System Overview

SteamVR™ Tracking

Introduction

Valve developed SteamVR™ Tracking technology to address the rigorous pose tracking requirements of virtual reality (VR) applications. It is hard to trick the brain into believing it has entered another world. Our eyes, motor system, and brain are working in concert nearly every waking moment from the time we are born. Replacing the image that our eyes perceive as reality is the goal of VR. However, if the motion of that image does not agree with the motion of our body, we quickly become nauseated. Displaying the correct image requires a precise, real-time understanding of the pose of a person's head in three dimensional space. SteamVR™ Tracking does exactly that by locating a VR headset in three dimensional space to submillimeter accuracy. But, SteamVR™ Tracking does not stop there. Our realities are not limited to a seat in front of a computer screen. We spin, jump, sit, stand, and walk around in our three dimensional reality. SteamVR™ Tracking was designed to allow that same freedom in virtual reality. The submillimeter accuracy of SteamVR™ Tracking is achieved within a cubic volume approximately five meters on a side, enabling the most immersive VR experience available today. Valve is excited to offer SteamVR™ Tracking technology to other product developers, so they may enrich that experience by developing new trackable objects. Understanding SteamVR™ Tracking system is the beginning of that design process.

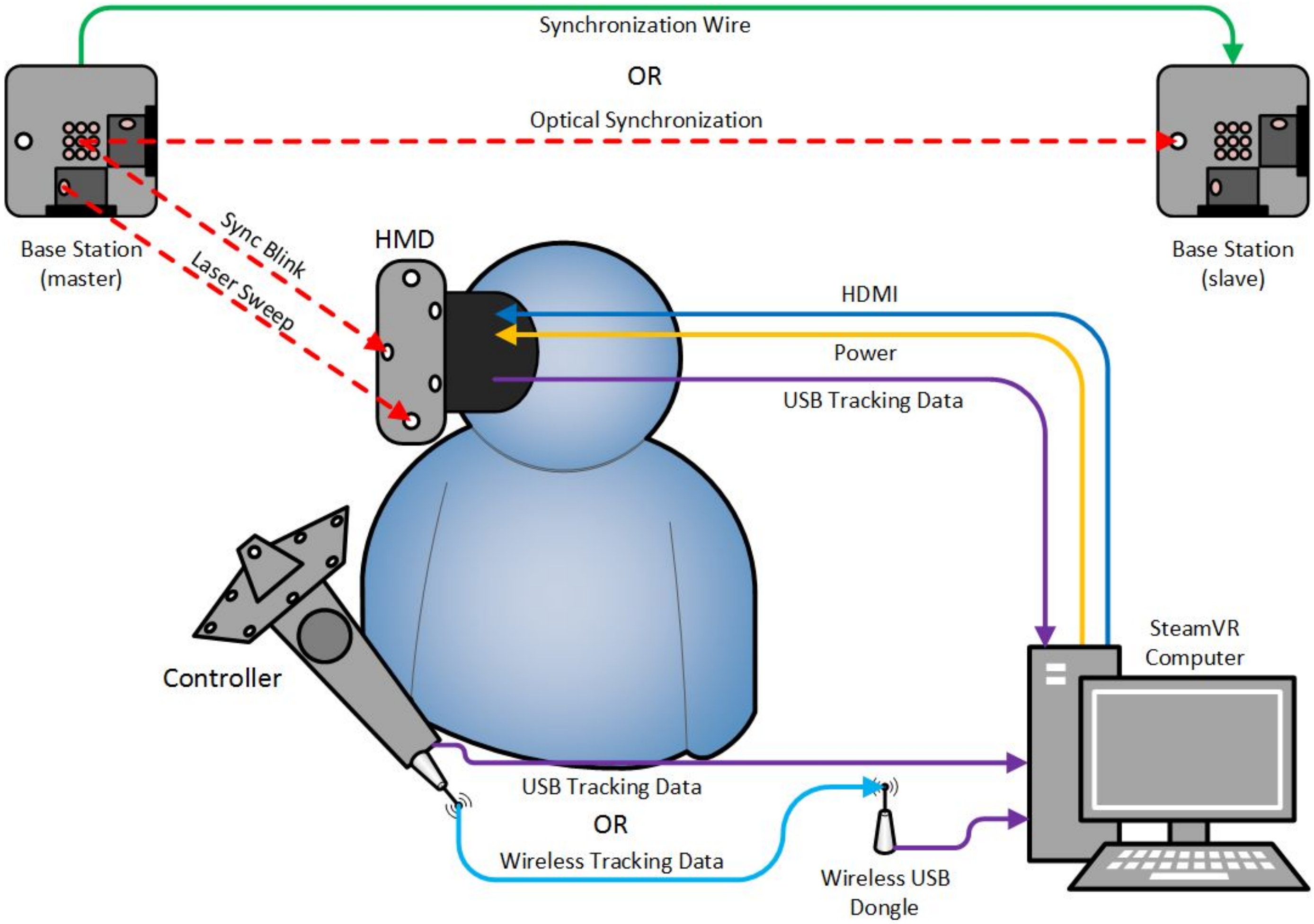
SteamVR™ Tracking documentation provides a complete description of SteamVR™ Tracking technology, the process of designing objects that are trackable in the 3D volume; testing, calibrating, and troubleshooting those devices; and integrating them into SteamVR™ software.

Application Space

Although SteamVR™ Tracking was developed to handle the demands of virtual reality systems, SteamVR™ Tracking is a pose tracking system with submillimeter accuracy, which could be valuable in applications outside of VR. Different position and orientation tracking systems have been around for decades, and are continually improving. Aeronautical tracking systems are used in autopilots. Automotive dead reckoning systems attempt to track vehicles when they are unable to detect GPS satellites. Asset tracking systems in large offices and warehouses use ambient RF signals to locate equipment and people within a facility. All of these systems include a common approach. First, they obtain an absolute location using GPS or some other positioning system. Then, they use inertial and pressure measurements to maintain an accurate position in the absence of the absolute positioning system. SteamVR™ takes a similar approach.

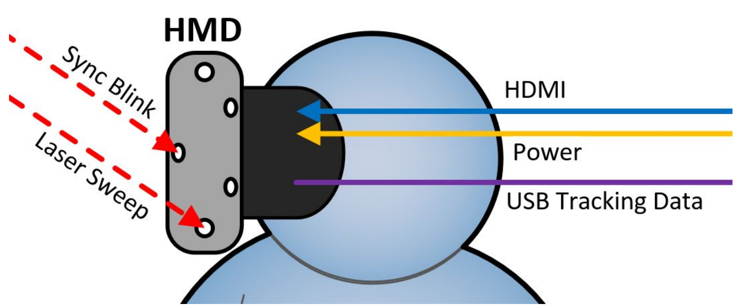
SteamVR™ Tracking embodies a position and orientation tracking system with incredibly high resolution, speed, and accuracy. SteamVR™ Tracking obtains absolute position information by triangulation using optical sensors that detect infrared (IR) reference signals sent from a base station to a tracked object. The absolute position is calculated up to 60 times per second. Meanwhile, an inertial measurement unit (IMU) inside the tracked object reports accelerometer and gyroscope measurements at a rate of 250-1000 Hz. These measurements are used to further refine the position, even when some of the optical sensors are occluded by the user or other obstructions. The performance of SteamVR™ Tracking enables VR applications, but it is important to note that this technology is not limited to VR and could be applied in any solution that would benefit from submillimeter positional accuracy within a cubic volume approximately five meters on a side.

