P2048.101TM/D3.0  
Draft Standard for Augmented Reality on Mobile Devices-General Requirements for Software Framework, Components and Intergration

Developed by the

Virtual Reality and Augmented Reality Standards Committee

of the

IEEE Consumer Technology Society

Approved

IEEE SA Standards Board

Copyright © 2022 by The Institute of Electrical and Electronics Engineers, Inc.

Three Park Avenue

New York, New York 10016-5997, USA

All rights reserved.

This document is an unapproved draft of a proposed IEEE Standard. As such, this document is subject to change. USE AT YOUR OWN RISK! IEEE copyright statements SHALL NOT BE REMOVED from draft or approved IEEE standards, or modified in any way. Because this is an unapproved draft, this document must not be utilized for any conformance/compliance purposes. Permission is hereby granted for officers from each IEEE Standards Working Group or Committee to reproduce the draft document developed by that Working Group for purposes of international standardization consideration. IEEE Standards Department must be informed of the submission for consideration prior to any reproduction for international standardization consideration ([stds.ipr@ieee.org](mailto:stds.ipr@ieee.org)). Prior to adoption of this document, in whole or in part, by another standards development organization, permission must first be obtained from the IEEE Standards Department ([stds.ipr@ieee.org](mailto:stds.ipr@ieee.org)). When requesting permission, IEEE Standards Department will require a copy of the standard development organization's document highlighting the use of IEEE content. Other entities seeking permission to reproduce this document, in whole or in part, must also obtain permission from the IEEE Standards Department.

IEEE Standards Department

445 Hoes Lane

Piscataway, NJ 08854, USA

Abstract: This standard specifies the general technical framework, components, integration, and main business processes of augmented reality systems applied to mobile devices, and defines its technical requirements, including functional requirements, performance requirements and the corresponding test methods.

The target applications and services include but are not limited to the design, development, and management of augmented reality and mixed reality enabled applications or features of applications on mobile devices.

Keywords: augmented reality, mobile device, real-time tracking and localization, scale estimation, target recognition and tracking, 3D reconstruction, illumination estimation, face alignment, gesture recognition, rendering engine, function requirement, performance requirement

[[1]](#footnote-2)·

Important Notices and Disclaimers Concerning IEEE Standards Documents

IEEE Standards documents are made available for use subject to important notices and legal disclaimers. These notices and disclaimers, or a reference to this page (<https://standards.ieee.org/ipr/disclaimers.html>), appear in all standards and may be found under the heading “Important Notices and Disclaimers Concerning IEEE Standards Documents.”

Notice and Disclaimer of Liability Concerning the Use of IEEE Standards Documents

IEEE Standards documents are developed within the IEEE Societies and the Standards Coordinating Committees of the IEEE Standards Association (IEEE SA) Standards Board. IEEE develops its standards through an accredited consensus development process, which brings together volunteers representing varied viewpoints and interests to achieve the final product. IEEE Standards are documents developed by volunteers with scientific, academic, and industry-based expertise in technical working groups. Volunteers are not necessarily members of IEEE or IEEE SA, and participate without compensation from IEEE. While IEEE administers the process and establishes rules to promote fairness in the consensus development process, IEEE does not independently evaluate, test, or verify the accuracy of any of the information or the soundness of any judgments contained in its standards.

IEEE makes no warranties or representations concerning its standards, and expressly disclaims all warranties, express or implied, concerning this standard, including but not limited to the warranties of merchantability, fitness for a particular purpose and non-infringement. In addition, IEEE does not warrant or represent that the use of the material contained in its standards is free from patent infringement. IEEE standards documents are supplied “AS IS” and “WITH ALL FAULTS.”

Use of an IEEE standard is wholly voluntary. The existence of an IEEE Standard does not imply that there are no other ways to produce, test, measure, purchase, market, or provide other goods and services related to the scope of the IEEE standard. Furthermore, the viewpoint expressed at the time a standard is approved and issued is subject to change brought about through developments in the state of the art and comments received from users of the standard.

In publishing and making its standards available, IEEE is not suggesting or rendering professional or other services for, or on behalf of, any person or entity, nor is IEEE undertaking to perform any duty owed by any other person or entity to another. Any person utilizing any IEEE Standards document, should rely upon his or her own independent judgment in the exercise of reasonable care in any given circumstances or, as appropriate, seek the advice of a competent professional in determining the appropriateness of a given IEEE standard.

IN NO EVENT SHALL IEEE BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO: THE NEED TO PROCURE SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE PUBLICATION, USE OF, OR RELIANCE UPON ANY STANDARD, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE AND REGARDLESS OF WHETHER SUCH DAMAGE WAS FORESEEABLE.

Translations

The IEEE consensus development process involves the review of documents in English only. In the event that an IEEE standard is translated, only the English version published by IEEE is the approved IEEE standard.

Official statements

A statement, written or oral, that is not processed in accordance with the IEEE SA Standards Board Operations Manual shall not be considered or inferred to be the official position of IEEE or any of its committees and shall not be considered to be, nor be relied upon as, a formal position of IEEE. At lectures, symposia, seminars, or educational courses, an individual presenting information on IEEE standards shall make it clear that the presenter’s views should be considered the personal views of that individual rather than the formal position of IEEE, IEEE SA, the Standards Committee, or the Working Group.

Comments on standards

Comments for revision of IEEE Standards documents are welcome from any interested party, regardless of membership affiliation with IEEE or IEEE SA. However, **IEEE does not provide interpretations, consulting information, or advice pertaining to IEEE Standards documents**.

Suggestions for changes in documents should be in the form of a proposed change of text, together with appropriate supporting comments. Since IEEE standards represent a consensus of concerned interests, it is important that any responses to comments and questions also receive the concurrence of a balance of interests. For this reason, IEEE and the members of its Societies and Standards Coordinating Committees are not able to provide an instant response to comments, or questions except in those cases where the matter has previously been addressed. For the same reason, IEEE does not respond to interpretation requests. Any person who would like to participate in evaluating comments or in revisions to an IEEE standard is welcome to join the relevant IEEE working group. You can indicate interest in a working group using the Interests tab in the Manage Profile & Interests area of the [IEEE SA myProject system](https://development.standards.ieee.org/myproject-web/public/view.html#landing). An IEEE Account is needed to access the application.

Comments on standards should be submitted using the [Contact Us](https://standards.ieee.org/content/ieee-standards/en/about/contact/index.html) form.

Laws and regulations

Users of IEEE Standards documents should consult all applicable laws and regulations. Compliance with the provisions of any IEEE Standards document does not constitute compliance to any applicable regulatory requirements. Implementers of the standard are responsible for observing or referring to the applicable regulatory requirements. IEEE does not, by the publication of its standards, intend to urge action that is not in compliance with applicable laws, and these documents may not be construed as doing so.

