LEXICON

Version 1.0 2017-07-01

VR Industry Forum

5177 Brandin Court

Fremont, CA 94538, USA

www.vr-if.org

Disclaimer

The VR Industry Forum accepts no liability whatsoever for any use of this document.

Copyright Notification

1 Introduction

This lexicon of terms applies to the fields of virtual reality (VR), augmented reality (AR), mixed reality (MR), and 360-degree video.

The list of terms is organized as follows

Term	Word or phrase
Acronym	Abbreviated form of the term (including initialisms).
Category	Classification of the term. (See list below.) This is intended to categorize the list and make it possible to select subsets of the list applicable to different needs. Entries are color-coded by category.
Definition	Meaning and use of the term
Workflow	Categorization of the phase(s) of ecosystem workflow the term is most relevant to. (See list below.)

There are two taxonomies: general category and workflow (usage). The list is color-coded by category.

Category	
Concept	High-level notion or abstract idea.
Audio	Technology or device specifically related to audio.
Camera	Video or audio capture device, or associated optical or aural system.
Display	Technology or device that presents video (and audio) to a user.
Interaction	System or technique enabling a user to interact and control a simulated environment.
Metric	A system, unit, or frame of reference for measurement.
Physiology	A human factor or human response.
Sensor	A device for tracking motion or position. (Excludes devices for capturing video, audio, or morphology.)
Software	Computer application or software library.
Technology	General technical knowledge or its application.
Video	Technology or device specifically related to video.

Workflow	
Capture	Convert an input signal into digital format. Includes cameras, microphones, sensors, and real-time video stitching.
Produce	Combine and manipulate captured and generated media elements into a desired final form. Includes data conversion, post-production, stitching, pre-rendering, 3D mapping, planar projection, editing, and QC
Encode	Convert digital media into a compressed format intended for distribution and playback. Includes transcoding, multiplexing, DRM license generation, and encryption.
Distribute	Deliver encoded digital media to end devices, typically by streaming or downloading over the Internet. Includes storage, CDN, streaming, download, and broadcast.
Decode	Decompress and convert encoded digital media into a form ready for rendering and playback. Includes DRM license verification and decryption.

Render	Generate an audio and video image in real time. Includes compositing from multiple sources, rendering audio and rasterizing video from 2D or 3D models, and extracting an image for a point of view.
Display	Present audio and video using devices such as audio headsets, video screens, and video projection. Includes HMD, HUD, and light-field display.
Interact	Control and affect an audio, visual, and possibly tactile experience using methods such as body motion, speaking, and manipulating input devices. Includes user input, latency
Experience	Perceive and participate in an environment that augments or overrides human senses and responds to human input. Includes physiology, user acceptance, haptics, and other human factors.

1.1 Version History

Date	Version	Notes
2017-07-01	v1	Initial version

Contents

2 Lexicon

Term	Acronym	Category	Definition	Workflow
360° video		Concept	Also known as spherical video, 360° video refers to capturing a very wide field of view (between a hemisphere and a full sphere), usually with multiple lenses whose independent streams are merged through the process of stitching. A key characteristic of 360° video is that it is usually intended to be viewed on a display device such as a tablet or HMD that shows only a subset of the panorama, the selection of which is normally governed by head tracking or device orientation to create an immersive experience. The viewing experience has three degrees of freedom — although the user can control where they look, they have no control over the positioning of the camera. See Panoramic single-view video and Panoramic stereoscopic video.	Capture, Render
3D cursor		Interaction	A direct interaction technique based on a 3D cursor which position is mapped according to the user hand according to a linear or non-linear mapping (go-go technique)	Interact
3D motion tracking		Technology	Process of determining from 2D image sequences the motion of objects or of the camera in the 3D real space. Knowing the pose of objects or of the camera in a given coordinate system of the real space for the first image of the sequence provides with their 3D pose for each image of the sequence (3D pose estimation).	Capture
3D pose estimation		Technology	Determining from a sequence of 2D images acquired by a vision device (camera, depth sensor) the 3D transformations or pose (position and orientation) of real objects or of the camera in a given coordinate system from the real space.	Capture
Accelerometer		Sensor	A device for detecting acceleration. Usually implemented by sensing change in velocity in the x, y, and z directions relative to some frame of reference (e.g., the object itself or a cubic volume in room-scale VR) to generate an acceleration vector. Found in smartphones and HMDs.	Capture, Interact
Accommodation		Physiology	The ability of a person to focus on objects at different distances. This declines with age, the end stage of which is a condition known as presbyopia.	Experience
Ambisonics		Audio	A system for encoding directional information in a 3D sound field by capturing x-, y-, z-, and omnidirectional components using special microphones. This allows a soundfield to localized when rendered through (e.g.) headphones to match the FOV being displayed on an HMD. Developed in the 1970's in the UK.	Encode
Analytics		Technology	In a VR context, analytics is employed to extract insights and construct visualizations such as heat maps, that are based on the input streams received from the user, specifically the position of the HMD, which can be used to infer (approximately) which elements of a VR scene the user is interacting with at a given instant in time.	Interact
Astigmatism		Physiology	A type of optical aberration that prevents a person from keeping the entirety of an object in focus at one time, even if all parts are the same distance away. A common manifestation is only being able to focus on the horizontal or vertical arm of a cross (but not both) at the same time. It is an issue with HMDs that don't work with prescription eyeglasses, as there is no simple correction for astigmatism.	Experience
Asynchronous time warp	ATW	Technology	A technique for masking video artifacts caused when the next frame has not finished being rendered at the end of the current frame. Without	Render

