

基于 Kinect 的上肢康复网络系统的设计与实现

孙光民, 李俊, 张俊杰, 闫正祥, 马北川

(北京工业大学 信息学部, 北京 100124)

摘要: 目前上肢运动功能障碍患者的康复训练一般为人工指导的方式, 针对此训练方式的人员、场地资源限制, 以及训练数据难以记录管理的问题, 设计并实现了一套可进行自主训练, 可详细记录、查询训练数据的网络康复训练系统; 系统通过 Kinect 传感器采集人体的运动信息, 计算人体骨骼向量, 并结合 dynamic time warping (DTW) 算法, 将患者的实际动作与标准动作进行匹配, 根据匹配结果指导患者进行自主康复训练, 节省了指导人员与场地资源, 并与游戏相结合, 提高了趣味性; 基于 Spring、SpringMVC、MyBatis (SSM) 框架的康复训练数据管理系统, 提供了记录、查询、统计训练数据的功能, 将病人的训练时间与完成情况绘制成图表, 方便专业人员跟踪病人的训练情况, 使康复训练更加自动化、数据化、高效化。

关键词: Kinect; 骨骼向量; DTW; 上肢康复训练; SSM

Design and Implementation of Upper Limb Rehabilitation Internet System Based on Kinect

Sun Guangmin, Li Jun, Zhang Junjie, Yan Zhengxiang, Ma Beichuan

(Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China)

Abstract: At present, the rehabilitation training for patients with upper extremity motor function disorder is generally a manual instruction, to deal with the limitations of personnel and site resource and the difficulties recording and managing training data of this training method, an online rehabilitation training system is designed and realized in detail, which can do autonomous trainings and records searchable data. This system gathers human motion information by the Kinect sensor, calculates vectors of human skeletons, matches patients' actual motions with standard motions using the dynamic time warping (DTW), instructs patients to do autonomous trainings according to the matching results. It saves instructors and site resources, and its combination of games improves the fun. Also, its data management system based on the Spring, SpringMVC, MyBatis (SSM) framework provides the functions of recording, querying, and doing statistics of training data, finally graphs the data of training time and completion which makes it easier for professionals to track patients' trainings. This system enables a more automated, data-based and high efficient rehabilitation training.

Keywords: Kinect; skeleton vector; DTW; upper limb rehabilitation; SSM

0 引言

近些年, 因事故、疾病等因素, 导致上肢运动功能障碍的人数逐年提升^[1]。这些患者出现了不同程度的劳动能力丧失问题, 有些甚至影响到了日常生活。对于这些运动功能障碍者, 除了初期的治疗外, 仍需要后续的长期康复训练。

研究表明, 通过大量密集、重复的康复训练, 病患可以最大限度的恢复机体功能, 提高工作和生活自理能力^[2]。因此, 康复训练是必要的治疗过程, 在上肢运动障碍疾病的治疗中具有非常重要的作用。

目前传统的康复训练方法主要是在医护人员的指导下一对一进行的, 虽然这是效果很好的训练方式, 但是它有明显的缺点: 1) 受限于医护人员资源; 2) 训练过程枯燥; 3) 训练数据难以记录。

这些问题困扰着医生与病患^[3-5]。因此, 一种可自主进

行训练、具有较高趣味性、对训练数据可以详细记录与查询的康复训练系统是当前的一个重要研究课题。

Saposnik 等人基于任天堂 Wii 游戏系统开发了一个康复训练系统^[6], 用于轻度偏瘫患者的上肢功能康复训练。阿尔斯特大学的 J. W. Burke、P. J. Morrow 等人提出了一个基于视觉游戏的中风病人上肢康复系统^[7]。汉诺威莱布尼兹大学的 Bruckner H、Bartels C, 等人提出了基于惯性传感器的人体运动康复系统^[8]。河北工业大学设计了结合主动和被动运动训练的踝关节康复系统用于踝关节的康复训练^[9]。