Data privacy

Users of IEEE Standards documents should evaluate the standards for considerations of data privacy and data ownership in the context of assessing and using the standards in compliance with applicable laws and regulations.

Copyrights

IEEE draft and approved standards are copyrighted by IEEE under US and international copyright laws. They are made available by IEEE and are adopted for a wide variety of both public and private uses. These include both use, by reference, in laws and regulations, and use in private self-regulation, standardization, and the promotion of engineering practices and methods. By making these documents available for use and adoption by public authorities and private users, IEEE does not waive any rights in copyright to the documents.

Photocopies

Subject to payment of the appropriate licensing fees, IEEE will grant users a limited, non-exclusive license to photocopy portions of any individual standard for company or organizational internal use or individual, non-commercial use only. To arrange for payment of licensing fees, please contact Copyright Clearance Center, Customer Service, 222 Rosewood Drive, Danvers, MA 01923 USA; +1 978 750 8400; https://www.copyright.com/. Permission to photocopy portions of any individual standard for educational classroom use can also be obtained through the Copyright Clearance Center.

Updating of IEEE Standards documents

Users of IEEE Standards documents should be aware that these documents may be superseded at any time by the issuance of new editions or may be amended from time to time through the issuance of amendments, corrigenda, or errata. An official IEEE document at any point in time consists of the current edition of the document together with any amendments, corrigenda, or errata then in effect.

Every IEEE standard is subjected to review at least every 10 years. When a document is more than 10 years old and has not undergone a revision process, it is reasonable to conclude that its contents, although still of some value, do not wholly reflect the present state of the art. Users are cautioned to check to determine that they have the latest edition of any IEEE standard.

In order to determine whether a given document is the current edition and whether it has been amended through the issuance of amendments, corrigenda, or errata, visit [IEEE Xplore](https://ieeexplore.ieee.org/browse/standards/collection/ieee/) or [contact IEEE](https://standards.ieee.org/content/ieee-standards/en/about/contact/index.html). For more information about the IEEE SA or IEEE’s standards development process, visit the IEEE SA Website.

Errata

Errata, if any, for all IEEE standards can be accessed on the [IEEE SA Website](https://standards.ieee.org/standard/index.html). Search for standard number and year of approval to access the web page of the published standard. Errata links are located under the Additional Resources Details section. Errata are also available in [IEEE Xplore](https://ieeexplore.ieee.org/browse/standards/collection/ieee/). Users are encouraged to periodically check for errata.

Patents

IEEE Standards are developed in compliance with the [IEEE SA Patent Policy](https://standards.ieee.org/about/sasb/patcom/materials.html).

Attention is called to the possibility that implementation of this standard may require use of subject matter covered by patent rights. By publication of this standard, no position is taken by the IEEE with respect to the existence or validity of any patent rights in connection therewith. If a patent holder or patent applicant has filed a statement of assurance via an Accepted Letter of Assurance, then the statement is listed on the IEEE SA Website at <https://standards.ieee.org/about/sasb/patcom/patents.html>. Letters of Assurance may indicate whether the Submitter is willing or unwilling to grant licenses under patent rights without compensation or under reasonable rates, with reasonable terms and conditions that are demonstrably free of any unfair discrimination to applicants desiring to obtain such licenses.

Essential Patent Claims may exist for which a Letter of Assurance has not been received. The IEEE is not responsible for identifying Essential Patent Claims for which a license may be required, for conducting inquiries into the legal validity or scope of Patents Claims, or determining whether any licensing terms or conditions provided in connection with submission of a Letter of Assurance, if any, or in any licensing agreements are reasonable or non-discriminatory. Users of this standard are expressly advised that determination of the validity of any patent rights, and the risk of infringement of such rights, is entirely their own responsibility. Further information may be obtained from the IEEE Standards Association.

IMPORTANT NOTICE

IEEE Standards do not guarantee or ensure safety, security, health, or environmental protection, or ensure against interference with or from other devices or networks. IEEE Standards development activities consider research and information presented to the standards development group in developing any safety recommendations. Other information about safety practices, changes in technology or technology implementation, or impact by peripheral systems also may be pertinent to safety considerations during implementation of the standard. Implementers and users of IEEE Standards documents are responsible for determining and complying with all appropriate safety, security, environmental, health, and interference protection practices and all applicable laws and regulations.

Participants

At the time this draft <gde./rec. prac./std.> was completed, the <Working Group Name> Working Group had the following membership:

<Chair Name>, Chair

<Vice-chair Name>, Vice Chair

Participant1

Participant2

Participant3

Participant4

Participant5

Participant6

Participant7

Participant8

Participant9

The following members of the <individual/entity> Standards Association balloting group voted on this <gde./rec. prac./std.>. Balloters may have voted for approval, disapproval, or abstention.

[To be supplied by IEEE]

Balloter1

Balloter2

Balloter3

Balloter4

Balloter5

Balloter6

Balloter7

Balloter8

Balloter9

When the IEEE SA Standards Board approved this <gde./rec. prac./std.> on <Date Approved>, it had the following membership:

[To be supplied by IEEE]

<Name>, Chair

<Name>, Vice Chair

<Name>, Past Chair

**Konstantinos Karachalios**, Secretary

SBMember1

SBMember2

SBMember3

SBMember4

SBMember5

SBMember6

SBMember7

SBMember8

SBMember9

\*Member Emeritus

Introduction

This introduction is not part of P<designation>/D<draft\_number>, Draft <Gde./Rec. Prac./Std.> for <Complete Title Matching PAR>.

<Select this text and type or paste introduction text>

Contents

<After draft body is complete, select this text and click Insert Special->Add (Table of) Contents>

Draft Standard for Augmented Reality on Mobile Devices-General Requirements for Software Framework, Components and Intergration

1. Overview
   1. Scope

This standard specifies the general technical framework, components, integration, and main business processes of augmented reality systems applied to mobile devices, and defines its technical requirements, including functional requirements, performance requirements and corresponding test methods.

This standard is applicable to the design, development, and management of augmented reality enabled applications or features of applications on mobile devices.

* 1. Purpose

This document will not include a purpose clause.

* 1. Word usage

***<This subclause is mandatory and shall appear after the Scope and Purpose (if included).>***

The word *shall* indicates mandatory requirements strictly to be followed in order to conform to the standard and from which no deviation is permitted (*shall* equals is *required to*).[[2]](#footnote-3),[[3]](#footnote-4)

The word *should* indicates that among several possibilities one is recommended as particularly suitable, without mentioning or excluding others; or that a certain course of action is preferred but not necessarily required (*should* equals is *recommended that*).

The word *may* is used to indicate a course of action permissible within the limits of the standard (*may* equals is *permitted to*).

