

ICS: 35.240  
CCS: L81

# 世界超高清视频产业联盟标准

T/UWA 022.1-2023

---

## 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

(V1.0)

2023-11-15 发布

2023-11-15 实施

---

世界超高清视频产业联盟



# 目 次

1 范围 .....	1
2 规范性引用文件 .....	1
3 术语和定义 .....	1
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 .....	5
6.3 反馈模块 .....	5

## 前 言

本文件按照 GB/T 1.1—2020《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。

本文件是 T/UWA 022-2023《信息技术 面向虚拟现实应用的人体动作捕捉系统》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：北京理工大学、厦门大学、信通院（江西）科技创新研究院有限公司、中国电子技术标准化研究院、上海市多媒体行业协会、上海青瞳视觉科技有限公司、凌云光技术股份有限公司、深圳市瑞立视多媒体科技有限公司、北京元客视界科技有限公司、北京元客方舟科技有限公司、合肥工业大学、北京圣威特科技有限公司、歌尔股份有限公司、深圳市奥拓电子股份有限公司、北京市博汇科技股份有限公司、数字电视国家工程研究中心、广东博华超高清创新中心有限公司、北京虚拟动点科技有限公司、深圳市洲明科技股份有限公司、咪咕文化科技有限公司、山东浪潮超高清智能科技有限公司、利亚德光电股份有限公司、青岛海尔多媒体有限公司、深圳市酷开网络科技股份有限公司、海信视像科技股份有限公司、海信电子信息集团、暨南大学、中央广播电视总台、工业和信息化部电子第五研究所、四川新视创伟超高清科技有限公司。

本文件主要起草人：宋维涛、郭诗辉、胡开拓、李婧欣、耿一丹、端木海婴、张海威、熊伟、张梦妮、许秋子、杜华、蒋再毅、王鲁宁、李琳、杨涛、迟小羽、王勇、王荣芳、殷惠清、陈智敏、崔超、谭胜淋、郭佩佩、王敏、刘莉、王宗华、曾有兰、刘卫东、王玉乾、毕蕾、高博宇、范晓轩、张黎敏、邹双泽。

## 引 言

T/UWA 022 拟由五个部分构成。

——第 1 部分：通用技术要求。目的在于描述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。

——第 2 部分：光学动作捕捉系统软件接口标准。目的在于规定光学动作捕捉系统的外部软件接口。

——第 3 部分：惯性动作捕捉系统软件接口标准。目的在于规定惯性动作捕捉系统的外部软件接口。

——第 4 部分：数据手套软件接口标准。目的在于规定数据手套的外部软件接口。

——第 5 部分：脸部捕捉系统软件接口标准。目的在于规定脸部捕捉系统的外部软件接口。



# 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。

本文件适用于面向虚拟现实应用的人体动作捕捉系统的设计、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 022.3-2023 《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

**可穿戴式设备** `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

T/UWA 022.1-2023

**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方米以上，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。

本文件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：

- 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；
- 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；
- 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。

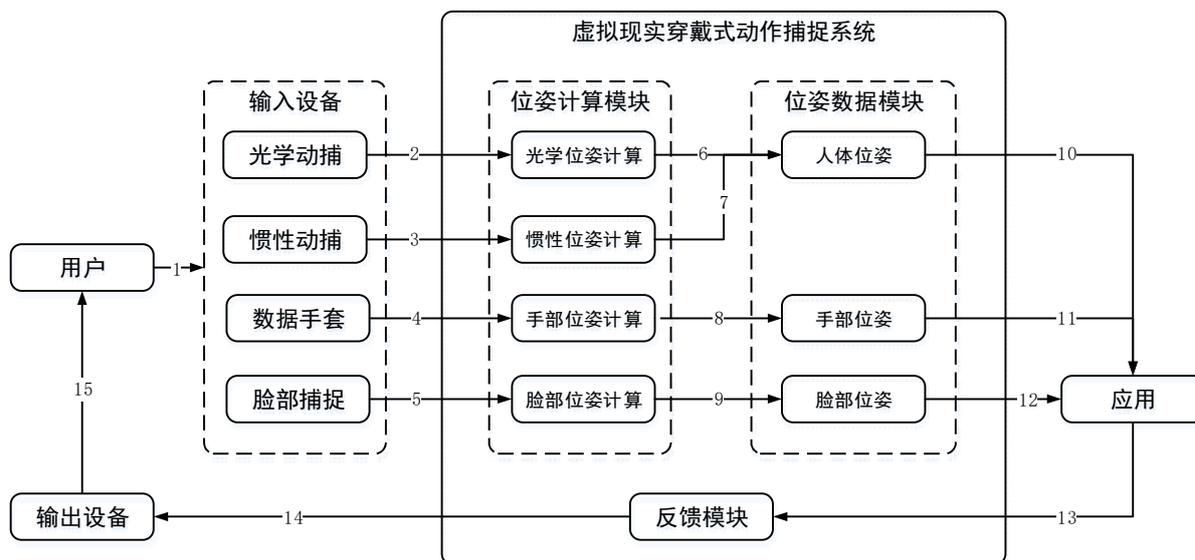


图1 面向虚拟现实应用的人体动作捕捉系统框架

图中：

- 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；
- 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 022.3-2023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；
- 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 022.3-2023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；

## T/UWA 022.1-2023

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储存介质;
- 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点;手部动捕设备关键点包括手指关节点;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保存至本地存储介质，数据从开始存储到结束存储的时间应 $\leq 500\text{ms}$ ；
- 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 $\leq 1000\text{ms}$ 。