



VIVE 3DSP Audio SDK Guide (Unity Plugin)

Release version: 1.3.10

©2024 HTC Corporation. All Rights Reserved. HTC, the HTC logo, Vive, the Vive logo, and all other HTC product and services names are the trademarks or registered trademarks of HTC Corporation and its affiliates in the U.S. and other countries.

All other trademarks and service marks mentioned herein, including company names, product names, service names and logos, are the property of their respective owners and their use herein does not indicate an affiliation with, association with, or endorsement of or by HTC Corporation.

Table of Contents

System Requirements	5
Why is 3D Audio Needed in VR?	6
Localization	6
Room Reverberation	7
Distance	7
Spatialization	8
VIVE 3DSP Audio SDK	8
Higher Order Ambisonics (HOA)	9
Head-Related Transfer Function (HRTF)	10
Room Audio	11
Sound Decay Model	11
Occlusion	12
Ambisonic Decoder	12
VIVE 3DSP Audio SDK Unity Plugin	13
Import Latest Package into New Project	13
Upgrade to Latest Package	13
Set Spatializer Plugin	14
Set Ambisonic Decoder Plugin	14
Audio Mixer Settings	16
Audio Source Settings	16
Audio Room Settings	20
Audio Occlusion Settings	23
Audio Listener Settings	25
Decoding watermarks	27
Hi-Res Audio Settings	28
Troubleshooting	31
Appendix	32

Material preset list 32

VIVE 3DSP Audio SDK Forum Shortcut..... 33

System Requirements

- Unity version 2017 or higher
- Windows® 7 SP1, Windows® 8 or later, Windows® 10
- Android 8.1 or later
- Intel® Core™ i5-4590 or AMD FX™ 8350 equivalent or better
- NVIDIA® GeForce® GTX 1060 or AMD™ Radeon™ RX 480, equivalent or better graphics card

Why is 3D Audio Needed in VR?

For many VR applications, there may be awesome visuals in the scene, but it will be less immersive without accompanying audio. Therefore, the 3D audio experience is crucial to increase true immersion and sense of presence in VR environments.

There are a lot of factors related to audio experiences in VR applications, such as sound Localization, Room Reverberation, Distance, Spatialization, etc.

Localization

Localization is the inherent ability of human beings to figure out where a sound is coming from based on human auditory system. By understanding the human auditory system, real-world spatialized sound can be simulated in a VR environment.

Here, the factors are listed as:

- **Lateral**
 - **Interaural time difference**

The direction of a sound source can be determined by the time difference that a sound arrives to our ears, i.e. a sound wave is transmitted to the right ear before the left ear, and so the listener can be aware that the sound wave comes from the right-hand side.
 - **Interaural level difference**

There is also the sound level difference between both ears of the listener during sound propagation, i.e. the sound wave reaches the right ear before the left ear, the sound energy decays while sound is transmitting.
- **Elevation and Front/Back**

In order to distinguish the direction of the sound which has the same arrival time and level, direction selective cues (pinna, head, shoulder, and torso) are used as the main reliable features. These cues are used to help determine where the sound comes from. Moreover, people can also cock their head up or turn their heads left or right to determine the direction of the sound source more precisely.

- **Pinna**

When the sound arrives at the head, it will then be collected and filtered by the outer ears. These filters provide the feeling to the brain to determine where the sound is coming from.

- **Head/Shoulder/Torso**

A sound wave is partially blocked or absorbed by the body of a listener during sound propagation. That means there is an inference of the body on sound feeling.

Room Reverberation

If the user is inside the room, the walls and objects in that room will reflect or scatter the sound waves. The sound waves will bounce again and again, and form a particular audio perceptual experience in that room. It not only builds a realistic feeling of the room, but also improves the accuracy of the directional perception.

Distance

The distance between a sound source and a listener is also an important factor on spatial audio application. There are some ways to describe the sound level decay that corresponds to the sound transmitting distance.

- **Loudness**

Louder sounds are usually closer while quieter sounds are usually farther. In some cases, it is hard to judge the distance by the loudness if people are not familiar with the sound. The relative loudness is also considered. When a sound wave approaches, it sounds louder and vice versa.