The word *can* is used for statements of possibility and capability, whether material, physical, or causal (*can* equals is *able to*).

1. Normative references

The following referenced documents are indispensable for the application of this document (i.e., they must be understood and used, so each referenced document is cited in text and its relationship to this document is explained). For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments or corrigenda) applies.

ISO IEC 18039 Information technology — Computer graphics, image processing and environmental data representation — Mixed and augmented reality (MAR) reference model

ETSI GS Augmented Reality Framework (ARF) 003

1. Definitions, acronyms, and abbreviations
   1. Definitions

For the purposes of this document, the following terms and definitions apply. The IEEE Standards Dictionary Online should be consulted for terms not defined in this clause. [[4]](#footnote-5)

3D feature map: 3D digital representation for a physical scene, including a set of images with known global positions and orientations, the feature connectivity among the images, as well as sparse 3D feature points of the scene and their 2D feature observations in the images in a world coordinates of the scene with real scale

6DoF tracking: Action for calculating the 6DoF pose of mobile device in the world coordinates of the physical scene in real time

pose: the 6DoF transformation of mobile device, including its position in 3DoF and orientation in 3DoF

absolute position error: The deviation between the true position and the estimated position of the mobile device at current time of the augmented reality system while the position of mobile device is changing

absolute rotation error: The deviation between the true orientation and the estimated orientation of the mobile device at current time of the augmented reality system while the orientation of mobile device is changing

augmented reality: Ability to mix in real-time spatially-registered digital content with the real world (ETSI GS ARF 003)

illumination estimation: Action for calculating the illumination distribution information of the physical scene

localization: Action for getting the pose of a specified object or device in the world coordinates of the physical scene

mobile device: A device equipped with camera, display screen, computational unit, and network functions, of which the pose and motion information are detectable

Note 1— Mobile devices include wearable devices, phones, AR glasses or tablets with GPS, accelerometers, magnetometers and gyroscopes.

relative position error: The deviation between the true variation and the estimated variation between the positions of the mobile device at the current time and the previous time of the augmented reality system while the position of mobile device is changing

relative rotation error: The deviation between the true variation and the estimated variation between the orientations of the mobile device at the current time and the previous time of the augmented reality system while the orientation of mobile device is changing

relocalization: Action of obtaining and resetting the correct pose of the mobile device when the mobile device loses its pose information due to unexpected motion or visual mutation

scale estimation: Action of obtaining the real scale information for the 3D digital representation of the physical scene in its world coordinate system

simultaneous localization and mapping: Action for locating the pose of a mobile device in the world coordinate system of a physical scene in real-time by incrementally observing and mapping the scene while inferring its pose

target image: Target object represented by a 2D image(ISO/IEC 18039-2019-02)

target object: A physical or virtual object that is designated, designed or chosen to allow detection, recognition, tracking, and augmentation of it in the world coordinate system of a physical scene

virtual object: Computer-generated entity that is designated for augmentation in association with a physical scene or object data representation

Note1—In the context of MAR, it usually has perceptual (e.g. visual, aural) characteristics and, optionally, dynamic reactive behavior

