BS ISO/IEC 18520:2019

BSI Standards Publication

# Information technology — Computer graphics, image processing and environmental data representation — Benchmarking of vision-based spatial registration and tracking methods for mixed and augmented reality (MAR)

Documents
Document Sharing Hub

Full document details: BS ISO/IEC 18520:2019

# National foreword

This British Standard is the UK implementation of ISO/IEC 18520:2019.

The UK participation in its preparation was entrusted to Technical Committee IST/31, Computer graphics, image processing and environmental data representation.

A list of organizations represented on this committee can be obtained on request to its secretary.

This publication does not purport to include all the necessary provisions of a contract. Users are responsible for its correct application.

© The British Standards Institution 2019
Published by BSI Standards Limited 2019

ISBN 978 0 580 95811 3

ICS 35.140

**Compliance with a British Standard cannot confer immunity from legal obligations.**

This British Standard was published under the authority of the Standards Policy and Strategy Committee on 31 January 2019.

**Amendments/corrigenda issued since publication**

| Date | Text affected |
|---|---|

Documents
Document Sharing Hub

Full document details: BS ISO/IEC 18520:2019

INTERNATIONAL STANDARD

# ISO/IEC 18520

First edition
2019-01-31

# Information technology — Computer graphics, image processing and environmental data representation — Benchmarking of vision-based spatial registration and tracking methods for mixed and augmented reality (MAR)

*Technologies de l'information — Infographie, traitement d'images et représentation des données environnementales — Étalonnage des méthodes d'enregistrement géométriques et de suivi basées sur la vision pour le MAR*

Reference number
ISO/IEC 18520:2019(E)

AVA Documents
Document Sharing Hub
© ISO/IEC 2019

**COPYRIGHT PROTECTED DOCUMENT**

© ISO/IEC 2019, Published in Switzerland

All rights reserved. Unless otherwise specified, no part of this publication may be reproduced or utilized otherwise in any form or by any means, electronic or mechanical, including photocopying, or posting on the internet or an intranet, without prior written permission. Permission can be requested from either ISO at the address below or ISO's member body in the country of the requester.

ISO copyright office
Ch. de Blandonnet 8 • CP 401
CH-1214 Vernier, Geneva, Switzerland
Tel. +41 22 749 01 11
Fax +41 22 749 09 47
copyright@iso.org
www.iso.org


© ISO/IEC 2019 – All rights reserved

Documents
Document Sharing Hub

Full document details: BS ISO/IEC 18520:2019

# Contents

Page

**Foreword** ... **iv**
**Introduction** ... **v**
**1 Scope** ... **1**
**2 Normative references** ... **1**
**3 Terms, definitions, acronyms and abbreviated terms** ... **1**
3.1 Terms and definitions ... 1
3.2 Acronyms and abbreviated terms ... 3
**4 Overview of the framework** ... **3**
**5 Benchmarking processes** ... **4**
5.1 Overview ... 4
5.2 Process and process flow ... 5
5.3 Stakeholders ... 6
**6 Benchmark indicators** ... **7**
6.1 Overview ... 7
6.2 Reliability indicators ... 8
6.3 Temporality indicators ... 8
6.4 Variety indicators ... 9
**7 Trial set for benchmarking** ... **10**
7.1 Overview ... 10
7.2 Dataset for on- and off-site benchmarking ... 11
7.3 Physical object instances for on- and off-site benchmarking ... 11
**8 Conformance** ... **12**
**Annex A** (informative) **Benchmarking activities** ... **14**
**Annex B** (informative) **Usage examples of conformance checklists** ... **39**
**Annex C** (informative) **Conceptual relationship between this document and other standards** ... **60**
**Bibliography** ... **62**

© ISO/IEC 2019 – All rights reserved


Documents
Document Sharing Hub
Full document details: BS ISO/IEC 18520:2019

# Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular the different approval criteria needed for the different types of document should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO and IEC shall not be held responsible for identifying any or all such patent rights. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents) or the IEC list of patent declarations received (see http://patents.iec.ch).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation on the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT) see www.iso.org/iso/foreword.html.

This document was prepared by Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 24, *Computer graphics, image processing and environmental data representation*.

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html.


© ISO/IEC 2019 – All rights reserved
Documents
Document Sharing Hub
Full document details: BS ISO/IEC 18520:2019

# Introduction

In the development of mixed and augmented reality (MAR) applications, one of the most important technologies involves spatial registration and spatial tracking methods, especially vision-based methods. The research and development on registration and tracking based on computer vision technologies is flourishing and many new algorithms have been proposed every year.

Therefore, this document aims at fostering objective evaluation and comparison of diverse registration and tracking methods, in order to facilitate fairer competition among small and major companies/institutes involved in MAR technologies, applications and services.

Moreover, this document can be the baseline to standardize spatial registration and tracking methods which not only utilize a video camera but combine a video camera with other sensors such as another video camera, a depth camera, inertial sensors and infrastructure-based positioning technologies, and which utilize technologies such as IoT (Internet of Things) and GNSS (Global Navigation Satellite System).

The target audience of this document includes stakeholders of benchmarking activities. The following are examples of how this document can be used directly or indirectly:

— by a benchmarking service provider, a benchmark provider or a benchmarking competition organizer who wishes to align their benchmarking activities including self-benchmarking and open/closed competitions to be consistent with this document;

— by a technology developer/supplier who wishes to estimate and evaluate the performance of a vision-based spatial registration and tracking (vSRT) method for MAR appropriately with a benchmarking service provider, a benchmark provider or a benchmarking competition organizer who aligns their benchmarking activities to be consistent with this document; or

— by a technology user who wishes to obtain benchmarking results based on a benchmarking activity, which is consistent with this document, or to compare the existing vSRT methods for MAR in terms of their performance.

