

IEEE Recommended Practice for the Quality Assessment of Light Field Imaging

IEEE Computer Society

Developed by the
Standards Activities Board

IEEE Std 3333.1.4™-2022

STANDARDS

IEEE Recommended Practice for the Quality Assessment of Light Field Imaging

Developed by the

Standards Activities Board
of the
IEEE Computer Society

Approved 8 November 2022

IEEE SA Standards Board

Abstract: This recommended practice establishes methods of quality assessment of light field imaging based on psycho-physical studies. It also defines metrics for the quality assessment and establishes criteria for subjective assessment of light field imaging, including human factors and judgments, and identifies and quantifies quality degradation including the impact of visual contents, camera settings, compression distortion, interpolation distortion by intermediate view rendering, and structural distortion. The recommended practice addresses a series of visual phenomena that can degrade visualization specifically for light field imaging. Visual environment characteristics and viewing conditions are also part of the scope of the standard, including viewing distance, viewer position, viewing freedom, and display characteristics.

Keywords: datasets, IEEE 3333.1.4, light field imaging, objective quality, psycho-physical studies, quality assessment, subjective tests

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

Copyright © 2023 by The Institute of Electrical and Electronics Engineers, Inc.
All rights reserved. Published 31 May 2023. Printed in the United States of America.

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

PDF: ISBN 978-1-5044-9262-1 STD25827
Print: ISBN 978-1-5044-9263-8 STDPD25827

IEEE prohibits discrimination, harassment, and bullying.

For more information, visit <https://www.ieee.org/about/corporate/governance/p9-26.html>.

No part of this publication may be reproduced in any form, in an electronic retrieval system or otherwise, without the prior written permission of the publisher.

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 IEEE Societies and subcommittees of IEEE Standards Association (IEEE SA) Board of Governors. 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. Statements made by volunteers may not represent the formal position of their employer(s) or affiliation(s).

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 subcommittees of the IEEE SA Board of Governors 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](#).¹ An IEEE Account is needed to access the application.

Comments on standards should be submitted using the [Contact Us](#) form.²

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

¹Available at: <https://development.standards.ieee.org/myproject-web/public/view.html#landing>.

²Available at: <https://standards.ieee.org/content/ieee-standards/en/about/contact/index.html>.

by public authorities and private users, neither IEEE nor its licensors 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](#).³ 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](#).⁴ 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](#).⁵

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.

³Available at: <https://ieeexplore.ieee.org/browse/standards/collection/ieee>.

⁴Available at: <https://standards.ieee.org/standard/index.html>.

⁵Available at: <https://standards.ieee.org/about/sasb/patcom/materials.html>.

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 recommended practice was completed, the P3333.1.4 Working Group had the following membership:

Maria Martini, *Chair*
Federica Battisti, *Vice Chair*
Peter A. Kara, *Vice Chair*
Edris Shaflee, *Secretary*

Sewoong Ahn	Mary Guindy	Suiyi Ling
Ali Ak	Jesus Gutierrez	Pramit Mazumdar
Shaymaa Al-Juboori	Kim Jaekyung	Chunli Meng
Ping An	Soumya Jana	Choi Seonghwa
Appina	Mingyu Jang	Barnita Sharma
Balasubramanyam	Huh Jungwoo	Liquan Shen
Nabajet Barman	Woojae Kim	Aniko Simon
Sridhar Borapureddi	Lee KyungJune	Roopak R. Tamboli
Marco Carli	Patrick Le Callet	Toinon Vigier
Nguyen Anh Duc	Changhon Lee	Chao Yang
Waqas Ellahi	Jeong Haeng Lee	Park Yeseung
Lu Ge	Sanghoon Lee	Yu Yuan

The following members of the individual Standards Association balloting group voted on this recommended practice. Balloters may have voted for approval, disapproval, or abstention.