然而这些系统并没有与互联网紧密结合。应用人机交互技术, 结合互联网构建康复训练系统, 是一种新颖的思路, 具有广阔的应用场景。

本文设计的康复训练系统, 利用动作采集技术, 通过对人体动作的匹配, 使患者在康复训练中与计算机交互, 由计算机引导患者进行康复训练, 摆脱了需要训练场所与专业人员的限制。通过互联网, 病人的训练记录以及训练过程中采集到的运动数据可以方便的存储与查询, 训练教练等专业人员可以随时查看病人的训练数据, 追踪病人的训练情况。

收稿日期: 2018-04-02; 修回日期: 2018-05-30。

基金项目: 国家自然科学基金(11527801); 国家自然科学基金(41706201); 北京市教委项目(JZ042001201701)。

作者简介: 孙光民(1960-), 男, 教授, 博士生导师, 主要从事神经网络、图像处理、模式识别方向的研究。

1 系统结构与功能设计

1.1 系统结构设计

系统的使用人员主要为两类: 患者与医护人员。系统为患者提供 PC 端软件, 可用它进行自主康复训练; 为医护人员提供康复训练数据管理系统, 可通过浏览器查询病人的训练记录与相关数据。PC 端软件会将患者训练过程中的数据发送到训练数据管理系统进行记录。整个系统的结构如图 1 所示。

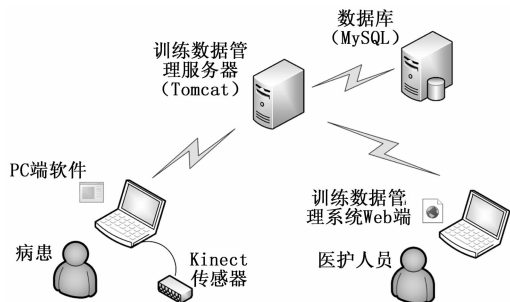


图 1 系统结构

PC 端软件通过 Kinect 传感器采集患者训练过程中的运动数据, 并通过 USB 传输数据。PC 端软件利用运动数据计算出姿势、动作的匹配结果, 通过图形界面指导病人进行康复训练以及康复游戏。对于病人的训练记录数据, PC 端软件通过 HTTP 协议将数据发送到训练数据管理系统, 存储到数据库。

通过浏览器可以访问训练数据管理系统。浏览器向系统服务端发起 HTTP 请求, 接收到请求后, 服务端查询数据库中满足条件的数据, 并将结果数据渲染为 HTML 代码, 响应给浏览器。浏览器将 HTML 代码渲染为可视化页面, 展示给用户。

1.2 系统功能设计

系统功能模块如图 2 所示。

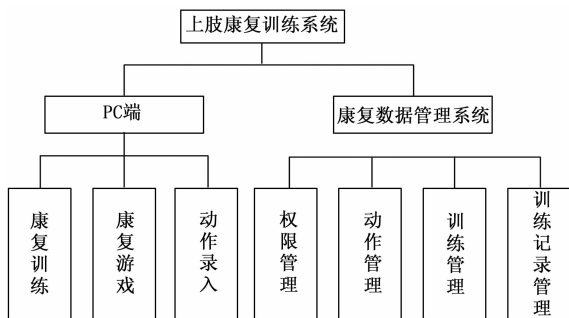


图 2 系统功能

整个系统主要包括两部分: PC 端与康复训练数据管理系统。

PC 端功能包括康复训练、康复游戏、动作录入。康复训练是计算机指导患者进行自主康复的功能模块, 训练过程中系统播放标准动作, 患者需做出相同的姿势动作。模块计算出患者动作与标准动作的匹配度, 若匹配度满足要求则进行下一动作的训练, 否则需重试; 康复游戏模块构

建 3D 游戏场景, 利用康复训练的匹配结果作为输入信号, 对游戏内的物体进行驱动, 使患者与游戏场景进行交互。动作录入模块用于采集标准动作的数据, 提供给康复训练与康复游戏。

康复训练数据管理系统包括动作管理、训练管理、训练记录管理 3 个模块。动作管理模块提供标准动作数据的查看、删除操作, 从 PC 端动作录入模块录入的标准动作可在本模块中管理。训练管理模块用于编辑训练信息, 利用已有标准动作构建训练。训练记录管理模块用于查看病患的训练记录数据, 提供了表格、折线图、柱状图 3 种可视化方式。