SteamVR™ Ecosystem



The SteamVR™ ecosystem includes base stations, head mounted displays (HMDs), and controllers that are connected to a central computer running SteamVR™ software.

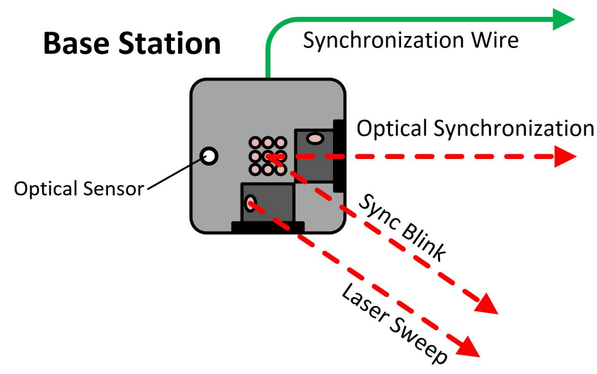
Head Mount Display



The head mounted display (HMD) is the iconic device present in any VR system. In essence, the HMD is a computer display worn in front of a user's eyes, but it is really composed of three distinct systems: binocular display, optics, and SteamVR™ Tracking. The binocular display is a monitor that is sized to fit in front of the eyes. It and divided into two halves, to show a stereoscopic image to the eyes. The binocular display is viewed through two lenses, one for each eye. These optics refocus the display to appear farther from the eyes, allowing the eyes to relax to a natural focal length, avoiding stress and fatigue. Finally, SteamVR™ Tracking allows SteamVR™ software to track the HMD within the 3D volume. Up to 32 optical sensors are placed on the HMD to detect the reference signals sent from a base station. By determining the position of the HMD to submillimeter accuracy, SteamVR™ can precisely represent virtual reality on the binocular display.

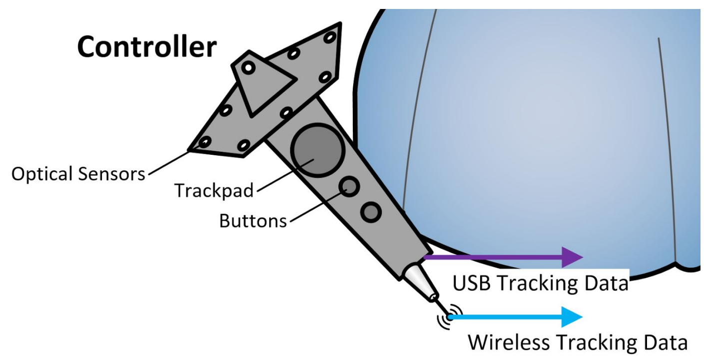
Base Station

The base station is the foundation of the SteamVR™ Tracking system. Base stations transmit IR signals to tracked objects in the VR system. Tracked devices have multiple optical sensors on each object and require only one base station to triangulate their position. Once the system is configured and the base stations are mounted to a wall, placed on a tripod, or set on a shelf they require no user interaction.



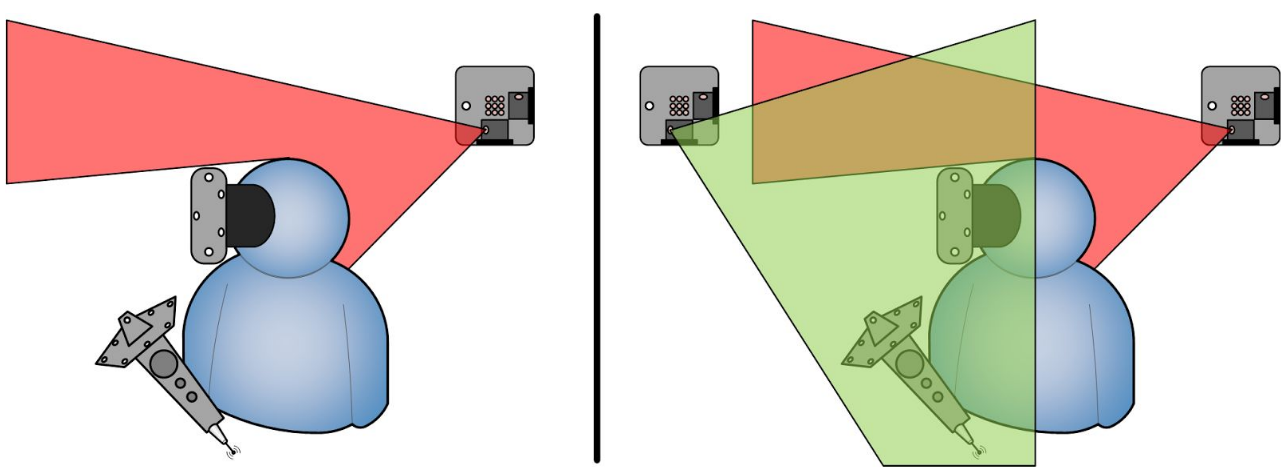
Controller

The minimum system required for tracking includes one base station, one HMD, and the computer; however, improving the user experience requires one or two controllers. Controllers allow users to interact with their virtual world. Held in each hand, they appear in VR as hands, paintbrushes, guns, bow and arrow, or anything a software developer can imagine. Buttons, triggers, and track pads on the controller allow users to select options, grab and hold objects, fire weapons, or open and navigate menus.



Multiple Base Stations

The SteamVR™ Tracking experience is greatly enhanced by adding a second base station. The second base station improves the chance that the HMD and controllers are always visible to at least one base station. Base stations emit infrared light, which is detected by the HMD and controllers. If a user turns their back to a base station, it is quite possible that the sensors are occluded by the user's body. When the IR reference signals are blocked by an obstruction the computer can no longer track the object. If another base station is placed on the opposite side of the room, the HMD and controllers enter the visual field of the second base station when the use turns around. Keeping tracked objects like HMDs and controllers in view of at least one base station maintains tracking within the entire volume regardless the user's orientation.