Ping An	Werner Hoelzl	Sanghoon Lee
Charles Barest	Peter A. Kara	Maria Martini
Thomas Coughlin	Piotr Karocki	Yu Yuan
David Fuschi	Patrick Le Callet	

When the IEEE SA Standards Board approved this recommended practice on 8 November 2022, it had the following membership:

David J. Law, *Chair*
Ted Burse, *Vice Chair*
Gary Hoffman, *Past Chair*
Konstantinos Karachalios, *Secretary*

Edward A. Addy	Johnny Daozhuang Lin	Mark Siira
Ramy Ahmed Fathy	Kevin Lu	Dorothy V. Stanley
J. Travis Griffith	Daleep C. Mohla	Lei Wang
Guido R. Hiertz	Andrew Myles	F. Keith Waters
Yousef Kimiagar	Damir Novosel	Karl Weber
Joseph L. Koepfinger*	Annette D. Reilly	Sha Wei
Thomas Koshy	Robby Robson	Philip B. Winston
John D. Kulick	Jon Walter Rosdahl	Daidi Zhong

*Member Emeritus

Introduction

This introduction is not part of IEEE Std 3333.1.4-2022, IEEE Recommended Practice for the Quality Assessment of Light Field Imaging.

Light field technology, besides being a hot topic in the research area of future content visualization, has a very high potential in a number of application areas. According to a recent report, the light field market is expected to reach USD 1822.3 million by 2023. The growth is expected to be driven by the increasing demand for light field technology in visual effects of movies, marketing, prototyping, defense, and healthcare. Although light field displays are not yet in the consumer market, they are indeed commercially available, as horizontal-only parallax (HOP) displays can be purchased, and their usage is spreading in the industry. No additional viewing equipment is required in order to experience the immersive visual content in 3D, a feature that makes them desirable from the perspective of the users.

The assessment of the quality of experience (QoE) of glasses-free light field visualization is a challenge mainly for two reasons: (1) Specific visual phenomena that do not apply to other forms of visualization; in particular, horizontal motion parallax is affected by reduced angular resolution that can result in a crosstalk effect (when adjacent source views interfere with each other) or discrete view jumps (when there is a perceivable sudden shift between source views, without proper transition). Insufficient spatial resolution results in blur that is not uniform across the screen of the display, as light rays hit irregular positions. (2) The corresponding subjective quality evaluation techniques are not yet standardized. Hence, there is a need for extensive efforts in developing on one hand objective light field imaging quality metrics to evaluate the visual quality in agreement with subjective human judgments; on the other hand, the specific procedures for performing subjective tests for light field imaging should be selected and standardized/recommended.

In order to address the aforementioned needs, the work toward this recommended practice started at the end of 2019 within working group P3333.1.4, and the first draft was completed in December 2021.

This document is organized in five main sections. In the first one, application scenarios, available acquisition devices, rendering methods, and characteristics of the light field displays and content are analyzed with respect to the use cases. The following clause discusses the factors that influence the quality of the light field content. Impairments are categorized based on the source of the influencing factor in the imaging pipeline. Subjective quality assessment for light field visualization is discussed in the following clause. After initial considerations, a list of guidelines and recommendations for subjective tests on light field displays is provided. Visualization on other display types used for visualizing light field content is also addressed. Objective metrics for quality assessment (QA) of LF media are presented and compared in the following clause, where relevant guidelines and recommendations are also provided. Finally, existing open source datasets that can support the development of new metrics for the quality assessment of light field imaging are presented and compared. The advantages and limitations of the existing datasets in different use cases are presented, and relevant guidelines and recommendations are provided.

Contents

1. Overview	10
1.1 Scope	10
1.2 Word usage	10
2. Normative references	10
3. Definitions, acronyms, and abbreviations	11
3.1 Definitions	11
3.2 Acronyms and abbreviations	12
4. Use cases, acquisition, visualization, and context characteristics	14
4.1 Introduction	14
4.2 Use cases	14
4.3 Characteristics of acquisition devices	15
4.4 Display characteristics	15
4.5 Content characteristics	16
4.6 Consideration	19
5. Influencing factors and impairments for the quality of light field imaging	20
5.1 Introduction	20
5.2 Acquisition-related factors	20
5.3 Impairments due to light field processing	21
5.4 Compression related factors	22
5.5 Effect of rendering techniques	22
6. Subjective assessment of light field visualization	23
6.1 Introduction	23
6.2 Subjective quality assessments on light field displays	23
6.3 Subjective quality assessment on other display types	26
7. Objective assessment of light field visualization quality	28
7.1 Introduction	28
7.2 Objective quality assessment of light field content	28
7.3 Summaries of quality assessment frameworks	30
7.4 Performance assessment of objective quality metrics	38
7.5 Considerations	39
7.6 Guidelines and recommendations	42
8. Datasets	43
8.1 Introduction	43
8.2 Review of existing light field datasets	43
8.3 A new LF video dataset acquired via plenoptic video camera	56
8.4 Other datasets	56
8.5 Considerations	56
8.6 Guidelines for the design of light field quality datasets	57
Annex A (informative) Bibliography	58

