

世界超高清视频产业联盟标准

T/UWA xxxx-xxxx

信息技术 面向虚拟现实应用的人体动作捕捉系统 第1部分：通用技术要求

Information technology-Wearable motion capture system in virtual reality-

Part1: General technical requirements

(征求意见稿)

xxxx-xx-xx 发布

xxxx-xx-xx 实施

世界超高清视频产业联盟

目 次

前 言	I
1 范围	1
2 规范性引用文件	1
3 术语和定义	1
3.1 虚拟现实 virtual reality	1
3.2 穿戴式设备 wearable device	1
3.3 动作捕捉 motion capture	1
3.4 面向虚拟现实应用的人体动作捕捉系统 wearable motion capture system in virtual reality	1
3.5 人体骨骼系统 human skeleton	1
3.6 人体位姿 human pose	1
3.7 手部位姿 hand pose	2
3.8 脸部位姿 face pose	2
3.9 大范围光学动作捕捉 wide-area optical motion capture	2
3.10 惯性动作捕捉 inertial motion capture	2
3.11 脸部捕捉 face capture	2
3.12 数据手套 data glove	2
4 面向虚拟现实应用的人体动作捕捉系统框架	2
5 功能要求	4
5.1 位姿计算模块	4
5.2 位姿数据模块	4
5.3 反馈模块	4
6 性能要求	4
6.1 位姿计算模块	4
6.2 位姿数据模块	4
6.3 反馈模块	5

前 言

本文件按照 GB/T 1.1—2021《标准化工作导则 第1部分：标准化文件的结构和起草规则》的规定起草，共分为5个部分：

- 第1部分：通用技术要求；
- 第2部分：光学动作捕捉系统软件接口标准；
- 第3部分：惯性动作捕捉系统软件接口标准；
- 第4部分：数据手套软件接口标准；
- 第5部分：脸部捕捉系统软件接口标准；

本部分为 T/UWA XXXXX 的第1部分。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别这些专利的责任。

本文件由世界超高清视频产业联盟提出并归口。

本文件起草单位：XXXXX。

本文件主要起草人：XXXX。

信息技术 面向虚拟现实应用的人体动作捕捉系统

第 1 部分：通用技术要求

1 范围

本文件规定了面向虚拟现实应用的人体动作捕捉系统通用的功能要求和性能要求。

本部分适用于影视、游戏面向虚拟现实应用的人体动作捕捉系统的研发和测试，其应用场景包括但不限于影视、游戏等领域。

2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中，注日期的引用文件，仅该日期对应的版本适用于本文件；不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

GB/T 38247-2019 《信息技术 增强现实 术语》

GB/T 38258-2019 《信息技术 虚拟现实应用软件基本要求和测试方法》

T/UWA XXXXX.2 《信息技术 面向虚拟现实应用的人体动作捕捉系统 第2部分：光学动作捕捉系统软件接口标准》

T/UWA XXXXX.3 《信息技术 面向虚拟现实应用的人体动作捕捉系统 第3部分：惯性动作捕捉系统软件接口标准》

3 术语和定义

下列术语以及定义适用于本文件。

3.1 虚拟现实 virtual reality

采用以计算机为核心的现代高科技手段生成的逼真的视觉、听觉、触觉、嗅觉、味觉等多感官一体化的数字化人工环境，用户借助一些输入、输出设备，采用自然的方式与虚拟世界的对象进行交互，相互影响，从而产生亲临真实环境的感觉和体验。