- **Initial time delay**

This is the time difference between the arrival of the direct sound and the first significant reflection, reflected from the surface of the room. When somebody is talking to a listener in a room, the listener will hear the direct voice right away then the reflected voice a bit later. If the person is standing farther away, the listener will probably hear the voice and reflections almost at the same time.

- **Mix of direct and indirect sounds**

The human auditory system unconsciously mixes direct and indirect sounds. When the direct sound is perceived louder than the indirect sound, the sound is emitted near to the listener, otherwise the sound is far away from the listener.

- **Interaural level difference**

When there is an obvious difference of loudness between the two ears of an audience, the sound is most likely to be close to the audience. When the loudness difference between the two ears of the audience is negligible, the sound is probably far away.

Spatialization

Spatialization is an audio effect that makes listeners feel immersed in a virtual 3D environment. Spatialization is a key factor for an immersive VR experience. A lot of features are accomplished to simulate the real spatial feeling in 3DSP Audio SDK, such as:

- Head-related transfer functions (HRTF) recording and improvement
- Higher Order Ambisonic simulation of sound direction
- Room audio simulation
- Adding background noise floor
- Real-world acoustic properties of distance
- Sound occlusion simulation
- Hi-Res audio support

and many other features

VIVE 3DSP Audio SDK

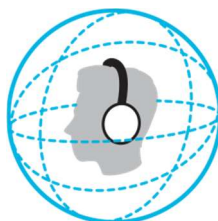
VIVE 3DSP Audio SDK is an audio solution for simulating realistic sounds in the virtual world. The SDK could be used to create engaging and immersive spatial audio for VR apps or experiences. A host of features are designed for developers.

Some of the unique features of VIVE 3DSP Audio SDK:

- [Higher Order Ambisonics \(HOA\)](#)
Highly optimized 3rd order HOA is implemented with very low computing power.
- [Head-Related Transfer Function \(HRTF\)](#)
The precise HRTF is modelled in this SDK for every 5 degree angles.
- [Room Audio](#)
The reflection and reverberation of a real space are simulated in the room audio model.
- [Hi-Res Audio Settings](#)
The quality is kept when Hi-Res audio is processed with VIVE 3DSP Audio Features.
- [Sound Decay Model](#)
The different models are adopted to create the different sound decay phenomena.
- [Geometric Occlusion](#)
The super-efficient occlusion model works well without any Unity collider. Moreover, the virtual sound wave could be partially occluded according to the obstacle geometry.
- [Ambisonic Decoder](#)
The AmbiX file is supported in the SDK.

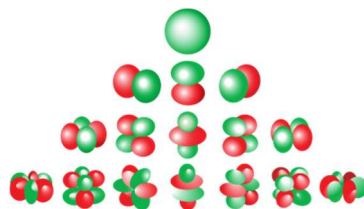
VIVE 3DSP Audio SDK is hardware-independent, high-quality, user-friendly, and powerful. Outstanding sound experiences can be provided especially for VIVE devices.

Higher Order Ambisonics (HOA)



Ambisonics is the technology that uses a full-sphere surround sound

technique to simulate spatial sound. Spherical harmonics is presented as the sound pressure changes around the listener.



The order of the spherical harmonics is also called the order of ambisonics. The order of ambisonics has a big influence on the realistic depiction of the direction of sound. Therefore, 3rd order ambisonics is used in the spatial audio model as it generates the better directivity of a spatial sound.

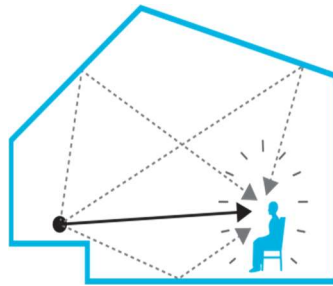
When the higher order is used, higher computing power is required for the ambisonic process. However, as a patent-pending technique is used on the VIVE 3DSP ambisonic process, it just requires a low computing power similar to the 1st order ambisonic.