IEEE Recommended Practice for the Quality Assessment of Light Field Imaging

1. Overview

1.1 Scope

This recommended practice establishes methods of quality assessment of light field imaging based on psycho-physical studies. It also defines metrics for the quality assessment and establishes criteria for subjective assessment of light field imaging, including human factors and judgments, and identifies and quantifies quality degradation including the impact of visual contents, camera settings, compression distortion, interpolation distortion by intermediate view rendering, and structural distortion. The recommended practice addresses a series of visual phenomena that can degrade visualization specifically for light field imaging. Visual environment characteristics and viewing conditions are also part of the scope of the standard, including viewing distance, viewer position, viewing freedom, and display characteristics.

1.2 Word usage

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*).^{6,7}

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

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

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

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

⁶The use of the word *must* is deprecated and cannot be used when stating mandatory requirements; *must* is used only to describe unavoidable situations.

⁷The use of *will* is deprecated and cannot be used when stating mandatory requirements; *will* is only used in statements of fact.

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.

ITU-T P910, Subjective video quality assessment methods for multimedia application.⁸

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.⁹

angle of view: Angular extent of a given scene that is imaged by a camera.

angular resolution: Minimal angle between distinct light rays that can be captured or visualized (for a display, with respect to a single point of the screen). Angular resolution quantifies the smoothness of the parallax effect. For microlens-based LF displays, when a microlens relays the image (formed by the objective lens on an intermediate image plane) to the image sensor, the number of pixels in the image sensor behind each of the lenslets defines the angular resolution. For contents that are generated from series of 2D images, angular resolution can be derived from the ratio between the number of source images (along an axis) and the FOV.

autostereoscopic display: Display showing stereoscopic content (with binocular perception of 3D depth) without the use of special glasses or other type of headgear on the part of the viewer.

dense light field: Light field (LF) where the transition between viewpoints is smooth without view interpolation.

depth of field: Distance between the nearest and the farthest extent of a scene that are in acceptably sharp focus.

disparity: Difference of location of the same 3D point in the images seen by the left and right eyes, resulting from the eyes' horizontal separation.

field of view: Angle in which 3D space is captured and visualized. For acquisition, it defines the portion of 3D space that is observable by every camera of the capture system; for visualization, it is the angle measured from a single point of the screen in which valid light field is observable.

focal length: Distance from the center of the lens to the principal foci (or focal points) of the lens.

full-parallax display: Display that can recreate the complete light field. It simultaneously supports horizontal and vertical parallax.

ground truth: Information known to be real or true, provided by direct observation and measurement.

horizontal-only parallax display: Display that is configured to provide the different views of the light field content in different view directions confined to the horizontal plane.

interpupillary distance: The distance between the centers of the pupils of the left and the right eyes.

⁸ITU publications are available for the International Telecommunications Union. (<https://www.itu.int/>).

⁹IEEE 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.

lenslet light field: Imaging representation model obtained placing a microlens (lenslet) array in the optical path of a conventional monocular camera, capturing the light for each spatial position (x, y) and coming from any angular direction.

parallax: Displacement or difference in the apparent position of an object viewed along two different lines of sight, measured by the angle or semi-angle of inclination between those two lines.

saliency: Pre-attentive stage of human visual perception that enables an observer to gain awareness about the different relevance of the regions appearing in an observed scenario. Saliency determines a spatially circumscribed region of the visual field (the so-called “focus of attention”).

sparse light field: Light field (LF) where the transition between viewpoints is not smooth without view interpolation.

super resolution: Light field (LF) imaging super-resolution aims at reconstructing high-resolution LFs from their low-resolution counterparts. In certain scientific contexts, super resolution refers to an angular resolution that is so high (i.e., the light field is so dense) that at least two distinct light rays may address a single pupil, enabling actual depth focusing within the visualized content.