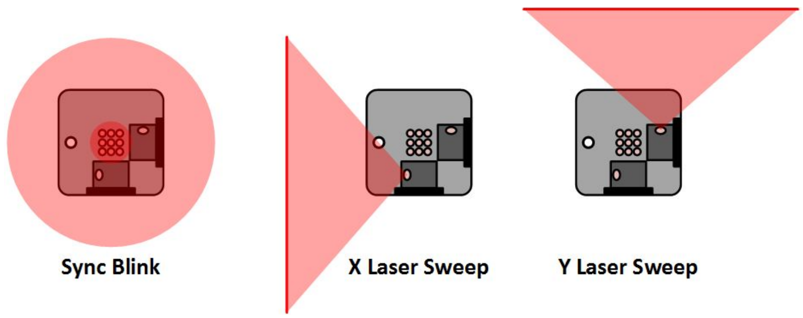


SteamVR™ Communication Paths

The devices comprising the SteamVR™ ecosystem communicate in a variety of ways to send tracking data downstream to the computer and video and control feedback upstream to the HMD and controllers.

IR Reference Signaling

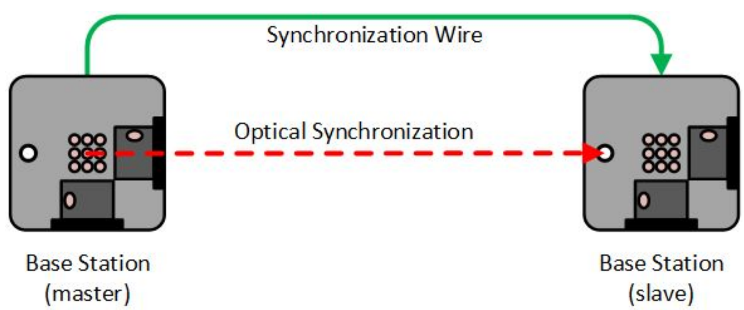
Base stations emit 830 nm infrared light to send reference signals to tracked objects. There are two IR signals sent from the base station. The first signal is a blink from a bank of IR LEDs, which indicates the beginning of a frame. The synchronization blink communicates to tracked objects that a laser sweep is coming. The second signal is a line of 830 nm laser light that sweeps through the volume and strikes the optical sensors on the tracked objects. There is a blink and a laser sweep for both the X and Y dimensions. The tracked objects record the time arrivals of these IR signals and send them to the computer over the wireless or USB link. SteamVR™ uses the timing information to calculate the position of the tracked object.



Base Station Synchronization

Base station synchronization signals are passed between the two base stations via a wire or optically. All base stations use LEDs and lasers that emit 830nm IR light. Because all base stations transmit on the same wavelength, they must synchronize their output signals in time to allow the system to associate the reference signals the tracked objects are receiving with the correct base station. To facilitate the synchronization, the base station designated as a master sends a 60 Hz pulse over a wire to the base station designated as a slave. The slave base station locks to the 60 Hz signal and synchronizes its IR transmissions to alternate cycles. This results in each base station sending reference signals at 30 Hz, multiplexed in time at a 50% duty cycle. However, running the wire between base stations does not provide the best user experience, so an optical path is also available.

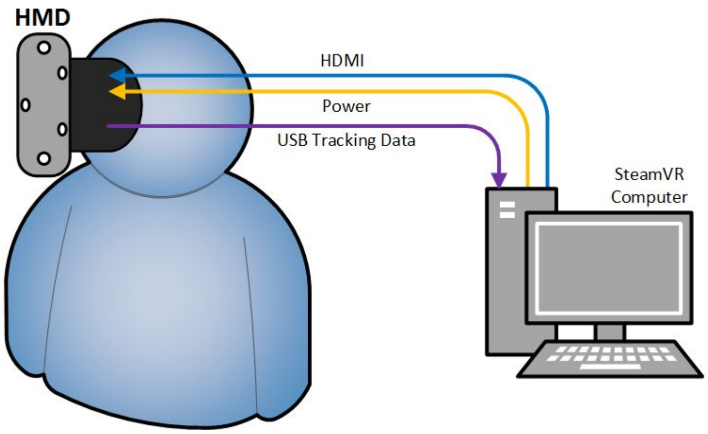
The base stations emit an IR blink at the beginning of each frame, and an optical receiver in the slave base station is able to detect the synchronization blink of the master base station. Using the sync-blink of the master, the slave locks to the 60 Hz frame rate of the master and multiplexes its own output with that of the master.



HMD Communications

SteamVR™ Tracking and Control Data

HMD Tracking and control data is communicated to the SteamVR™ computer over a USB 2.0 Full Speed link.



Audio/Video Signals

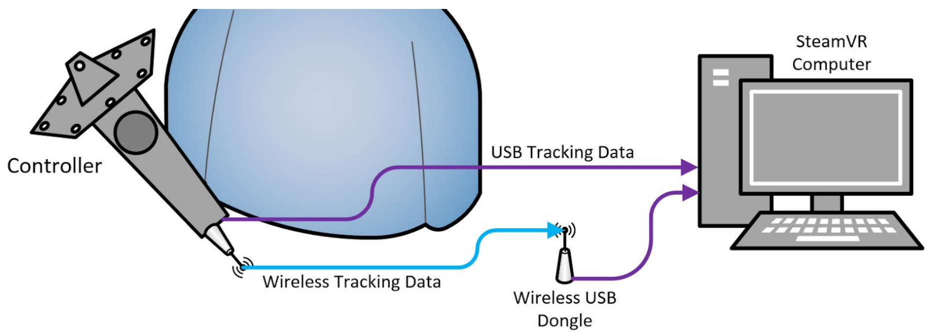
HMD binocular displays get their video data from an HDMI connection to the SteamVR™ computer. Audio is also delivered through the HDMI port, and is output to a stereo headphone connection on the HMD.

Power

Power is supplied to the HMD from a wall mounted DC power supply.

Controller Communications

SteamVR™ controllers are connected to the SteamVR™ computer over USB or wirelessly through the 2.4 GHz wireless USB dongle. The wireless environment and requirements are documented in **Antenna Design and RF Guidance**.



Theory of Operation

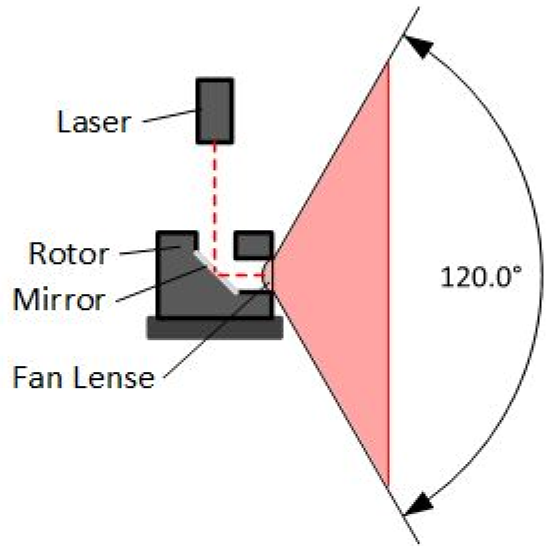
The base stations, tracked objects and computer work together to triangulate the position of tracked objects using the time arrivals of the IR reference signals from the base stations. The reference signals from the base stations are received by the optical sensors on the tracked objects. Not all optical sensors may receive a reference signal, because some may be occluded by the tracked object itself or another obstruction. The system requires four sensors to receive a reference signal to triangulate the position of the tracked object. The arrival of reference signals at each sensor are timestamped using a 32 bit counter running at 48 MHz. The timestamps associated with each sensor are transmitted to the computer. Finally, the computer uses the timing information and the known geometry of the object to triangulate the position of all the visible sensors, and uses those positions to solve for the overall position and orientation of the tracked object. However, the reference signals alone are not sufficient for high performance tracking.