© ISO/IEC 2019 – All rights reserved

Documents
Document Sharing Hub
Full document details: BS ISO/IEC 18520:2019

*This page deliberately left blank*

Documents
Document Sharing Hub

Full document details: BS ISO/IEC 18520:2019

INTERNATIONAL STANDARD ISO/IEC 18520:2019(E)

# Information technology — Computer graphics, image processing and environmental data representation — Benchmarking of vision-based spatial registration and tracking methods for mixed and augmented reality (MAR)

## 1 Scope

This document identifies the reference framework for the benchmarking of vision-based spatial registration and tracking (vSRT) methods for mixed and augmented reality (MAR).

The framework provides typical benchmarking processes, benchmark indicators and trial set elements that are necessary to successfully identify, define, design, select and apply benchmarking of vSRT methods for MAR. It also provides definitions for terms on benchmarking of vSRT methods for MAR.

In addition, this document provides a conformance checklist as a tool to clarify how each benchmarking activity conforms to this document in a compact form by declaring which benchmarking processes and benchmark indicators are included and what types of trial sets are used in each benchmarking activity.

## 2 Normative references

There are no normative references in this document.

## 3 Terms, definitions, acronyms and abbreviated terms

For the purposes of this document, the following terms and definitions apply.

ISO and IEC maintain terminological databases for use in standardization at the following addresses:

— ISO Online browsing platform: available at https://www.iso.org/obp

— IEC Electropedia: available at http://www.electropedia.org/

### 3.1 Terms and definitions

**3.1.1**
**benchmark**
reference point against which comparisons can be made

Note 1 to entry: In the context of this document, the performance of vSRT methods for MAR is the object of comparison.

Note 2 to entry: See ISO/IEC 29155-1.

**3.1.2**
**benchmark indicator**
indicator that qualitatively shows a particular aspect of a *benchmark* (3.1.1) with appropriate metrics

**3.1.3**
**benchmarking**
activity of comparing objects of interest to each other or against *benchmarks* (3.1.1) to evaluate relevant characteristics

Note 1 to entry: In the context of this document, the object of interest is the performance of vSRT methods for MAR, and the characteristics are particular aspects of the performance such as reliability, temporal consistency, etc.

© ISO/IEC 2019 – All rights reserved


AVA Documents Document Sharing Hub
Full document details: BS ISO/IEC 18520:2019

Note 2 to entry: See ISO/IEC 29155-1.

**3.1.4**
**benchmarking instrument**
tool, method or guide used to support every activity within the *benchmarking* (3.1.3) framework

**3.1.5**
**benchmarking method**
particular procedure for conducting *benchmarking* (3.1.3) and obtaining *benchmarking results* (3.1.7)

**3.1.6**
**benchmarking repository**
repository which is designated for retaining information necessary for *benchmarking* (3.1.3) such as *datasets* (3.1.9) for benchmarking and *benchmarking results* (3.1.7)

Note 1 to entry: Some benchmarking repositories might contain all information, and some might contain a subset.

Note 2 to entry: See ISO/IEC 29155-1.

**3.1.7**
**benchmarking results**
*benchmarks* (3.1.1) as primary results, and intermediate results and reports on *benchmarking* (3.1.3) including findings, issues and lessons learned as secondary results

Note 1 to entry: In the context of this document, a typical intermediate result is the output of vSRT methods such as the estimated position of a camera the accuracy of which is evaluated with a benchmark indicator.

**3.1.8**
**competition organizer**
person or organization that hosts an *on-site* (3.1.14) or *off-site* (3.1.13) benchmarking competition

**3.1.9**
**dataset**
collection of data that contains target images and the *ground truth* (3.1.11) regarding the target images for *benchmarking* (3.1.3)

**3.1.10**
**extrinsic camera parameters**
parameters that define the position and orientation of a camera reference frame with respect to a known world reference frame such as the translation vector and the rotation matrix

**3.1.11**
**ground truth**
collection of measurements that is much more accurate as a whole than measurements by technologies which are the targets of *benchmarking* (3.1.3)

**3.1.12**
**intrinsic camera parameters**
parameters that define the relationship between the pixel coordinates of an image point and the corresponding coordinates in the camera reference frame

Note 1 to entry: Intrinsic camera parameters contain the focal length, the scale factors, the skew, the principal point, the lens distortion, etc.

**3.1.13**
**off-site benchmarking**
*benchmarking* (3.1.3) that is conducted with target images in datasets which were prepared beforehand

**3.1.14**
**on-site benchmarking**
*benchmarking* (3.1.3) that is conducted by executing vSRT programs while capturing target images on the spot


© ISO/IEC 2019 – All rights reserved
Documents
Document Sharing Hub
Full document details: BS ISO/IEC 18520:2019

**3.1.15**
**physical object**
object, which exists in the real world, used as a target of *spatial registration* (3.1.16), *spatial tracking* (3.1.17) and/or augmentation

Note 1 to entry: Especially for off-site benchmarking with physical objects, the physical object instances shall be easily available or deliverable objects such as paper crafts and toy bricks, or they shall be made accessible by providing information on how to find them for capturing the images. The physical object instances in off-site benchmarking are utilized to gather and acquire information necessary for vSRT methods such as visual features and 3D models of the objects.

