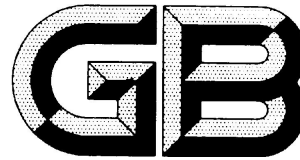


ICS 35.240.01

CCS L70



中华人民共和国国家标准

GB/T XXXXX—XXXX

信息技术 移动设备增强现实系统技术规范

Information technology—Technical specification of the augmented reality system for mobile devices

点击此处添加与国际标准一致性程度的标识

(征求意见稿)

(在提交反馈意见时,请将您知道的相关专利连同支持性文件一并附上)

××××-××-××发布

××××-××-××实施

国家市场监督管理总局
国家标准化管理委员会

发布

目 次

目 次	I
前 言	III
信息技术 移动设备增强现实系统技术规范	1
1 范围	1
2 规范性引用文件	1
3 术语和定义	1
4 缩略语	3
5 系统结构	4
6 功能要求	4
6.1 在线跟踪定位	4
6.2 尺度估计	5
6.3 标志物识别跟踪	5
6.4 三维重建	5
6.5 光照估计	6
6.6 人脸识别跟踪	6
6.7 手势识别跟踪	6
6.8 渲染引擎	6
6.9 虚实遮挡	7
7 性能要求	7
7.1 在线跟踪定位	7
7.2 尺度估计	8
7.3 标志物识别跟踪	8
7.4 三维重建	8
7.5 光照估计	9
7.6 人脸识别跟踪	9
7.7 手势识别跟踪	9
7.8 渲染引擎	9
7.9 虚实遮挡	9
7.10 运行	9
8 测试方法	10
8.1 测试环境	10
8.1.1 室内小场景	10
8.1.2 室内外大场景	10
8.2 功能测试方法	10
8.2.1 在线跟踪定位	11
8.2.2 尺度估计	11
8.2.3 标志物识别跟踪	11
8.2.4 三维重建	12
8.2.5 光照估计	12
8.2.6 人脸识别跟踪	12

8.2.7 手势识别跟踪	12
8.2.8 渲染引擎	12
8.2.9 虚实遮挡	13
8.3 性能测试方法	13
8.3.1 在线跟踪定位	13
8.3.2 尺度估计	13
8.3.3 标志物识别跟踪	13
8.3.4 三维重建	14
8.3.5 光照估计	15
8.3.6 人脸识别跟踪	15
8.3.7 手势识别跟踪	15
8.3.8 渲染引擎	15
8.3.9 虚实遮挡	15
8.3.10 运行	15
附录 A（规范性） 增强现实系统相关手持式移动设备传感器	17
A.1 曝光参数	17
A.2 惯性测量单元	17
A.3 对焦模式	17
A.4 时间对准	17
A.5 摄像头图像质量	17
A.6 深度摄像头	17
附录 B（规范性） 面部和手势关键点分布	19
B.1 面部关键点分布	19
B.2 手部关键点分布	19

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。

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。

本文件由全国信息技术标准化技术委员会（SAC/TC28）提出并归口。

本文件起草单位：浙江商汤科技开发有限公司、中国电子技术标准化研究院。

本文件主要起草人：

信息技术 移动设备增强现实系统技术规范

1 范围

本文件给出了移动设备的增强现实系统的结构，规定了功能要求、基本性能要求和测试方法。本文件适用于移动设备增强现实系统的设计、生产、应用和维护。

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

3 术语和定义

GB/T 38247-2019界定的以及下列术语和定义适合于本文件。

3.1

手持式移动设备 handheld mobile device

具备摄像头、显示屏、网络功能以及可检测其方位和运动信息的元件的手持式终端设备。

注：包括具备六自由度定位功能的手机或平板电脑等智能设备。

3.2

增强现实 augmented reality

采用以计算机为核心的现代高科技手段生成的附加信息对使用者感知到的真实世界进行增强的环境，生成的信息以视觉、听觉、味觉、嗅觉、触觉等生理感觉融合的方式叠加至真实场景中。

[来源：GB/T 38247-2019，2.1.2]

3.3

真实环境 actual environment

人类所处的不需要借助装备就能感知的包含多物理对象的真实世界。

[来源：GB/T 38247-2019，2.2.4]