The reference signals have a best case update rate of 60 Hz, which is augmented by acceleration data from the tracked object’s inertial measurement unit (IMU) to achieve the best tracking performance. The absolute position of the tracked object is calculated at the 60 Hz rate of the reference signals. However, this rate could decrease under some circumstances. Two base stations in a system will time domain multiplex (TDM) their reference signals. This reduces the reference rate from one base station to 30 Hz. Additionally, there may be times when an obstacle occludes enough sensors to reduce the number of visible sensors below five. Each tracked object includes a six-axis IMU, which implements an accelerometer and gyroscope. The IMU constantly measures the linear and angular acceleration of the tracked object. Acceleration data is transmitted to the computer at a rate of 1 kHz. The computer uses the IMU data to continue calculating the change in position of the tracked object between absolute measurements. Dead reckoning the position of an object using IMU data allows the system to maintain high performance tracking when the reference rate reduces. Also, although five sensors are required to begin tracking an object, IMU data allows the computer to track an object with as few as two visible sensors.

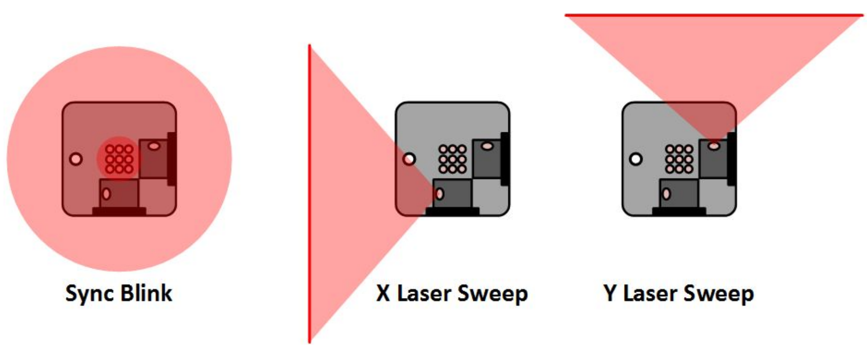
The combination of triangulating the absolute position of an object using optical reference signals and smoothing that information using relative changes measured by the IMU, allows SteamVR™ Tracking to determine the position of tracked objects to submillimeter accuracy, within a cubic volume approximately five meters on a side.

Generating Reference Signals

The reference signals from the base stations comprise two different IR signals, an LED blink and a laser sweep. The base station has two lasers and two spinning motors inside, one for each axis. When the laser beam shines into the hole in the top of the motor's rotor, it hits a mirror that deflects the beam by 90° and sends it out the side of the rotor through a fan lens that turns the laser beam into a line of laser light that is oriented parallel to the axis of rotation. The fan lens spreads the laser light through 120°.



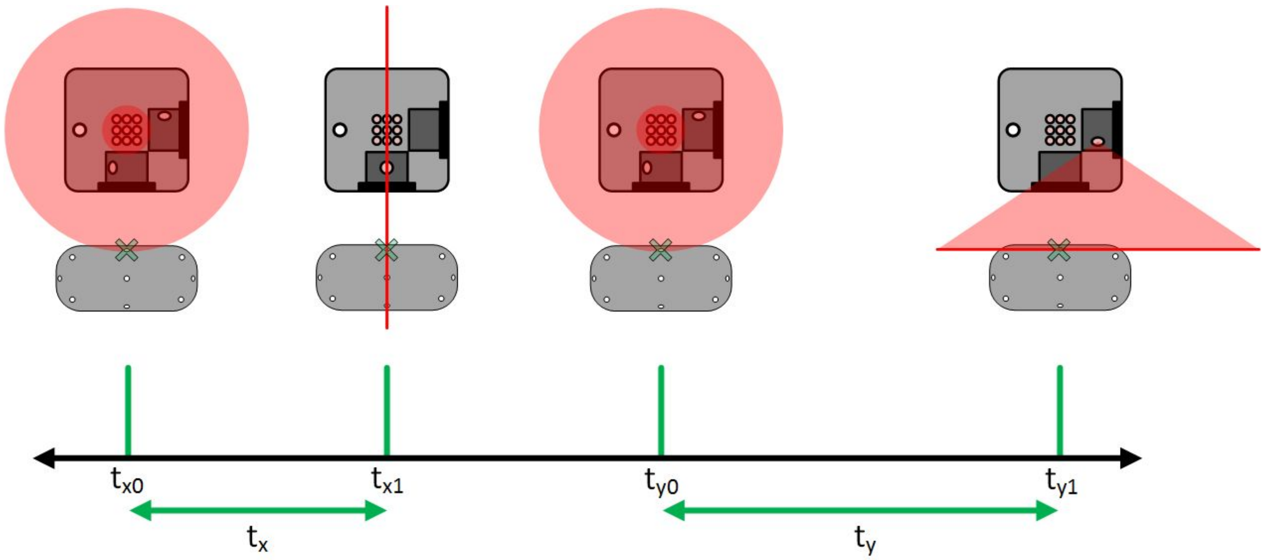
Orienting the rotors 90° from each other creates two lines of laser light, one the sweeps in the X direction (left to right), the other sweeping the the Y direction (down to up).



A 0° reference signal is required to calculate the angle of the laser sweep. The base station uses a bank of IR LEDs which blink at the beginning of a laser sweep. The blink provides a timestamp indicating an angle of 0°.

Comparing the timestamp of the laser sweep to the 0° timestamp yields the angle from the base station to the sensor.

To create reference signals in the X and Y dimensions, the base station phase aligns its rotors with a 180° offset. A synchronization blink occurs at X rotor 0°. Then, the X laser sweeps through the volume. Another synchronization blink occurs when the Y rotor reaches 0°, followed by the Y laser sweeping through the volume. This pattern is repeated at 60 Hz, unless there are two base stations in the system.



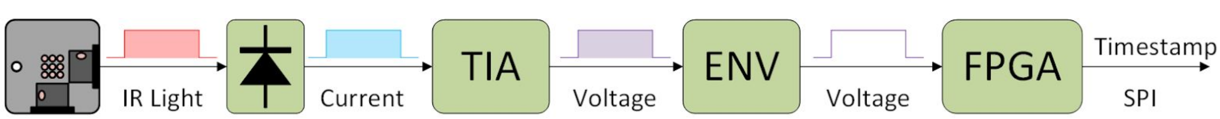
If two base stations are present, then the slave base station phase aligns its motors to the master, and the two base stations alternate sending IR signals. The motors of each base station continue spinning, but they shutter their IR LEDs and lasers during alternate frames. The slave base station is able to phase align to the master by detecting the synchronization signal sent over the wire between the base stations, or by using an optical receiver in the base station to detect the synchronization blink of the master base station.