Term	Acronym	Category	Definition	Workflow
			ATW, this will result in the current frame appearing twice, which the user perceives as a video glitch that undermines immersiveness. With ATW, if the next frame isn't ready the current frame will be warped (affine transform) to approximate the next frame (assuming the scene hasn't changed and the head motion is relatively small). This has been shown to mask the users perception of video glitches in many cases, though success is not guaranteed.	
Augmented reality	AR	Concept	The inclusion of synthetic objects into a user's live or indirect real-world environment. This includes a spectrum of capability, progressing from a simple overlay (e.g. in a HUD), to simulating partial occlusion of synthetic and real imagery, to full global illumination (e.g. light reflecting from the synthetic objects affects real objects, and vice versa) and audio integration with sound-cancellation techniques. In its most advanced form, AR is more challenging than VR because of the requirement to seamlessly blend real and synthetic objects.	Render
Autostereoscopic display		Display	The ability to reproduce stereoscopic images without requiring the user to wear special gear.	Display
Avatar		Concept	A representation of a person inside a virtual environment It may or may not be based on how a person actually looks. In many MMORPG's, an avatar is chosen from a list, or is a graphical representation derived from a photograph.	Experience
Beam-splitter		Camera	A system for splitting a beam of light. It can be used to dynamically change the interaxial distance in a system of lenses.	Capture
Binaural microphones		Audio	A pair of microphones to be placed in or near the ears to capture a sound field as perceived by a person.	Capture
Binaural Rendering		Audio	Concept of processing audio signals for headphone reproduction. Binaural rendering mimics the source transmission from a point in space to the ears of users. It can be used for presentation of immersive audio material over headphones maintainung spatial cues. Binaural rendering is part of the MPEG-H 3D Audio standard.	Render
Binocular rivalry		Physiology	An uncomfortable ocular effect that appears when parts of objects in negative parallax are clipped by the edge of the stereoscopic display, causing parts of objects that should be visible by both eyes but that are visible by only one eye due to display edge clipping.	Experience
Bullet time		Concept	Extreme slow motion. The phrase originated from the 1999 film "The Matrix".	Produce
Calibration		Technology		
Camera array		Camera	A set of cameras that are mounted in a harness to assume fixed positions relative to each other. Depending on the camera geometry, the array may be used to capture parallax (stereo), 360° video, or a light field.	Capture
Camera rig		Camera	A piece of equipment to which a camera is attached when used for professional image capture.	Capture
CG VR		Concept		
Chromatic aberration		Physiology	A class of visual artifacts caused by light with longer wavelengths being refracted less strongly than light with shorter wavelengths. HMDs containing lenses will cause severe chromatic aberration unless corrected by applying the inverse amount of aberration to the image displayed.	Experience
Cinematic VR		Concept		Produce

Term	Acronym	Category	Definition	Workflow
Codec		Technology	A contraction of "code" and "decode", codec refers to a pair of complementary transforms applied to a stream of digital data. Codecs can achieve many goals, but the primary one in a VR context is to compress the data. For example, H.264 is a video codec that supports coding an uncompressed video stream for more bit-efficient transport, and decoding it in a playback device for display.	Encode, Decode
Compositing		Technology	The process of combining visual or aural elements from separate sources into a single image or sound field. Generally requires a registration process for good alignment.	Produce
Computer-generated imagery	CGI	Video	A large class of software and hardware systems that support the creation of realistic imagery based on an internal model that captures both the geometry and lighting attributes of an object (that may or may not exist in reality).	Produce, Render
Cube map		Video	A format for 360° video that is based on projecting a spherical image onto the faces of a surrounding cube. This is usually represented as a cross-shape (3x4 or 2x3 rectangles) created by unrolling 4 parallel faces into the same plane, and then folding down the remaining two faces into that plane to form the arms of the cross. Compared to Equirectangular, cube maps exhibit less distortion but are more complex to store and render owning to the non-rectangular shape and discontinuities at the edges.	Produce
CUDA		Technology	Originally an acronym for (Compute Unified Device Architecture), CUDA is an API that abstracts the capabilities of parallel computing hardwware such as GPUs. This simplifies the development of algorithms and provides a degree of device independence.	Render
Data glove		Sensor	An input device that is worn like a glove and records the relative positions of the fingers.	Interact
Degree of freedom	DOF	Metric	Refers to the freedom of movement of a rigid body in a three dimensional space. It is categorized by the independent translation and rotation of the x-, y-, and z-axes. Thus the most flexible system is often referred to as six degrees of freedom. This term is commonly applied to camera mounting systems, e.g. on drones.	Capture, Render
Depth image-based rendering	DIBR	Technology	A family of techniques for synthesizing a view of a scene (i.e. from a position not captured by a camera) based on one or more reference images that include depth information for each pixel site. A central problem of DIBR is that of disocclusion: the synthesized image may include regions now exposed to view that were not captured in the reference image(s). A variety of methods exist for estimating pixel values for those regions based on localized information and the depth values.	Render
Depth of field	DOF	Metric	In an optical system, the range between the nearest and farthest objects that are in focus. The Depth of Field is positively correlated with the f-stop (higher f-stop results in greater DOF).	Capture, Render
Diegetic Sound		Concept	Sound elements that are understood to originate from within a scene being depicted in a cinematic, theatrical, or VR experience. They may be off-screen, but if so would be audible to the characters in the scene. Dialog spoken by a character would be an example of Diegetic Sound. Conversely, sound track music heard only by the audience would be an example of non-Diegetic sound. Also known as "Diegetic Audio" or "Source Music".	Produce
Diopter		Physiology	A measurement of how strongly a lens bends light.	Capture