2 运动数据采集与匹配

2.1 Kinect 运动数据采集

Kinect 是微软发布的一款体感设备^[10-11], 集成了许多计算机视觉技术。Kinect 传感器包含一个红外发射器、红外接收器、RGB 摄像头, 当人进入 Kinect 的视角范围时, Kinect 会以每秒 30 帧的速率采集彩色视频流、深度视频流。Kinect 把人体分为 32 部分, 使用随机决策森林的分类算法与概率近似度的匹配方法, 识别人体关节点, 生成人体骨骼数据^[12-13]。每帧骨骼数据包含 20 个人体关键关节点的空间坐标信息, 如图 3 所示。

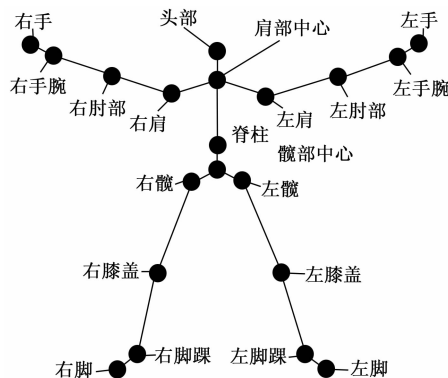


图 3 人体骨架关节点

20 个关节点分别为: 头部、肩部中心、左手、右手、左手腕、右手腕、左肘部、右肘部、左肩、右肩、脊柱、腕部中心、左髌、右髌、左膝盖、右膝盖、左脚踝、右脚踝、左脚、右脚。每个关节点的数据包含 x 、 y 、 z 三维空间坐标, Kinect 的世界坐标系是右手坐标系, 其中, Z 轴正方向为 Kinect 传感器的朝向方向, Y 轴正方向为正上方方向, X 轴正方向为 Kinect 的左方向。

2.2 上肢姿态特征提取

静态姿势为人在某一帧的骨骼状态, 基于 Kinect 采集到的人体骨骼关节点坐标数据, 本文提出了一种上肢静态姿势特征的计算方法。

本文采用 4 个向量来描述人体上肢, 即左肩—左肘、左肘—左手、右肩—右肘、右肘—右手腕这 4 个向量作为特征来描述一个静态姿势。两个动作的相似度, 由四组对应向量的余弦相似度来表示。

采用这种方式的优点为:

1) 向量的平移不变性, 可保证不同的位置不会影响相似度。

2) 两向量的余弦相似度仅与两向量的夹角有关, 与向量的模无关, 此性质可保证不同的身材不会影响相似度。

通过 Kinect 获取左肩、左肘、左手腕、右肩、右肘、右手腕这 6 个点的坐标。设某个向量起点关节、终点关节的坐标分别为 (x_1, y_1, z_1) 、 (x_2, y_2, z_2) , 则对应的向量为:

$$v = (x_2 - x_1, y_2 - y_1, z_2 - z_1) \quad (1)$$

分别计算四组向量的余弦相似度, 对于两个向量 $a(x_1, y_1, z_1)$ 与 $b(x_2, y_2, z_2)$, 余弦相似度为:

$$\cos\alpha_i = \frac{x_1 \cdot x_2 + y_1 \cdot y_2 + z_1 \cdot z_2}{\sqrt{x_1^2 + y_1^2 + z_1^2} \cdot \sqrt{x_2^2 + y_2^2 + z_2^2}} \quad (2)$$

余弦相似度的取值范围为 $[-1, 1]$, 当值为 1 时两向量方向相同, 相似度最高; 值为 -1 时两向量方向相反, 相似度最低。为了方便求和运算, 本文采用 $1 - \cos$ 的值作为两向量的相似度, 将余弦相似度映射到 $[0, 2]$ 上, 其值越小相似度越高。故最终相似度的计算公式为:

$$\text{sim}\alpha_i = 1 - \frac{x_1 \cdot x_2 + y_1 \cdot y_2 + z_1 \cdot z_2}{\sqrt{x_1^2 + y_1^2 + z_1^2} \cdot \sqrt{x_2^2 + y_2^2 + z_2^2}} \quad (3)$$

