P2048.101™/D1.0 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 23 Sep 2020

IEEE SA Standards Board

Abstract: <Select this text and type or paste Abstract—contents of the Scope may be used>

Keywords: <Select this text and type or paste keywords>

The Institute of Electrical and Electronics Engineers, Inc.
3 Park Avenue, New York, NY 10016-5997, USA

PDF: ISBN 978-0-XXXX-XXXX-X STDXXXXX
Print: ISBN 978-0-XXXX-XXXX-X STDPDXXXXX

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. An IEEE Account is needed to access the application.

Comments on standards should be submitted using the Contact Us 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 or contact IEEE. 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. 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. Users are encouraged to periodically check for errata.

Patents

IEEE Standards are developed in compliance with the IEEE SA Patent Policy.

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

1.1 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, safety 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.

1.2 Purpose

This document will not include a purpose clause.

1.3 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*).[1,2]

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*).

---

[1] The use of the word *must* is deprecated and cannot be used when stating mandatory requirements, *must* is used only to describe unavoidable situations.
[2] The use of *will* is deprecated and cannot be used when stating mandatory requirements, *will* is only used in statements of fact.

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

2. 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

3. Definitions, acronyms, and abbreviations

3.1 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. [3]

3.1.1 mobile device

component with camera, display screen and network functions, of which the location and motion information are detectable

Note: Smart devices such as mobile phones or tablet computers of GPS, accelerometer, magnetometer and gyroscope.

3.1.2 augmented reality

environment that augments the real world perceived by users using additional information generated by modern high-tech means with computer as the core,

wherein, the generated information is superimposed on the real scene by fusion of physiological feelings such as vision, auditory sense, gustatory sense, olfactory sense and tactile sense

3.1.3 actual environment

real world containing multiple physical objects, where human beings live in and can perceive without the aid of equipment

3.1.4 virtual object

computer-generated object with geometric shape, specific format or specific behavior

Note: Its prototype may be either a real object or a completely fictional object.

[3] *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.

3.1.5 2 dimensional marker

pre-set 2 dimensional image with special code or pattern, which may be detected, tracked or recognized to estimate its position and orientation relative to the mobile device in a scene, image or video

3.1.6 3 dimensional marker

pre-set 3 dimensional image with special shape or texture, which may be detected, tracked or recognized to estimate its position and orientation relative to the mobile device in a scene, image or video

3.1.7 anchor point

benchmark used in the process of putting virtual object in the environment

3.1.8 cloud end

system with computing and storage functions for remote access of mobile devices with access permission

3.1.9 localization

action for getting the spatial position of the specified object

3.1.10 relocalization

process of obtaining correct pose information and resetting the current pose of the mobile device by triggering the relocalization mechanism thereof according to the previously tracked historical position and orientation when the mobile device loses its location due to unexpected position mutation

3.1.11 six degree of freedom tracking

action for calculating the six degree of freedom pose of mobile device relative to actual scene in real time

3.1.12 simultaneous localization and mapping

action for device (mobile terminal or robot, etc.) locating its pose in an unknown environment by observing the external environment, and then incrementally mapping the unknown environment through its position

3.1.13 illumination estimation

process of analyzing and calculating source illumination distribution information of physical scene from sensor or camera view

3.1.14 scale estimation

process of obtaining the length information of physical world in the system

3.1.15 absolute position error

average deviation between the true value and the estimated value of current position measured by the augmented reality system when the position of mobile device changes

3.1.16 absolute rotation error

average deviation between the true value and the estimated value of current rotation angle measured by the augmented reality system when the rotation angle of mobile device changes

3.1.17 relative position error

average deviation between the true variation and the estimated variation of the current position from the previous position measured by the augmented reality system when the position of mobile device changes

3.1.18 relative rotation error

average deviation between the true variation and the estimated variation of the current rotation angle from the previous rotation angle measured by the augmented reality system when the rotation angle of mobile device changes