Term	Acronym	Category	Definition	Workflow
Direct Mode		Technology	Driver technology that allows an HMD to be treated as an independent display that VR applications can render directly to. Normally all displays attached to a PC are assumed to be part of a larger virtual display; this is undesirable because having the window system involved is a source of avoidable latency.	Display
Direct3D	D3D	Software	A low-level 3D graphics rendering library, created by Microsoft for Windows platforms.	Render
Dome cinemas		Display	A system for projecting video onto a spherical or hemispherical screen that is viewed from the inside. It is suitable for projecting 360° video but has no motion tracking capability.	Display
Double buffering		Technology	A technique for masking video artifacts (e.g. tearing) by supporting two independent frame buffers. One contains the currently displayed buffer, and the other (not visible) is drawn into to create the next image. At the end of the current frame time the buffers are swapped and the cycle begins anew.	Display
Equirectangular		Video	A format for 360° video that is based on projecting a spherical image onto a surrounding cylinder (similar to a Mercator projection in cartography). The benefit of the format is the cylinder can be represented as a rectangle, allowing reuse of existing rectangle-based video codecs. The principle disadvantage is that more pixels are used to represent polar regions, which typically contain the least interesting content to the viewer (e.g. sky, ground).	Produce
Eye tracking		Technology	(See Gaze tracking)	Interact
f-stop	f/n	Metric	The f-stop is the ratio between the focal length and the diameter of the aperture through which light enters the lens.	Capture
Field of view	FOV	Metric	The angle formed by the left and right edges of the viewing area of an HMD. There is a tradeoff between the Field of View (higher is better) and Resolution (which is inversely proportional).	Display
First-person View	FPV	Concept	A view from the perspective of an individual, with the video and audio representing what a person would experience. Also refers to a method of remote control (e.g. of a vehicle, drone, etc.) where the control interface presents a view from the perspective of the device being controlled, often referred to as a Remote Person View (RPV).	Capture, Display
Focal length		Metric	The distance (measured in millimeters) between the center of a lens and the point at which light entering a lens will converge	Capture
Frame cancellation		Physiology	An uncomfortable ocular effect that appears when objects in negative parallax are clipped by the edge of the stereoscopic display, causing depth perception conflicts between occlusion and parallax cues: Objects are perceived behind the display plane as they are considered as occluded by its edges, while they are perceived in front of the display plane as their parallaxes are negative.	Experience
Frames per second	FPS	Metric	In a video track, the rate at which frames are encoded. In a display system, the rate at which frames of video are presented to the user. The persistence of the video system determines what portion of a frame the image is actually visible. In VR systems, high frame rates (e.g. 90–120 Hz) are required to avoid motion sickness.	Display
Free-viewpoint video		Display	A video that a user may view from an arbitrary position and direction By comparison, in a 360° video only the direction can be selected on playback; the location is determined by the camera at the moment that frame was captured.	Display