3.4

虚拟对象 virtual object

计算机生成的具有几何形状、特定格式或特定行为的对象。

注：其原型可以是现实对象，也可以是完全虚构的对象。

[来源：GB/T 38247-2019，2.2.1]

3.5

二维标志物 2 dimensional marker

预先设定的具有特殊编码或图案的二维图像，通过检测、跟踪或识别该二维图像，可在场景、图像或视频中估计出其相对于手持式移动设备的位置和朝向。

3.6

三维标志物 3 dimensional marker

预先设定的具有特殊形状或纹理的三维物体，通过检测、跟踪或识别该三维图像，可在场景、图像或视频中估计出其相对于手持式移动设备的位置和朝向。

3.7

锚点 anchor point

将虚拟物体放置在环境过程中所使用的基准点。

[来源：GB/T 38247-2019，2.2.45]

3.8

云端 cloud end

供具备访问权限的手持式移动设备远程接入、具备计算和存储功能的系统。

3.9

定位 localization

获取确定指定对象的空间六自由度位置和姿态信息。

3.10

重定位 relocalization

当跟踪定位丢失时，手持式移动设备根据之前跟踪过的历史位置和朝向，通过触发对手持式移动设备的重新定位机制来获取正确的位姿信息，并重置手持式移动设备当前的位姿。

3.11

六自由度跟踪 six degree of freedom tracking

实时计算手持式移动设备相对真实场景的六自由度位置和姿态。

3.12

同步定位与地图构建 simultaneous localization and mapping

设备（移动终端或者机器人等）在未知环境里通过观察外部环境来定位自身的位置和姿态，再通过自身的位置增量式地构建未知环境的地图。

[来源：GB/T 38247-2019，2.2.67]

3.13

光照估计 illumination estimation

从传感器或摄像机视图中分析计算物理场景的光源照度分布信息的过程。

[来源：GB/T 38247-2019，2.2.50]

3.14

尺度估计 scale estimation

在系统中获得物理世界的长度尺寸信息的过程。

[来源：GB/T 38247-2019，2.2.51]

3.15

绝对位置误差 absolute position error

增强现实系统所测得的当前时刻手持式移动设备位置的真实值与估计值之间的平均偏差。

3.16

绝对旋转误差 absolute rotation error

增强现实系统所测得的当前时刻手持式移动设备旋转角度的真实值与估计值之间的平均偏差。

3.17

相对位置误差 relative position error

手持式移动设备位置发生变化时，增强现实系统所测得的当前时刻与前一时刻位置的真实值变化量与估计值变化量之间的平均偏差。

3.18

相对旋转误差 relative rotation error

手持式移动设备旋转角度发生变化时，增强现实系统所测得的当前时刻与前一时刻旋转角度的真实值变化量与估计值变化量之间的平均偏差。

3.19

关键点平均召回率 keypoints mean recall

手持式移动设备拍摄的图片中，正确跟踪的面部或手部的关键点与实际存在的关键点数量的平均比值。

注1：面部和手部的关键点分布应符合附录B的规定。

注2：正确跟踪是指所跟踪的关键点与实际存在的关键点的距离小于待检测区域宽度十分之一的状态。

4 缩略语

下列缩略语适用于本文件。

6DoF： 六自由度 (Six Degrees of Freedom)

APE： 绝对位置误差 (Absolute Position Error)

AR： 增强现实 (Augmented Reality)

ARE： 绝对旋转误差 (Absolute Rotation Error)

CPU： 中央处理器 (Central Processing Unit)

FPS： 每秒帧数 (Frames Per Second)

IMU： 惯性测量单元 (Inertial Measurement Unit)

RPE： 相对位置误差 (Relative Position Error)

RRE： 相对旋转误差 (Relative Rotation Error)

SLAM： 同步定位与地图构建 (Simultaneous Localization And Mapping)