Head-Related Transfer Function (HRTF)

The head-related transfer function represents the spatial sound model. There are many factors that could change the audio experience coming from an audio source travelling to the human ears, and it is too complicated to simulate the phenomena by dedicated functions. Therefore, a database has been created where all acoustic responses have been modelled. The audio response can be demonstrated by using the recorded data from this database.

In the VIVE 3DSP Audio SDK, there is an HRTF database that was measured with fine directional segments, both horizontally and vertically. This means that thousands of HRTF directions have been measured. For HRTF directions that have not been measured, these could also be generated by the measured directions nearest to it, with the quality benefitting from the fine measurement of the HRTF database. The quality degradation caused by the environment, the measuring equipment, and other factors that affect HRTF have also been improved in our database.

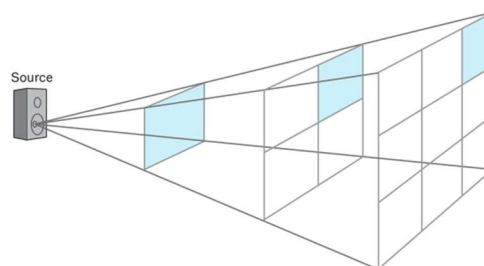
Room Audio



Room Audio is the technology that simulates the sound feelings in a room such as the early reflection, late reverberation, ambient sound, and others. The material (e.g. wood, concrete, glass, etc.) of the walls in a room affects the sound reflections in the early stage of the traveling sound. After that, the sound encounters more reflections and mixes with the background sound and finally turns into late reverberation. Furthermore, binaural acoustics should be simulated especially when a human is located at the corner of the room. Moreover, there should be a lot of low-level sound like the sound from a refrigerator or an air conditioner, therefore the phenomena is also presented in Room Audio technology.

Note: Room audio only works when the audio source and listener are in the same room.

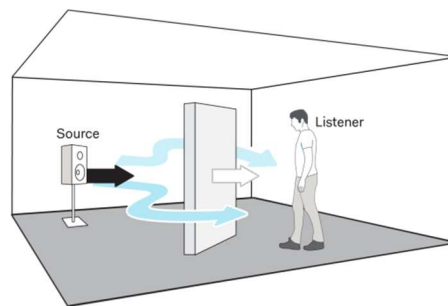
Sound Decay Model



During the sound wave propagation, the sound energy changes according to the distance. The inverse square law is a conventional method for this scenario to determine the change in the sound levels. However, the capability

of the sound decay varies depending on the frequency. Point source, line source and linear curve attenuation models are provided.

Occlusion



The occlusion effect is adopted to simulate a phenomena when a sound wave encounters an obstacle. The resulting sound is produced depending on the properties of the obstacles, like the material and thickness.

In some cases, the obstacle asymmetrically occludes the sound transmission between a sound source and a human head. For more realistic experiences, binaural occlusion is considered in VIVE 3DSP Audio SDK to produce different sound feelings for both ears.

In this SDK, a high-precision analytical geometry technique is present in occlusion effect. The occluded-sound effect can be presented according to some geometric factors of the obstacles, such as positions, rotations, sizes, and shapes. It also performs efficiently without using any amount of physics engine of other platforms like Unity or Unreal Engine.

Ambisonic Decoder

Audio recordings for 360 and VR videos are now mostly achieved by sound field microphones in the 4-channel B-format. Here, these channels are not translated to a fixed representation of sound. Instead, the ambisonic decoder can interpret and reshape the sound recordings such that it can dynamically render a stereo audio according to a listener's rotation.