3.2 Acronyms and abbreviations

2D	two-dimensional
3D	three-dimensional
4D	four-dimensional
bpp	bits per pixel
CF	colorfulness
CNN	convolutional neural network
CR	compression ratio
CSC	convolutional sparse coding
D-EPI	discriminative epipolar plane image
DCT	discrete cosine transform
DIBR	depth-image-based rendering
DMOS	difference mean opinion score
DQ	quantized depth map
DSCQS	double stimulus continuous quality scale
DeLFIQE	deep learning-based light field image quality evaluator
EPI	epipolar plane image
FOV	field of view
FP	full parallax
FR	full reference
GAUSS	Gaussian blur
HEVC	high efficiency video coding
HOG	histogram of oriented gradients

HOP	horizontal-only parallax
HRC	hypothetic reference circuit
HVS	human visual system
IQA	image quality assessment
IW-SSIM	information content weighted structural similarity measure
JND	just noticeable difference
JOD	just objectionable difference
KROCC	Kendall's rank-order correlation coefficient
KTCC	Kendall's tau rank correlation coefficient
LCC	linear correlation coefficient
LF-IQA	light field image quality assessment
LF-OCC	light field occlusion
LFDD	light field image dataset for performance evaluation of objective quality metrics
LFI	light field image
LFR	light field raw
LF	light field
LI	linear interpolation
MOS	mean opinion score
MLI	microlens image
MS-SSIM	multiscale structural similarity index
MSE	mean square error
MV	multiview
NNI	nearest neighbor interpolation
NR-LFIQA	no-reference light field image quality assessment
NR	no reference
NSS	natural scene statistics
QA	quality assessment
OPT	optical flow estimation
PCSC	principal component spatial characteristic
PLCC	Pearson's linear correlation coefficient
PNG	portable network graphics
PSNR	peak signal-to-noise ratio
PVS	processed video sequence
QP	quantisation parameter
QoE	quality of experience
RMSE	root mean square error

RR	reduced-reference
SAI	sub-aperture image
SINQ	S3D INtegrated Quality
SI	spatial information
SMV	super-multiview
SRC	source reference content
SROCC	Spearman rank order correlation coefficient
SSDC	sparse set and disparity coding
SSIM	structural similarity
SVR	support vector regression
TAVI	tensor angular variation index
UGGD	univariate generalized Gaussian distribution
VOP	vertical-only parallax

4. Use cases, acquisition, visualization, and context characteristics

4.1 Introduction

In this clause, application scenarios, available acquisition devices, rendering methods, and characteristics of the light field displays and content are analyzed with respect to the use cases.

4.2 Use cases

There are many scenarios in which light field systems (e.g., acquisition, transmission, rendering, compression, etc.) could be utilized. To this aim, it is important to take into account the potential issues arising from both the specific use cases and light field technology itself. In the scientific literature, multiple efforts have been devoted to classify the most relevant use case scenarios for light field visualization (see Kara, et al. [B1] and Kara, et al. [B46]). These efforts distinguish the context of medical, industrial, commercial, educational, cultural, and communication use cases. As light field still has not reached the consumer market and is in an early emerging phase in the context of professional utilization, the majority of the use case scenarios covered by this document are projected by the experts of the field. The success of the adoption of light field systems is strictly dependent on the offered visual quality. In fact, in all the discussed contexts, the degradation of light field visualization quality can compromise the efficiency and the effectiveness of the respective purpose.

The severity of the degradation can have different impacts depending on the application scenario (including type of display, number of simultaneous users, environmental conditions, etc.).

Medical use cases are prominent examples, as even slight degrees of quality degradation can result in inaccurate diagnosis. Among the possible degradations, limited spatial and angular resolutions, the presence of the crosstalk effect, insufficient depth and contrast values should be avoided.

Certain industrial use cases can tolerate limited angular and spatial resolution if the visualized content is not characterized by fine details. On the contrary, brightness and contrast might have a major impact on these types of applications, particularly for the purpose of prototyping or commercial demonstration.

Commercial scenarios are less demanding with respect to spatial resolution requirements as long as the information can be effectively delivered to the intended user. Two crucial aspects are the depth information and

the contrast for numerous commercial contexts, especially in the case of utilization in outdoor environments (e.g., digital signage).

In the field of education, the level of quality required for the transfer of knowledge may vary depending on the specific learning scenario and the display device. For example, in the case of specialized training, such as in medical and engineering courses, the spatial resolution is a relatively strict requirement. Regarding head up displays, brightness and contrast may significantly affect the effectiveness of training.

In the case of cultural use cases, a wide field of view is required to accommodate a large number of simultaneous viewers, and high levels of brightness and contrast can be relevant depending on the location of the display (e.g., indoor and outdoor exhibitions). Requirements toward angular resolution are typically not strict in the case of scenarios of static observation (e.g., cinema systems).