**3.1.16**
**spatial registration**
establishment of the spatial relationship or mapping between two models, typically between virtual objects and target *physical objects* (3.1.15)

[SOURCE: ISO/IEC 18039:2019, 3.1.20]

**3.1.17**
**spatial tracking**
update of the spatial relationship or mapping between two models, typically between virtual objects and target *physical objects* (3.1.15) over time

**3.1.18**
**trial set**
combination of a dataset and a collection of *physical object* (3.1.15) instances for *off-site* (3.1.13) and *on-site* (3.1.14) benchmarking of vSRT methods for MAR

**3.1.19**
**vision-based spatial registration and tracking**
*spatial registration* (3.1.16) and *spatial tracking* (3.1.17) based on image processing and computer vision technologies

Note 1 to entry: The term "spatial registration and tracking" is also called "geometric registration and tracking" in Annex A and in some of the bibliography references.

## 3.2 Acronyms and abbreviated terms

| | |
|---|---|
| 3DEVO | 3D error of a virtual object |
| MAR | Mixed and augmented reality[1] |
| PEVO | Projection error of a virtual object |
| SLAM | Simultaneous localization and mapping[1] |
| vSRT | Vision-based spatial registration and tracking |

# 4 Overview of the framework

This clause outlines the reference framework of benchmarking of vSRT methods for MAR, the details of which are described in Clauses 5, 6, and 7. As shown in Figure 1, the reference framework is composed of the following three core components.

— **Benchmarking processes**, which include how to produce benchmarking outcomes such as benchmarking results, benchmark surveys and benchmarking instruments with benchmark indicators and trial sets and how to share benchmarking outcomes.

© ISO/IEC 2019 – All rights reserved

Documents
Document Sharing Hub
Full document details: BS ISO/IEC 18520:2019

— **Benchmark indicators**, which quantify the performance of vSRT methods in MAR by taking into account not only characteristics of vSRT methods in MAR such as reliability and temporal characteristics, but also fair comparisons.

— **Trial set elements**, which are composed of datasets and physical object instances to provide each benchmarking attempt with the same conditions.

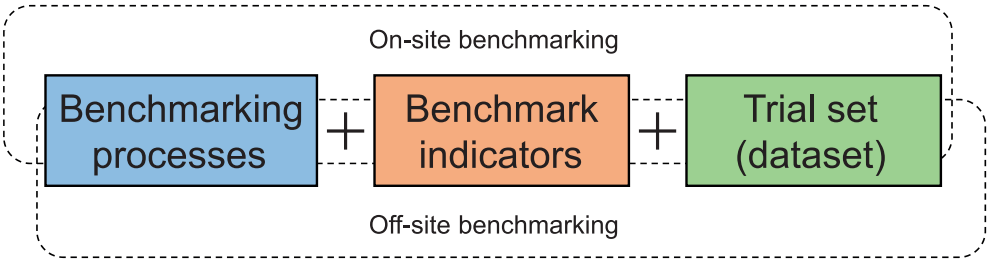


**Figure 1 — Core components of on- and off-site benchmarking framework**

The above three components are identified and defined in accordance with grass-roots activities for standardizing benchmarking schemes and for conducting on-site or off-site comparison of vSRT methods and MAR systems which are often held as contests and are introduced in Annex A.

On-site benchmarking methods are used to conduct benchmarking on the spot while capturing images of physical objects with working MAR systems. Compared with off-site benchmarking methods mentioned afterwards, it is inevitable that human factors affect on-site benchmarking results due to time and cost limitations, and constraints in preparation and operation. Therefore, it is highly recommended to simplify the implementation of benchmarking frameworks for on-site benchmarking.

By contrast, off-site benchmarking methods are used to conduct benchmarking with target images in datasets prepared beforehand. Compared with on-site benchmarking methods, the stakeholders have more time for preparing and conducting benchmarking. However, additional effort on the implementation of benchmarking frameworks is needed for alleviating issues related to fine tuning and cheating.

The on-site or off-site competition is one of the most concrete cases of on-site or off-site benchmarking methods, respectively. A.6, A.7 and A.8 introduce several case examples of on- and off-site competitions.

Typical processes of on- and off-site benchmarking are extracted from the grass-roots activities as in Annex A, and they are schematically described in Clause 5 by referring to the ISO/IEC 29155 series, especially ISO/IEC 29155-1. The conceptual relationship among the ISO/IEC 29155 series, ISO/IEC 18039 and this document is graphically indicated in Annex C. Each benchmark indicator and trial set element is also extracted from outcomes and discussions in the grass-roots activities as in Annex A. Clause 6 describes three major types of benchmark indicators which correspond to reliability, temporality and variety indicators, and Clause 7 describes reference elements in a trial set which contains dataset elements and physical object instances.

# 5 Benchmarking processes

## 5.1 Overview

This clause outlines benchmarking processes and related components necessary to produce and share benchmarking outcomes. Figure 2 illustrates the basic benchmarking process flow.