* 1. Acronyms and abbreviations

2D: 2 dimensional

3D: 3 dimensional

6DoF: six　degrees　of　freedom

APE: absolute　position　error

AR: augmented　reality

ARE: absolute　rotation　error

CPU: central　processing　unit

FPS: frames　per　second

IMU: inertial　measurement　unit

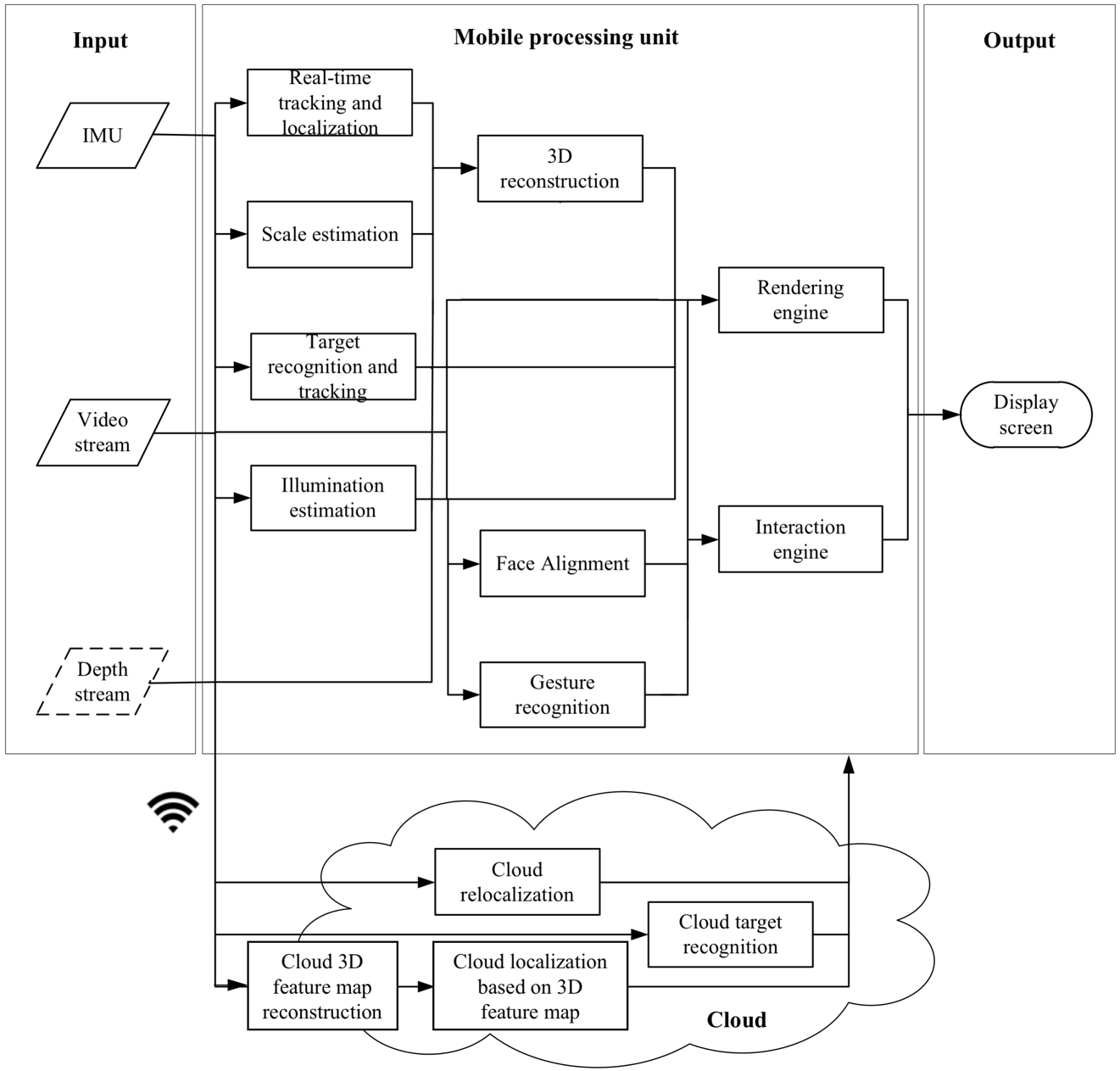
RPE: relative　position　error

RRE: relative　rotation　error

SLAM: simultaneous　localization　and　mapping

1. System structure

The augmented reality system on mobile devices consist of modules such as real-time tracking and localization, scale estimation, target recognition and tracking, 3D reconstruction, illumination estimation, face alignment, gesture recognition and rendering engine. The real-time tracking and localization module uses the input data of the mobile device to calculate the 6DoF pose. The scale estimation module uses the input data to calculate the scale information of the physical world. The 3D reconstruction module uses the 6DoF pose, scale information and depth data to reconstruct the 3D model of the scene. The target recognition and tracking module realizes the recognition and tracking of markers based on the video stream. The illumination estimation module uses the video stream to realize the calculation of illumination information. The face alignment module predicts a set of facial key points based on the video stream. The gesture recognition module detects hand bounding boxes and classifies hand gestures based on the video stream. If the mobile device has cloud computing services, it also supports the transmission of data to the cloud server through the network to achieve more efficient cloud relocation, cloud target recognition and cloud 3D feature map reconstruction, and transmits the obtained 6DoF pose and target information through the network to mobile device. The 6DoF pose, 3D model, target recognition and tracking, lighting are aggregated into the rendering engine to achieve a realistic augmented reality effect and presented on the display screen of the mobile device. Figure 1 is a workflow diagram of an augmented reality system on mobile device. The sensor requirements of augmented reality system on mobile devices shall comply with the provisions of Appendix A.



1. Workflow diagram of augmented reality system on mobile device
2. Functional requirements
   1. Real-time tracking and localization
      1. 6DoF real-time tracking

The augmented reality system on mobile device shall estimate the 6DoF pose of the device with respect to the real world coordinate reference system in real time.

* + 1. Initialization

The augmented reality system on mobile device shall automatically initialize with true scale.

* + 1. Local relocalization

When the mobile device loses its 6DoF pose due to unexpected position mutation, occlusion and so on, the augmented reality system on mobile device shall relocalize the device and reset its current pose.

* + 1. Cloud relocalization

When the mobile device loses its 6DoF pose due to unexpected position mutation, occlusion and so on, the augmented reality system on mobile device with device-cloud collaboration shall relocalize the device and reset its current pose at the cloud.

* 1. Cloud localization based on a 3D feature map

The augmented reality system on mobile device shall meet the following requirements:

a) construct a 3D feature map for a defined scene on the cloud;

b) calculate the 6DoF pose in the coordinate system of the 3D feature map on the cloud for a query image which is uploaded to the cloud;

* 1. Target recognition and tracking
     1. Recognition and tracking of target image

The augmented reality system on mobile device shall meet the following requirements:

a) recognize the target image when scanning it for a while;

b) perform 6DoF pose tracking for the target image;

c) trigger the augmented reality effect when the predefined specific target image is recognized.

* + 1. Recognition and tracking of target object

The augmented reality system on mobile device shall meet the following requirements:

a) recognize the target object when scanning it for a while;

b) perform 6DoF pose tracking for the target object;

c) trigger the augmented reality effect when the predefined specific target object is recognized.

* + 1. Cloud target recognition

The augmented reality system on mobile device with device-cloud collaboration shall meet the following requirements:

a) recognize the target image from the image frame uploaded to the cloud by the mobile device system, and return the position and orientation information of the target image in the frame to the user end;

b) recognize the target 3D object from the image frame uploaded to the cloud by the mobile device system, and return the position and orientation information of the 3D object in the frame to the user end.

* 1. 3D reconstruction
     1. Plane detection
        1. Single plane reconstruction

The augmented reality system on a mobile device shall meet the following requirements:

a) detect the position of the main 3D horizontal plane that is perpendicular to the gravity of the physical scene in its world coordinate system;

b) determine the 3D polygonal boundary of the plane;

c) expand the 3D polygonal boundary and update the position of the plane in real time as the visible scene region is extended while the mobile device is moving.

* + - 1. Multi-plane reconstruction

For a scene containing multiple planar structures, the augmented reality system on a mobile device should meet the following requirements:

a) detect the positions of the multiple planes of the physical scene in its world coordinate system;

b) support the detection of horizontal planes perpendicular to the gravity and vertical planes parallel to the gravity of the physical scene;

c) determine the 3D polygonal boundaries of the multiple planes;

d) expand the 3D polygonal boundaries and update the positions of the multiple planes in real time as the visible scene region is extended while the mobile device is moving.

* + 1. Dense point cloud reconstruction

For a scene containing non-planar complex structures, the augmented reality system on a mobile device shall meet the following requirements:

a) reconstruct the dense 3D point cloud of the scene, each point contains the information of position, normal direction and color;

b) expand the dense 3D point cloud and update the positions of the 3D points in real time as the visible scene region is extended while the mobile device is moving.

* + 1. Dense mesh reconstruction

For a scene containing non-planar complex structures, the augmented reality system on a mobile device shall meet the following requirements:

a) reconstruct the dense 3D triangular mesh of the scene, with each vertex indexed by the triangles contains the information of position, normal direction and color;

b) expand the dense 3D mesh and update the vertex positions and mesh topology in real time as the visible scene region is extended while the mobile device is moving.

* 1. Cloud 3D feature map reconstruction

The augmented reality system shall meet the following requirements:

a) capture images of a specified scene with mobile devices.

b) send the images or the extracted visual features of the images to the cloud and use them to construct a 3D feature map of the scene.

* 1. Illumination estimation

The augmented reality system on mobile device shall estimate the lighting information of the physical scene from the input data, which can be used for the rendering of virtual objects and virtual scenes.

* 1. Face alignment

The augmented reality system on mobile device shall meet the following requirements:

a) predict a series of facial key points including eyes, eyebrows, nose, mouth and face contour;

b) achieve at least 85% accuracy rate in complex conditions such as large head poses, different illumination conditions or extreme expressions.

* 1. Gesture recognition

The augmented reality system on mobile device shall meet the following requirements:

a) detect hand bounding boxes in the defined scene;

b) classify specific hand gestures based on image classification.