Finally, in the use cases of communication (e.g., telepresence), angular resolution may be more relevant for mobile scenarios (i.e., the participant of the call is moving), which also requires a larger field of view. Static scenarios are more tolerant in both aspects.

4.3 Characteristics of acquisition devices

Light field cameras are based on the integral imaging technique proposed by Lippmann [B63]. Modern versions of such devices were implemented by Adelson [B4], and light field cameras recently appeared on the consumer market. The two main types of acquisition devices are the afocal [B65] and the focused [B65] configurations, based on the distance between the microlens and the sensor. Camera arrays are also used as acquisition devices ([B112], [B120], and [B112]). As described in 4.1, different use cases have different requirements, depending on the application at hand. Therefore, the acquisition device should be selected based on the best match between its hardware characteristics and the considered use case.

- Multiplexed acquisition devices capture the four-dimensional (4D) LF on the two-dimensional (2D) sensor plane by multiplexing the angular domain into the spatial domain. Commonly, a microlens array is used between the image sensor and the optical system of the camera. Each microlens captures the scene from a slightly different position. A trade-off between spatial and angular resolution is inevitable for these systems. In other words, increased angular resolution may stem reduced spatial resolution.
- Multi-sensor methods rely on an array of image sensors to simultaneously capture the 4D LF from different positions. The number of image sensors on the array dictates the angular resolution of the captured LF. The spatial resolution of the captured LF depends on the spatial resolution of the individual image sensors. Image sensors are often distributed uniformly on a planar or spherical surface.
- Time-sequential approaches utilize a single image sensor and capture the 4D LF by changing the position of the image sensor over time. Commonly, the image sensor is mounted on a mechanical gantry or a robotic arm. The angular dimensions of the captured LF depends on the gantry stop positions, while the spatial resolution is dictated by the image sensor. It is not possible to capture a LF instantaneously with time-sequential approach. Therefore, only static scenes can be captured.

An analysis of the characteristics of acquisition devices is in [B121].

4.4 Display characteristics

The majority of QoE-related research (i.e., subjective studies on real light field displays) is carried out on projection-based displays. These systems can be classified into two broad categories: back-projection and front-projection systems (see [B46] and [B43]). The attributes of such devices and their effects on QoE are detailed in Clause 6. In the case of back-projection systems, the projectors and the observers are on the different sides of the screen, while in the case of front-projection systems, they are located on the same side.

From a perceptual point of view, light field displays are designed to replicate the field of light by representing both angular and spatial information. They provide additional perceptual cues, such as focus (accommodation and retinal blur), convergence, motion parallax, and binocular disparity.

In the following list, displays that may accommodate multiple viewers are reported. In such scenarios, multiple individuals may simultaneously view the visualized content within the valid viewing area, specified by the display's field of view and angular resolution.

- Wave optics refer to holographic displays in which the interference pattern is computer-generated by modeling the object reference beam interaction. The resulting fringe pattern is sent to a spatial light modulator for object reconstruction. Using wave optics suffers from severe limitations in terms of field of view and rendering capabilities due to the high frame rate required by the holographic video signal.
- Integral imaging utilizes arrays of microlenses over a set of light emitters to provide the light emitted by each pixel with a specific and distinct direction. The main limitation of integral imaging designs is the trade-off between spatial and angular resolution, considering the fixed size and number of the underlying pixels. Despite the relatively good angular resolutions shown by the microlens-based approaches, a small depth of field may result in noticeable blur at a shallow depth into the scene.
- Diffraction-based displays are created through the modifications of conventional 2D displays. For example, the directional control of individual pixels on existing LCD displays is achieved through the diffractive backlighting of commercial LCDs. Flexibility is supported by the ability to tune each grating to an arbitrary direction by varying the direction and frequency of the diffraction grating. However, given the reduced amount and size of pixels in existing LCD screens, the resulting angular resolution may be greatly limited.
- Compressive displays are multi-layered light field displays based on stacking optics and multiple light modulators. This technique allows a directional resolution greater than the total number of pixels due to the superposition principle, thus copying with the spatio-angular trade-off of integral imaging. However, the computational cost of the modulation patterns for the individual layers is significantly higher; therefore, it is recommended to use these displays particularly for static content.
- Volumetric displays are non-compressive layered displays based on additive layers, resulting in high-quality 3D images, adequate accommodation, and full-parallax cues. However, such solutions are not capable of rendering occlusions properly.
- Projection-based displays incorporate projector arrays and a holographic screen without internal mechanisms. Due to the characteristics of this latter component, implementing larger screens is a more straightforward challenge.