© ISO/IEC 2019 – All rights reserved
Documents
Document Sharing Hub
Full document details: BS ISO/IEC 18520:2019

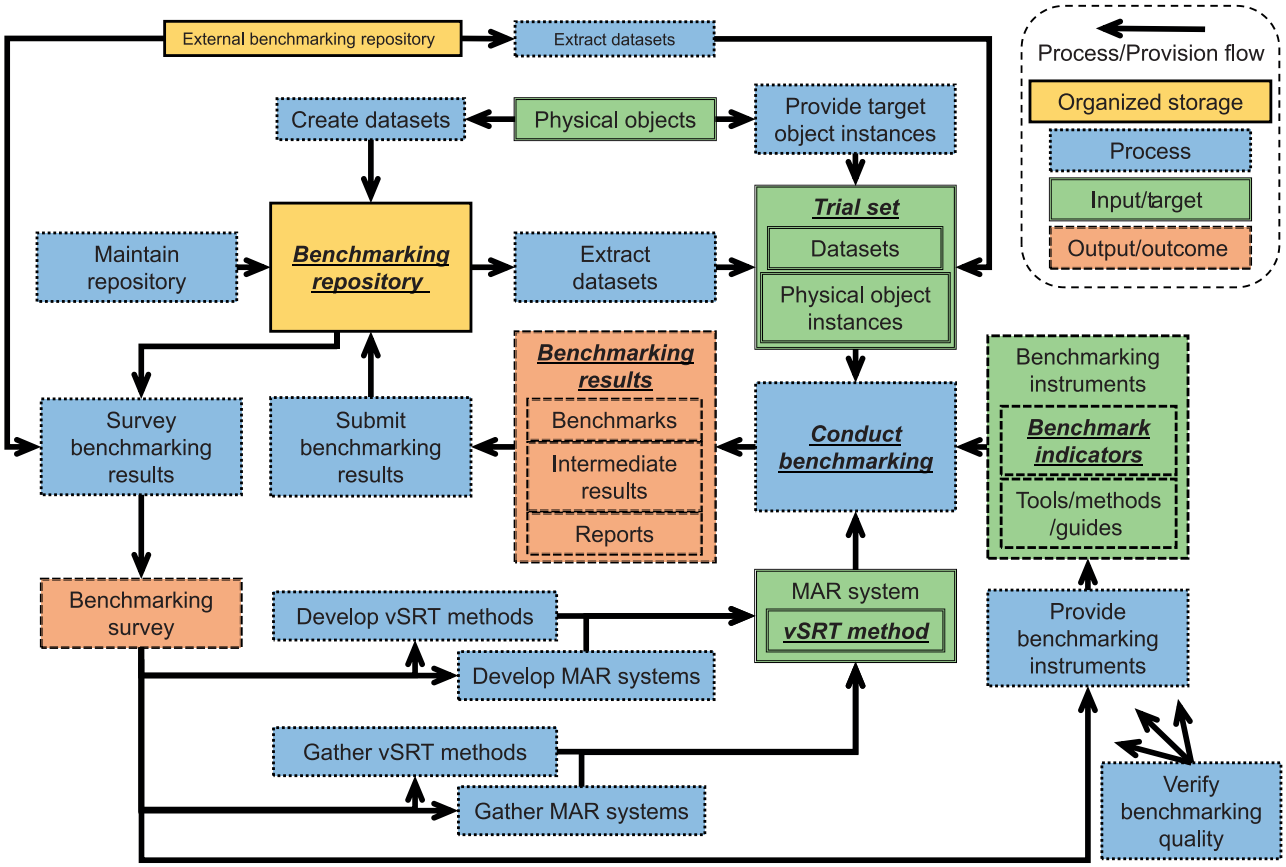


**Figure 2 — Basic benchmarking process flow**

## 5.2 Process and process flow

Although the details of the process flow in Figure 2 can differ in each specific benchmarking, it generally consists of process, target, input, output/outcome and organized storage, described in detail as follows:

— **Process**, which consists of one or more micro processes:

  — to develop or gather vSRT methods and MAR systems,

  — to prepare or conduct benchmarking,

  — to provide or maintain benchmarking instruments and repositories,

  — to share benchmarking results,

  — to manage or verify benchmarking quality.

— **Target**, which is a vSRT method or MAR system used as a benchmarking target.

— **Input**, which includes:

  — trial sets,

  — physical objects,

  — benchmarking instruments such as benchmark indicators, tools, methods and guides.

NOTE Trial sets and benchmarking instruments are also regarded as important outcomes of benchmarking activities.

© ISO/IEC 2019 – All rights reserved

Documents
Document Sharing Hub
Full document details: BS ISO/IEC 18520:2019

— **Output/outcome**, which includes:

  — benchmarking results such as benchmarks, intermediate results and reports,

  — benchmarking surveys.

— **Organized storage**, which is a benchmarking repository or other external repository.

## 5.3 Stakeholders

Various stakeholders are involved in processes on benchmarking of vSRT methods for MAR. Figure 3 illustrates a typical example of the correspondence between stakeholders and their roles in benchmarking processes. Based on roles, stakeholders can be logically classified into the following groups[2]:

— **Benchmark provider**, who creates and gathers datasets, maintains benchmarking repositories and provides benchmark surveys;

— **Benchmarking service provider**, who develops and provides benchmarking instruments, prepares trial sets, conducts benchmarking at the request of technology users and submits benchmarking results to a benchmarking repository;

— **Quality verifier**, who verifies benchmarking quality;

— **Technology developer**, who develops vSRT methods or MAR systems;

— **Technology supplier**, who supplies vSRT methods or MAR systems that technology developers have developed;

— **Technology user**, who chooses and utilizes vSRT methods or MAR systems based on the outcomes of benchmarking.

Targets of benchmarking are vSRT methods or MAR systems developed by technology developers or gathered by technology suppliers. To conduct benchmarking, benchmarking service providers shall prepare benchmarking instruments and trial sets. Datasets in the trial sets are extracted from a benchmarking repository. For on-site benchmarking, physical objects including rooms and spaces shall also be prepared. Benchmarking of vSRT methods or MAR systems is conducted with those inputs, and the results of benchmarking are submitted in a benchmarking repository by benchmarking service providers.

