SMPTE ENGINEERING GUIDELINE

Stereoscopic Distribution Master — Glossary

Table of Contents

Approved
July 29, 2016

Foreword

SMPTE (the Society of Motion Picture and Television Engineers) is an internationally-recognized standards developing organization. Headquartered and incorporated in the United States of America, SMPTE has members in over 80 countries on six continents. SMPTE's Engineering Documents, including Standards, Recommended Practices, and Engineering Guidelines, are prepared by SMPTE's Technology Committees. Participation in these Committees is open to all with a bona fide interest in their work. SMPTE cooperates closely with other standards-developing organizations, including ISO, IEC and ITU.

SMPTE Engineering Documents are drafted in accordance with the rules given in its Standards Operations Manual.

SMPTE EG 2061 was prepared by Technology Committee 30MR.

Introduction

This section is entirely informative and does not form an integral part of this Engineering Document.

SMPTE created a Task Force on 3D to the Home which began its work in August 2008 and published its recommendations in March 2009. SMPTE committees performed further work on specific outputs of this Task Force, ultimately resulting in creation of a Glossary of Terms applying to Stereoscopic television. These committees requested creation of a standalone Glossary document for the benefit of the SMPTE community.

1 Scope

This document provides a centralized Glossary of Terms related to the distribution and display of Stereoscopic content for the home.

2 Conformance Notation

This Engineering Guideline is purely informative and meant to provide tutorial information to the industry. It does not impose Conformance Requirements and avoids the use of Conformance Notation.

Engineering Guidelines frequently provide tutorial information about a Standard or Recommended Practice and when this is the case, the user should rely on the Standards and Recommended Practices referenced for interoperability information.

3 Glossary

To assist with comprehension and consistency of documents relating to Stereoscopic television, the following definitions are provided. Some of the terms have been adopted from previously published glossaries and documents. Other terms have been generated in order to be specifically unambiguous, especially in cases where ambiguous terms have been popularly used with multiple definitions.

3.1 Basic Concepts

2D Extraction: A process of deriving a 2D image from a Stereoscopic source.

2D Overlay: A single instance of a two-dimensional closed caption, subtitle, or graphics overlay that contains no information embedded in it regarding depth. See also: **3D Overlay**, **Overlay Essence Track**, **Overlay Offset Value**.

2D Overlay With Offset: A 2D Overlay with Overlay Offset information. See also: **2D Overlay**, **Overlay Essence Track**, **Overlay Offset Value.**

3D-Compatible or 3D-Compliant Device: A device (such as a receiver or set-top box (STB)) that can identify Stereoscopic 3D content and pass-through that Stereoscopic 3D content data stream (without change) to downstream devices**.**

3D-Enabled System: A system that can decode standardized 3D distribution formats and properly display them in 3D. This system can include a display along with additional devices (receiver, STB, and/or IR transmitter), or be integrated into a single display. A 3D-Compliant Device, a 3D-Ready Display, and additional devices (such as a Stereoscopic selection device) can combine to form a 3D-Enabled System.

3D-Exclusive Content: 3D TV content designed or intended only to be viewed in 3D, and not designed or intended to be converted to 2D for 2D display. Examples of 3D-Exclusive content might include movies, concerts or other productions that are produced specifically for 3D viewing that are either not made available for 2D or that require separate mastering for a 2D viewing version.

3D-Ready Display: A 3D display that requires the addition of a decoder before it can properly display Stereoscopic 3D content encoded using a standardized Stereoscopic distribution format**.**

3DHM: An acronym for 3D Home Master**.**

3D Distribution Data: Compressed and possibly encrypted data file or stream derived from the Stereoscopic Distribution Master used for actual distribution to the home**.**

3D Home Master: An uncompressed and unencrypted image format or file package derived from a 3D Source Master

3D Overlay: A single instance of a closed caption, subtitle, or graphics overlay that includes either**:**

> **3D Overlay Image Pair**: A Stereoscopic 3D left eye/right eye image pair; or
>
> **3D Image Content**: Image content which results in two or more simultaneous views of a scene to create the perception of depth at the display**.**

3D Source Master: The digital master of a Stereoscopic feature.