3.2 Acronyms and abbreviations

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

4. System structure

The processing unit of augmented reality system on mobile device consists of modules such as real-time tracking and localization, scale estimation, marker recognition and tracking, 3D reconstruction, virtual occlusion module, marker recognition and tracking, illumination estimation and rendering engine. The real-time tracking and localization module calculates 6DoF pose using the input data of mobile device; the scale estimation module calculates scale information using input data; the 3D reconstruction module reconstructs the 3D model of the scene using 6DoF pose, scale information and deep flow data; the virtual occlusion module realizes occlusion effect between virtual object and real environment using video flow and deep flow; the marker recognition and tracking module realizes marker recognition and tracking based on video flow; the illumination estimation module solves illumination information using video flow; the mobile device, if equipped with cloud computing service, also supports transmitting data to the cloud server via a network to realize cloud relocalization, cloud marker recognition and cloud 3D reconstruction with higher efficiency, and transmitting the calculated 6DoF pose, marker information and 3D model back to the processing unit of mobile device via the network; the information such as 6DoF pose, 3D model, virtual occlusion, marker recognition and tracking and illumination are summarized in the rendering engine to achieve a realistic augmented reality effect, which is presented on the display screen of mobile device. See Figure 1 for the structural diagram of augmented reality system on mobile device. The requirements for sensor of mobile device related to augmented reality system shall meet the requirements of Annex A.

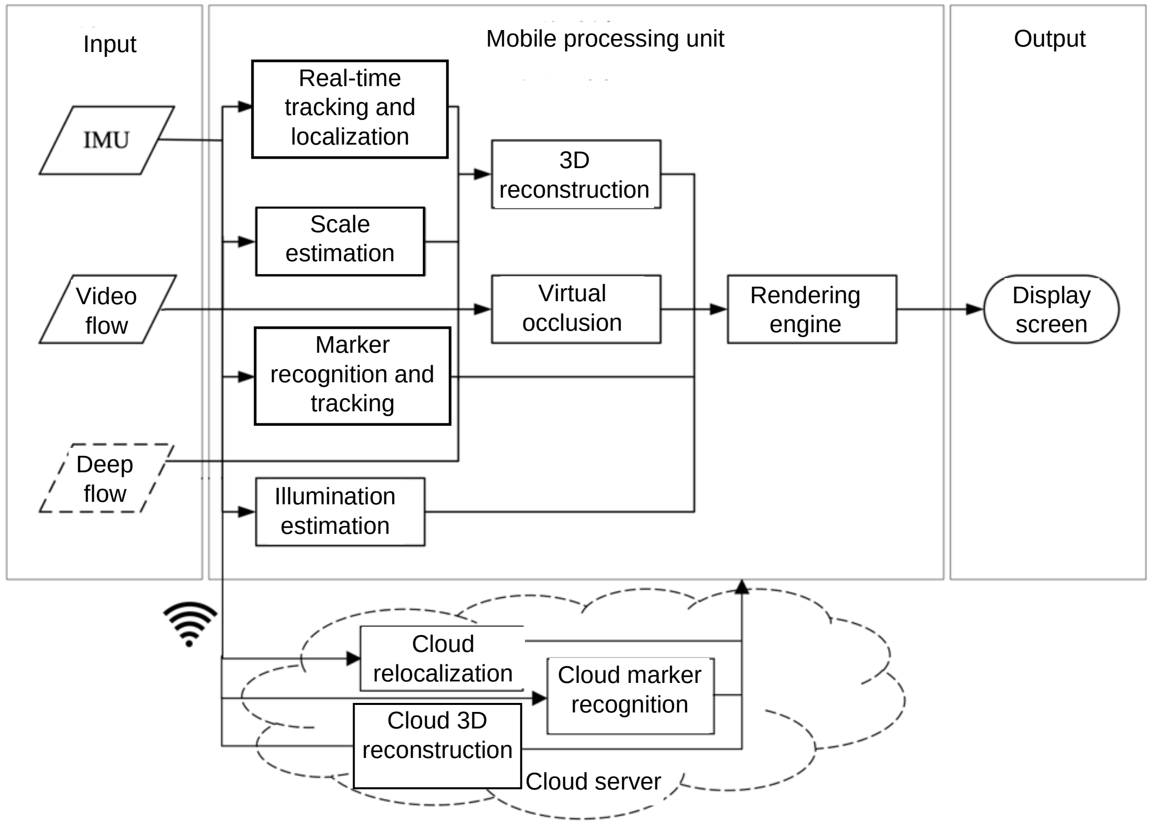

Figure 1Structural diagram of augmented reality system on mobile device

5. Functional requirements

5.1 Real-time tracking and localization

5.1.1 Six degree of freedom real-time tracking

The augmented reality system on mobile device shall estimate the pose of six degree of freedom in real time.