The LED blink and the laser sweeps are modulated at 1.8 MHz. This modulation allows the optical receiver in the tracked object to reject unwanted ambient IR light to more reliably detect the reference signal.

Receiving Reference Signals

The IR reference signals from the base station are received and timestamped by the tracked object. Each tracked object uses 20 to 32 optical receivers to detect the LED blinks and laser sweeps sent from the base station. This number of optical receivers is necessary to ensure that reference signals are received regardless the orientation of the tracked object. Once received, the IR pulses are timestamped using a counter running at 48 MHz, with a resolution of 32 bits. These timestamps are transmitted to the computer for further processing.

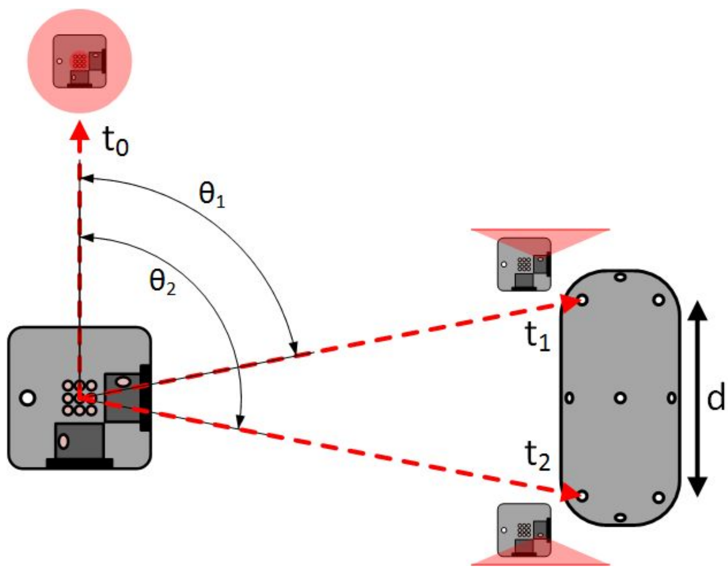
The IR reference signals from the base station are detected by optical receivers using a photodiode, transimpedance amplifier (TIA), and envelope detector. Incident IR light causes a current to flow in the photodiode. That current is converted to a voltage using a high-gain transimpedance amplifier. The 1.8 MHz modulation is removed from the signal using an envelope detector. Finally, the digital signal representing the period of time that the photodiode was detecting the reference signal is transmitted to an FPGA in the tracked object. The FPGA timestamps the rising and falling edges of the envelope and those measurements are transmitted to the computer.



Rising and falling edges are used to calculate the centroid of the photodiode. It is important to calculate the centroid, not only use the leading edge of the reference signal, because the diode has a square photosensitive area of 2.65 mm × 2.65 mm. Submillimeter precision can only be achieved by calculating the centroid to account for the shape of the photosensitive area and its orientation to the sweeping laser line.

Calculating Position

The position of a tracked object is calculated by triangulating the relative positions of the optical sensors and then fitting those measurements to the known distances between the sensors on the object. Triangulation begins by measuring the angle from 0° to each of the sensors. 0° is marked by a synchronization blink from the base station. The time reference for 0° is captured using the 48 MHz, 32 bit counter in the tracked object. Then, as the laser sweeps through space, the laser hits at each sensor are captured by the tracked object's counter. At the end of the laser sweep, SteamVR™ Tracking has a timestamp for 0° and for laser hits at all visible sensors. At that point, it is easy to calculate the angle from the base station to each visible sensor. Once the angles are calculated, SteamVR™ Tracking solves the geometry for a translation and rotation that fits the known dimensions of the sensor placement on the tracked object.



Calculating θ from the recorded timestamps is possible given the number of elapsed counter ticks from the synchronization blink, the period of the timestamp counter, and the speed of the motor.

$$\theta = t \times T_{counter} \times \omega_{motor}$$

SteamVR™ Tracking uses motors running at 60 Hz and timestamp counters running at 48 MHz, which yields the following equation for θ.

$$t = t_1 - t_0 \; ticks$$

$$f_{counter} = \frac{48\times10^6 \; ticks}{1 \; s}$$

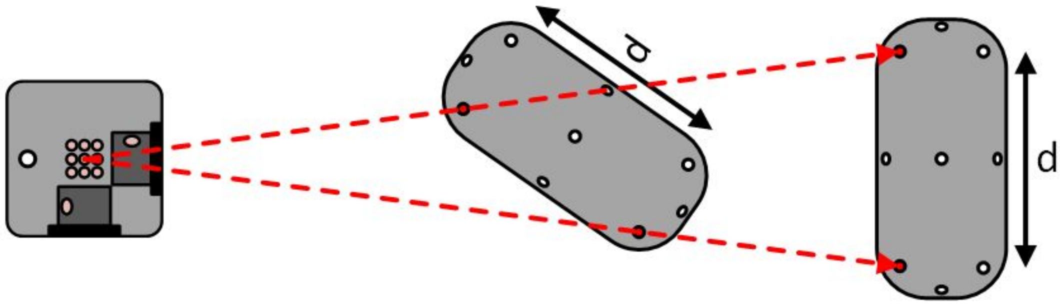
$$T_{counter} = \frac{1 \; s}{48\times10^6 \; ticks}$$

$$\omega_{motor} = \frac{2\pi \; rad}{1 \; cycle} \times \frac{60 \; cycles}{1 \; s} = \frac{120\pi \; rad}{1 \; s}$$

$$\theta = t \; ticks \times \frac{1 \; s}{48\times10^6 \; ticks} \times \frac{120\pi \; rad}{1 \; s}$$

$$\theta = t \times \frac{\pi}{400{,}000} \; rad$$

After SteamVR™ collects timestamps from the tracked object and calculates the angle of each sensor from the base station, it solves for a translation and rotation of the known object geometry that fits the measured set of angles. SteamVR™ knows the geometry of every tracked object, because each object contains a calibrated JSON file describing its sensor geometry. In the simplified two dimensional drawing shown above, SteamVR™ would solve for a distance from the base station that satisfies the measured angles, knowing that the two sensors are separated by a distance (d). However, this two dimensional example also demonstrates that using only two sensors for tracking is not sufficient. With only two sensors, it is possible to find many combinations of position and orientation that satisfy both the measured angles and the known geometry.



Tracking requires five visible sensors to locate the translation and rotation of a tracked object. Two visible sensors leaves many possible solutions, as shown above. Three visible sensors still does not constrain the object to a single position and orientation. Four sensors is the theoretical minimum to obtain a single answer; but, five sensors is the practical minimum. This limit suggests one of the fundamental requirements of tracked object design. An object designed for robust tracking requires five sensors to be in view at all object orientations and positions.