To choose appropriate vSRT methods or MAR systems, technology users refer to benchmark surveys or benchmarking results. The quality of the benchmark surveys or benchmarking results shall be ensured by verifying the quality of processes, inputs, outputs and organized storages in benchmarking activities. This is the main role of quality verifiers.

Various role-sharing schemes can be used in practice. Any person or organization can fulfil one or more roles. For example, benchmark providers can also have a role as benchmarking service providers. By contrast, one role can be fulfilled by several persons or organizations. For example, the competition organizer together with the contestants often fulfil the role of benchmarking service provider in conducting benchmarking. Many academic researchers do not maintain benchmarking repositories for a long term, but they often publish benchmarking surveys by conducting benchmarking for comparing several vSRT methods. In this case, they partially fulfil the benchmarking service provider's role and benchmark provider's role. In other cases, technology developers and suppliers often fulfil the partial roles of a benchmarking service provider in self-benchmarking or of a contestant of an on-site or off-site benchmarking competition.


© ISO/IEC 2019 – All rights reserved
Documents
Document Sharing Hub
Full document details: BS ISO/IEC 18520:2019

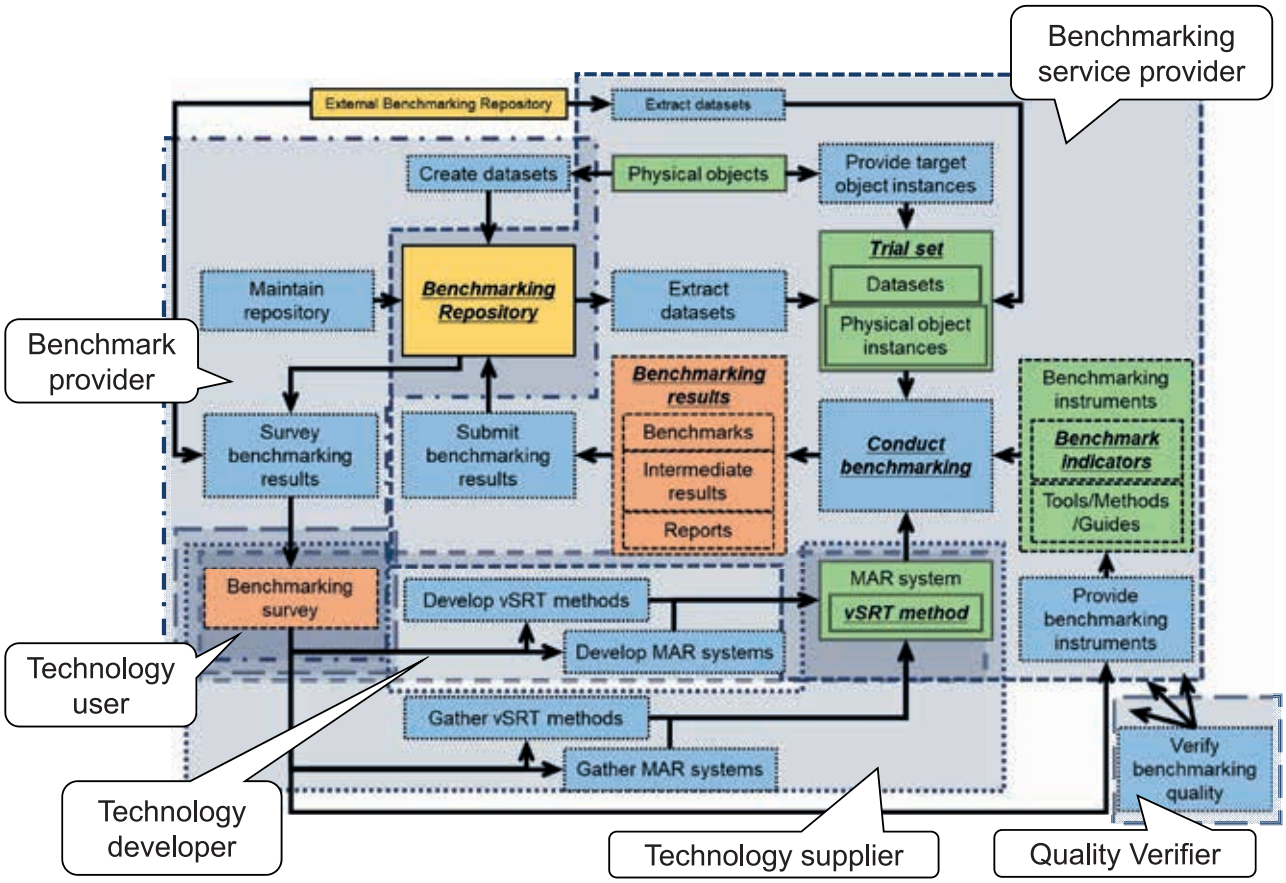


**Figure 3 — Example of the correspondence between stakeholders and their roles in benchmarking processes**

# 6 Benchmark indicators

## 6.1 Overview

This clause outlines three major types of benchmark indicators (reliability, temporality and variety), which should be utilized for fair comparison of vSRT methods in MAR. Table 1 shows representative benchmark indicators for both on- and off-site benchmarking.

**Table 1 — Benchmark indicators for off-site and on-site benchmarking**

| | Off-site | On-site |
|---|---|---|
| Reliability | — 3DEVO<br>— PEVO<br>— Re-projection error of image features<br>— Position and posture errors of a camera | — 3DEVO<br>— PEVO<br>— Re-projection error of image features<br>— Position and posture errors of a camera<br>— Completeness of a trial |

© ISO/IEC 2019 – All rights reserved

Documents
Document Sharing Hub
Full document details: BS ISO/IEC 18520:2019

| | Off-site | On-site |
|---|---|---|
| Temporality | — Throughput<br>— Latency | — Throughput<br>— Latency<br>— Time for trial completion |
| Variety | — Number of datasets<br>— Variety on properties of datasets | — Number of trials<br>— Variety on properties of trials |

## 6.2 Reliability indicators

This subclause presents reliability indicators. The following four indicators are for both off-site and on-site benchmarking.