5.1.2 Initialization

The augmented reality system on mobile device shall automatically initialize the scale information of the scene.

5.1.3 Local relocalization

When the mobile device loses its location due to unexpected position mutation, the augmented reality system on mobile device shall relocalize the device and reset its current pose locally.

5.1.4 Cloud relocalization

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

5.2 Scale estimation

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

a) obtain the scale information of the physical world;

b) register the virtual scene and the physical world in the same scale coordinate system;

c) achieve the 1:1 virtual-real fusion between virtual object and physical world.

5.3 Cloud localization based on a high-precision map

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

a) conduct high-precision remapping for a defined scene, wherein, the high-precision map has a real scale, an image set with a known global pose and a common view among the images, as well as sparse 3D points in the scene and observation in the image;

b) localize the image uploaded to the cloud end, and calculate the 6DoF pose in the high-precision map coordinate system;

c) extract image features at the mobile end, and upload them to the cloud end for localization.

5.4 Marker recognition and tracking

5.4.1 Recognition and tracking of 2 dimensional marker

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

a) recognize the 2 dimensional marker in the defined scene;

b) perform 6DoF pose tracking for the 2 dimensional marker in the defined scene;

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

5.4.2 Recognition and tracking of 3 dimensional marker

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

a) preprocess the texture or structure information of the 3 dimensional marker in a real scene;

b) compare the real-time information from the camera with the processed information to match the 2 dimensional information with 3 dimensional information;

c) understand the position and orientation of 3D object in a scene.

5.4.3 Cloud marker recognition

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

a) recognize the image frame with 2 dimensional marker uploaded to the cloud end by a user using computing resources of the cloud server, and return the position and orientation information of the 2 dimensional marker of the frame to the user end;

b) recognize the image frame with 3 dimensional object uploaded to the cloud end by a user using the computing resources of the cloud server, and return the position and orientation information of the 3 dimensional object of the frame to the user end.

5.5 3D reconstruction

5.5.1 Plane detection and anchor point localization

5.5.1.1 Single plane reconstruction and anchor point localization

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

a) support the detection of horizontal and vertical planes;

b) understand the main horizontal plane of the real scene, and localize the anchor point on it or render it;

c) recognize the 3 dimensional points on the plane, determine the plane position and extend it;

d) detect the single plane in the scene;

e) reconstruct the single plane and localize the anchor point.

5.5.1.2 Multiple-plane reconstruction and anchor point localization

For a scene containing multiple planes, the augmented reality system on mobile device shall meet the following requirements:

a) understand multiple planes (including horizontal plane or vertical plane) of the real scene, and localize the anchor points on them or render them;

b) recognize the 3 dimensional points on the plane, determine the plane position and extend it;

c) detect multiple planes in the scene;

d) reconstruct multiple planes and localize the anchor points.

5.5.1.3 Automatic adjustment of plane position

In case of a closed loop or other optimization behaviors triggering the historical frame pose, the augmented reality system on mobile device shall meet the following requirements:

a) automatically adjust the plane position according to pose optimization;

b) automatically adjust the position of anchor point on the plane.

5.5.2 Reconstruction of dense point cloud and anchor point localization

When placing virtual objects on a complex non-planar scene, the augmented reality system on mobile device shall meet the following requirements:

a) support incremental real-time extension;

b) reconstruct dense 3D point cloud information of the scene, wherein each 3D point in the point cloud contains information such as position, normal direction and color;

c) solve the anchor point position and normal direction of dense point cloud.

5.5.3 Reconstruction of dense mesh and anchor point localization

When constructing complex virtual-real fusion effects such as occlusion, shadow and collision on a real scene, the augmented reality system on mobile device shall meet the following requirements:

a) reconstruct the dense 3 dimensional mesh information of the scene;

b) solve the anchor point position and normal direction of dense mesh.