* 1. Rendering engine

The augmented reality system on mobile device shall meet the following requirements:

a) render dynamic/static virtual object in real time;

b) fusion of virtual of objects and the physical world with matching camera parameters, light information, and 3D reconstruction scene data;

* 1. Asynchronous time warp

For optical see-through devices, the augmented reality system on mobile device shall meet the following requirements:

a) predict the pose at the future time when user sees virtual content before sending the rendered image to the display screen;

b) warp the rendered image according to the predicted pose and send the warped image to the display screen;

1. Performance requirements
   1. Real-time tracking and localization
      1. 6DoF real-time tracking

The 6DoF real-time tracking of augmented reality system on mobile device shall meet the following requirements:

a) the frame rate is no less than 24fps;

b) the tracking accuracy rate is not less than 95%. That is, the number of frames which meet the following requirements is not less than 95% of the total number of frames:

1) APE is less than 10cm or less than 3% of the maximum moving distance of the camera;

2) ARE is less than 3°;

3) the relative position error does not exceed 3cm within 0.1s;

4) the angle error does not exceed 2° within 0.1s.

* + 1. Initialization

The initialization of augmented reality system on mobile device shall meet the following requirements:

a) the initialization time shall be less than 3s;

b) the scale error of initialization shall be less than 15%.

* + 1. Local relocalization

The local relocalization of augmented reality system on mobile device shall meet the following requirements:

a) the success rate of relocalization is no less than 80%. That is, the number of relocalizations that meet the requirements of the following two requirements is no less than 80% of the number of relocalizations that meet the requirement 1):

1) the user puts the mobile device back to the position and orientation after the tracking loss to complete the relocalization;

2) the deviation between the relocalized position and the estimated position before the tracking loss is less than 10cm and less than 5% of the median value of the depths in the image taken by the camera;

b) if successful, the time of relocalization does not exceed 2s.

* + 1. Cloud relocalization

The cloud relocalization of augmented reality system on mobile device shall meet the following requirements:

a) the success rate of relocalization is no less than 90%, that is, the error between the relocalized position in the test dataset and the true value is less than 10cm or less than 5% the median value of the depth of the picture taken by the camera (whichever is greater);

b) if successful, the time of localization does not exceed 2s.

* 1. Cloud localization based on a 3D feature map

The cloud localization based on a 3D feature map for augmented reality system on mobile device shall meet the following requirements:

a) the success rate of localization is no less than 90%. That is, a successful localization requires that the error between the estimated position and the true value is less than 10cm and less than 5% the median value of the depths in the image captured by the camera;

b) if successful, the time of localization does not exceed 2s.

* 1. Target recognition and tracking
     1. Recognition and tracking of target image

The recognition and tracking of target image for augmented reality system on mobile device shall meet the following requirements:

a) the recognition and tracking of no less than four plane targets shall be supported at the same time;

b) in the case of successful recognition, the recognition delay of single plane target does not exceed 0.5s;

c) the recognition accuracy is no less than 90%, that is, the average deviation between the projection contour of target image under the estimated pose in the test dataset and the true value is no greater than 5 pixels and no greater than 1% the larger value of the width and height of the true value contour, and the number of frames for target image is no less than 90% the total number of frames;

d) the tracking frequency of single-plane target is no less than 24fps;

e) the tracking frequency of multi-plane target is no less than 20fps.

* + 1. Recognition and tracking of target object

The recognition and tracking of target object for augmented reality system on mobile device shall meet the following requirements:

a) the frequency is no less than 24fps;

b) in case of successful recognition, the recognition time does not exceed 1s;

c) the error between the tracked position and the true value is no greater than 3cm/m, and no greater than 3% the maximum moving distance of object;

d) the recognition accuracy is no less than 90%, that is, the average deviation between the projection contour of target object under the estimated pose in the test dataset and the true value is no greater than 5 pixels and no greater 1% the larger value of the width and height of the true value contour, and the number of frames for target object is no less than 90% the total number of frames.

* + 1. Cloud target recognition

In case of successful recognition, the cloud target recognition time (excluding network transmission delay) of augmented reality system on mobile device shall not exceed 0.2s per query.

* 1. 3D reconstruction
     1. Plane detection

The plane detection of augmented reality system on mobile device shall meet the following requirements:

a) the process time for plane extension is no greater than 166ms/keyframe;

b) the position error of plane does not exceed 2cm/m.

* + - 1. Dense point cloud reconstruction

The dense point cloud reconstruction of augmented reality system on a mobile device shall meet the following requirements:

a) the process time for dense point cloud extension is no greater than 166ms/keyframe;

b) the position error of dense point cloud does not exceed 3cm/m.

* + - 1. Dense mesh reconstruction

The dense mesh reconstruction of augmented reality system on a mobile device shall meet the following requirements:

a) the process time for dense mesh extension is no greater than 166ms/keyframe;

b) the geometric error of dense mesh does not exceed 3cm/m.

* 1. Cloud 3D feature map reconstruction

The cloud 3D feature map reconstruction for augmented reality system shall meet the following requirements:

1. for a scene smaller than 10,000 square meters, the average error of the 3D feature map does not exceed 10cm. The average error refers to the average distance between the positions of some sampled points in the map and their true values.
2. For a scene smaller than 10,000 square meters, the time of the map reconstruction does not exceeds 2 hours.
   1. Illumination estimation

The illumination estimation of augmented reality system on mobile device shall meet the following requirements:

a) the response time to environmental changes does not exceed 1s;

b) the error between the estimated value and the true value of the scene illumination shall not exceed 0.3 (the illumination value is uniformly normalized to the range of 0~1).

* 1. Face alignment

The face alignment of augmented reality system on mobile device shall meet the following requirements:

a) the frequency is no less than 30 fps;

b) the prediction accuracy shall not be less than 99%;

* 1. Gesture recognition

The gesture recognition of augmented reality system on mobile device shall meet the following requirements:

a) in case of successful recognition, the recognition time does not exceed 20ms;

b) the recognition accuracy shall not be less than 95%.

* 1. Rendering engine

The rendering engine of augmented reality system on mobile device shall meet the following requirements:

a) the rendering frame rate is not lower than the frame rate of video captured by the camera;

b) the rendering resolution is not lower than the resolution of video captured by the camera.

* 1. Asynchronous time warp

The asynchronous time warp of augmented reality system on optical see-through device shall meet the following requirements:

a) the frame screen refresh rate is constant and no less than 60Hz;

b) the translation error of predicted pose is less than 1cm, and the rotation error is less than 0.5°.

* 1. Running
     1. CPU and memory occupancy

In the monocular scheme, the augmented reality system on mobile device shall meet the following requirements when running basic tracking and sparse point cloud mapping:

a) the CPU occupancy rate is no greater than 50%;

b) in a 5m×5m room, the memory occupancy does not exceed 500MB.

* + 1. Algorithm running efficiency

The running frame rate of augmented reality system on mobile device shall not be less than 20fps.