VIVE 3DSP Audio SDK Unity Plugin

Import Latest Package into New Project

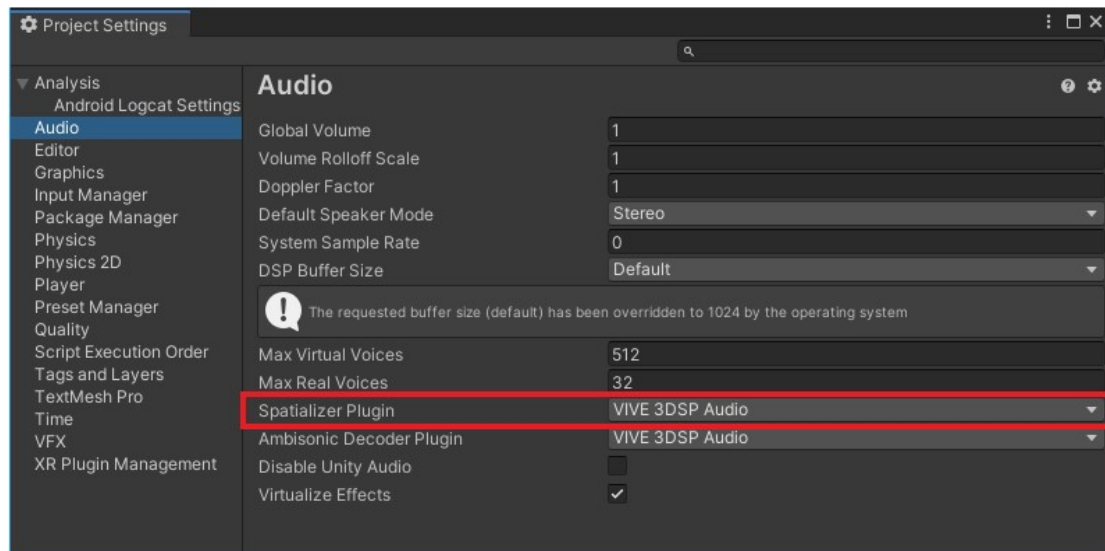
1. Download the latest version of VIVE 3DSP Audio Plugin in [website](#).
2. Create a new 3D project in Unity.
3. Click **Assets > Import Package > Custom Package** .
4. Select the latest version of **Vive3DSP .unitypackage** file and click **Import**.

Upgrade to Latest Package

1. Close Unity project, if open.
2. In the Unity project, go to the **Assets** folder and delete the **Vive3DSP** folder.
3. Open the project and import the latest version of VIVE 3DSP Audio Plugin (follow step 3 and 4 from the [Import Latest Package into New Project](#) section).

Set Spatializer Plugin

Select **VIVE 3DSP Audio** in **Spatializer Plugin** .

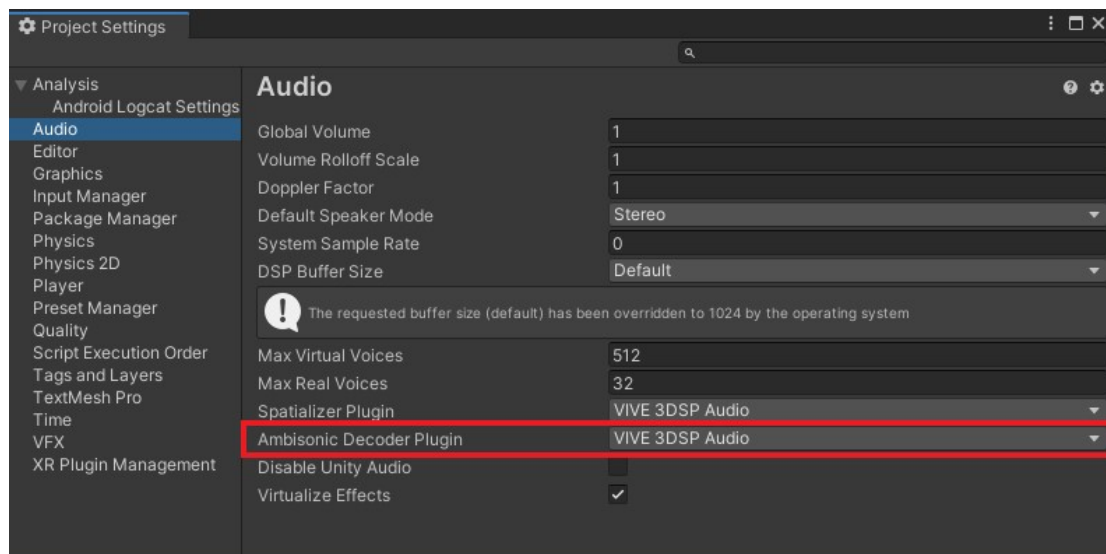


1. Click **Edit > Project Settings > Audio** .
2. Select **VIVE 3DSP Audio** in **Spatializer Plugin** .

Note: VIVE 3DSP support 3rd order ambisonics for spatializer.