5.5.4 Cloud 3D reconstruction

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

a) support incremental real-time extension;

b) carry out 3D reconstruction for the key frame image uploaded by the user end using the computing resources of the cloud server;

c) align the reconstruction result with the unified coordinate system;

d) read or display the reconstruction information.

5.6 Illumination estimation

The augmented reality system on mobile device shall estimate the global illumination environment information of the physical scene used for virtual object and virtual scene rendering from the sensor or camera view.

5.7 Rendering engine

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

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

b) closely fit the virtual object and the real background;

c） correctly respond to the illumination changes in the real world;

d) support rendering of static model and dynamic model.

5.8 Virtual occlusion

The augmented reality system on mobile device shall realize the occlusion between virtual object and real environment, including the occlusion between virtual object and static background environment and that between virtual object and dynamic foreground.

5.9 Asynchronous time distortion

For penetrating rendering device, before sending the rendered image to the display screen, the augmented reality system on mobile device shall predict the pose upon display, so as to deform the rendered image and correct the head movement.

6. 6 Performance requirements

6.1 Real-time tracking and localization

6.1.1 Six degree of freedom real-time tracking

Under the movement types and test scenes covered by the benchmark test dataset described in this document, the six degree of freedom real-time tracking of augmented reality system on mobile device shall meet the following requirements:

a) the frame rate is not less than 24fps;

b) the tracking accuracy is not less than 95%, that is, the number of frames that meet the following requirements in the test dataset is not less than 95% of the total number of frames:

1) APE is less than 10cm or less than 5% the maximum moving distance of the camera (whichever is greater);

2) ARE is less than 6°;

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

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

6.1.2 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 initialization quality error shall be less than 10, the defined initialization duration is tinit, the initialization scale estimation error is €scale, and the initialization quality €init=tinit*(€scale+0.01)0.5.

6.1.3 Local relocalization

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

a) the success rate of relocalization is not less than 90%, that is, the number of relocalizations that meet the requirements of the following items 1) and 2) in the test dataset is not less than 90% the number of relocalizations that meet the requirement of item 1):

1) the user puts the mobile device back to the position before 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 5cm or less than 5% the median value of the depth of the picture taken by the camera (whichever is greater);

b) the relocalization time does not exceed 2s if succeed;

c) the tracking robustness error shall be less than 10, the defined tracking loss time percentage is αlost, the relocalization error is €RL and the position error is €APE, and the robustness error is €R=(αlost+0.05)(€RL+0.1*€APE).

6.1.4 Cloud relocalization

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

a) the success rate of relocalization is not 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) the relocalization time does not exceed 2s if succeed.

6.2 Scale estimation

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

a) the deviation between the estimated value of actual environment and its true value does not exceed 15%;

b) the time does not exceed 2s.

6.3 Cloud localization based on a high-precision map

The cloud relocalization based on a high-precision map for augmented reality system on mobile device shall meet the following requirements:

a) the success rate of cloud relocalization based on a high-precision map is not 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) the relocalization time does not exceed 2s if succeed.

6.4 Marker recognition and tracking

6.4.1 Recognition and tracking of 2 dimensional marker

The recognition and tracking of 2 dimensional marker for augmented reality system on mobile device shall meet the following requirements:

a) the recognition and tracking of not less than 4 plane marks shall be supported at the same time;

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

c) the recognition accuracy is not less than 90%, that is, the average deviation between the projection contour of 2 dimensional marker under the estimated pose in the test dataset and the true value is not greater than 5 pixels or 1% the larger value of the width and height of the true value contour (whichever is greater), and the number of frames for 2 dimensional marker is not less than 90% the total number of frames;

d) the tracking frequency of single-plane marker is not less than 24fps;

e) the tracking frequency of multiple-plane marker is not less than 20fps.

6.4.2 Recognition and tracking of 3 dimensional marker

The recognition and tracking of 3 dimensional marker for augmented reality system on mobile device shall meet the following requirements:

a) the frequency is not 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 not greater than 3cm/m, or not greater than 3% the maximum moving distance of object (whichever is greater);

d) the recognition accuracy is not less than 90%, that is, the average deviation between the projection contour of 3 dimensional marker under the estimated pose in the test dataset and the true value is not greater than 5 pixels or 1% the larger value of the width and height of the true value contour (whichever is greater), and the number of frames for 3 dimensional marker is not less than 90% the total number of frames.