Absolute Parallax Transition Limit: The maximum rate of parallax change (dZ/dt) that allows a viewer to fuse views into a Stereoscopic 3D image.

Accommodation: The ability of the eyes' lenses to change shape in order to focus.

Accommodation/Vergence Relationship: The relationship between the focusing of the eyes and verging of the eyes when looking at a particular object point in the visual world. When accommodation and vergence match, the time required to fuse Stereoscopic stimulus is reduced, distortions in perceived depth are reduced, and viewer fatigue and discomfort are reduced. Also called the accommodation convergence relationship, or the convergence accommodation relationship.

Anaglyph: A Stereoscopic moving or still image where each view has been superimposed with a complimentary color that produces a 3D effect when viewed through similarly colored glasses.

Binocular Symmetries: Identical visual parameters between a stereo pair of left and right eye images with the sole exception of horizontal disparity (parallax). The symmetrical parameters consist of illumination, focus, geometry, color, registration and temporal symmetry.

Blur Gradient: In real scenes, the retinal image is sharpest for objects in focus and blurred for objects not in focus. In many Stereoscopic 3D display systems with flat screens, the entire image on the screen is in focus at all times, reducing the Blur Gradient depth cue.

Camera: As defined here for this document includes the lens and all associated equipment required to create digital representations of motion picture images. The camera is generally capable of stand alone operation.

Camera Parameters: These parameters completely determine the transformation between 3D coordinates in the world reference frame of an object being imaged by a camera and the 2D image plane coordinates of the resulting image.

> **Intrinsic Camera Parameters**: A set of parameters that relate the geometry of ideal perspective projection to the physics of the camera. These parameters consist of the camera constant f (the distance between the image plane and the lens), the principal point which is the intersection point of the optic axis with the image plane in the measurement reference plane located on the image plane, the geometric distortion characteristics of the lens, and the horizontal and vertical scale factors, i.e., distances between rows and between columns. In the field of Analytic Photogrammetry, this set of parameters is also called the Interior Orientation or the Inner Orientation of the camera.
>
> **Extrinsic Camera Parameters**: A set of parameters that determine the pose of the camera in the 3D world reference frame. The parameters determine the rigid body transformation between the world coordinate frame and camera-centered 3D coordinate frame. In the field of Analytic Photogrammetry, this set of parameters is also called the Exterior Orientation of the camera.

Camera Systems: As defined here for this document consists of all camera's and associated equipment to create S3D content or multi-view content. In the context of this document a camera system is also considered an imaging system.

Coincident In Time: With respect to dual video signals for Stereoscopic television, this means that not only are the two video signals genlocked, but that they represent the same moment in time for the scene displayed.

Color Rivalry: A condition where the color information of homologous points creates a depth cue that is inconsistent with other depth cues, notably parallax. Note that this does not imply that the color information of multiple points must be the same.

Compatible 3D Content: Stereoscopic 3D TV content which is designed to include or enable conversion to and viewing of the content in 2D. Users with 3D-Enabled equipment supporting the distribution format will display compatible Stereoscopic 3D Content in 3D, while those without such equipment will still enjoy the same program as a 2D representation.

Compression: A lossless or lossy process on Stereoscopic 3D content to reduce the size of the image data in order to be able to store or transmit data in an efficient manner**.**

Compression Encoding: The process of converting data into a specialized format for efficient transmission or storage.

Convergence: The inward rotation of the eyes, in the horizontal direction, producing fusion. The more general term is vergence, which includes inward and outward rotation. Convergence is also sometimes used to also refer to the toe-in of stereo cameras.