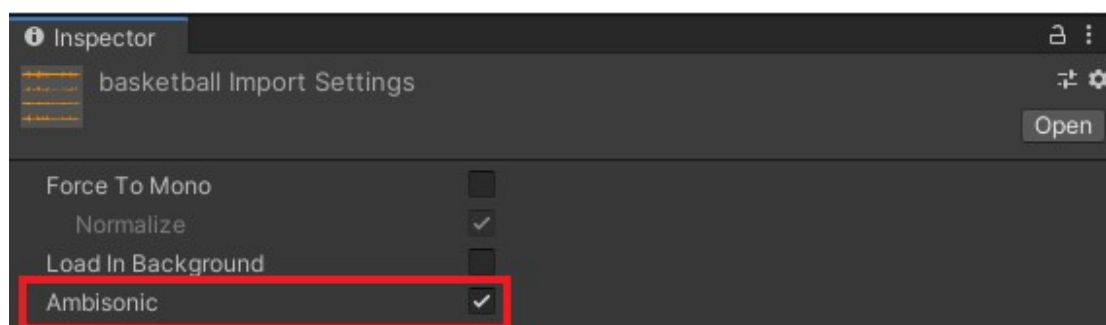
Set Ambisonic Decoder Plugin

Select **VIVE 3DSP Audio** in **Ambisonic Decoder Plugin** .



1. Click **Edit > Project Settings > Audio** .
2. Select **VIVE 3DSP Audio** in **Ambisonic Decoder Plugin** .

Note: VIVE 3DSP support 1st order ambisonic for decoder.



1. Put the AmbiX format audio file into your Unity Assets folder.

Note: The ambisonic file should be in **ACN/SN3D** format.

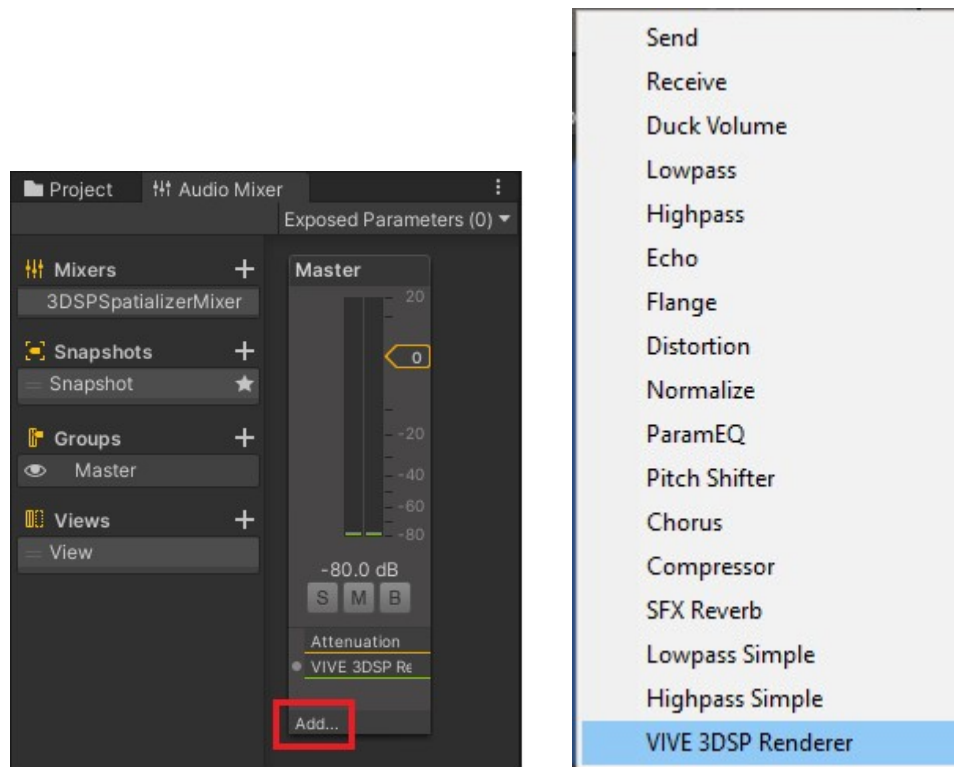
2. Find the file in the **Project** window and click it.
3. Select **Ambisonic** in the **Inspector** window and click **Apply**.
4. Add the ambisonic file as an **Audio Clip** to the Unity Audio Source.