6.4.3 Cloud marker recognition

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

6.5 3D reconstruction

6.5.1 Plane detection

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

a) the frame rate for processing is consistent with that for 6DoF tracking;

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

6.5.2 Reconstruction of dense point cloud

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

a) the frame rate for processing is consistent with that for 6DoF tracking;

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

6.5.3 Reconstruction of dense mesh

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

a) the frame rate for processing is consistent with that for 6DoF tracking;

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

6.6 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 color shall not exceed 0.3 (the illumination color value is uniformly normalized to the range of 0~1 in the brightness color space).

6.7 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.

6.8 Virtual occlusion

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

a) the frame rate of image depth map acquisition and occlusion processing is the same as that of 6DoF tracking;

b) the deviation of the occlusion edge does not exceed 5 pixels or 1% the larger value of the width and height of the picture (whichever is greater);

c) the error rate of the occlusion relationship in the picture does not exceed 10%.

6.9 Asynchronous time distortion

The asynchronous time distortion of augmented reality system on mobile device shall meet the following requirements:

a) the frame screen refresh rate is constant and not less than 60Hz;

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

6.10 Running

6.10.1 CPU and memory occupancy

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

a) the CPU occupancy rate is not greater than 50%;

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

6.10.2 Algorithm running efficiency

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

7. Test method

7.1 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 the same 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 wallpaper on the wall surface and patterned tablecloths on the square table;

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

d) arrange five 2 dimensional markers and five 3 dimensional markers in the scene;

e) refer to the public datasets such as ZJU - SenseTime VISLAM Benchmark and NEAR-VI-Dataset for benchmark dataset.

7.2 Function test method

7.2.1 Real-time tracking and localization

7.2.1.1 Six degree of freedom real-time tracking

The six degree of freedom 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 100lx;

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.

7.2.1.2 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 100lx;

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.

7.2.1.3 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 100lx;

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.

7.2.1.4 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 100lx;

b) allow the algorithm to fully construct the scene map information and place it at the cloud end 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.

7.2.2 Scale estimation

The scale estimation function of mobile device is tested by observing whether the size of virtual object is close to the real size after successful scale estimation, thus judging whether the augmented reality system has the function of scale estimation.

7.2.3 Cloud localization based on a high-precision map

The function of localization based on a high-precision map for 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 100lx;

b) scan the scene in advance to construct a high-precision map of the scene which meets those specified in 6.3.a, and place it at the cloud end;

c) ensure the smooth network of mobile device, and send images or features to the cloud end for localization. The successful localization (virtual object appears in the correct position of the scene) means that the augmented reality system has the function of cloud localization based on a high-precision map; otherwise, it does not have the function of cloud localization based on a high-precision map.

7.2.4 Marker recognition and tracking

7.2.4.1 Recognition and tracking of 2 dimensional marker

The recognition and tracking function of 2 dimensional marker is tested according to the following method:

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

b) scan and shoot the 2 dimensional marker with mobile device in the test environment, and judge whether the augmented reality system is able to successfully recognize and track 2 dimensional marker by observing the presence of virtual AR content and its position on the device display screen.

7.2.4.2 Recognition and tracking of 3 dimensional marker

The recognition and tracking function of 3 dimensional marker is tested by scanning and shooting 3 dimensional marker 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 3 dimensional marker.

7.2.4.3 Cloud marker recognition

The recognition function of cloud marker is tested by ensuring the smooth network of mobile device, scanning and shooting the marker in the test environment, uploading the scene video stream 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 2 dimensional marker and 3 dimensional marker at the cloud end.

7.2.5 3D reconstruction

7.2.5.1 Plane detection

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

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

b) carry out plane detection on the mobile platform for the test environment, thus testing whether the augmented reality system has the function of plane detection, whether the plane detection has increasing real-time scalability and whether multiple planes can be detected.

7.2.5.2 Reconstruction of dense point cloud

The reconstruction function of dense point cloud is tested by reconstructing dense point cloud on mobile platform in the test environment, thus testing whether the augmented reality system has the function of dense point cloud reconstruction and whether the reconstruction of dense point cloud has increasing real-time scalability.

