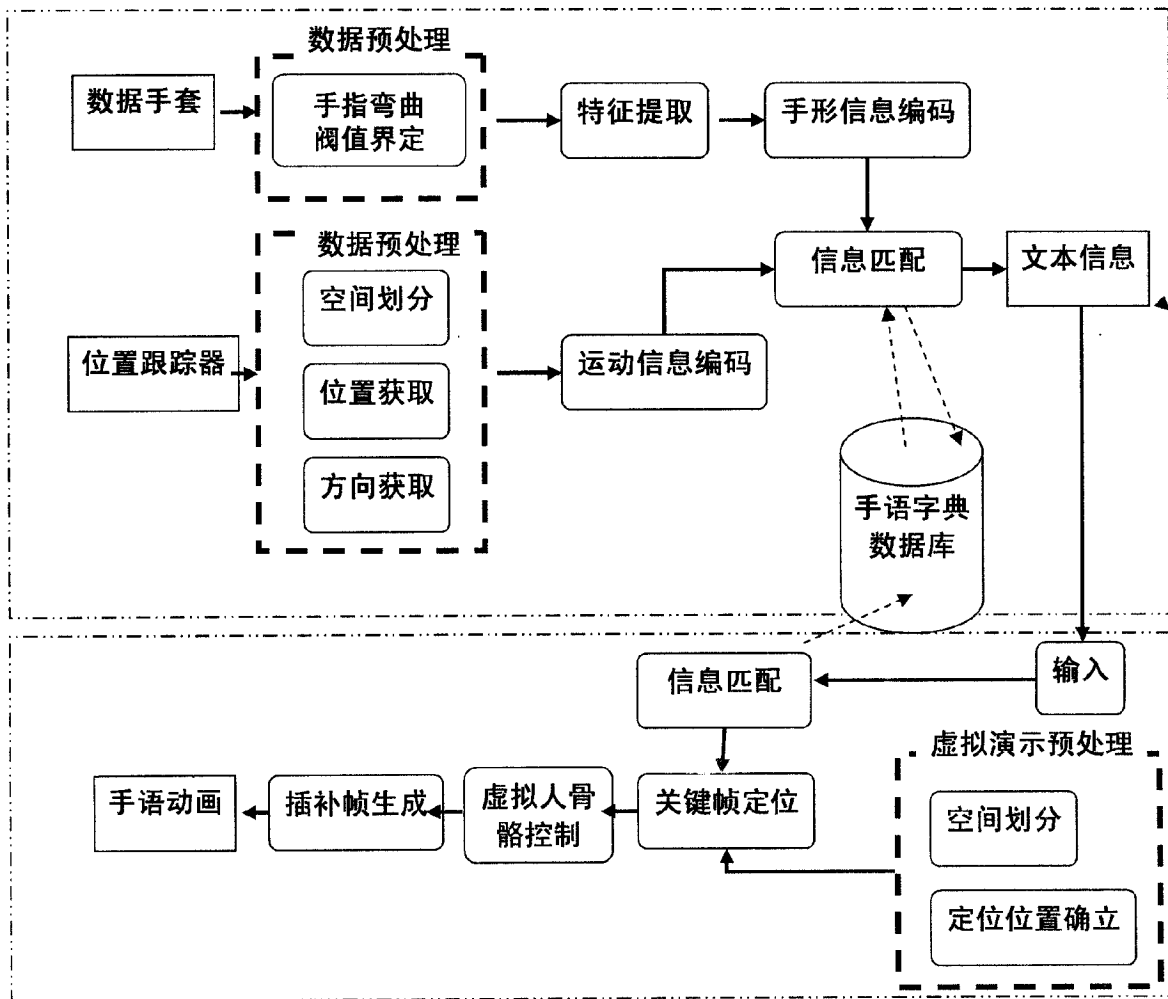


图 1

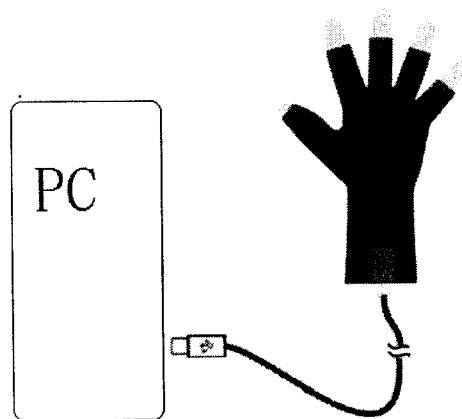


图 2

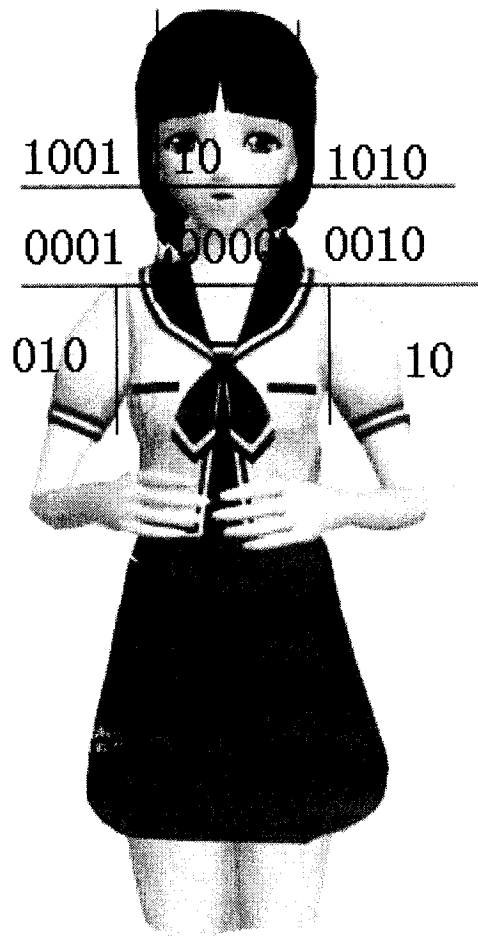