4.5 Content characteristics

Based on the works in the state-of-the-art scientific literature, 2D contents can be classified based on sets of features that are commonly distinguished as low-level (e.g., brightness, color, textures) and high-level features (e.g., a sad person, lesson). When dealing with light field data, such classification is not sufficient. In fact, light field cameras can also record the angle information [B11], thus providing information about depth dependence and Lambertian lighting. Therefore, transparency, reflection, and depth of field (DoF) have to be considered as additional relevant features. For this reason, the content characteristics depend on a number of features. This document addresses these features in the remainder of this clause as well as in [Clause 5](#).

4.5.1 Light field features

A set of image and video attributes is reported in [Table 1](#) and can be categorized into general and light-field-specific ones.

Spatial information (SI): It can be used for estimating the perceptual impact of coding artifacts and transmission impairments. It is typically computed by estimating long edge detection. Separate horizontal and vertical filters are applied, and the total edge energy is computed as Euclidean distance. SI filters are applicable to light field data (see [B80]).

Colorfulness (CF): It is used as a perceptual indicator of the variety and intensity of colors in the content. In the work of Hasler et al. [B38], a CF metric for natural images, based on the pixel distribution of the image in CIE Lab color space, is presented. It is applied to light field in the work of Paudyal et al. [B80] by considering the CF value of the central view.

Temporal perceptual information (TI): It can be used for estimating the amount of temporal changes of a video sequence. It is usually higher for high motion sequences.

Contrast: It conveys meaningful visual information. It has been demonstrated that quality evaluation is a physiological procedure, and the psychological phenomenon called interaction also has a very strong correlation with the contrast [B73]. In addition, contrast is very important textural information. In the context of light field, contrast is usually measured in the central view. However, it could be beneficial to consider the contrast of the entire light field, or the consistency of the central view with the other sub-aperture images, in the case of the evaluation of image enhancement methods based on the selective modification of the contrast (see [B16]).

Depth map: Several approaches have been proposed to estimate the depth information from data captured by different acquisition systems (e.g., camera arrays, plenoptic cameras). Furthermore, multiple approaches should be considered when dealing with sparse (e.g., captured by camera arrays) or dense (e.g., captured by plenoptic cameras) light fields, since the different acquisition technologies may vary greatly in terms of properties, such as baseline and spatial aliasing [B10]. In the case of sparse light fields, depth can be estimated by using traditional multi-view methods [B29], or specific techniques based on plane sweeping [B18] or multi-resolution matching [B24]. Regarding dense light fields acquired by using lenslet-based cameras, several techniques have been proposed to cope with the significantly narrow baseline between sub-aperture images [B4]. According to the adopted light field representation, it is possible to classify these algorithms into the following categories:

- Multi-view stereo methods: These methods are the direct evolution of stereo matching algorithms for depth estimation, considering horizontal, vertical, and diagonal disparity when minimizing the cost volume [B74];
- EPI-based methods: The slope of the epipolar lines in the epipolar plane images (EPIs) can be estimated by shearing the EPI volume [B21], or by means of the first order structure tensor [B116];
- Angular-patches-based methods: The distribution of non-occluded pixels is estimated by analyzing the orientations of discontinuities in the angular patches, enforcing the LF regularity [B94];
- Focal-stack-based methods: Such techniques use depth-from-defocus [B105] and focal-stack-symmetry [B100].

Disparity range: Content characterization [B110] can benefit from absolute disparities in terms of pixels. Depending on the acquisition system, disparity range can change significantly. As an example, the disparity range between sub-aperture images of a light field acquired by using a specific lenslet camera is between $[-1, 1]$ pixels [B126], while it can be much higher for a sparse, wide-baseline setup, such as the one proposed in the work of Sabater et al. [B86]. Information regarding either the depth range in the scene in terms of distances from the nearest and farthest objects, or the camera calibration parameters, is required.

Occlusions: They represent a notable component of the depth estimation for light field contents. Handling occlusions is particularly important for EPI-based methods, where failing to comply causes the incorrect ordering of the depth layers. At the time of this standard, only a few depth estimation algorithms take into

account and model occlusions. The work of Wang et al. [B114] proposes an algorithm that can be applied over the light field data structure obtained by decoding the light field images. The amount of potentially occluded pixels is computed and considered in content characterization.