7.2.5.3 Reconstruction of dense mesh

The reconstruction function of dense mesh is tested by reconstructing the dense mesh on mobile platform in the test environment, thus testing whether the augmented reality system has the function of dense mesh reconstruction and whether the reconstruction of dense mesh has increasing real-time scalability.

7.2.6 Illumination estimation

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

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

b) observe whether the color of virtual object is consistent with the illumination of test environment, and whether the color will change with the illumination change, thus judging whether the augmented reality system has the function of illumination estimation.

7.2.7 Rendering engine

The rendering engine function is tested by wearing the above-mentioned mobile device in a laboratory to move in different ways and enable the augmented reality system, and observing the rendering effect displayed on the screen of mobile device, thus judging whether the augmented reality system has the following functions:

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

b) closely fit the virtual object and the real background;

c) correctly respond to the illumination changes in the real world;

d) support the rendering of common static and dynamic AR models (such as skeleton animation).

7.2.8 Virtual occlusion

The virtual occlusion function is tested according to the following method:

a) look for a scene with complex depth of field hierarchy;

b) shoot the scene with a mobile device, and enable the augmented reality system, then place dynamic virtual objects in the picture of mobile phone scene to realize the occlusion effect between virtual objects and real scenes (including static background and dynamic foreground) through augmented reality system, thus judging whether the augmented reality system has the function of virtual occlusion.

7.2.9 Asynchronous time distortion

The function of asynchronous time distortion is tested by wearing the above-mentioned mobile device in a laboratory, enabling the augmented reality system, opening the scene experience with different complexities, recording the predicted track and comparing it with the six degree of freedom tracking track, and observing the rendering effect displayed on the mobile device screen, thus judging whether the augmented reality system has the following functions:

a) whether the frame rate is stable and the picture is smooth;

b) whether the picture flutters, and the occasional picture is not updated during the movement.

7.3 Performance detection method

7.3.1 Real-time tracking and localization

7.3.1.1 Six degree of freedom real-time tracking

The 6DoF real-time tracking performance 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; true value 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 data acquisition 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.

7.3.1.2 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; true value 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 data acquisition 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.

7.3.1.3 Local relocalization

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

a) acquire the following data with the mobile data acquisition tool in 8.3.1.1:

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.

7.3.1.4 Cloud relocalization

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

a) acquire the following data with the mobile data acquisition tool in 8.3.1.1:

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 end for relocalization.

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

7.3.2 Scale estimation

The scale estimation performance is tested as follows: generate the required 6DoF parameters and true value data of algorithm by reference to the data acquisition method in 8.3.1.1, and evaluate the 6DoF pose scale error of the algorithm by comparing the algorithm pose parameters with the true value data.

7.3.3 Cloud localization based on a high-precision map

The performance of cloud localization based on a high-precision map is tested according to the following method:

a) construct a benchmark dataset (covering different scenes) for evaluating the cloud localization performance based on a high-precision map, construct a high-precision map meeting those specified in 6.3.a for each scene, and acquire test pictures with a mobile device, the position true values of which can be obtained by measuring devices such as motion capture system or total station;

b) locate all test pictures based on a high-precision map;

c) count the localization success rate, localization precision and localization time based on above data.

7.3.4 Marker recognition and tracking

7.3.4.1 Recognition and tracking of 2 dimensional marker

The recognition and tracking performance of 2 dimensional marker 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 2 dimensional marker per frame within the execution time of not less than 5min, including:

1) respectively test the recognition time of single 2 dimensional marker and the average recognition time of five 2 dimensional markers;

2) respectively test the tracking time of single 2 dimensional marker and the average tracking time of five 2 dimensional markers.

b) test the recognition success rate of 2 dimensional marker, including the success rate under different angles, different distances and different illumination intensities.

7.3.4.2 Recognition and tracking of 3 dimensional marker

The recognition and tracking performance of 3 dimensional marker 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 3 dimensional marker per frame within the execution time of not less than 5min, including:

1) respectively test the recognition time of single 3 dimensional marker and the average recognition time of five 3 dimensional markers;

2) respectively test the tracking time of single 3 dimensional marker and the average tracking time of five 3 dimensional markers.

b) test the recognition success rate of 3 dimensional marker, including the success rate under different angles, different distances and different illumination intensities.