5 系统结构

手持式移动设备增强现实系统处理单元由在线跟踪定位、尺度估计、标志物识别跟踪、三维重建、虚实遮挡模块、标志物识别与跟踪、光照估计、人脸识别跟踪、手势识别跟踪、渲染引擎、交互引擎等模块组成。在线跟踪定位模块利用手持式移动设备的输入数据计算6DoF位姿；尺度估计模块利用输入数据计算尺度信息；三维重建模块利用6DoF位姿、尺度信息和深度流数据重建场景的三维模型；虚实遮挡模块利用视频流和深度流实现虚拟物体与现实环境的遮挡效果；标志物识别跟踪模块基于视频流实现标志物的识别与跟踪；光照估计模块利用视频流实现光照信息的解算；人脸识别跟踪模块对视频流中的人脸关键点进行识别与跟踪；手势识别跟踪模块对视频流中的手部关键点进行识别与跟踪；如果手持式移动设备具备云计算服务，还支持将数据通过网络传输到云服务器端实现更效率的云端重定位、云端标志物识别和云端三维重建，并将解算获得的6DoF位姿、标志物信息与三维模型通过网络传输回手持式移动设备处理单元；6DoF位姿、三维模型、虚实遮挡、标志物识别与跟踪、光照这些信息汇总至渲染引擎实现真实感的增强现实虚实融合效果，人脸关键点、手部关键点这些信息汇总至交互引擎实现真实感的增强现实互动效果，并在手持式移动设备的显示屏上呈现。手持式移动设备增强现实系统的结构图见图1。增强现实系统相关手持式移动设备传感器要求应符合附录A的规定。

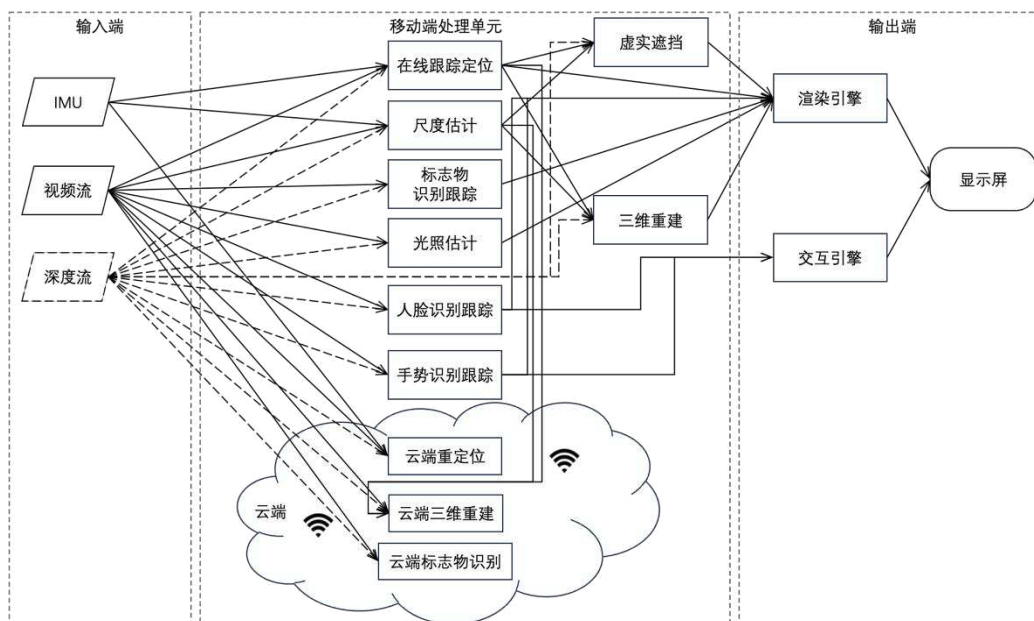


图1 手持式移动设备增强现实系统结构图

6 功能要求

6.1 在线跟踪定位

6.1.1 六自由度在线跟踪

手持式移动设备增强现实系统应在线估计设备的六个自由度位姿。

6.1.2 本地重定位

当跟踪定位丢失时，手持式移动设备增强现实系统应在本地重新定位并重置设备当前的位姿。

6.1.3 云端重定位

当跟踪定位丢失时，端云协同的手持式移动设备增强现实系统应在云端重新定位并重置设备当前

的位姿。

6.2 尺度估计

手持式移动设备增强现实系统应符合以下要求：

- a) 得到物理世界的尺度信息；
- b) 使得虚拟场景与物理世界注册在同一尺度的坐标系下；
- c) 实现虚拟对象与物理世界 1:1 虚实融合的效果。

6.3 标志物识别跟踪

6.3.1 二维标志物识别跟踪

手持式移动设备增强现实系统应符合以下要求：

- a) 对定义好的场景二维标志物进行识别；
- b) 对定义好的场景二维标志物进行 6DoF 位姿跟踪；
- c) 在识别到预定义的特定标志时，触发增强现实效果。