Note: Make sure **Spatialize** in Audio Source is not selected.

Audio Mixer Settings

Add an audio mixer

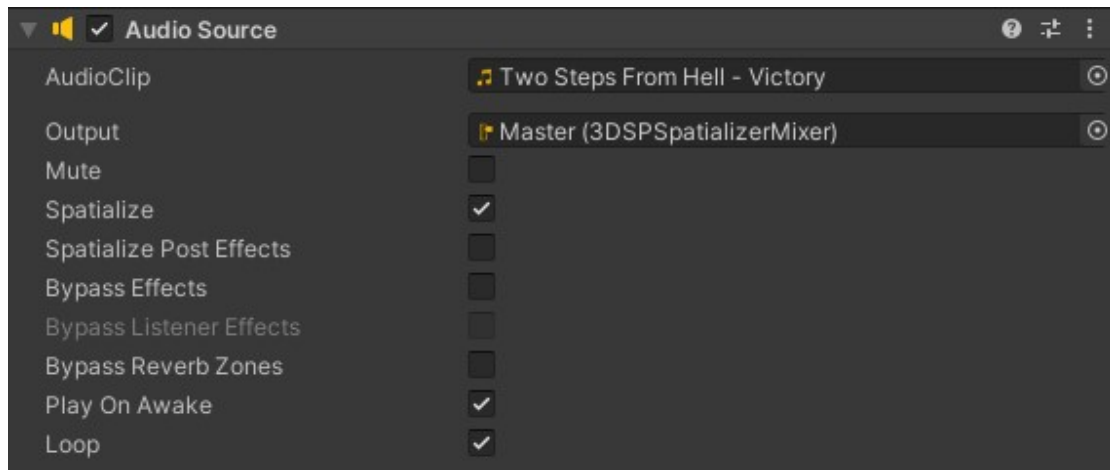
1. Click **Window > Audio > Audio Mixer** .
2. Click **+** to add a new audio mixer.
3. In the **Audio Mixer** tab, click **Add**, and then select **VIVE 3DSP Renderer** .