1. Test method
   1. Test environment

In a 5m×5m laboratory, the standard test scene is arranged according to the following requirements:

a) arrange four light source modes on the ceiling of room, i.e. red, green, blue and white, and keep each light source diffused to ensure an uniform illumination in the scene; the wavelength of red light, green light and blue light is respectively 700nm, 550nm and 460nm, and the white light is the mixture of red, green and blue lights in average proportion. The illumination of each light source is allowed to be adjusted within the range of 20~200lx;

b) in the scene, three walls and a square table are available, with patterned wallpapers on the wall surface and patterned tablecloths on the square table;

c) place sundries on the corners of three walls and on the table;

d) place five target images and five target objects in the scene;

Beside the standard test scene, related public datasets can also be referred as benchmark dataset for test. For example, ZJU - SenseTime VISLAM Benchmark (<http://www.zjucvg.net/eval-vislam>) can be referred for testing real-time tracking and localization. ZJU - SenseTime 3DR dataset (<https://zju3dv.github.io/3DR-dataset.github.io>) can be referred for testing 3D reconstruction.

* 1. Function test method
     1. Real-time tracking and localization
        1. 6DoF real-time tracking

The 6DoF real-time tracking function of mobile device is tested according to the following method:

a) adjust the illumination color of the test environment to white light with the illumination of 400lx;

b) move in different ways by holding the mobile device and enable the augmented reality system;

c) observe the visual track of the device and the position of the virtual object on the display screen.

* + - 1. Initialization

The initialization function of mobile device is tested according to the following method:

a) adjust the illumination color of the test environment to white light with the illumination of 400lx;

b) move in different ways by holding the mobile device and enable the augmented reality system;

c) in the initial phase, observe the visual track of the device and the position of the virtual object on the device display screen.

* + - 1. Local relocalization

The local relocalization function of mobile device is tested according to the following method:

a) adjust the illumination color of the test environment to white light with the illumination of 400lx;

b) allow the algorithm to fully construct the scene map information by scanning the scene;

c) quickly shake the device or occlude the camera for a long time, so that the algorithm enters the tracking failure state;

d) re-scan the scene for relocalization. The successful localization (the virtual object is restored to its original position) means that the augmented reality system has the function of local relocalization; otherwise, it does not have the function of local relocalization.

* + - 1. Cloud relocalization

The cloud relocalization function of mobile device is tested according to the following method:

a) adjust the illumination color of the test environment to white light with the illumination of 400lx;

b) allow the algorithm to fully construct the scene map information and place it at the cloud by scanning the scene in advance;

c) quickly shake the device or block the camera for a long time, so that the algorithm enters the tracking failure state;

d) ensure the smooth network of mobile device, and re-scan the scene for relocalization. The successful localization (the virtual object is restored to its original position) means that the augmented reality system has the function of cloud relocalization; otherwise, it does not have the function of cloud relocalization.

* + 1. Cloud localization based on a 3D feature map

The function of cloud localization based on a 3D feature map for mobile device is tested according to the following method:

a) adjust the illumination color of the test scene to white light with the illumination of 100lx;

b) scan the scene in advance to construct a 3D feature map of the scene , and deploy it at the cloud;

c) ensure the smooth network of mobile device, and send images or visual features to the cloud for localization. The successful localization (i.e., pre-placed virtual objects appear in the correct position of the scene) means that the augmented reality system has the function of cloud localization based on a 3D feature map; otherwise, it does not.

* + 1. Target recognition and tracking
       1. Recognition and tracking of target image

The recognition and tracking function of target image is tested according to the following method:

a) adjust the illumination color of the test environment to white light with the illumination of 400lx;

b) scan and shoot the target image with mobile device in the test environment, and judge whether the augmented reality system is able to successfully recognize and track target image by observing the presence of virtual AR content and its position on the device display screen.

* + - 1. Recognition and tracking of target object

The recognition and tracking function of target object is tested by scanning and shooting target object with mobile device in the test environment, and observing the correct presence of virtual AR content and its position on the device display screen, thus judging whether the augmented reality system is able to successfully recognize and track target object.

* + - 1. Cloud target recognition

The recognition function of cloud target is tested by ensuring the smooth network of mobile device, scanning and shooting the target in the test environment, uploading the scene images to the cloud platform through the mobile device, and observing the correct presence of virtual AR content and its position on the device display screen, thus judging whether the augmented reality system is able to successfully recognize target image and target object at the cloud.

* + 1. 3D reconstruction
       1. Plane detection

The plane detection function is tested according to the following method:

a) take a mobile device with the augmented reality system running on it while walking around all the walkable paths in the test scene until walking back to the start point;

b) perform plane detection in the test environment on the mobile device to confirm if this function is available, and check whether the augmented reality system can detect multiple planes and these planes should be incrementally expanded in real-time.

* + - 1. Dense point cloud reconstruction

The function is tested according to the following method:

a) take a mobile device with the augmented reality system running on it while walking around all the walkable paths in the test scene until walking back to the start point;

b) perform dense point cloud reconstruction on the mobile device in the test environment, and confirm if this function is available. The reconstructed dense point cloud should be incrementally expanded in real-time.

* + - 1. Dense mesh reconstruction

The function is tested according to the following method:

a) take a mobile device with the augmented reality system running on it while walking around all the walkable paths in the test scene until walking back to the start point;

b) perform dense mesh reconstruction on the mobile device in the test environment, and confirm if this function is available. The reconstructed dense mesh should be incrementally expanded in real-time.

* + 1. Cloud 3D feature map reconstruction

The function of cloud 3D feature map reconstruction is tested according to the following method:

1. capture images of a test scene with a mobile device
2. send the images or the visual features extracted from the images to the cloud and run the program of 3D feature map reconstruction
3. if a 3D feature map is obtained after the implementation of the program, the function is available, otherwise not.
   * 1. Illumination estimation

The illumination estimation function is tested according to the following method:

a) switch red, green, blue and white light sources in the test environment, and adjust different illumination;

b) observe whether the color of the virtual object is consistent with the test environment illumination, and whether it will change with the illumination change, and judge whether the augmented reality system has the illumination estimation module through these observations.

* + 1. Face alignment

The face alignment function shall be tested according to the following methods:

a) record face videos in different scenes;

b) the successful prediction of a series of face landmarks means that the augmented reality system has the function of face alignment.

* + 1. Gesture recognition

The gesture recognition function shall be tested according to the following methods:

a) record hand videos in different scenes;

b) if the mobile device can predict the hand gesture correctly in the video, the augmented reality system has the function of gesture recognition.

* + 1. Rendering engine

The rendering engine is tested by moving mobile device in different ways, and observing the rendering effect displayed on the screen of mobile device:

a) render dynamic/static virtual objects in real time;

b) virtual objects and the physical world are closely aligned;

c) correctly respond to lighting changes in the physical world.