6.3.2 三维标志物识别跟踪

手持式移动设备增强现实系统应符合以下要求：

- a) 对真实场景中三维标志物体的纹理或结构信息进行预处理；
- b) 将来自摄像头的实时信息与处理过的信息进行比较，实现二维和三维信息的匹配；
- c) 理解场景中三维物体的位置和朝向。

6.3.3 云端标志物识别

端云协同的手持式移动设备增强现实系统应符合以下要求：

- a) 对用户上传到云端的带二维标志物的图像帧，利用云端服务器计算资源进行识别，并返回该帧二维标志物的位置和朝向信息至用户端；
- b) 对用户上传到云端的带三维物体的图像帧，利用云端服务器计算资源进行识别，并返回该帧三维物体的位置和朝向信息至用户端。

6.4 三维重建

6.4.1 平面检测及求锚点

6.4.1.1 单一平面重建及求锚点

手持式移动设备增强现实系统应符合以下要求：

- a) 需支持水平面和竖直平面的检测；
- b) 理解真实场景的主水平面，并在其上求出锚点或对其进行渲染；
- c) 对平面上的三维点进行识别、判定平面位置并进行拓展；
- d) 对场景中单一平面的检测；
- e) 对单一平面进行重建并求锚点。

6.4.1.2 多个平面重建及求锚点

对于包含多个平面的场景，手持式移动设备增强现实系统应符合以下要求：

- a) 理解真实场景的多个平面（包括水平平面或垂直平面），并在其上求出锚点或对其进行渲染；
- b) 对平面上的三维点识别、判定平面位置并进行拓展；

- c) 对场景中多个平面进行检测;
- d) 对多个平面进行重建并求锚点。

6.4.1.3 平面位置自动调节

当发生回路闭合或触发其它历史帧位姿的优化行为时,手持式移动设备增强现实系统应符合以下要求:

- a) 根据位姿的优化,自动调整平面的位置;
- b) 自动调整平面上锚点的位置。

6.4.2 稠密点云重建及求锚点

当在较为复杂的非平面场景上置放虚拟物体时,手持式移动设备增强现实系统应符合以下要求:

- a) 支持渐增式的实时拓展;
- b) 重建场景的稠密三维点云信息,点云中的每个三维点包含位置、法向、颜色等信息;
- c) 求解稠密点云的锚点位置和法向。

6.4.3 稠密网格重建及求锚点

当在真实场景上构建遮挡、阴影、碰撞等复杂的虚实融合效果时,手持式移动设备增强现实系统应符合以下要求:

- a) 重建场景稠密的三维网格信息;
- b) 求解稠密网格的锚点位置和法向。

6.4.4 云端三维重建

端云协同的手持式移动设备增强现实系统应符合以下要求:

- a) 支持渐增式的实时拓展;
- b) 利用云端服务器计算资源对用户端上传的关键帧图像进行三维重建;
- c) 将重建结果与统一坐标对齐;
- d) 读取或显示重建信息。

6.5 光照估计

手持式移动设备增强现实系统应从传感器或者摄像头视图中估计出用于虚拟对象、虚拟场景渲染的物理场景的全局光照环境信息。

6.6 人脸识别跟踪

手持式移动设备增强现实系统人脸识别跟踪应符合以下要求:

- a) 在手持式移动设备的镜头视野下能够检测人脸的外接框;
- b) 检测眼睛、眉毛、鼻子、嘴巴、脸部轮廓等一系列不少于 68 个面部关键点;
- c) 在大的头部姿态、不同的光照条件或极端表情等复杂条件下能够实现准确的关键点跟踪。