2.3 静态姿势匹配

对于两个动作的匹配, 本文采用的方法为: 计算两个动作对应的四组上肢向量的相似度 Sim1、Sim2、Sim3、Sim4, 当且仅当 4 个的相似度均满足某个阈值条件时, 两个姿势匹配成功, 否则匹配失败。

此阈值越低则对两个姿势的相似度要求越高, 经过试验表明, 将此阈值设置为 0.1 能取得较好的效果。

2.4 动作匹配

动作是一组静态姿势的序列, 每一帧都是一个静态姿势。本系统的训练策略为训练者模仿标准动作, 将病人的实际动作与标准动作进行匹配, 故采用帧匹配的方法, 每组对应的帧之前计算相似度, 最后计算所有对应帧相似度的平均值作为两个动作的相似度。

然而对于同一个动作, 由于动作快慢的因素, 两个序列可能长度不同, 且不同的运动节奏也会导致两个动作序列在时间轴上有一定的差异, 帧不是简单的按顺序一一对应关系。采用动态时间规整算法 DTW (dynamic time warping) 可以找到两个序列最佳的对齐方式。

DTW 是在孤立词语音识别中最简单有效的算法, 基于动态规划 (DP) 的思想, 解决了长短发音不同的模板匹配问题, 是语音识别中出现较早、较为经典的一种算法, 有着广泛的应用。算法利用动态规划的思想, 寻找两个不同长度序列的最小匹配路径, 匹配路径上的点就是两个序列的对应关系。DTW 算法可以消除时间轴上的差异, 规整时间轴, 找到最佳对齐方式^[14]。

设有两序列 R, T , 长度分别为 M, N :

$$\begin{cases} R = R_1, R_2, \dots, R_m, \dots, R_M \quad \{R_m, m = 1, \dots, M\} \\ T = T_1, T_2, \dots, T_n, \dots, T_N \quad \{T_n, n = 1, \dots, N\} \end{cases} \quad (4)$$

记 R_m 与 T_n 之间的距离为 $d(R_m, T_n)$, 简写为 $d(m, n)$, $1 \leq m \leq M, 1 \leq n \leq N$ 。对齐的 R_m 与 T_n , 表示为 $a(m, n)$, 则连个序列对齐的路径 A 可表示为:

$$A = a_1, a_2, \dots, a_k, \dots, a_K$$

$$\max(M, N) \leq K \leq M + N \quad (5)$$

则两个序列的匹配距离为 A 的累积距离为:

$$D(R, T) = \sum_{k=1}^K d_k(m, n) \quad (6)$$

对于路径 A , 有以下 3 个约束条件:

1) 边界约束: 路径必须从起点 $(1, 1)$ 到达终点 (M, N) 。

2) 单调性约束: 路径保证时间顺序单调不减, 对于两个匹配 $a_{k-1}(m, n)$ 、 $a_k(x, y)$, 则存在 $m \leq x, n \leq y$ 。

3) 连续性约束: m 和 n 只能依次增加 0 或者 1, 即 (m, n) 后的点必须是 $(m+1, n)$ 、 $(m, n+1)$ 或 $(m+1, n+1)$ 中的其中一个。

具有最小累积距离的路径是最优规整路径, 满足上述约束条件, DTW 的递推公式为:

$$D(R, T) = d(M, N) + \min \begin{cases} D(M-1, N-1) \\ D(M, N-1) \\ D(M-1, N) \end{cases} \quad (7)$$

两个序列的匹配路径与对齐关系如图 4 所示。

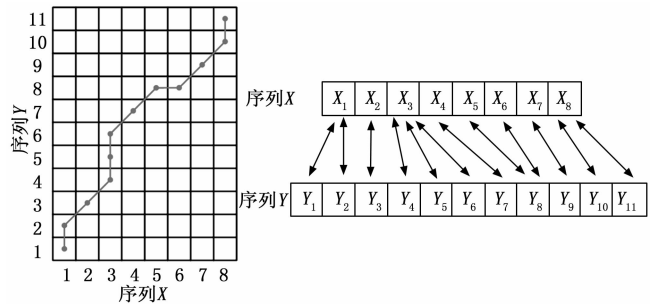


图 4 匹配路径与对齐关系

Kinect 采集的动作序列为 30 帧/秒, 由于每两帧之间相隔时间过短, 差异过小, 故每 3 帧抽取 1 帧, 即每秒抽取 10 帧。两姿势帧之间的相似度 $d(m, n)$ 用 4 对四肢向量的相似度平均值来表示, 即:

$$d(m, n) = \frac{1}{4} \sum_{i=1}^4 \text{sim}\alpha_i \quad (8)$$

最终求得路径上的累积相似度和 $D(R, T)$, 路径长度 $Len(R, T)$, 则路径平均相似度为:

$$\text{Mean}_a = D(R, T) / \text{Len}(R, T) \quad (9)$$

若路径的平均相似度小于阈值时, 则两个动作匹配成功。经过试验表明, 将此阈值设置为 0.05 时, 匹配精度与完成难度能达到较好的平衡。

系统会实时采集训练者的运动数据作为样本数据与模板动作进行匹配。模板动作由康复教练录入到系统, 动作的起点和终点可人工进行选择。而对于在训练过程中实时采集的样本数据, 则需要找到动作起点和终点。

在本系统中, 起点和终点的判定策略如下: 当前的实时

姿势与模板动作的起始姿势相匹配时, 则将此姿势标记为起点; 在起点已标记的情况下, 若 3 秒内姿势的变化程度很小, 即认定为没有变化时, 则将 3 秒前的姿势标记为终点。

终点标记成功后, 样本动作数据为起点、终点间的姿势序列数据。

3 系统功能实现与展示

3.1 PC 端的功能实现与展示

PC 端的功能模块与数据流如图 5 所示。

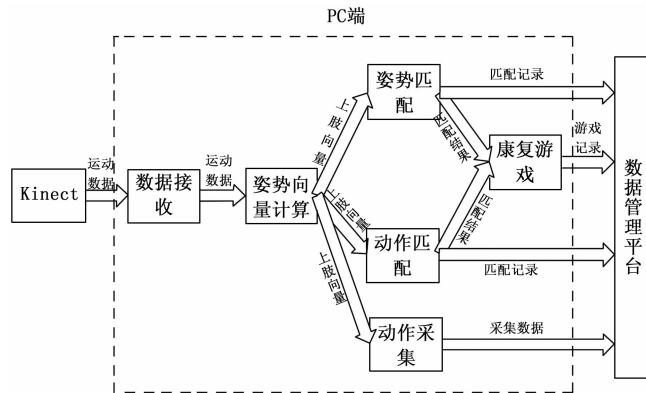


图 5 PC 端功能模块与数据流

其中, Kinect 通过 USB 连接到计算机; 数据接收模块采用 Kinect SDK 提供的事件模型获取骨骼数据; 姿势向量模块计算出上肢姿势向量后, 将向量数据发送到姿势匹配、动作匹配、动作采集模块, 通过消息队列、采用生产者-消费者模式进行数据的传输与处理, 生产者为姿势向量计算模块, 消费者为姿势匹配、动作匹配、动作采集模块; 由于姿势匹配、动作匹配与康复游戏之间存在环境隔离, 且只需要发送识别结果, 数据量较小, 因此采用 TCP 协议, 通过 Socket 进行数据的传递; PC 端各模块与数据管理平台之间通过 Http 协议, 采用 json 数据格式进行数据的传递。

PC 端为患者提供了康复训练、康复游戏两种方式。康复训练方式, 会在窗口上显示要完成的标准姿势、动作, 当病患完成的动作与标准动作匹配成功或超时, 则会进行下一个动作的练习; 训练的所有动作都完成后训练结束, 如图 6 (a) 所示。

在动作训练过程中, Kinect 实时采集姿势帧数据, 在某个时刻会有 3 种状态: 初始、开始、结束。其中, 初始状态存在于开始动作采集且还没有标记到动作起点的时刻; 开始状态存在于已标记动作起点且还没有标记到动作终点的时刻; 结束状态存在于标记终点后的时刻。