Sources and Consequences of Error

The idealized SteamVR™ system described above provides the basis for SteamVR™ Tracking, but the challenges and limitations of the system are defined by inherent errors in a real embodiment of the system. There are several types of error in the system. Some errors may be introduced into the system by design, and cannot be calibrated away. These errors must be avoided. Other errors may be introduced through manufacturing and may be calibrated out. Finally, there are random errors inherent in the physical implementation of the system, which must be minimized by design to guarantee performance.

Sources of Error

Sensor Covering

Protecting the photodiodes of the optical sensors from physical and electrostatic damage necessitates placing them behind some kind of covering. However, sensor coverings have optical properties that affect the performance of the sensor. One property is transmissivity, which reduces the overall strength of the optical signal. Any obstacle that reduces the strength of the optical signal ultimately limits the effective range of the SteamVR™ system. The optical range of the system is detailed in the **Optical Link Budget**. Another important effect of sensor coverings is refraction.

As incident IR light penetrates the sensor covering, the light bends. This refraction is described by Snell's Law, and may change the performance of the optical system if it is not accounted for in the design. Because the angle of refraction is dependent on the angle of incidence, this error changes with the orientation of the tracked object to the base station. Therefore, errors introduced by sensor covering cannot be calibrated out of the system. As a result, it is very important to consider sensor covering early in the design process, especially because sensor covering is also a cosmetic feature important in the industrial design. A full primer on sensor covering requirements and techniques is available in the **Sensor Covering** document.

Sensor Placement

Placing sensors during the manufacturing process inevitably introduces some amount of error. There is a tolerance stackup between the photodiode placement on the PCB/FPC, registration of the sensor PCB/FPC to the window in the device chassis, and variation in alignment between different sections of molded chassis parts. The depth of sensors may also change from device to device, within the tolerances defined for adhesive and plastic thicknesses. Fortunately, these errors are largely constant once a tracked object is assembled. As a result, the errors introduced by non-ideal sensor placement may be calibrated out. However, this calibration is required for each device, which adds a step to the manufacturing process. Valve has developed the tools to perform optical calibration, and they are described in the **Optical Calibration** document.

Jitter

Jitter is the measure of the statistical variation in period from one cycle of a periodic signal to the next. In the SteamVR™ system, jitter refers to the variation between successive time measurements between the synchronization pulse and the arrival of a laser sweep at a given sensor. If a tracked object is perfectly still, an ideal system would always report the same time difference (t) between a synchronization pulse and the arrival of a laser sweep. However, in a real system, several factors contribute to the random variation of t from cycle to cycle. As a result, it can be represented statistically as a gaussian distribution about the ideal value. This effect is such an important factor in the performance of the SteamVR™ system that a jitter budget was developed. The jitter budget is described in the **Jitter Budget** document. Two of the main sources of jitter error are briefly described below.

Sampling Error

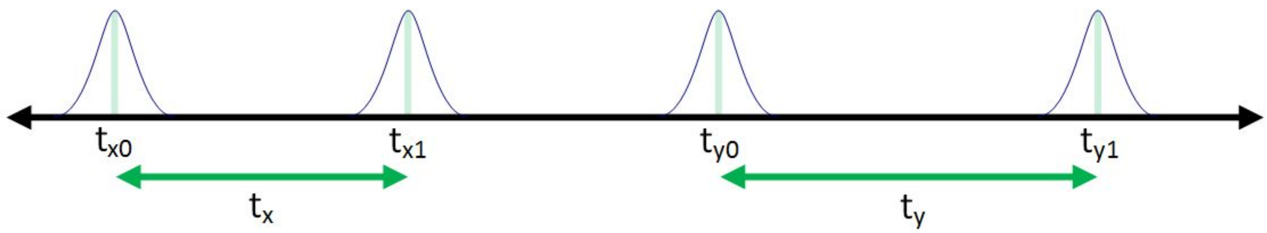
The reference signal transmitted from the base station is asynchronous to the 48 MHz clock in the tracked object. As a result, the timestamp for a rising or falling edge of the reference signal must by synchronized with the 48 MHz clock before timestamping. This resynchronization results in a quantization error with a resolution of one 48 MHz period or ~21 ns. Reducing this error requires increasing the sample rate of the optical sensor. However, the sampling error is not the dominant source of jitter at 48 MHz. Therefore, there is not enough benefit to justify increasing the sampling rate.

Motor Jitter

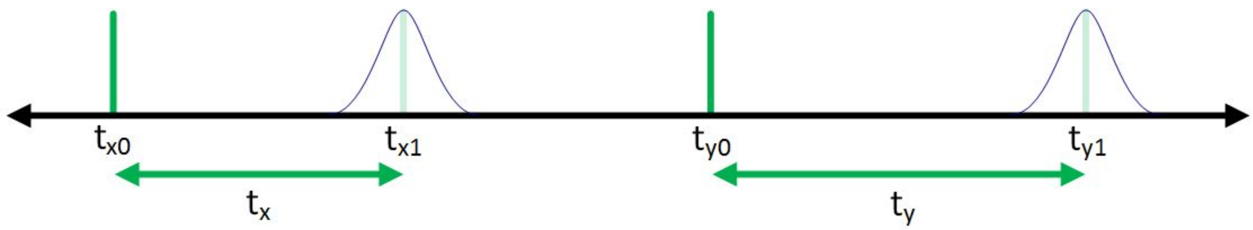
Motor jitter is one of the dominant sources of jitter in the SteamVR™ system. Motor jitter refers to the cycle to cycle variation in the period of the motor’s rotation. Ideally, each motor would rotate at exactly 60 Hz and have a period of 16.667 ms. However, there are several factors that change the period on a cycle by cycle basis. Variations the the friction encountered during rotation due to bearing imperfections, air turbulence, and vibration momentarily slow the motor during rotation. Sampling and measurement errors in the sensorless detection of rotor position affects motor commutation timing which leads to variations in motor velocity within a rotation. Analog noise and sampling error in the detection of the 0° marker on the rotor add to the error associated with the 0° synchronization pulse. All of these errors sum to create an overall jitter in the motor system that varies with a gaussian distribution. Currently, the design goal of the system is to control the jitter of the motor system such that one standard deviation of the motor jitter is 10 ppm of the nominal period. At 60 Hz, this means a jitter target of $1\sigma \leq 167$ ns.

Accumulation of jitter

The errors above accumulate, ultimately limiting certain aspects of the of the system. Jitter in the base stations create variation in the timing of the synchronization pulse and the laser sweeps. Jitter in the tracked object receiving and timestamping, introduces more timing error.



When measuring the time difference between the synchronization LED blink and the arrival of a laser sweep, the errors in the base stations and the tracked object add together. For the sake of analysis, it can be useful to picture all of this error accumulating at the time the laser sweep is captured.