6.7 手势识别跟踪

手持式移动设备增强现实系统手势识别跟踪应符合以下要求:

- a) 在手持式移动设备的镜头视野下能够检测手部的外接框;
- b) 检测手指尖、手关节、手掌等一系列不少于 21 个手部关键点;
- c) 在不同手势动作、手部运动、光照条件下能够实现准确的关键点跟踪。

6.8 渲染引擎

手持式移动设备增强现实系统应符合以下要求：

- a) 实时将动/静态虚拟对象渲染在现实画面中；
- b) 紧密贴合虚拟对象与现实背景；
- c) 正确响应现实世界的光照变化；
- d) 支持静态模型和动态模型的渲染。

6.9 虚实遮挡

手持式移动设备增强现实系统应实现虚拟物体与真实环境之间的遮挡，包括虚拟对象与静态背景环境之间的遮挡和虚拟对象与动态前景之间的遮挡。

7 性能要求

7.1 在线跟踪定位

7.1.1 六自由度在线跟踪

在本文件所述基准测试数据集所涵盖的运动类型和测试场景下，手持式移动设备增强现实系统6DoF在线跟踪应符合以下要求：

- a) 帧率不低于24FPS；
- b) 跟踪准确率不低于95%，即测试数据集中的同时符合以下要求的帧数不低于总帧数的95%：
 - 1) APE 小于10cm 或小于摄像头运动轨迹最大距离的5%（以较大值为准）；
 - 2) ARE 小于 6° ；
 - 3) 与前0.1s时刻的相对位置误差不超过3cm；
 - 4) 与前0.1s时刻的角度误差不超过 2° 。
- c) 初始化质量指标平均不超过5；

注：初始化指标定义如下： $\varepsilon_{init} = t_{init} * (\varepsilon_{scale} + 1)^{0.5}$ ，其中 t_{init} 表示初始化完成的时间， ε_{scale} 表示初始化的尺度误差，可以基于公式 $\varepsilon_{scale} = 0.5 * (|s - 1| + |1/s - 1|)$ 计算得到 ε_{scale} ，其中 s 是待测试的六自由度在线跟踪模块计算得到的尺度系数。

- d) 跟踪鲁棒性指标平均不超过2。

注：跟踪鲁棒性指标定义如下： $\varepsilon_R = (\alpha_{lost} + 0.05) * (\varepsilon_{RL} + 0.1 * \varepsilon_{APE})$ ，其中 α_{lost} 表示跟踪丢失的帧数占整个输入视频帧序列的比例， ε_{RL} 表示整个视频帧序列中每次重定位的平均误差， ε_{APE} 表示整个视频帧序列的绝对位置误差。

7.1.2 本地重定位

手持式移动设备增强现实系统本地重定位应符合以下要求：

- a) 重定位成功率不低于90%，即测试数据集中的同时符合以下要求1、2的重定位次数不低于符合要求1的重定位次数的90%：
 - 1) 用户将手持式移动设备置回到未跟踪丢失之前的位置完成重定位；
 - 2) 重定位的位置与未跟踪丢失之前估计的位置偏差小于5cm 或小于摄像头所拍摄画面深度的中值的5%（以较大值为准）；
- b) 在重定位成功的情况下，重定位时间不超过2s。

7.1.3 云端重定位

手持式移动设备增强现实系统云端重定位应符合以下要求：

- a) 重定位成功率不低于90%，即测试数据集中的重定位位置与真值的误差小于10cm 或小于摄像

头所拍摄画面深度的中值的 5%（以较大值为准）；

b) 在重定位成功的情况下，重定位时间不超过 2s。

7.2 尺度估计

手持式移动设备增强现实系统对真实环境的尺度估计值与其真实值的偏差不超过 5%。

7.3 标志物识别跟踪

7.3.1 二维标志物识别跟踪

手持式移动设备增强现实系统二维标志物识别跟踪应符合以下要求：

- a) 支持不少于 4 个平面标志的同时识别与跟踪；
- b) 在识别成功的情况下，单个平面标志物识别的延迟不超过 0.5s；
- c) 识别准确率不低于 90%，即测试数据集中的符合估计位姿下二维标志物的投影轮廓与真值之间的平均偏差不超过 5 个像素或真值轮廓宽高较大值的 1%（以较大值为准）的二维标志物帧数不低于总帧数的 90%；
- d) 单个平面标志物的跟踪帧率不低于 24FPS；
- e) 多个平面标志物的跟踪帧率不低于 20FPS。