Term	Acronym	Category	Definition	Workflow
Gamma correction	γ	Metric	An empirical change to correct for the human visual system's non-linear response to changes in luminance.	Display
Gaze tracking		Technology	Any of a variety of techniques for gathering detailed information on the movement of the eyes of a viewer while in an immersive VR experience. This is not the same as assuming the center of an HMD is always where a user is looking. The most common data collected are the direction vectors for each eye (known as pupillometry or gaze tracking), and small involuntary motions known as saccades. More sophisticated systems also may capture the diameter of the pupil, which can be indicative of ambient lighting or internal emotional state. Current HMDs do not have eye tracking capability, though this may change in the future. A few current HMDs can be retrofitted with eye tracking hardware.	Interact
Global illumination		Video	In 3D graphics, global illumination seeks to realistically light a scene by accounting for not just direct light sources, but indirect lighting caused by objects recursively reflecting light between each other. These calculations are extremely compute intensive, and generally require ray-tracing techniques that can't be done in real-time, though GPUs help a great deal.	Produce
Graphics Pipeline		Technology	A sequence of computational steps to create a raster graphics image from an internal 3D representation, including the following: transformation from local coordinate system to camera(s) coordinate system(s); camera projection of the elements of the 3D scene according to the intrinsic parameter; view frustum clipping; rasterization; texturing; lighting; fragment shading. On recent graphics hardware, some of these steps are implemented by programmable shaders, including: pixel or fragment shaders (relative to the pixel of the output raster image); vertex shaders (to modify properties relative to input 3D models' vertices); geometry shaders (to generate new primitives such as 3D points, lines, triangles, etc); tessellation shaders (to subdivide a mesh to generate on-the-fly levels of detail). User interactions, physics, and management of object behavior are not part of the rendering process. These computations update the scene graph in a pre-rendering step of a global loop. Although the distortion correction adapted to each HMD lens can be done through a dedicated pixel shaders, the time-warp process is done in a post-rendering step.	Render
Graphics processing unit	GPU	Technology	A collection of processors that can speed up algorithms by performing computation in parallel. The computations are not independent but conform loosely to the SIMD (single-instruction multiple data) paradigm. Typically the same program (set of instructions) is performed independently by each GPU core, but the input parameters can vary programmatically by a series of iterators which can be used to index tabular arrays. GPUs were originally developed in 3D graphics accelerators to perform shading (per-pixel) calculations, but since then have evolved to be usable with a larger class of applications. GPUs are sometimes referred to as GPGPUs (general-purpose GPUs) when applied to non-graphics computations.	Render
Gyroscope		Sensor	A class of device capable of detecting and measuring changes in the orientation (roll, pitch, and yaw) of an object. Originally implemented as a rapidly spinning disk attached to a cage with three degrees of freedom of motion, new versions use long fiber loops and have no moving parts.	Interact
Haptic		Concept	Literally, relating to the sense of touch. In a VR context. Refers to a class	Interact

Term	Acronym	Category	Definition	Workflow
			of input controllers that can provide programmatic feedback to a user, usually based on the location of the controller in the scene. Haptic feedback may be realistic (e.g. to simulate the physical effect of interacting with an object) or just to provide cues to the user that they should perform (or stop performing) an action when certain criteria are met.	
Head tracking		Sensor	A system for capturing the orientation of an HMD. This is not the same as gaze tracking. The HMD direction is needed to determine the viewport into a spherical image to display. In walk-around systems it is necessary to determine the location of the HMD as well as its orientation (see Positional head tracking).	Interact
Head-mounted display	HMD	Display	A class of display device characterized by: high frame rate; high-resolution display; a lens system to make the display viewable to a person; and a head tracking system to report back the current orientation of the display.	Display
Head-related transfer function	HRTF	Audio	The HRTF is measured at a spot near the entrance of each person's ear. It is parameterized by a direction (e.g. latitude-longitude) and frequency; the value is the actual intensity of the sound received. The HRTF is required to produce binaural (spatially accurate) sound.	Produce, Render
Heads-up display	HUD	Display	A class of display device capable of overlaying text and graphical information in the user's field of view. Google Glass is an example.	Display
High dynamic range	HDR	Video	The ability to represent a large variation in the lightest and darkest areas of an image. Depending on context, HDR may refer to the capabilities of the image sensor, of the number bits used to represent intensity values, of functions mapping pixel values to intensity, or of the range of intensities an output device can reproduce. SMPTE has developed a family of HDR specifications; pixels are typically 10 bits or higher.	Capture, Display
High-Efficiency Video Codec	HEVC	Video	A video compression standard that offers better coding efficiency compared to MPEG-2 or H.264. Also known as H.265.	Encode
High-order ambisonics	HOA	Audio	An enhancement to ambisonics that improves the size of the region over which a soundfield can be reproduced, but at the cost of encoding additional channels, which in turn require a larger array of microphones to capture.	Encode
Holo-cinema		Concept	A term coined by ILMxLAB to refer to a shared VR experience where multiple users are immersed into the same environment (often based on a movie franchise), are aware of each other, and can interact with the environment.	Display
Holography		Concept	The science of recording and displaying light fields. A holographic capture differs from a conventional photograph in that it captures an entire light field. In practice, this means recording not just a single color at a pixel site, but a set of values corresponding incident light (color & intensity) from every direction. Holographic display means reconstructing the light field at the wavefront level; a subject moving while viewing a hologram will experience shifts in parallax (and occlusion changes) as if they had witnessed the original scene. Holography requires capturing an order of magnitude more data than conventional photography, and the field is still in its infancy. Traditional holography is based on recording laser interference patterns on photographic film. More recently, so-called light field cameras use per-pixel image sensor lens to record incident light at multiple angles. These images can be manipulated computationally to apply effects such as re-focusing in post	Display