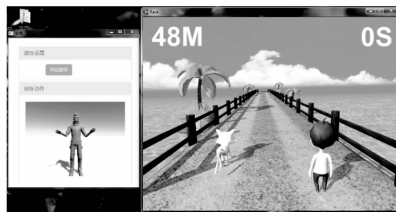
最初, 系统处于初始状态; 在初始状态下, 当动作起点标记完成后进入开始状态; 在开始状态下, 当动作终点标记完成后进入结束状态; 在结束状态下, 起点与终点之间的动作数据会与模板动作进行匹配, 匹配完成后则再次进入初始状态, 开始新一轮的动作数据采集。

康复游戏采用 Unity3D 平台进行开发, 采用赛跑的游戏策略, 控制人物与猫进行赛跑。游戏可以选择简单、一

般、困难 3 个难度, 难度越高猫的移动速度越快。游戏接收姿势、动作匹配模块的匹配结果, 当匹配成功时会控制游戏人物角色前进 3 秒, 时间可累加。若控制的人物角色先到达终点则游戏胜利, 否则游戏失败, 如图 6 (b) 所示。



(a) 上肢训练界面



(b) 康复游戏界面

图 6 PC 端软件界面

角色的前进包括两部分: 空间位置的移动与角色动画的播放。在改变角色坐标的同时播放角色的奔跑动画, 便可实现角色的前进效果。

累积前进时间每秒都会减 1, 若累积时间大于 0 角色会前进, 当累积时间减至 0 则停止前进。累积时间的自动减少通过 Unity 中的协程 (IEnumerator) 来实现, 协程执行的程序会提交到主线程中运行, 因此不会导致线程不安全的问题。

本游戏的跑道在 Unity 场景的世界坐标系中是沿着 Z 轴方向的, 因此可通过判断角色位置的 Z 坐标是否超过终点线的 Z 坐标来判别是否抵达终点。

医护人员可以利用动作采集功能, 向系统录入标准姿势、动作。点击采集页面上的开始按钮会开始采集数据, 点击停止按钮会停止录制。录制完成后可以调节两个滑动条分别选择起点、终止帧, 在输入框输入名称后上传姿势、动作。上传的姿势是起点帧对应的姿势, 上传的动作为起点、终点帧之间的动作。

3.2 数据管理系统的功能实现

数据管理系统采用 MVC 模式, 将程序分为三部分: 模型 (Model)、视图 (View)、控制器 (Controller)。Model 层用于实现系统功能的处理逻辑; View 层用于实现用户与页面的交互逻辑; Controller 层是 Model 与 View 之间的桥梁。MVC 模式实现了系统的职能分工, 降低系统的耦合性^[15]。

数据管理系统使用 SSM 框架进行开发。SpringMVC 框架负责请求的转发, 根据请求的 URL 调用相应的处理方法处理请求; MyBatis 框架负责对象关系映射, 使得数据库的操作更加的便捷; Spring 框架作为 Bean 容器, 提供了 IOC 与 AOP 的特性, 接管了对象的创建、组装工作, 使得程序的开发更加的便捷。3 个框架整合后简化了程序的开发工作。

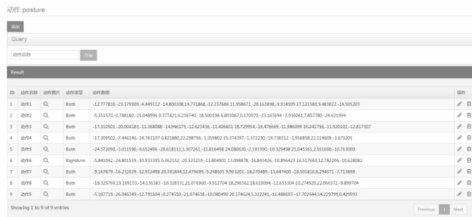
数据管理系统提供了权限管理、动作管理、训练管理、

训练记录管理的功能。

权限管理是管理系统的基础模块，用于区分限制不同用户的功能与数据的操作权限。在本系统中，用户的角色分为 3 种：系统管理员、医护人员、病患。系统管理员能够创建用户、分配用户角色、为病患用户分配医护人员；医护人员可以查看病患信息、姿势数据、动作数据、训练数据、训练记录数据。对于病患信息、训练记录数据，医护人员只能查看分配给自己的病患的数据。

动作管理用于管理录入的标准姿势、动作，提供了查看、编辑、删除的功能，如图 7 (a) 所示。

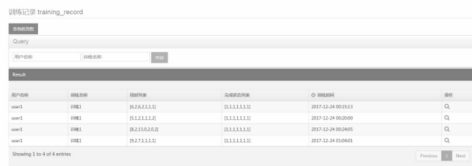
训练管理可以让医护人员利用已有的标准姿势、动作，自由组合搭配，构建康复训练，如图 7 (b) 所示。