— **3D error of a virtual object (3DEVO)**, which is the difference between the estimated position of a virtual object and the ground truth. 3DEVO is one of the most direct and intuitive indicators for vSRT methods for MAR, as one of the principal functions of MAR systems is to align virtual objects in 3D space based on the results obtained by the target vSRT method.

— **Projection error of a virtual object (PEVO)**, which is also one of the most direct and intuitive indicators for vSRT methods for MAR, as one of the most important functions of MAR systems is to render virtual objects based on the results obtained by the target vSRT method[3]. Assuming the simplest case in which a virtual point is projected as a virtual object to an estimated image plane, the distance between the projected and correct points is calculated as a PEVO value. The PEVO value can be measured in degrees or in pixels. Angular distance measure can provide a uniform measure in a screen space, whereas pixel number varies depending on positions in a screen space.

— **Re-projection error of an image feature**, which is the distance between a detected image feature in an image plane and the re-projection to the image plane with the 3D coordinates of the image feature that are recovered based on the target vSRT method. Assuming the simplest case in which the image feature is a feature point, the re-projection error can be the distance between the detected feature point and the re-projected point, and can be measured in degrees or in pixels as with PEVO.

— **Position and posture errors of a camera**, which is the difference between the estimated position and posture of a camera and the ground truth.

In addition to the aforementioned four reliability indicators, completeness of a trial should be employed especially for on-site benchmarking. This is because, in many on-site competitions, many MAR systems cannot help but stop performing spatial registration and tracking in the middle of the trial. Completeness of a trial involves evaluating the extent of a trial completion. It is regarded as the robustness of the target vSRT method.

## 6.3 Temporality indicators

This subclause presents temporality indicators which are necessary to discuss real-time issues in MAR.

The following two indicators as shown in Figure 4 are generally suitable for off-site benchmarking.

— **Throughput**, which is the rate at which a target image is processed through a target vSRT method or target MAR system during a specific period. It is often called frame rate.


© ISO/IEC 2019 – All rights reserved
Documents
Document Sharing Hub
Full document details: BS ISO/IEC 18520:2019

— **Latency**, which is the time delay produced by a target vSRT method or target MAR system.

  — For MAR-system benchmarking, the latency might be the length of time from when starting to capture a target image with the system to when rendering a virtual object based on the estimated position and posture of a camera with which the target image was captured.

  — For vSRT-method benchmarking, the latency might be the length of time from when starting to load a target image into the target method for input to when returning the estimated position and posture of a camera with which the target image was captured.

For on-site benchmarking, the time for trial completion can also be used as a temporality indicator because it is easy to measure. It is the length of time from starting a trial to finishing it.

The time for trial completion is often used in on-site competitions and can represent the overall performance of a target MAR system. However, it generally includes other aspects of the MAR system such as image capturing, virtual object rendering and human factors regarding the operator being considered.

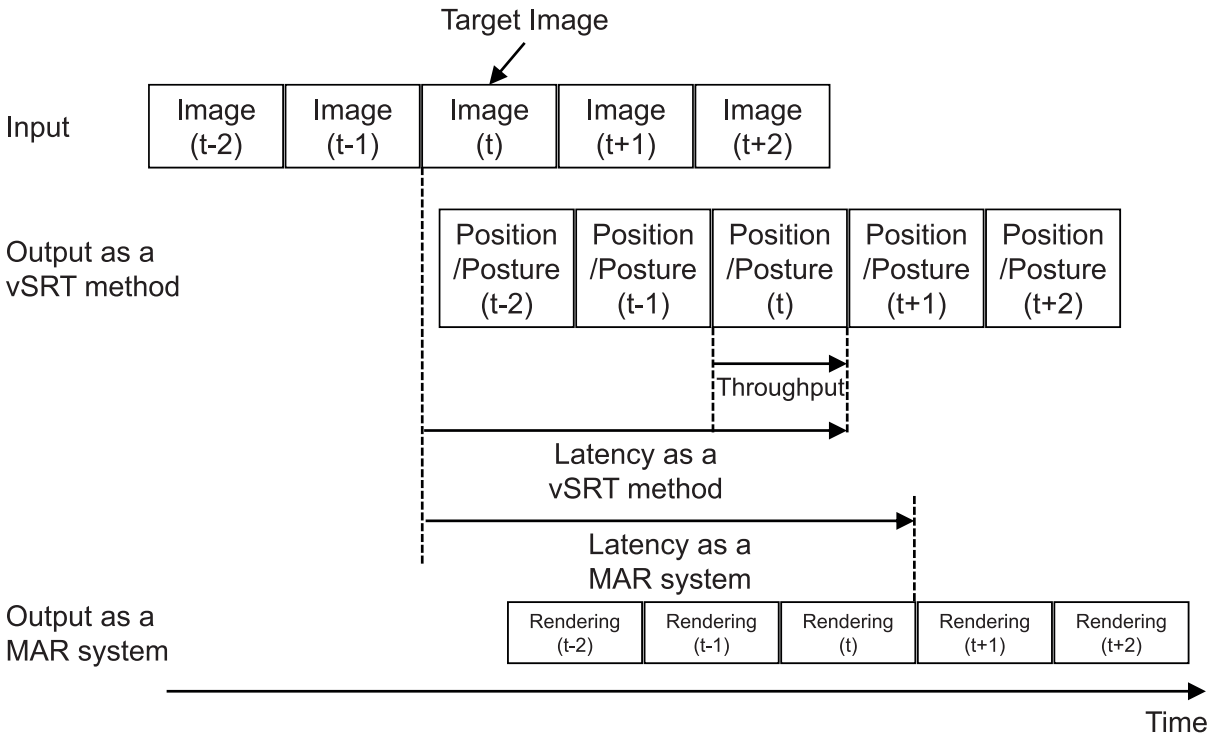


**Figure 4 — Temporality indicators — Latency and throughput**

## 6.4 Variety indicators

This subclause presents two variety indicators required to prevent fine tuning and cheating with some specific datasets for benchmarking.