* + 1. Asynchronous time warp

The function of asynchronous time warp is tested by wearing the above-mentioned optical see-through device in a laboratory with arbitrary head motion, launching the augmented reality system, loading the virtual scenes with different complexities, recording the 6DoF predicted trajectory and comparing it with the SLAM trajectory, and observing the rendering effect displayed on the mobile device screen, in order to judge whether the augmented reality system has the following functions:

a) whether the frame rate is stable and the rendered image is smooth without delay;

b) whether the rendered image jitter, or is not updated during the movement.

* 1. Performance test method
     1. Real-time tracking and localization
        1. 6DoF real-time tracking

The 6DoF real-time tracking performance is tested according to the following method:

a) construct a benchmark dataset for AR evaluation, wherein, the dataset shall cover data for different scenes and motions, including basic SLAM operation data such as image data, IMU sensor data and corresponding calibration parameters; ground-truth data (which can be obtained by motion capture system) shall be provided; a standard public dataset can also be used as a benchmark dataset;

b) build a mobile tool supporting off-line running algorithm, which can read the benchmark dataset and run the algorithm normally;

c) run the algorithm based on the benchmark data, and record the 6DoF poses of all image frames and single frame processing time;

d) evaluate indicators such as APE, ARE, RPE, and RRE of the algorithm with the aid of precision evaluation tools to measure the algorithm precision;

e) evaluate the accuracy and frame rate of the algorithm.

* + - 1. Initialization

The initialization performance of mobile device is tested according to the following method:

a) construct a benchmark dataset for AR precision evaluation, wherein, the dataset shall cover data for different scenes and motions, including basic SLAM operation data such as image data, IMU sensor data and corresponding calibration parameters; ground-truth data (which can be obtained by motion capture system) shall be provided; a standard public dataset can also be used as a benchmark dataset;

b) build a mobile tool supporting off-line running algorithm, which can read the benchmark dataset and run the algorithm normally;

c) run the algorithm based on the benchmark data, and record the 6DoF poses of all image frames and single frame processing time;

d) evaluate indicators such as initialization time and initialization quality of the algorithm with the precision evaluation tools to measure the initialization performance.

* + - 1. Local relocalization

The local relocalization performance is tested according to the following method:

a) acquire the following data with the mobile data acquisition tool:

1) allow the algorithm to fully construct the scene map information by scanning the scene;

2) quickly shake the device or occlude the camera for a long time, so that the algorithm enters the tracking failure state;

3) re-scan the scene for relocalization.

b) evaluate the local relocalization success rate, relocalization time and relocalization precision based on above data.

* + - 1. Cloud relocalization

The cloud relocalization performance is tested according to the following method:

a) acquire the following data with the mobile data acquisition tool:

1) allow the algorithm to fully construct the scene map information by scanning the scene;

2) quickly shake the device or occlude the camera for a long time, so that the algorithm enters the tracking failure state;

3) re-scan the scene and transmit it to the cloud for relocalization.

b) evaluate the cloud relocalization success rate, relocalization time and relocalization precision based on above data.

* + 1. Cloud localization based on a 3D feature map

The performance of cloud localization based on a 3D feature map is tested according to the following method:

a) construct 3D feature maps for a set of test scenes..

b) capture test images with a mobile device for each test scene, and acquire the ground truth positions of the mobile device for each test image. The ground-truth can be obtained by a motion capture system, e.g. Vicon, or other surveying devices like total stations. Make a statistic of the median depth of each test image using the 3D map points that should be observed by this image.

c) locate all the test images;

c) count the success rate of localization, and the localization time.

* + 1. Target recognition and tracking
       1. Recognition and tracking of target image

The recognition and tracking performance of target image is tested according to the following method:

a) record the execution time of recognition and tracking algorithm for each frame in a log, and count the average recognition and tracking time of target image per frame within the execution time of no less than 5min, including:

1) respectively test the recognition time of single target image and the average recognition time of five target images;

2) respectively test the tracking time of single target image and the average tracking time of five target images.

b) test the recognition success rate of target image, including the success rate under different angles, different distances and different illumination intensities.

c) test the precision of recognition and tracking algorithm.

* + - 1. Recognition and tracking of target object

The recognition and tracking performance of target object is tested according to the following method:

a) record the execution time of identification and tracking algorithm for each frame in a log, and count the average recognition and tracking time of target object per frame within the execution time of no less than 5 min, including:

1) respectively test the recognition time of single target object and the average recognition time of five target objects;

2) respectively test the tracking time of single target object and the average tracking time of five target objects.

b) test the recognition success rate of target object, including the success rate under different angles, different distances and different illumination intensities.

c) test the precision of recognition and tracking algorithm.

* + - 1. Cloud target recognition

The recognition performance is tested as follows: upload the scene images to the cloud platform through the mobile device under the condition of smooth network of tested mobile device, and then test the recognition speed and success rate of target images and target objects by cloud, including:

a) test the average network time, algorithm time ,success rate and precision of target image recognition;

b) test the average network time, algorithm time , success rate and precision of target object recognition.

* + 1. 3D reconstruction
       1. Plane detection

The performance of plane detection is tested according to the following method:

a) take an accurate 3D model of a scene scanned by a 3D laser scanner as ground truth;

b) perform plane detection on a mobile device for the scene, and evaluate the position error between each detected plane and the ground truth;

c) record the single frame execution time of single plane detection and multi-plane detection algorithms in the log, and count the average detection time of each frame with the total execution time no less than 5 min.

* + - 1. Dense point cloud reconstruction

The performance of dense point cloud reconstruction is tested according to the following method:

a) take an accurate 3D model of a scene scanned by a 3D laser scanner as ground truth;

b) perform dense point cloud reconstruction on a mobile device for the scene, and evaluate the position error between the reconstructed dense point cloud and the ground truth;

c) record the single frame execution time of dense point cloud reconstruction algorithm in the log, and count the average reconstruction time of each frame with the execution time no less than 5 min.

* + - 1. Dense mesh reconstruction

The performance of dense mesh reconstruction is tested according to the following method:

a) take an accurate 3D model of a scene scanned by a 3D scanner as ground truth;

b) perform dense mesh reconstruction on a mobile device for the scene, and evaluate the geometric error between the reconstructed dense mesh and the ground truth;

c) record the single frame execution time of dense mesh reconstruction algorithm in the log, and count the average reconstruction time of each frame with the total execution time no less than 5 min.