Term	Acronym	Category	Definition	Workflow
			processing. Holographic videography is in its infancy, and most displays (e.g. Hololens) are not performing wavefront reconstruction but rather using microprisms to inject imagery over the field of view.	
Horopter		Physiology	In stereoscopic vision, the horopter is informally defined as the set of points in a field of view that appear at the same place in the retina of both eyes. This will vary depending on whether a person is fixated on a nearer or farther object. Objects that are closer or further away than the horopter will appear as "double images": the former is known as a crossed disparity and the latter is known as an uncrossed disparity.	Display
Hyperstereo		Metric	Stereoscopic images where the Interaxial distance is larger than the IPD. Makes objects appear smaller than in real-life because of exaggerated parallax.	
Hypostereo		Metric	Stereoscopic images where the Interaxial distance is smaller than the IPD. Makes objects appear larger than in real-life.	
Image-based rendering		Technology	A set of techniques for synthesizing a 3D representation of an object from a series of photographs taken at different angles.	Produce
Immersive Audio		Audio	Concept and format for audio with the goal to capture, mix and reproduce 3D Audio, i.e. sound from all directions around a user. Immersive audio can be formatted as channels e.g. 7.1+4, scene-based audio (HOA), object-based audio, as well as combinations thereof. MPEG-H 3D Audio has been designed to these formats and their rendering for VR applications.	Capture, Produce, Render
Immersiveness		Metric	See "Presence".	Experience
Interaxial distance		Camera	The distance between two lenses in a multi-camera array. The interaxial distance determines the amount of parallax captured by this camera pair.	Capture
Interocular distance	IPD	Physiology	Synonym for Interpupillary Distance.	Experience
Interpupillary distance	IPD	Physiology	The distance between the pupils of a person's eyes. This in turn affects the amount of parallax they perceive. A large discrepancy between the parallax in stereoscopic 360° video and a person's IPD may result in a degraded VR experience.	Experience
Judder		Concept	A temporal aliasing artifact caused when converting an animated sequence from one frame rate to another that is not an integral multiple.	Display
Latency		Metric	In any system characterized by a processing pipeline, latency measures the time elapsed between entry to and exit from that pipeline. In a VR display system, minimizing latency between user action and video audio presentation is essential for preventing motion sickness. (See motion-to-photon.)	Interact
Light detection and ranging	LIDAR	Sensor	A device for creating a depth map by timing the round-trip time of a laser pointed in a specific direction.	Capture
Light field		Technology	Conceptually, a vector function describing the light flowing in every direction through every point in space. In photography and VR, a representation of an image from multiple points of view. A light-field camera (also known as a plenoptic camera) captures information about the light field emanating from a scene, including the intensity of light and the direction that the light rays are traveling in space. Light-field technology allows the point of focus and depth of field in a photograph to be selected by the viewer after a scene has been recorded.	Capture
Light-field camera		Camera	Generically, any system for capturing a light field. A synonym for	Capture

Term	Acronym	Category	Definition	Workflow
			plenoptic camera.	
Light-field display		Display	A light-field display fully reproduces parallax and occlusions from any position, with positioning constraints being nearly imperceptible.	Display
Localization		Concept	Determining the position in 3D space of real objects such as a camera for augmented reality purpose. This localization can provide the position only or the pose (position and orientation) of objects. For example, 3D pose estimation, infrared tracking device, and GPS are solutions to the localization problem providing different precisions.	Capture
Location-based entertainment	LBE	Technology	A class of VR experience —usually employing advanced hardware such as haptics— that occurs at a fixed venue outside the home (e.g. a theme park or cinema multiplex).	Experience
Magnetometer		Sensor	A device to measure orientation relative to a magnetic field. In conjunction with an accelerometer, this can be used for more accurate head tracking. Many smartphones contain magnetometers.	Interact
Mapping		Technology	1) Transforming image data from a planar surface to a two-dimensional plane or three-dimensional object. See Projection. 2) Scanning a 3D object to capture its shape and optical characteristics.	Capture, Produce
Markerless		Technology	In VR, refers to camera-based motion tracking systems that do not require the presence of distinguished shapes (markers) in the scene to establish the orientation of the camera.	Interact
Massively multiplayer online role-playing game	MMORPG	Concept	A category of multi-player game, usually featuring 3D graphics, that can support potentially millions of players world-wide. The games are hosted on a network of servers, which permits players to interact in a shared world. None of the major MMORPG's have announced support for HMDs, but this is anticipated, and some players have already modified existing games to support this.	Interact
Mixed reality	MR	Concept	A hybrid environment that contains elements of both VR and AR.	
Model-based 3D		Technology	An abstract representation of the salient characteristics of 3D objects. Currently, polygon meshes are most common representation of geometry. Associating attributes with each polygon such as texture and environment maps allows the object to be rendered realistically.	Produce, Render
Motion capture		Sensor	The ability to capture the 3D coordinates of an object (or constituent parts thereof) in real time and at high resolution over a time interval.	Capture
Motion sickness		Physiology	A feeling of discomfort (possibly intense) that arises from viewing a VR experience. A major cause is failure of sensory fusion caused by discontinuities between the senses (e.g. long latency responding to head movement). Another cause is content specific (abrupt movements, excessive parallax).	Experience
Motion-to-photon		Metric	A specialized latency metric that captures the time elapsed from the time a person wearing a HMD moves their head to when the first frame of video is displayed with the new field of view. A low value is essential for avoiding motion sickness.	Experience
Multi-view display		Display	A kind of autostereoscopic display, featuring tens of views. Because of the coarse spacing, a user has to position themselves carefully to obtain an acceptable view.	Display
Multi-view video		Display	A system where captured video, upon play back, can be viewed from multiple angles.	Display
Nadir		Metric	The point on a celestial sphere antipodal to the Zenith. More generally, the point (or direction) represented by -90° ("South") in any spherical coordinate system. See also "Zenith".	Capture, Display