The following two indicators are for off-site benchmarking.

— **Number of datasets**, which is the number of datasets used to obtain a benchmark indicator.

— **Variety on properties of datasets**, which is the variety on properties of datasets used to obtain a benchmark indicator. Typical examples of dataset properties are camera motion types, camera configurations, image quality and lighting conditions.

In general, performing fine tuning or cheating with many datasets is difficult. However, this difficulty diminishes if the properties of the datasets used for benchmarking are homogeneous.

© ISO/IEC 2019 – All rights reserved

Documents
Document Sharing Hub
Full document details: BS ISO/IEC 18520:2019

The following two indicators are for on-site benchmarking.

— **Number of trials**, which is the number of trials attempted during on-site benchmarking.

— **Variety on properties of trials**, which is the variety on the properties of trials for on-site benchmarking.

# 7 Trial set for benchmarking

## 7.1 Overview

This clause identifies the reference elements in a trial set used for benchmarking. Table 2 shows representative elements in a trial set for on- and off-site benchmarking.

**Table 2 — Trial set for benchmarking**

| | | Off-site | On-site |
|---|---|---|---|
| Dataset | Contents | — Image sequences<br>— Intrinsic/extrinsic camera parameters<br>— Challenge points<br>— Optional contents<br>  — 3D models for the target objects and for virtual objects<br>  — Image feature correspondences<br>  — Depth image sequences<br>  — Self-contained sensor data, etc. | — Challenge points<br>— 3D models for the target objects and for virtual objects |
| | Metadata | — Scenario<br>— Camera motion type<br>— Camera configuration<br>— Image quality | — Scenario |
| Physical object instances | Contents | — Physical objects | |
| | Metadata | — Information on how to find the physical objects | |


© ISO/IEC 2019 – All rights reserved
Documents
Document Sharing Hub
Full document details: BS ISO/IEC 18520:2019

## 7.2 Dataset for on- and off-site benchmarking

This subclause presents representative elements such as contents and metadata in datasets for off-site benchmarking. As shown in Table 2, elements in datasets for on-site benchmarking are a subset of them.

— **Contents**, which include:

  — **Image sequences**, which are target image sequences.

  — **Intrinsic/extrinsic camera parameters**, which are the ground truth of intrinsic and extrinsic parameters for one or more video cameras used to capture the image sequences.

  — **Challenge points**, which are points with the ground truth position for rating each MAR system based on a 3DEVO or PEVO measurement. The position of each challenge point is estimated and visualized by the target MAR system for rating purposes.

  — **Optional contents**, which can be included in datasets according to requirements of each specific benchmarking. The following elements are typical examples of optional contents:

    — **3D models for both target and virtual objects**, which are 3D model data for the target objects in image sequences and for virtual objects overlaid in benchmarking.

    — **Image feature correspondences**, which is the ground truth of image feature correspondences through the image sequences.

    — **Depth image sequences**, which are image sequences where each pixel value is measured with depth sensors using, for example, an active stereo method.

    — **Self-contained sensor data**, which are sensor data measured by self-contained sensors such as an accelerometer, gyro sensor, magnetometer or barometer.

— **Metadata**, which include:

  — **Scenario**, which is the description of an application of MAR such as indoor/outdoor navigation, tabletop MAR interface or industrial application.

  — **Camera motion type**, which is the description of camera motion such as translation only, rotation only, walking motion, handheld motion or vehicle motion.

  — **Camera configuration**, which is the description of the camera configuration such as white balance, shutter type (global or rolling) and shutter speed.

  — **Image quality**, which is the description of the target-image quality such as image resolution, defocusing and motion blur.

The image format can be JPEG, PNG or Raw. A lossless format is more appropriate. The ground truth data format can be XML (eXtensible Markup Language), X3D (eXtensible 3D), JSON (JavaScript Object Notation), etc. The 3D model format can be X3D, Collada, etc. The metadata can be used as a source of variety indicators and can also correspond to the properties of each dataset for ease of dataset retrieval. The metadata format can be XML, JSON, etc.

## 7.3 Physical object instances for on- and off-site benchmarking

This subclause presents representative elements in physical object instances for on- and off-site benchmarking.

In off-site benchmarking, physical objects shall be easily available or deliverable so that contestants can capture the images because such physical objects are necessary for technology developers to develop and adjust their vSRT method or MAR system before conducting off-site benchmarking with target image sequences in which the physical objects are observed. In actual benchmarking activities, paper models, toy bricks, cars, etc. can be employed as easily available or deliverable physical objects, as introduced in Annex A.

© ISO/IEC 2019 – All rights reserved

Documents
Document Sharing Hub
Full document details: BS ISO/IEC 18520:2019

In on-site benchmarking, physical objects are also used for benchmarking preparation and actual benchmarking. For developing and adjusting vSRT methods or MAR systems, the physical objects should be easily available or deliverable as in off-site benchmarking. However, when conducting the actual benchmarking, physical objects may be observable only during trials, as in Annex A.