(a) 动作管理界面



(b) 训练管理界面



(c) 训练记录数据表格



(d) 数据折线图

(e) 数据柱状图

图 7 数据管理系统 web 端界面

训练记录管理可以查看病患的训练记录，包括病患在某次训练中每个动作的用时、完成情况数据，如图 7 (c) 所示。此外，系统提供了柱状图与折线图的数据可视化方式，可以查看指定病患历次训练的训练总用时、完成度趋势折线图，以及某次训练中各动作的用时、完成度柱状图，分别如图 7 (d) 与图 7 (e) 所示。

4 结论

本文设计的系统包含 PC 端与训练数据管理系统两部

分，PC 端提供给患者，用于进行自主康复训练；训练数据管理系统可以记录、管理训练动作数据及病患的训练记录，并在 WEB 端提供多种可视化方式供医护人员查看数据。经实验，本系统可以供病患进行自主康复训练，并能够有效地解决医护人员资源限制、训练过程枯燥、训练数据与记录不易管理等问题。综上，本文所述基于 Kinect 的上肢康复网络系统，在病患康复训练方面有着广阔的应用场景。

参考文献:

[1] 高 聪, 蒲蜀湘, 朱德仪. 早期康复治疗对脑卒中偏瘫患者肢体功能及日常生活能力的影响 [J]. 中国康复医学杂志, 2001, 16 (1): 27 - 29.

[2] A. A. Rizzo. A SWOT Analysis of the Field of Virtual Reality Rehabilitation and Therapy [J]. Presence, 2005, 14 (2): 119 - 146.

[3] 励建安. 康复医学在 21 世纪的发展趋势 [J]. 实用医院临床杂志, 2007, 4 (4): 1 - 3.

[4] 王茂斌. 康复医学的崛起与发展 [J]. 中华全科医师杂志, 2005, 4 (12): 709 - 711.

[5] 刘克敏, 唐 涛, 王安庆, 等. 骨关节功能康复的现状与展望 [J]. 中国矫形外科杂志, 2009, 229 (11): 865 - 869.

[6] Sapoznik G, Teasell R, Mamdani M, et al. Effectiveness of virtual reality using Wii gaming technology in stroke rehabilitation; a pilot randomized clinical trial and proof of principle. [J]. Stroke: a journal of cerebral circulation, 2010, 41 (7): 1477 - 84.

[7] Burke J W, Morrow P J, Mcneill M D J, et al. Vision Based Games for Upper - Limb Stroke Rehabilitation [A]. Machine Vision and Image Processing Conference, 2008. IMVIP '08 [C]. IEEE, 2008: 159 - 164.

[8] Bruckner H, Bartels C, Blume H. PC - Based Real Time Sonification of Human Motion Captured by Inertial Sensors [J]. 2011.

[9] 王瑞利, 刘更谦, 于庆增, 张小俊. 基于虚拟现实踝关节康复系统的研究 [J]. 机电产品开发与创新, 2007, 20 (1): 16 - 17.

[10] N. M. DiFilippo. Characterization of Different Microsoft Kinect Sensor Models [J]. IEEE Sensors Journal, 2015, 15 (8): 4554 - 4564.

[11] Webb J, Ashley J. Beginning Kinect programming with the Microsoft Kinect SDK [M]. Apress, 2012.

[12] Tran J, Ufkes A, Ferworn A, et al. 3D Disaster Scene Reconstruction Using a Canine - Mounted RGB - D Sensor [A]. International Conference on Computer and Robot Vision [C]. IEEE Computer Society, 2013: 23 - 28.

[13] Zhou Ren. Robust Part - Based Hand Gesture Recognition Using Kinect Sensor [J]. IEEE Transactions on Multimedia, 2013, 15 (5): 1110 - 1120.

[14] Kruskal J B, Liberman M. The symmetric time - warping problem: from continuous to discrete [C] // Time Warps, String Edits, and Macromolecules: The Theory and Practice of Sequence Comparison. Addison - Wesley, 1983.

[15] 李守振, 张南平, 常国锋. Web 应用分层与开发框架设计研究 [J]. 计算机工程, 2006, 32 (22): 274 - 276.