Corresponding Points, Conjugate Points, or Homologous Points: The points imaged in the left and right eyes (or the points in the left and right Stereoscopic image channels) referring to the same point on the depicted object**.**

Cue Conflict: A condition that occurs when multiple depth cues indicate different depth ordering (e.g., when monocular depth cues such as size difference conflict with binocular depth cues such as parallax)**.**

Depth Map: A Depth Map is an array containing multiple Depth Values. A Depth Map corresponds to an image from a particular viewpoint and each Depth Value corresponds to a pixel in that image. A Depth Map shall have one Depth Value for each pixel in the corresponding image. No additional metadata is required for interpretation of a Depth Map.

Depth Range: The range of distances in camera space from the background point producing maximum acceptable positive parallax to the foreground point producing maximum acceptable negative parallax. This is also known as the depth budget, Maximum Allowable On-Film Deviation (MAOFD), or Fusible Range. It can be given as a ratio of screen width. When referring to the related parallax, the term parallax budget is used.

Depth Value: Depth Value shall represent the Physical Depth, *Z*, in the IEEE 754-2008 single-precision (32-bit) binary floating-point format.

Diplopia: also called 'Double vision', which in stereo viewing is a condition where the left and right homologues in a stereogram remain separate instead of being fused into a single image.

Disparity: The distance between conjugate points on a stereo imaging device or on recorded images. On human retinae, this is known as retinal disparity. The corresponding term for the display screen is parallax.

Disparity Map: An array representing multiple Disparity Values corresponding to a particular one of the images in the stereoscopic image pair.

Disparity Units (DU): An image-width referenced measure used to represent disparities in a stereoscopic image pair.

Disparity Value: the horizontal displacement between corresponding points in a stereoscopic image pair, measured in Disparity Units, from a point in the left-eye image plane to the corresponding point in the right-eye image plane.

Display Parallax Transition Limit: The maximum rate of parallax change for a particular display system that allows a viewer to fuse views into a 3D image. Note that this takes into account display characteristics, such a refresh rate, update frequency, image display multipliers, etc.

Edge Violation: When objects in front of the screen plane (negative parallax) intersect the edge of frame. This can lead to contradictory depth cues for the viewer. Causing an edge violation is sometimes called "Breaking the Frame". This problem can be reduced through the use of a Floating Window.

Exterior Orientation: In the field of Analytic Photogrammetry, Exterior Orientation refers to an image's position and orientation related to an exterior coordinate system.

Floating Window: A black or partially transparent mask that can be placed on the left of the left image and/or on the right of the right image in order to reduce the depth cue confusion due to edge violations. Floating Windows create the illusion of moving the physical borders of the screen surround forward or backward. A floating window can be a Dynamic Window by changing window depth or tilting the window at different times during playout, and can either be pre-rendered into the image or generated by the display using metadata.

Focal Volume: The volumetric space in a Stereoscopic 3D image that is in focus. This is the focal plane multiplied by the depth of field.

Formatting: A process for converting Stereoscopic content into another format/representation. It could involve spatial and/or temporal decimation, interpretation, extraction, etc., to create a new image structure with quality that is equal to or less than the original Stereoscopic content.

Fusion: The combination, by the mind, of the images seen by the left and right eyes into a single image.

Genlock: Abbreviation of "[sync] Generator Lock." Genlock is a technique for locking a device's internal sync structures (and thus image structures) to a common external reference (a "sync generator"). This is especially important for S3D imaging (camera) systems where both eye images need to be aligned precisely both vertically and horizontally in respect to each other.

Ghosting: Crosstalk between Stereoscopic image channels. This can impair Stereoscopic depth cues. Sometimes this is due to imperfect selection devices.

Horizontal Image Translation (HIT): Moving left and right eye images with respect to each other to adjust the zero parallax position for image points.

Horopter: The locus in space of zero binocular retinal disparity. The theoretical horopter in the horizontal plane (if eye fixation is directly ahead) is known as the Vieth-Müller Circle.

Image Formatting: A Generic term for the processing or transforming of image data from one format to another.