Term	Acronym	Category	Definition	Workflow
Negative parallax		Metric	In a stereoscopic display system, an object that appears closer to the user than a reference plane (e.g. a screen in a 3D theater) exhibits negative parallax. It results in eyes pointing inwards, and excessive amounts can cause eyestrain.	Display
Omnidirectional video		Concept	See 360° video.	
Open Source Virtual Reality	OSVR	Software	An open source, cross-platform software framework intended to support HMDs and controllers from multiple vendors. OSVR currently is hosted on Windows, OS/X, Android, and Linux platforms. Separate from the software, there is also a hardware development kit for HMDs known as the Hacker Development Kit.	
OpenCL		Software	A cross-platform library for programming GPUs.	Render
OpenGL		Software	A low-level cross-platform 3D graphics rendering library. It is the primary 3D API on Android and iOS smartphones, and on OS/X (there is also a Windows implementation).	Render
OpenVX: Software		Technology	A cross-platform acceleration of computer vision applications for face and gesture tracking, object and scene reconstruction, pose estimation for AR applications, …	
Organic light-emitting diode	OLED	Display	An emissive display technology that is commonly used in smartphones and HMDs. They are characterized by high density, superior contrast, and greater brightness compared to Liquid Crystal displays.	Display
Orthostereo		Metric	Stereoscopic images where 1) Interaxial distance is equal to the IPD; and 2) Camera focal length is equal to that of the human eye. Recreates human depth/scale perception	
Panoramic single-view video		Concept	Capture or display of monoscopic video with a very wide horizontal field of view and low distortion to create an enhanced perception of "being there." Most commonly this is achieved with multiple cameras that have overlapping fields of view that are stitched into a composite image. However single-camera devices such as slit-scan cameras and fisheye lens have also been used to capture panoramas.	Capture, Display
Panoramic stereoscopic video		Concept	Capture or display of stereoscopic video with a very wide horizontal field of view and low distortion to create an enhanced perception of "being there." Most commonly this is achieved with two equal-sized sets of multiple cameras. Each camera in one set is a fixed distance from a corresponding camera in the other (both pointing the same direction), allowing parallax to be directly captured. Cameras within a set have overlapping fields of view that are stitched into a composite image. However other arrangements are possible, e.g. using image-based rendering techniques on a single set of cameras with high overlap to synthesize parallax mathematically.	Capture, Display
Parallax		Metric	The change in relative positions of objects when seen from different locations. The magnitude of displacement is directly proportional to the distance between the measuring locations, and inversely proportional to distance from the observation point, which makes it possible to infer distance. Parallax is the basis for stereoscopic vision, and is also used in some autofocus systems. Measuring parallax is the basis of stereoscopic vision, and is used heavily in photogrammetry and image based rendering.	Capture, Display
Persistence		Metric	A measurement (typically in milliseconds) that captures the length of time the pixels of a frame are visible. A low value is essential for avoiding	Display