【来源：GB/T 38247-2019，2.1】

3.2 穿戴式设备 wearable device

用户可以直接穿在身上或是整合到用户的衣服或配件上的便携式设备。

3.3 动作捕捉 motion capture

测量、跟踪、记录人或者其他运动物体在三维空间中的运动轨迹的技术。

3.4 面向虚拟现实应用的人体动作捕捉系统 wearable motion capture system in virtual reality

通过让用户穿戴便携式设备以满足虚拟现实应用中动作捕捉需求的系统（区别于使用结构光、摄像机等无需接触人体的动作捕捉系统）。

3.5 人体骨骼系统 human skeleton

人体位姿描述的基础，定义了人体可活动关节的数量和它们之间的运动传递关系。

3.6 人体位姿 human pose

通过人体骨骼系统描述的人体运动关节的位置和朝向。

3.7 手部位姿 hand pose

人体手部各个关节的位置和朝向。

3.8 脸部位姿 face pose

人体脸部的位置、朝向和表情。

3.9 大范围光学动作捕捉 wide-area optical motion capture

工作范围可达到百平米以上，使用光学传感器得到人体位姿的动作捕捉技术。

3.10 惯性动作捕捉 inertial motion capture

使用惯性传感器得到人体位姿的动作捕捉技术。

3.11 脸部捕捉 face capture

获取用户脸部位置、朝向和表情的动作捕捉技术。

3.12 数据手套 data glove

获取用户手部各个关节的位置和朝向的设备。

4 面向虚拟现实应用的人体动作捕捉系统框架

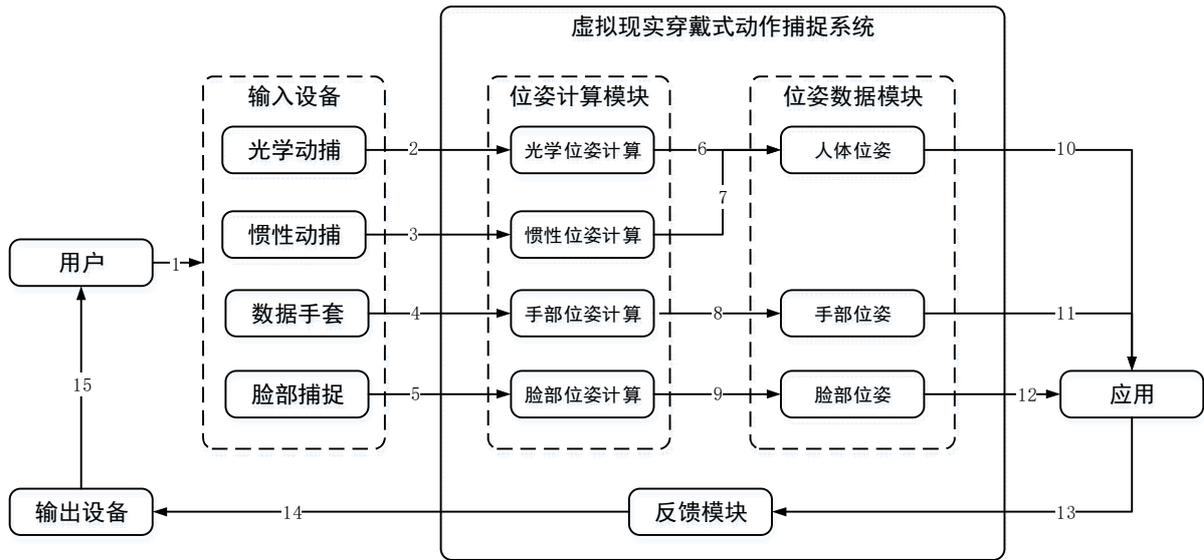
面向虚拟现实应用的人体动作捕捉系统通过输入设备采集原始数据，通过位姿计算后，向应用输出格式化的人体位姿数据，应用得到人体位姿数据后进行业务逻辑处理并向输出设备输出控制信号，从而给用户展示可视化的人体位姿结果。

面向虚拟现实应用的人体动作捕捉系统为应用提供人体位姿、手部位姿、脸部位姿和反馈相关的管理接口，如图1所示。图中连线表示模块之间互相操作的接口，带有箭头的连线既表示模块之间互相操作的接口，又通过箭头方向表示信息流的传输方向。

本部分只规定面向虚拟现实应用的人体动作捕捉系统的通用规范。

面向虚拟现实应用的人体动作捕捉系统包括位姿计算模块、位姿数据模块和反馈模块：

- a)位姿计算模块通过输入设备获得采集到的原始数据，根据数据类型分别使用光学位姿计算、惯性位姿计算、手部位姿计算、脸部位姿计算等子模块进行数据处理，获得人体位姿、手部位姿和脸部位姿，其计算结果将发送给位姿数据模块进行后续处理；
- b)位姿数据模块将位姿计算模块得到的计算结果转换成应用可识别的格式化位姿数据。该模块向应用提供人体位姿、手部位姿和脸部位姿三种位姿数据接口；
- c)反馈模块负责将应用提供的反馈信息转化为输出设备可处理的信息，以便最后向用户输出面向虚拟现实应用的人体动作捕捉系统的处理结果。