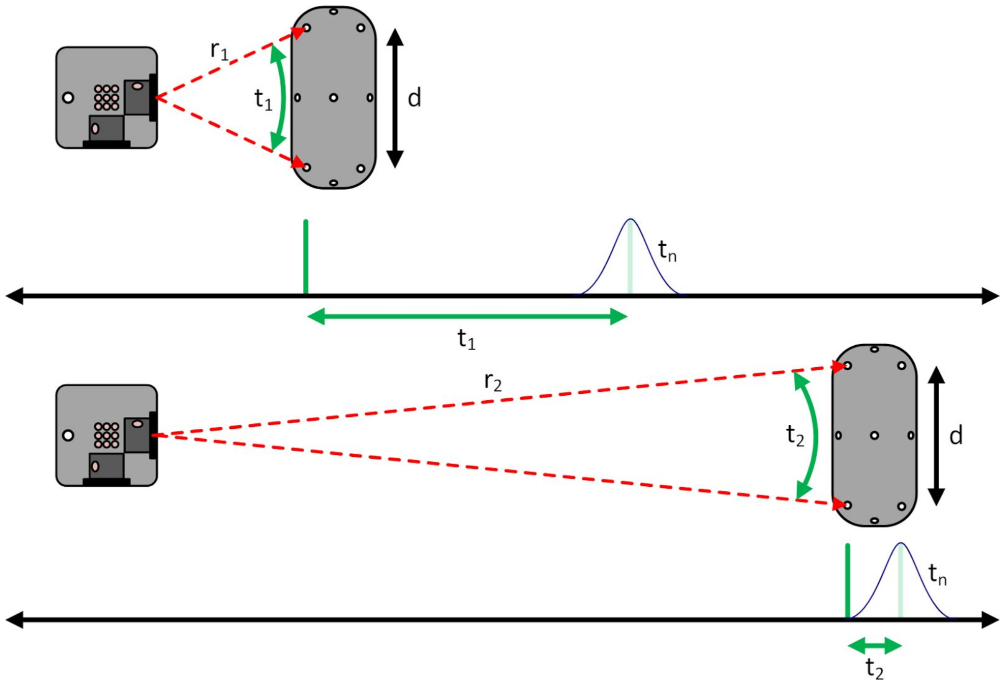
Effects of Error

It is important to recognize the impact of jitter on the limits of the system. Measuring the position of an object in three dimensional space can be described by its translation and rotation. Translation refers to the object's position in a three dimensional coordinate system. Rotation refers to an object's orientation in terms of roll, pitch, and yaw. Jitter in the system adds noise to translation and rotation measurements. System limits are reached at points where noise begins to dominate the timing measurements. Understanding these limits helps a designer understand how to optimize the placement of sensors to minimize the effects of jitter and get the best performance from SteamVR™ Tracking. A host of documentation on tracked object design describes the process of designing object geometries and placing sensors to minimize the effects of noise in the measurement system. The document **Object Design and Integration Overview** begins the topic of object design and optimization.

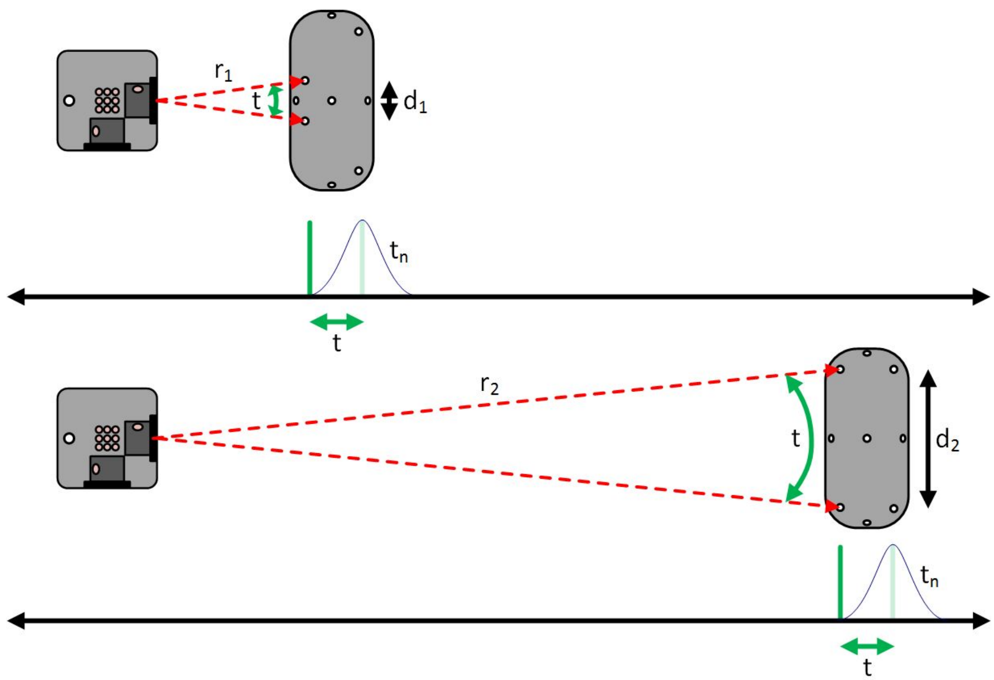
Translation Error

Translation error refers to the error in detecting the correct X, Y, Z location of an object in three dimensional space. To visualize the effects of translation on measurement, it is helpful to picture an object as it is "seen" by the base station. Remember that a base station is spinning a motor which sweeps a line of laser light through space, and the time difference between sensors is used to calculate the position of the object. The angular speed of the motor is constant, which means that the farther an object is located from the base station the faster the tangential velocity of the sweeping laser line. Imagine what happens to the elapsed time between sensors as the object moves farther from the base station. Increasing the distance from the base station (r) , increases the velocity of the laser line (v), and shortens the elapsed time between sensors (t). One can imagine that as the object moves farther from the base station the sensors appear closer together, just as railroad tracks appear closer together as they extend into the distance.

If the measurement of time (t) has an error (t_n) introduced by jitter in the system, one can easily imagine a translation where the noise t_n begins to dominate the measurement ($t + t_n$). From this analysis, it can be seen that translation error is affected most by the translation of the tracked object away from the base station, because the effect of jitter on the measurement increases as the angle between sensors decreases.