Term	Acronym	Category	Definition	Workflow
			motion sickness.	
Photogrammetry		Technology	In general, the extraction of geometric information from photographs. Many of the techniques have been around for decades. In the context of VR, is usually refers to either recovering 3D coordinates from a series of overlapping photographs, or the extraction of features from non-overlapping photographs for the purpose of creating high-resolution texture maps. Also referred to as Videogrammetry.	Produce
Photometric corrections		Technology	A variety of techniques for estimating sources of error in a photometric observation (e.g. positional jitter, temperature) and applying correction factors to achieve a more accurate measurement.	Produce
Plato's cave		Concept	The concept that one's perception of reality is limited (distorted) by the senses available to them. Plato describes people confined to a cave and only able to look at a wall illuminated by a fire behind them. Their inferences of reality are based on shadows cast by people and things walking between them and the fire.	Experience
Plenoptic camera		Camera	Generically, any system for capturing a light field. The term refers to the mathematical equation used to define a light field. A synonym for light-field camera.	Capture
Point cloud		Technology	A method of representing an object by a set points in a 3D coordinate system. The attributes of a point capture the visual appearance of the object at that point. Point clouds are commonly created by 3D scanning hardware.	Capture
Polygonal mesh		Technology	A method of representing an object by a set of connected polygons, sharing some edges and vertices. Material properties of the object (such as colors, surface normals, or textures) can be associated with vertices, edges, and faces. Rendering is the process by which polygon meshes are converted to visible pixels in a display.	Render
Positional head tracking		Sensor	Systems that measure the location and orientation of a subject's head. To be usable for VR the measurements must be continuous, in real time, with high accuracy and low latency. Head tracking does not perform eye tracking (where the eyes are looking).	Interact
Positive parallax		Metric	In a stereoscopic display system, an object that appears farther from the user than a reference plane (e.g. a screen in a 3D theater) exhibits positive parallax. It results in eyes pointing outwards, and even small amounts can cause eyestrain.	Display
Presence		Metric	A subjective measurement of the extent to which a subject is unaware of being immersed in a Virtual Reality experience. Nearly every aspect of a VR system contributes to or detracts from the sensation of Presence, including video, audio, latency, ability to interact with the environment. Synonymous with Immersiveness.	Experience
Procedural shader		Technology	A 3D technology that can be thought of as a generalization of texture mapping. Each pixel of a 3D polygon is computed by running a program that draws on a variety of inputs, including but not restricted to a texture map. These calculations are performed by a GPU.	Render
Projection		Concept	Representation of a three-dimensional scene as a two-dimensional image. For example, spherical-to-2D projections such as equirectangular and orthographic are used to transform omnidirectional video into rectangular video.	Produce
Proprioception		Physiology	The ability of a person to sense the relative position of different parts of their body. This is possible (even with the eyes closed) because of	Experience

Term	Acronym	Category	Definition	Workflow
			specialized receptors in the muscles, joints, and tendons.	
Quad buffering		Technology	The use of two double buffers for stereoscopic displays (see double buffering), one for the left eye and the other one for the right eye. The swap command is applied on both pairs of buffer.	Display
Ray casting		Interaction	An indirect interaction technique based on a 3D virtual light ray used to select a virtual object.	Interact
Redirected walking		Interaction	Redirected walking allows users to walk through large-scale Immersive Virtual Environments while physically remaining in a reasonable small workspace thanks to a non-linear mapping between user motion in real space and his/her corresponding motion in virtual space.	Interact
Registration		Technology	Process of aligning different set of data into a unique coordinate system. For example, Image registration is used by augmented reality technology to perceive virtual and real data as co-located.	Produce
Render		Concept	The process of converting between an internal representation of a media essence (e.g. 3D model, sound field) to display hardware (e.g. HMD, speakers). May specifically refer to generating a raster graphics image from 2D and 3D models, textures, light sources, camera views, and more structured in a scene graph defining their relative spatial relationships.	Render
Resolution		Metric	The number of pixels per degree of Field of View. This expresses the level of detail that can be reproduced in a HMD.	Display
Rolling shutter effects		Technology		
Room-scale system		Technology	A VR system allowing people to walk around within a limited area. Also known as "walk-around" VR, such a system must be capable of tracking not only the orientation of the headset but also the location (six coordinates versus three).	Interact
Rotoscoping		Technology	A VFX technique for compositing imagery over existing footage. It is now performed by computers, though originally this was achieved through mechanical means. Rotoscoping is very commonly used in post-production workflows, and in VR, non-linear projections such as equirectangular make rotoscoping more difficult, and as a result new plug-ins are being created.	Produce
Sensory fusion		Physiology	A measure of how closely the human senses are synchronized. Sensory fusion contributes directly to immersiveness of a visual environment. Discrepancies between different senses (e.g. a lag between what is displayed to the eyes in a HMD and the actual position of the head) is a major source of discomfort and motion sickness.	Experience
Simultaneous localization and mapping	SLAM	Technology	Simultaneous localization and mapping combines a 3D motion tracking technique with a structure-from-motion technique by simultaneously updating a 3D map of an unknown environment according to an estimation of a camera pose itself estimated according to this 3D map (chicken-and-egg problem). This technique is widely used by augmented reality SDK for the so-called markerless registration, but this loop has the disadvantage of drifting over time.	Capture
Six degrees of freedom	6DoF	Concept	Ability of a viewpoint to change orientation by rotating through three perpendicular axes, often termed pitch, yaw, and roll, and also change position by moving on three perpendicular axes forward/backward (surge), up/down (heave), left/right (sway).	Display
Stabilization		Technology	Techniques for reducing loss of resolution in imaging because of camera	Capture