图 3

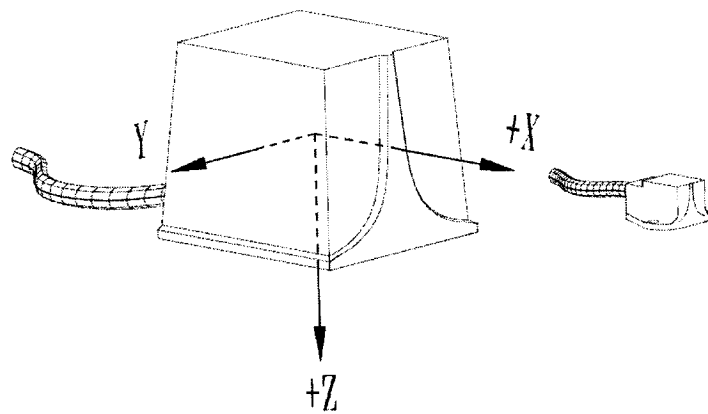


图 4

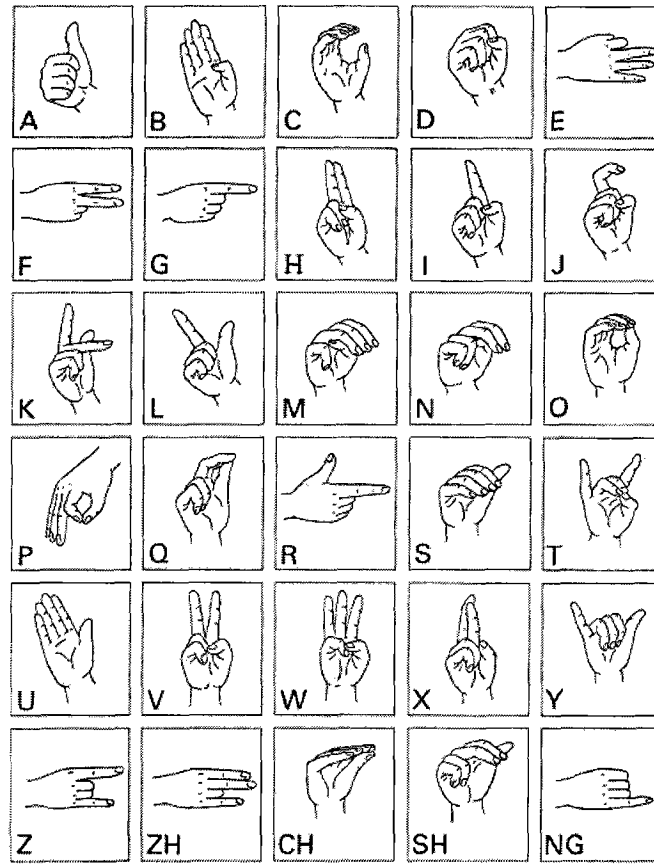


图 5



图 6



图 7