7.3.4.3 Cloud marker recognition

The recognition performance is tested as follows: upload the scene video flow 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 2 dimensional markers and 3 dimensional markers by cloud end, including:

a) test the average network time, algorithm time and success rate of 2 dimensional marker recognition;

b) test the average network time, algorithm time and success rate of 3 dimensional marker recognition.

7.3.5 3D reconstruction

7.3.5.1 Plane detection

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

a) construct multiple plane models as true values using modeling software (such as 3dsmax or Maya) according to the measured size of the scene;

b) evaluate the position precision error between each detected plane and the true value plane;

c) record the execution time of plane detection algorithm of each frame in the log, and count the average time for plane detection per frame within the execution time of not less than 5min, including the time for detecting single plane and average time for detecting five planes.

7.3.5.2 Reconstruction of dense point cloud

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

a) take an accurate 3D model of a scene scanned by a 3D scanner as the true value;

b) reconstruct the dense point cloud on the mobile platform for the scene, and evaluate the position precision error between the reconstructed dense point cloud and the true 3D model;

c) record the execution time of dense point cloud reconstruction algorithm of each frame in a log, and count the average time for reconstruction of dense point cloud per frame within the execution time of not less than 5min.

7.3.5.3 Reconstruction of dense mesh

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

a) take an accurate 3D model of a scene scanned by a 3D scanner as the true value;

b) reconstruct the dense mesh on the mobile platform for the scene, and evaluate the geometric precision error between the reconstructed dense mesh and the true 3D model;

c) record the execution time of dense mesh reconstruction algorithm of each frame in a log, and count the average time for reconstruction of dense mesh per frame within the execution time of not less than 5min.

7.3.6 Illumination estimation

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

a) turn off all lights, and measure the error between the estimated value and the true value of illumination color in the scene respectively;

b) turn on the white light only, and measure the error between the estimated value and the true value of illumination color under three conditions where the illumination of the white light is 20lx, 100lx and 200lx respectively;

c) turn on the red light only, and measure the error between the estimated value and the true value of illumination color under three conditions where the illumination of the red light is 20lx, 100lx and 200lx respectively;

d) turn on the green light only, and measure the error between the estimated value and the true value of illumination color under three conditions where the illumination of the green light is 20lx, 100lx and 200lx respectively;

e) turn on the blue light only, and measure the error between the estimated value and the true value of illumination color under three conditions where the illumination of the blue light is 20lx, 100lx and 200lx respectively.

7.3.7 Rendering engine

The rendering engine performance is tested as follows: provide static model in standard format and skeleton animation model, and use the estimated illumination value and color value as ambient light illumination and color 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 common frame cutting tools of mobile device, such as Arm or Qualcomm frame cutting tool.

7.3.8 Virtual occlusion

The virtual occlusion performance is tested according to the following method:

a) shoot the scene with a mobile device, and enable augmented reality system, then place dynamic virtual objects in the picture of mobile phone scene to realize the occlusion effect between virtual objects and real scenes (including static background and dynamic foreground) through augmented reality system, meanwhile, output the occlusion time of each frame to the log, and judge whether the frame rate can reach the frame rate of six degree of freedom tracking;

b) record five groups of real-time virtual occlusion frames with common frame cutting tools of mobile device, and give the frame information to five participants, each participant evaluates the edge accuracy and error rate of virtual occlusion in five groups of scenes, and calculate the average of these evaluation results to get the edge accuracy and correct rate of virtual occlusion of augmented reality system on mobile platform.

7.3.9 Asynchronous clock distortion

The asynchronous clock distortion performance is tested as follows: 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.

7.3.10 Running

7.3.10.1 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.

7.3.10.2 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 not less than 5min.

Annex A

(Normative)

Sensor of mobile device related to augmented reality system

A.1 Exposure parameters

The camera shall have exposure parameter setting function.

The exposure time should not exceed 20ms.

A.2 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.

A.3 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 should be selected for augmented reality. If continuous focusing is adopted, the focus change shall be smooth.

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

A.4 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.

A.5 Quality of camera image

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

A.6 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 not 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.