* + 1. Cloud 3D feature map reconstruction

the performance of cloud 3D feature map reconstruction is tested according to the following method:

1. Place several tags on a test scene, and acquire the ground truth positions of the tags by using a surveying device like a total station.
2. Capture the images of the scene and the tags with a mobile device
3. Construct the 3D feature map of the scene
4. evaluate the average error of the reconstructed map using the positions of the tags, count the time cost of the reconstruction process.
   * 1. Illumination estimation

The illumination estimation performance is tested according to the following method, and the estimated value is normalized to 0~1 by algorithm and the true value is normalized to 0~1 by system camera:

a) turn on the white light only, and measure the error between the estimated value and the true value of illumination color from camera image under 4 conditions where the illumination of the white light is 100lx , 200lx, 300lx and 400lx respectively;

b) turn on the red light only, and measure the error between the estimated value and the true value of illumination color from camera image under 4 conditions where the illumination of the red light is 100lx , 200lx, 300lx and 400lx respectively;

c) turn on the green light only, and measure the error between the estimated value and the true value of illumination color from camera image under 4 conditions where the illumination of the green light is 100lx , 200lx, 300lx and 400lx respectively;

d) turn on the blue light only, and measure the error between the estimated value and the true value of illumination color from camera image under 4 conditions where the illumination of the blue light is 100lx , 200lx, 300lx and 400lx respectively.

* + 1. Face Alignment

The face alignment performance is tested as follows:

a) construct a benchmark dataset covering different scenes;

b) compute the accuracy of key points on above data.

* + 1. Gesture recognition

The gesture recognition performance is tested as follows:

a) construct a benchmark dataset covering different scenes;

b) compute the recognition accuracy on above data.

* + 1. Rendering engine

The rendering engine performance is tested as follows: provide static model and skeleton animation model in standard format, and use the estimated illumination to render, test the loading of static model and dynamic model, animation update and rendering results, and test the rendering resolution and running frame rate of rendering engine by tools of mobile device, such as Arm or Qualcomm frame GPU profiler tool.

* + 1. Asynchronous time warp

The asynchronous time warp performance is tested as follows:

a) provide scenes with static model in standard format and skeleton animation model with different complexity, and use the estimated illumination value and color value as ambient light illumination and color to render, test whether the rendering frame rate can be stabilized at a fixed frame rate of >60Hz, and test the running frame rate of rendering by common frame cutting tools of mobile device, such as Arm or Qualcomm frame cutting tool;

b) record the 6DoF predicted trajectory and the SLAM trajectory, then compare the two numerically.

* + 1. Running
       1. CPU and memory occupancy test

After the augmented reality on mobile device platform is enabled, the built-in commands or system tools of operating system can be used to check the CPU and memory occupancy of the system process. The command or tool depends on the operating system on mobile device platform. For example, after the augmented reality system on the Android mobile device platform is enabled, the CPU occupancy rate of augmented reality system on mobile device can be checked by running the command line "adb shell top". The real-time memory occupancy of the process corresponding to the package name can be get by running the command line "adb shell dumpsys meminfo package name", and the CPU and memory occupancy of augmented reality system on mobile device process can be get by Profile tool of Android Studio; after the augmented reality system on the iOS mobile device platform is enabled, the CPU and memory occupancy of mobile device based augmented reality can be analyzed by using the Instruments tool of XCode.

* + - 1. Algorithm running efficiency test

After the augmented reality system on mobile device platform is enabled, the algorithm running time of each frame of data is recorded in the log, and the maximum, minimum and average time for algorithm per frame can be counted based on the record within the algorithm execution time of no less than 5 min.

(Informative)

**Sensor of mobile device related to augmented reality system**

## Exposure parameters

The camera shall have exposure parameter setting function.

The exposure time shall not exceed 20ms.

## Inertial measurement unit

IMU shall be able to measure the triaxial pose angle (or angular velocity) and acceleration of an object.

The IMU information frequency shall not be lower than 200Hz; clock shall be aligned with camera, and the time deviation shall not exceed 5ms.

## Focus mode

**Fixed focus mode**

The focus of a camera on the mobile device shall be at infinity by default where the automatic focusing is not enabled.

In single-shot focus mode, the focus will be determined by manual selection or at infinity by default and will not be changed upon enabling.

**Zoom mode**

In zoom mode, the camera shall be able to focus continuously for many times to improve the definition.

**Mode selection**

The camera in fixed focus mode shall be selected for augmented reality. If continuous focusing is adopted, the focus change shall be smooth.

The augmented reality system shall be equipped with a camera loopmotor.

## Time alignment

The mobile device shall be in alignment with the time on camera and sensors such as IMU, and the time stamp difference shall not exceed 5ms.

## Quality of camera image

The resolution of each frame of video captured by mobile device camera shall not be less than 720p.

## Depth camera

The mobile device depth camera shall be used to obtain spatial depth information in pictures, and is usually classified into the following two types with different implementation modes:

a) structured light camera

b) TOF camera

The frame rate of ordinary camera shall be an integral multiple of that of depth camera which shall not be less than 5fps.

The clocks of depth camera and ordinary camera shall be accurately calibrated and aligned, with a time deviation of no greater than 15ms.

The depth camera and ordinary camera shall be provided with accurate external parameter (relative pose) calibration. Under the image resolution of 640\*480, the deviation of pixel position respectively corresponding to the color and depth shall not exceed 2 pixels, and under the image resolution of 320\*240, the deviation shall not exceed 1 pixel.

1. The Institute of Electrical and Electronics Engineers, Inc.

   3 Park Avenue, New York, NY 10016-5997, USA

   Copyright © 2022 by The Institute of Electrical and Electronics Engineers, Inc.

   All rights reserved. Published <Date Published>. Printed in the United States of America.

   IEEE is a registered trademark in the U.S. Patent & Trademark Office, owned by The Institute of Electrical and Electronics   
   Engineers, Incorporated.

   PDF: ISBN 978-0-XXXX-XXXX-XSTDXXXXX

   Print: ISBN978-0-XXXX-XXXX-XSTDPDXXXXX

   IEEE prohibits discrimination, harassment, and bullying.

   For more information, visit [https://www.ieee.org/about/corporate/governance/p9-26.html](http://www.ieee.org/web/aboutus/whatis/policies/p9-26.html).

   No part of this publication may be reproduced in any form, in an electronic retrieval system or otherwise, without the prior written permission of the publisher. [↑](#footnote-ref-2)
2. The use of the word *must* is deprecated and cannot be used when stating mandatory requirements, *must* is used only to describe unavoidable situations. [↑](#footnote-ref-3)
3. The use of *will* is deprecated and cannot be used when stating mandatory requirements, *will* is only used in statements of fact. [↑](#footnote-ref-4)
4. IEEE Standards Dictionary Online is available at: <http://dictionary.ieee.org>. An IEEE Account is required for access to the dictionary, and one can be created at no charge on the dictionary sign-in page. [↑](#footnote-ref-5)