Interaxial: The distance between the lens axes of stereo camera heads, also denoted as t_c.

Interocular or **Interpupillary**: The distance between the eyes, also denoted as *t.*

Inversion: The visual effect achieved when the planes of depth in a Stereoscopic display are seen in reverse order; e.g., when the left-hand image is seen by the right eye, and vice-versa, often referred to as pseudostereo.

Lenticular: A system of many small lenses. Some Autostereoscopic displays use a fixed lenticular optical element to steer the images of pixels in order to create an appropriate light field for Stereoscopic viewing.

Light Field: The radiance as a function of position and direction, in regions of space free of occluders (free space). In free space, the light field is a 4-dimensional function.

Luminance Rivalry: A condition where the perceived brightness of homologous points in unbalanced (e.g., brighter in one eye than the other; common in many anaglyph processes).

Overlay Essence Track: A sequence of 2D Overlays or 3D Overlays. See also: **2D Overlay**, **3D Overlay**, **Overlay Offset Value**.

Overlay Offset Value: The horizontal distance (disparity) between the same point in the left eye and right eye presentations of a 2D Overlay. See also: **2D Overlay**, **3D Overlay**, **Overlay Essence Track**.

Panum's Fusion Area: The maximum retinal disparity angle that will allow for fusing together of left and right eye images, typically +/- 15 to 30 arc minutes.

Parallax: The distance between corresponding points in two Stereoscopic images as displayed. It can be measured with a ruler (especially in the case of a Stereoscopic display) or in terms of angular measure. In the latter case the parallax angle directly provides information about disparity.

> Positive parallax appears to be within a screen or behind the surface of a screen (e.g. in screen space). On Stereoscopic screens, the left eye image is to the left of the right eye image.
>
> Zero parallax appears to be in the plane of the screen. On Stereoscopic screens, the left eye image and right eye images are coincident.
>
> Negative parallax appears to be in front of a screen (e.g. in theater space also known as audience space). On Stereoscopic screens, the left eye image is to the right of the right eye image.

Parallax Barrier: A selection device for some Autostereoscopic displays that blocks the view of some pixels from a particular eye.

Percival's Zone of Comfort: The approximate range of vergence and accommodation responses for which the viewer can fuse left and right eye images without discomfort. This region is smaller than Panum's Fusion Area.

Physical Depth: The Physical Depth is the distance from the Reference Camera to a point on the surface of an object imaged by the camera. The distance is measured along the optical axis of the Reference Camera from the front of the Reference Camera lens to the point closest to the surface point.

Production Aperture: The production aperture is the image lattice that represents the maximum possible active image area in a given image format.

Quincunx: A term for the "checkerboard" pattern where left eye and right eye pixels are carried in the positions of the "white" and "black" squares of the checkerboard. The term Quincunx comes from an X-shaped pattern of five dots found on the five twelfths (quinque + uncia) of the "as" Roman bronze coin.

Reference Camera: The Reference Camera is the camera that corresponds to the particular viewpoint of the image corresponding to the Depth Map. The Reference Camera may be virtual, that is, the Reference Camera might not match a camera used for capture.

S3D: Acronym for Stereoscopic 3D.

S3D Imaging Systems: All camera systems and associated real-time equipment necessary to create S3D content or multi-view content.

Selection Device: The mechanism in a Stereoscopic display for providing a left image to the left eye and a right image to the right eye. In Stereoscopic displays, the selection device is a kind of glasses, in Autostereoscopic displays the selection device helps to produce a non-uniform light field (such as lenticular optical elements or parallax barriers).

Stereopsis: The binocular depth sense, from a Greek word meaning "solid seeing".

Stereoscopic: Relating to the use of binocular vision to create the perception of depth in an image.

Stereoscopic 3D: See **Stereoscopic**.

Stereoscopic Distribution Master: An unencrypted image format or file package derived from a 3D Source Master. The Stereoscopic Distribution Master is intended to be used in the creation of Stereoscopic Distribution Data.