Refocusing features: One of the main applications of light field imaging is providing the possibility to modify the focus of the elements in an image during post-processing procedures. Refocused images are commonly obtained with the traditional “shift and sum” technique [B65] (i.e., by summing multiple sub-aperture images). The resulting images are limited in terms of depth of field (i.e., they are commonly shallow). Moreover, light field acquisition devices enable overcoming the inherent trade-off between depth-of-field and the light-gathering capabilities of conventional imaging systems by appropriately designing the spatial and angular bandwidth of the capturing device.

Therefore, the availability of appropriate descriptors of the focus properties of the light field can support the characterization of light field content, providing an estimation of the performance in the specific use case. To this aim, the properties and the shape of the disparity histogram—or as in the works of Liu et al. [B64] and Brizzi et al. [B17], a measure of the blur—have been adopted. Other approaches compute the refocusing range of light field images by using the “shift and sum” algorithm. It directly computes the slope parameters used to obtain images refocused on the nearest and the farthest elements of the scene.

Table 1—Key quality attributes

Features	Relevance
Spatial information (SI)	Perceptual indicator of spatial information of the scene
Temporal information (TI)	Perceptual indicator of amount of motion in a scene
Colorfulness (CF)	Perceptual indicator of naturalness of the image
Contrast	Perceptual indicator of the degree of difference between the darker and lighter parts of an image
Depth map	Map containing information related to the distance of the surfaces of scene objects from a viewpoint
Disparity range	Minimum and maximum distance of the surfaces of scene objects from a viewpoint
Occlusion	A part of a scene whose viewing is blocked by objects at a shorter distance from the viewpoint
Refocusing feature	Possibility to modify the focal plane after exposure

4.5.2 Key performance indicators for light field visualization

Many factors may affect the quality of light field visualization in terms of both display and content. These parameters can be considered as key performance indicators (KPIs) since they define the objectively and subjectively measurable visual performance, and any of them with insufficient characteristics may severely degrade the overall quality of light field visualization.

- **Resolution.** For rendering a 3D mesh, rasterization and ray tracing are used. The quality of the output depends on the spatial and angular resolution of the display. Due to the higher computational complexity, ray tracing is used for static content while rasterization is used for real-time contents. Subjective experiments performed in [B56] show that reductions in the spatial resolution may be tolerable while the perceived quality of a one-degree angular resolution is low.
- **View frustum** is the region of space in the rendered world delimited by cutting planes that may appear on the screen. It can be used for describing the space in which the content resides. The cutting planes in the front and back define the depth of the content. The depth value of the frustum is relevant since if too small then the perception of the depth of the content is inaccurate. If the frustum is more narrow than what the projection system requires, the sides of the view become invalid or the content needs to be stretched and cropped, resulting in the loss of visual data and distorted aspect ratio.

- **Scalability** refers to the upscaling and downscaling methodologies for different types of content. Upscaling usually requires interpolation that may introduce artifacts in the rendered data. The visibility of the degradation may be reduced by using a display with high angular resolution (in fact, a much higher angular resolution) and will not necessarily result in the degradation of the subjectively perceived visual quality. It should also be noticed that projection-based light field displays always interpolate the source content during conversion.
- **Frame rate.** Displaying content with a high frame rate requires high computational complexity and memory, which is challenging for several reasons. First, it requires more data to be stored and transmitted. Second, even though the GPUs of the system may have a solidly high frame rate on their outputs, this does not guarantee a sufficiently high final frame rate. This is because the different GPUs may have different workloads due to the diversity of computational requirements based on the content itself; the content may very well be simple from a specific direction while it may be quite complex from another direction. Furthermore, these GPU outputs need to be synchronized; thus, the slowest one may become a bottleneck for the entire visualization process.

4.6 Consideration

In this subclause, several use cases for light field visualization are addressed: medical, industrial, commercial, educational, cultural, and communicational. Each use case has its own peculiarities that must be taken into account in order to exploit the capabilities of this new technology. At the time of this standard, light field technology is still under continuous development and currently only a limited number of light field displays are commercially available. For this reason, the scientific literature studies the theoretical and practical impacts of specific features on the quality and performance of selected use cases and the interconnection between such features.