7.3.2 三维标志物识别跟踪

手持式移动设备增强现实系统三维标志物识别跟踪应符合以下要求：

- a) 帧率不低于 24FPS；
- b) 在识别成功的情况下，识别时间不超过 1s；
- c) 跟踪位置与真值的误差不超过 3%，即不大于物体运动轨迹的最大距离的 3%；
- d) 识别准确率不低于 90%，即测试数据集中的符合估计位姿下三维标志物的投影轮廓与真值之间的平均偏差不超过 5 个像素或真值轮廓宽高较大值的 1%（以较大值为准）三维标志物帧数不低于总帧数的 90%。

7.3.3 云端标志物识别

在识别成功的情况下，手持式移动设备增强现实系统的云端标志物识别时间（不含网络传输延时）应不超过 0.2s。

7.4 三维重建

7.4.1 平面检测

手持式移动设备增强现实系统平面检测应符合以下要求：

- a) 处理帧率与 6DoF 跟踪帧率一致；
- b) 平面位置精度误差不超过 2%。

注：平面位置精度误差即检测平面与真值平面之间的平均距离与待检测平面最大轴向尺寸的比值。

7.4.2 稠密点云重建

手持式移动设备增强现实系统稠密点云重建应符合以下要求：

- a) 处理帧率与 6DoF 跟踪帧率一致；
- b) 稠密点云的位置精度误差不超过 3%。

注：稠密点云位置精度误差即重建点云与真值点云之间的平均距离与待重建点云模型最大轴向尺寸的比值。

7.4.3 稠密网格重建

手持式移动设备增强现实系统稠密网格重建应符合以下要求：

- a) 处理帧率与 6DoF 跟踪帧率一致；
- b) 稠密网格的几何精度误差不超过 3%。

注：稠密网格位置精度误差即重建网格与真值网格之间的平均距离与待重建网格模型最大轴向尺寸的比值。

7.5 光照估计

手持式移动设备增强现实系统光照估计应符合以下要求：

- a) 对于环境变化的响应时间应不超过 1s；
- b) 场景光照颜色的估计值与真值的误差应不超过 0.3（光照颜色值在亮度色彩空间中统一归一化为 0 至 1 的范围）。

7.6 人脸识别跟踪

手持式移动设备增强现实系统人脸识别跟踪应符合以下要求：

- a) 跟踪帧率不小于 20FPS；
- b) 关键点平均跟踪召回率不低于 90%。

7.7 手势识别跟踪

手持式移动设备增强现实系统手势识别跟踪应符合以下要求：

- a) 跟踪帧率不小于 20FPS；
- b) 关键点平均跟踪召回率不低于 90%。

7.8 渲染引擎

手持式移动设备增强现实系统渲染引擎应符合以下要求：

- a) 渲染帧率不低于摄像头输入视频流的帧率；
- b) 渲染分辨率不低于摄像头输入视频流分辨率。

7.9 虚实遮挡

手持式移动设备增强现实系统虚实遮挡应符合以下要求：

- a) 画面深度图的获取和遮挡处理的帧率与 6DoF 跟踪帧率相同；
- b) 遮挡边缘的偏差不得超过 5 个像素或画面宽高较大值的 1%（以较大值为准）；
- c) 画面中遮挡关系错误占比不得超过 10%。

7.10 运行

7.10.1 CPU 和内存占用

在单目方案中，手持式移动设备增强现实系统在运行基本的跟踪和稀疏点云地图构建时，应符合以下要求：

- a) CPU 占用率不超过 50%；
- b) 在 5m×5m 的房间内，内存占用量不超过 500MB。

7.10.2 算法运行效率

手持式移动设备增强现实系统的运行帧率应不小于 20FPS。

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：<https://d.book118.com/975321131202011232>