Metadata shall be provided with physical objects, and it shall contain information on how to find or obtain the physical objects.

## 8 Conformance

Table 3 shows an example of a conformance checklist. This checklist or customized ones shall be used to clarify how each benchmarking activity conforms to this document in a compact form. This checklist is useful to summarize and declare which benchmarking processes and benchmark indicators are included, and what types of trial sets are used in each benchmarking activity. Annex B shows examples of how to use the conformance checklist shown in Table 3.

**Table 3 — Conformance checklist**

| | | Check | Item | Remarks |
|---|---|---|---|---|
| Process flow | Process | [ ] | Develop vSRT methods and/or MAR systems: | |
| | | [ ] | Gather vSRT methods and/or MAR systems: | |
| | | [ ] | Prepare and conduct benchmarking: | |
| | | [ ] | Provide and maintain benchmarking instruments: | |
| | | [ ] | Provide and maintain benchmarking repositories: | |
| | | [ ] | Share benchmarking results: | |
| | | **Check** | | |
| | Target/ input/ output/ organized storage | [ ] | vSRT method: | |
| | | [ ] | MAR system: | |
| | | [ ] | Trial sets and physical objects: | |
| | | [ ] | Benchmarking instruments: | |
| | | [ ] | Benchmarking results: | |
| | | [ ] | Benchmarking surveys: | |
| | | [ ] | Benchmarking repository: | |
| | | [ ] | External repositories: | |
| | | **Check** | | |
| Indicator | Reliability | [ ] | 3DEVO: | |
| | | [ ] | PEVO: | |
| | | [ ] | Re-projection error of image features: | |
| | | [ ] | Position and posture errors of a camera: | |
| | | [ ] | Completeness of a trial: | |
| | Temporality | [ ] | Throughput: | |
| | | [ ] | Latency: | |
| | | [ ] | Time for trial completion: | |
| | Variety | [ ] | Number of datasets/trials: | |
| | | [ ] | Variety on properties of datasets/trials: | |
| | | **Check** | | |


© ISO/IEC 2019 – All rights reserved
Documents
Document Sharing Hub
Full document details: BS ISO/IEC 18520:2019

| | | | Check | Item | Remarks |
|---|---|---|---|---|---|
| Trial set | Dataset | Contents | [ ] | Image sequences: | |
| | | | [ ] | Intrinsic/extrinsic camera parameters: | |
| | | | [ ] | Challenge points: | |
| | | | [ ] | Optional contents: | |
| | | Metadata | [ ] | Scenario: | |
| | | | [ ] | Camera motion type: | |
| | | | [ ] | Camera configuration: | |
| | | | [ ] | Image quality: | |
| | Physical object instances | Contents | [ ] | Physical objects: | |
| | | Metadata | [ ] | How to find the physical objects: | |

For on- and off-site benchmarking, in general cases such as self-benchmarking, repeatability shall be guaranteed as much as possible by the stakeholders such as benchmarking service providers and benchmark providers. For this purpose, they shall declare the specified benchmarking framework including processes, targets, inputs, outputs, organized storages, the specified formulas of benchmark indicators and the specified format and contents of trial sets to other stakeholders. The declaration with the checklist shall be based on the reference framework in Clauses 5, 6, and 7. In addition, the stakeholders shall consistently provide trial sets and benchmarking surveys while continuously maintaining repositories as in Figure 2.

For on- and off-site benchmarking in competitions, it is more difficult to guarantee repeatability completely owing to human factors. For on-site benchmarking, condition changes over time are also inevitable. However, the competition organizers at least shall make every effort to declare the specified processes and benchmark indicators, and the format of trial sets, to other stakeholders such as the contestants as early as possible. The declaration with the checklist shall be also based on the reference framework in Clauses 5, 6, and 7.

© ISO/IEC 2019 – All rights reserved

Documents
Document Sharing Hub
Full document details: BS ISO/IEC 18520:2019

# Annex A
(informative)

# Benchmarking activities[1)]

## A.1 Open datasets by an academic community

In the TrakMark working group[4][5][6], various activities involving evaluations of tracking methods for MAR have been held. One of the main activities is to generate and provide datasets for benchmarking of vSRT methods. The datasets are composed of camera images and ground truth data that include intrinsic and extrinsic camera parameters of each image. Three packages were prepared as a trial, referred to as the image sequence set No.1 (Figure A.1). This set consists of indoor, outdoor and CG-based sequence packages. For example, the upper images are samples of the sequence in the film studio package. The special feature of this package is that these sequences were recorded with an HD camera and reference data were measured by physical sensors such as a rotary encoder. The lower images are samples of the sequence in the conference venue package, which includes the ground truth of the camera pose. The venue of ISMAR 2009 was recreated as a 3D CG model, and computer-generated images and the related ground truth were made from it.

**Figure A.1 — Examples in TrakMark datasets No.1**

Based on the experience of making the first set, five more packages as the second image sequence set were prepared. In making this set, a variation of the image sequences was focused on. For example, NAIST campus package 02 (Figure A.2) includes 18 sequences. Although these sequences were taken in the same location, there are a variety of properties as in Table A.1.

1) Any names of specific applications, file formats, etc. are given for the convenience of users of this document and do not constitute an endorsement by ISO/IEC.


© ISO/IEC 2019 – All rights reserved
Documents
Document Sharing Hub
Full document details: BS ISO/IEC 18520:2019