Term	Acronym	Category	Definition	Workflow
			motion. These can generally be divided into techniques that counteract motion by physically changing the image path to the sensor (optical stabilization), and techniques that counteract motion artifact after the image has been captured (digital stabilization).	
Stereoscopic		Physiology	A stereoscopic display system has separate video for each eye in order to create a system of parallax. Many 360° video cameras are monoscopic (each eye is presented with the same video). Stereoscopic capture may be achieved by having two cameras for each viewpoint, or may be synthesized computationally with certain single camera geometries provided every point in the scene is captured by at least two cameras. Matching the parallax of the capture system (i.e. interaxial distance) with human IPD is important for a good VR experience.	Display
Stitching		Technology	The process of transforming a set of overlapping videos into a single unified sequence. Stitching is very compute-intensive, and if not done carefully can introduce artifacts such as visible seams, or misalignment of objects caused by misidentification of overlapping features.	Produce
Structure from motion		Technology	Process of estimating 3D structures from 3D images sequences. These 3D structures can be sparse (features of interest) or dense (surfaces of real objects). In case of a dense structure, we will talk about 3D reconstruction.	Capture
Super-multi-view display		Display	A multi-view display characterized by hundreds of views, which makes the user less aware of having to position themselves correctly.	Display
Tactical haptics reactive grip		Sensor	A motion controller that employs haptic feedback that can be used to create the sensation of resistance to motion to guide user actions.	Experience
Telenaut		Concept	NASA terminology for a VR system that immerses a person in the environment of another planet.	
Telepresence VR		Concept	A class of VR experience that places a person (or their avatar) into a different location in real-time. It includes the ability to interact with the remote environment.	Experience
Tethered HMD		Display	A HMD that is physically connected via cables to a base station (usually a PC).	Display
TetraMic		Audio	An Ambisonics sound field capture microphone array featuring four microphones in a tetrahedral arrangement.	Capture
Texture mapping		Technology	A 3D technology for mapping an image (texture) or subset thereof onto the surface a polygon (usually a triangle), and dealing with properly handling rendering it (e.g. perspective correction).	Produce, Render
Three degrees of freedom	3DoF	Concept	Ability of a viewpoint to change orientation by rotating through three perpendicular axes, often termed pitch, yaw, and roll, but not change position on those axes.	Display
Tone mapping		Video	Techniques for mapping high dynamic-range (and/or wide color-gamut) images to an output device with more limited capabilities.	Display
Uncanny valley		Physiology	The perception (by a human) that a rendered CGI object (especially another human) is "unreal" as its quality approaches photographic realism.	Experience
Vergence		Physiology	The movement of both eyes to track an object. The result is the object is centered on each retina, and as a corollary each eye is looking in a different direction (e.g. left eye looking rightward and right eye leftward).	Experience
Vertigo		Physiology	The sensation of falling caused by either a disruption of the vestibular system or discrepancies between the vestibular and visual/audio senses.	Experience

Term	Acronym	Category	Definition	Workflow
			Strongly correlated with motion sickness.	
Vestibular		Physiology	A specialized part of the human auditory system that captures the sense of balance.	Experience
Virtual reality	VR	Concept	New combined definition: A rendered environment overriding human senses (visual, acoustic, tactile, and other) with captured or synthetically generated data, providing an immersive experience to a user who can interact with it in a seemingly real or physical way using special electronic equipment (e.g. head-mounted display, audio rendering, and sensors/actuators). Also called immersive media.	
Vision processing unit		Technology	A class of microprocessor designed to accelerate machine vision task (called Holographic Processing Unit by Microsoft)	Capture
Visual effects	VFX	Technology	Any of a large variety of techniques (including animation, matting, compositing) for creating imagery that appears realistic despite incorporating synthetic elements. Because of advances in computer-based tools, most contemporary movies incorporate VFX to some extent, not just the expected genres such as Action or Science Fiction.	Produce
Volumetric capture		Technology	A technique that records a subject with multiple cameras (image and depth data) positioned at different perspectives. Image-based rendering techniques are then applied to the videos to construct a high resolution 3D mesh. Each constituent polygon of the mesh is associated with a texture map corresponding to the captured pixel values. This allows an object (or person) to be placed inside an immersive VR environment.	capture
Walk-around system		Interaction	A system that allows a user to move around within a confined space, usually rectangular. Walk-around capability requires measuring the six-tuple position and orientation of the user's head. An example is the HTC Vive. Walk-around capability enables experiences similar to those provided by walk-in-place systems.	Interact
Walk-in-place system		Interaction	A system that holds a user's torso in place, allowing it to rotate as the user's feet to walk in any direction on a free-axis moving tread-mill like surface or a low-friction surface, enabling the user to walk through a large-scale immersive virtual environment while physically remaining in place in the real world. An example is Virtuix Omni.	Interact
Wand		Sensor	A generic name for controllers with joysticks, buttons, triggers or pads tracked in the 3D space such as the Playstation Move, Vive or Oculus controller as well as the Flystick	Interact
Workbench		Display	Device of type "drafting table" made up of one (slanted) or two (horizontal and vertical) displays with a tracking system allowing designing at scale 1:1 virtual prototypes of size generally smaller than 1m3.	Display
Zenith		Metric	The highest point on a celestial sphere, antipodal to the Nadir. More generally, the point (or direction) represented by +90° ("North") in any spherical coordinate system. See also "Nadir".	Capture, Display