Stereoscopic Image Pair: A pair consisting of a left eye image and the corresponding right eye image.

Super-Multiview: A multiview display with more than two view images passing through each pupil of a viewer's eye. This can lead to the reduction of accommodation/vergence mismatch because multiple views are incident to each eye with the parallax angle corresponding to the focus accommodation.

Synthesized From 2D: Content that originated in a 2D format, and was converted into 3D.

Temporal Alignment: With respect to the capture mechanisms of digital motion picture cameras, the capture sensor systems need to be synchronized in order to maintain proper representation of motion when using more than one sensor or device to capture S3D or multi-view images for presentation.

Toe-in: The angle between the lens axes of stereo camera heads in a side-by-side rig**.**

Track: With respect to captions, subtitles or graphics overlay elements — a particular collection of text or graphical information with a particular time line relating the information to the time line of the Stereoscopic program. There can be more than one track in a Stereoscopic Distribution Master. During a specific distribution event, none, one or more of these tracks can be put to use.

Vertical Misalignment: Vertical deviation between corresponding points, also known as Vertical Alignment Error.

Volumetric: Two images with spatially varying disparity, resulting in the illusion of volume.

Z Distortion: Geometrical mismatch between two or more views, commonly caused by asymmetrical toe-in angles, zoom mismatch, camera rotation, etc., causing a warp along the Z-axis most commonly noticeable at the horizontal borders of the image.

3.2 Image Format Representations

2D+Delta: A single image along with data that represents the difference between that image view and a second eye image view along with other additional metadata. The delta data could be spatial temporal stereo disparity, temporal predictive or bidirectional motion compensation.

2D+Depth: A single image along with depth data of a scene, also known as 2D+Z.

2D+DOT: A single image along with depth, occlusion, and transparency data of a scene.

Frame Compatible: A data stream where the left-eye and right-eye images of a stereoscopic image pair are arranged in a spatial or temporal multiplex which results in a single image stream that can be treated like a conventional 2D serial data stream. The frame compatible serial data stream may be progressive, interlaced or progressive segmented frame in accordance with the format of the originating stereoscopic image pair.

Frame Packing: See Frame Compatible

Image: A digitized representation of the 2D optical projection of a 3D scene — comprising a 2D set of image points or pixels. The representation includes scene brightness as a scalar valued function at each point at a minimum and can include color information as a vector valued function at each point.

Image Field: A component of an image frame. One or more image fields comprise an image frame.

Image Frame: A 2D representation of an image resulting from a specific sampling structure. This representation allows the Frame to be divided into multiple Image Fields.

Image Pair: A pair consisting of a left eye image and the corresponding right eye image, also called Stereoscopic Image Pair.

Le/Re: Left eye and right eye images of a scene, also known as stereo images.
Stereoscopic Image Pair: same as Image Pair.

Multiview: More than two images of a scene (not to be confused with MPEG MVC).

Side by Side (SbS): A Frame Compatible format where each image of a Stereoscopic Image Pair is subsampled to fit in half of the horizontal space and then combined within a single frame.

Stereo+2Z: Images obtained and depth maps measured from two separate camera images of a scene.

Stereo+Depth: Le/Re images along with depth data of a scene, also known as Stereo+Z.

Stereoscopic Content: Particular programming that includes a Stereoscopic Image Pair.

Top and Bottom (TaB): A Frame Compatible format where each image of a Stereoscopic Image Pair is subsampled to fit in half of the vertical space and then combined within a single frame.

3.3 Display Types

Autostereoscopic: A display that does not require the viewer to wear any type of glasses in order to view Stereoscopic content.

HMBD: Head Mounted Binocular Display.

Holographic Display: A 3D display that is based on spatial light modulation using a pattern of light wave interference.

Stereoscopic Display Systems: Systems that provide a binocular depth cue, the sense of depth due to different images reaching the left and right eye.