说明：

- 1——用户使用输入设备，用户的位姿信息通过该接口被输入设备采集；
- 2——输入设备和光学位姿计算模块间的软件接口，输入设备通过该接口向位姿计算模块输出光学设备数据流，该接口需要满足T/UWA XXXXX.2的要求；
- 3——输入设备和惯性位姿计算模块间的软件接口，输入设备通过该接口向位姿计算模块输出惯性设备数据流，该接口需要满足T/UWA XXXXX.3的要求；
- 4——输入设备和手部位姿计算模块间的软件接口，输入设备通过该接口向位姿计算模块输出数据手套数据流；
- 5——输入设备和脸部位姿计算模块间的软件接口，输入设备通过该接口向位姿计算模块输出脸部捕捉设备数据流；
- 6——光学位姿计算模块和人体位姿模块间的软件接口，光学位姿计算模块通过该接口向人体位姿模块输出人体位姿计算结果，该接口需要满足T/UWA XXXXX.2的要求；
- 7——惯性位姿计算模块和人体位姿模块间的软件接口，惯性位姿计算模块通过该接口向人体位姿模块输出人体位姿计算结果，该接口需要满足T/UWA XXXXX.3的要求；
- 8——手部位姿计算模块和人体位姿模块间的软件接口，手部位姿计算模块通过该接口向人体位姿模块输出手部位姿计算结果；
- 9——脸部位姿计算模块和人体位姿模块间的软件接口，脸部位姿计算模块通过该接口向人体位姿模块输出脸部位姿计算结果；
- 10——人体位姿模块和应用模块间的软件接口，人体位姿模块通过该接口向应用输出格式化的人体位姿数据；
- 11——手部位姿模块和应用模块间的软件接口，手部位姿模块通过该接口向应用输出格式化的手部位姿数据；
- 12——脸部位姿模块和应用模块间的软件接口，脸部位姿模块通过该接口向应用输出格式化的脸部位姿数据；
- 13——应用模块和反馈模块间的软件接口，应用通过该接口向反馈模块提供应用相关的反馈信息；
- 14——反馈模块和输出设备间的软件接口，反馈模块通过该接口向输出设备输出硬件设备可处理的反馈信息；
- 15——用户使用输出设备，输出设备通过该接口向用户输出面向虚拟现实应用的人体动作捕捉系统的最终反馈。

5 功能要求

5.1 位姿计算模块

位姿计算模块应具有功能如下功能

- a)接收来自动捕设备的原始数据;
- b)原始数据记录包括数据捕捉的时间戳;
- c)原始数据保存至本地磁盘;
- d)处理来自动捕设备的原始数据,将位姿作为计算结果输出给位姿数据模块。

示例:惯性动捕设备通过无线网络将原始惯性传感器数据发送给惯性位姿计算模块,该模块计算结束后,将人体活动关节的位置和朝向作为人体位姿发送给位姿数据模块进行后续处理。

5.2 位姿数据模块

位姿数据模块应提供相应(包括动作、手部、脸部)关键点的相对三维空间点坐标,并提供格式要求。

示例:全身动捕设备关键点包括四肢关节;手部动捕设备关键点包括手指关节;脸部动捕设备关键点包括脸部表情控制关键点;或者按照根据应用场景的特定需要给出的相应关键点。三维空间点坐标包括给定笛卡尔坐标系条件下X、Y、Z坐标。

5.3 反馈模块

反馈模块功能如下:

- a)在需要反馈结果的应用中,应以视觉、听觉、触觉或其他反馈方法中的至少一种将当前位姿状态通知用户;
- b)应能够接收应用提出的反馈请求,通知用户的位姿在应用中的响应结果。

示例:用户做出某个姿势后,将用户当前姿态通过图像显示的方式,以人体骨骼位姿的形式显示在屏幕上。

6 性能要求

6.1 位姿计算模块

位姿计算模块性能如下:

- a)位姿信号采样率应不低于20Hz;
- b)位姿时间戳的精度应不低于30ms;
- c)针对单帧位姿,从位姿输入的结束时刻到到位姿计算模块完成计算步骤,所持续的时间。动作捕捉系统的响应时间应不高于1000ms;
- d)针对单帧位姿,输出的位姿与真实位姿的单方向角度自由度平均误差不超过10°;
- e)针对单帧位姿,输出的位姿与真实位姿平均误差不超过3%。

示例:手部动作捕捉数据误差不能超过手部最大尺寸的3%;脸部动作捕捉误差不能超过脸部最大尺寸的3%;全身动作数据误差不能超过身体最大尺寸的3%。

6.2 位姿数据模块

位姿数据模块性能如下:

- a)应将位姿计算模块输出数据完整、准确保存至本地磁盘,数据保存的延迟时间应 $\leq 500\text{ms}$;
- b)针对单帧位姿,从位姿输入的结束时刻到到位姿计算模块完成计算步骤,动作捕捉系统完成上述计算的响应时间应 $\leq 1000\text{ms}$ 。

6.3 反馈模块

反馈模块性能如下：

- a)响应时间：从位姿输入的结束时刻到动作捕捉系统输出反馈所持续的时间。动作捕捉系统的响应时间应 $\leq 2000\text{ms}$ ；
- b)反馈延时：从位姿数据模块输出识别结果到反馈模块输出反馈信息的时间长度。反馈延时应 $\leq 1000\text{ms}$ 。