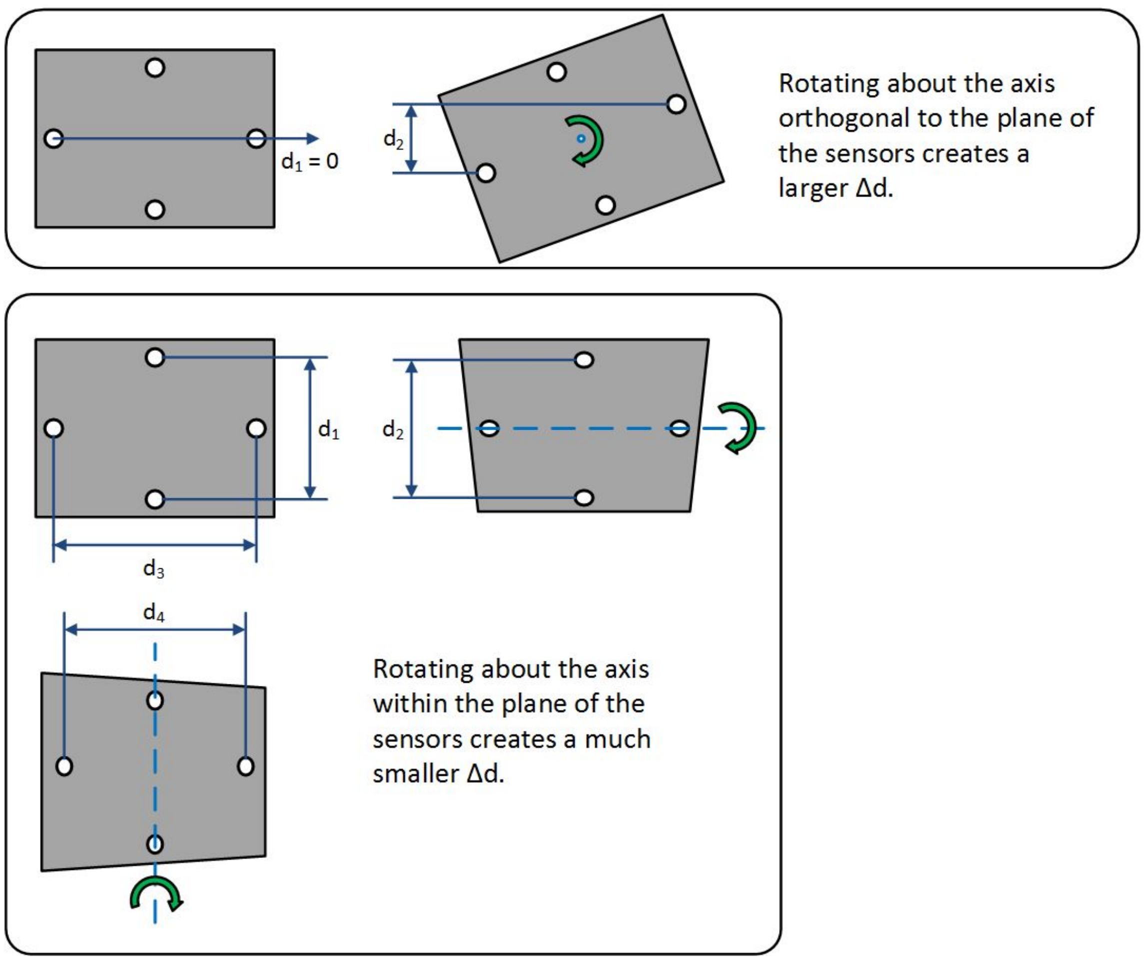


Further, it is important to note that increasing the distance between sensors is one way to extend the range of the system; because, the tracked object could accept more translation in the Z direction before the distance between sensors is foreshortened to a point where error dominates the measurement.



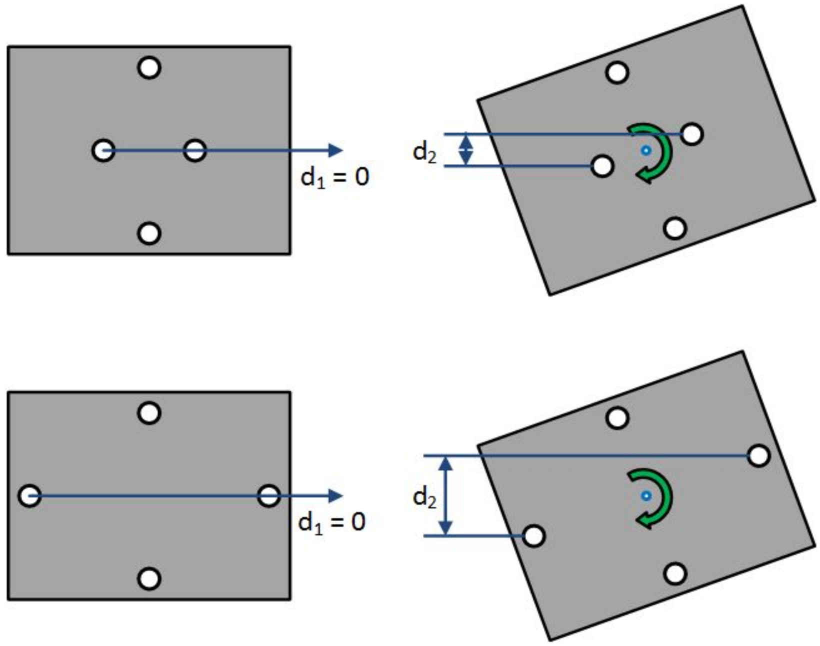
Rotation Error

Rotation error refers to the error in detecting the correct roll, pitch, and yaw of an object in three dimensional space. To visualize the drivers of rotation error, it is helpful to remember that SteamVR™ Tracking is measuring the change in distance between two sensors as an object rotates. Rotation error becomes dominant if a rotation does not produce a relative change in position between sensors that is significantly larger than the error in the measurement system. Therefore, rotation error is maximized when sensors are in a plane, and the object rotates about an axis in the plane.



Based on this understanding of rotation error, it is easy to see the importance of arranging sensors in all three planes, XY, XZ, and YZ. When visible sensors exist in all three planes, rotation on any axis may produce a significant change in sensor position.

As with translation error, increasing the distance between sensors reduces effects of jitter. Maximizing the distance between sensors, in all three dimensions, maximizes the relative change of position for a given angle of rotation.



Summary

SteamVR™ Tracking is a pose tracking system capable of submillimeter accuracy within a cubic volume approximately five meters on a side volume. Although SteamVR™ Tracking is not limited to VR and gaming, the SteamVR™ ecosystem creates a platform for the most immersive VR experience available today. The SteamVR™ ecosystem includes base stations that emit IR reference signals that are received and timestamped by tracked objects. Those timestamps are transmitted over a wireless or USB link to the computer. In the computer, timestamps are analyzed to track the position of the object in three dimensions. SteamVR™ requires a minimum of five visible sensors in all positions and orientations for robust tracking.The position calculations are affected by errors in the base station and tracked objects. Some of these errors, like those from IR refraction through sensor covering, must be avoided. Other errors, like the misalignment of sensors, may be calibrated out during manufacturing. Error in the system due to jitter in the periodic measurement of reference signals cannot be calibrated away and must be minimized through good design practice. Jitter contributes to translation and rotation error, and it may limit the distance and resolution of SteamVR™ Tracking. Designing objects with sensors in all three axes that can receive reference signals under any rotational pose while maximizing the distance between sensors enables tracked objects to realize the submillimeter accuracy promised by SteamVR™ Tracking and users to realize the most fluid and immersive VR experience.