As an example, the spatial and angular resolution of light field displays are tightly-interconnected features—regardless of use case—that are important and relevant in general. This consideration is valid in the medical use case, in which limited spatial resolution can lead to a masking effect of visual medical indicators. This can prevent medical staff and students from performing an accurate diagnosis, resulting in a false negative or positive judgment. Spatial resolution is similarly relevant in scenarios in which the reproduction of high-frequency signal components (i.e., sharp edges) is needed, such as in archaeological or pictorial reproductions. Analogous issues arise in the case of insufficient angular resolution. This may severely affect the perceived visual quality, thus negatively impacting the selected use cases, such as medical ones. However, insufficient spatial and angular resolution may be less relevant for other use cases, such as certain commercial ones. For example, in the case of cinematography, due to the lack of observer motion in the conventional viewing model of cinemas, disturbances regarding the parallax effect are less apparent. In the case of digital signage, lower spatial resolution can be tolerated if the visualized content contains only large font sizes. In fact, angular resolution is significantly more important for digital signage, since observers are typically moving and the attractive nature of such form of visualization originates from the glasses-free 3D experience itself (i.e., the immersion created by the smoothly changing parallax, correctly in accordance with the viewing angle).

Another example of interconnectivity among features is the one between angular resolution and contrast. The method generally used to verify the contrast of light field displays (i.e., the alternation of thin black and white stripes) can be distorted by an inappropriate angular resolution of the display or content, which can prevent the correct visualization of the stripes (i.e., the adjacent stripes perceptually interfere with each other through the crosstalk effect).

One must note that considerations for use cases may vary not only between specific applications, but between given content as well. For example, in the case of the visualization of an industrial prototype, spatial and angular resolution is less relevant if the displayed entity does not contain highly-detailed portions. On the other hand, if the prototype has small components that are vital to its visual communication, then higher extents of spatial and angular resolution are required.

5. Influencing factors and impairments for the quality of light field imaging

5.1 Introduction

This clause discusses the factors that influence the quality of the light field content. Impairments are categorized based on the source of the influencing factor in the imaging pipeline.

5.2 Acquisition-related factors

A traditional camera captures only a 2D projection of a scene on its sensor. In other words, at a fixed time, the captured image has a fixed image plane and a fixed sensor plane. To capture a 4D light field, more samples in the angular dimension are required. Methodologies and/or hardware to acquire a light field with higher dimensionality shall be divided into three main categories as multiplexed, sequential, and multi-sensor.

5.2.1 Multiplexed acquisition

This approach encodes the light field via a 2D sensor with the help of an array of microlenses in front of the image sensor. Spatially multiplexed cameras are often referred to as plenoptic cameras (or light field cameras). There is a constant trade-off between the spatial and angular resolution for a given sensor size. In other words, increasing spatial resolution results in a sparsely arranged microlens array (thus, lower angular resolution), or increasing the density of the microlens array results in lower spatial resolution. Due to the spatial/angular resolution trade-off, one might require additional processing steps (e.g., spatial/angular super resolution) in order to capture the desirable light field for the task defined. These processing steps often cause impairments as introduced in 5.3.

Another consideration when it comes to spatially multiplexed acquisition is the lower frame rate compared to traditional 2D photography or multi-sensor acquisition methods. Currently available plenoptic cameras are not capable of handling high frame rate acquisition. Therefore, temporal super resolution might be necessary as a processing step and might introduce specific artifacts on the temporal dimension.

Light field acquired with plenoptic cameras often suffers from lack of information on the corner sub-aperture views due to microlens array arrangements.

5.2.2 Sequential acquisition

Time sequential acquisition often includes a robotic arm which governs the movement of a single sensor in a predefined grid. Each point on the grid corresponds to a sub-aperture view. This methodology allows to capture a light field with high spatial and angular resolution as well as a flexible baseline. Therefore, it does not suffer the same artifacts with spatial multiplexing acquisition. On the other hand, sequential acquisition is limited to static shots. Dynamic scenes cannot be captured with this approach. Ghosting artifacts might occur on the acquired light field due to time sequential capturing.

5.2.3 Multi-sensor acquisition

The multi-sensor capture array requires a set of arrays ordered on a planar or spherical surface. Sensors are used simultaneously to capture a light field sampled from different points in the space. Since each viewpoint requires an individual sensor, setup is generally expensive and not portable. Although this methodology does not lead to any specific impairments, low baseline values are hard to achieve and limited by the individual camera sizes.