Audio Source Settings

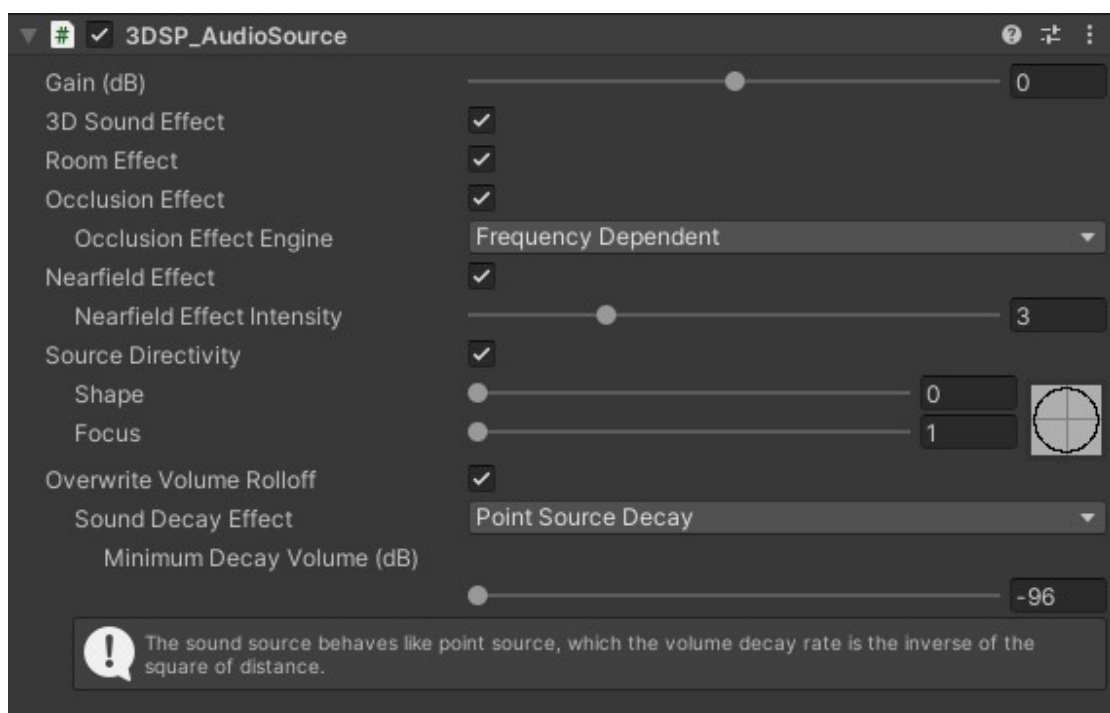
Set an object as the audio source

1. Do any of the following:
 - a. Attach the VIVE 3DSP Audio Source component to a Unity audio source object. Click **Add Component > VIVE > 3DSP_AudioSource**.
 - b. Attach the VIVE 3DSP Audio Source component to an object that contains a Unity Audio Source component.
2. Add an **Audio Clip** to the Unity Audio Source.
3. Set the audio mixer that includes **VIVE 3DSP Renderer** as the Unity Audio Source output.



Note: Don't select **Bypass Effects** and **Bypass Listener Effects** in the Unity Audio Source.

There are a lot of Unity default features in the Audio Source object. The VIVE 3DSP Audio Source component offers more features for the spatializer effect.



- **Gain (dB):** This option sets the audio clip volume level (in dB).
- **3D Sound Effect :** This option turns on or off the 3D Sound effect, which includes the directional sound feeling and sound decay properties.

- **Room Effect:** The option turns on or off the room audio application.
- **Occlusion Effect:** This option turns on or off the sound occlusion application.
- **Occlusion Effect Engine :** There are two types of occlusion engines that can be selected:
 - a. **Frequency Dependent :** The different material corresponds to the different sound frequency responses during sound transmission. The sound decay caused by objects is different based on frequencies. The real-world material response is modelled in this application.
 - b. **Frequency Free:** The sound decay influence of an obstacle is simulated by the volume decay only and the computing cost is minimal.
- **Nearfield Effect:** This option turns on or off the nearfield effect—which performs the sound response—when the sound source is near the listener. The nearfield effect is only active when the distance between the source and listener is **less than one meter**.
 - Nearfield Effect Intensity:** This option represents the strength of the nearfield effect. The higher the value, the stronger the nearfield effect will be.
 - Note:** DRC is enabled by default to help prevent audio signal clipping.
- **Source Directivity:** This option turns on or off the source directivity effect which simulates the sound direction and response when the pattern changes.
 - a. **Shape:** This option presents the shape of the source directivity pattern.
 - b. **Focus:** This option presents the sharpness of the source directivity pattern.
- **Overwrite Volume Rolloff :** If this feature is selected, the Unity built-in volume roll off effect will be disabled, and the VIVE 3DSP Sound Decay Effect will be used instead. Several types of **Sound Decay Effect** can be selected:
 - a. **Point Source Decay :** This is the sound decay effect that is based on inverse-square law. It behaves as the point source where the sound energy decreased ratio is based on the propagation area (the square of the propagation distance).
 - b. **Line Source Decay :** This is the sound decay effect that is based on inverse law. It behaves as the line source where the sound energy

decreased ratio is based on the propagation distance.

- c. **Linear Decay:** It is the customized sound decay effect. The volume decay is linearly proportional to the distance between the source and listener. These parameters can be adjusted:
 - **Minimum Distance (M)** : When the distance is lower than the minimum distance, the volume decay rate is 1.0.
 - **Maximum Distance (M)** : The volume decay rate will be the **Minimum Decay Volume (dB)** if the distance is longer than the **Maximum Distance** .
- d. **No Decay:** No sound decay effect will be applied in this audio source.

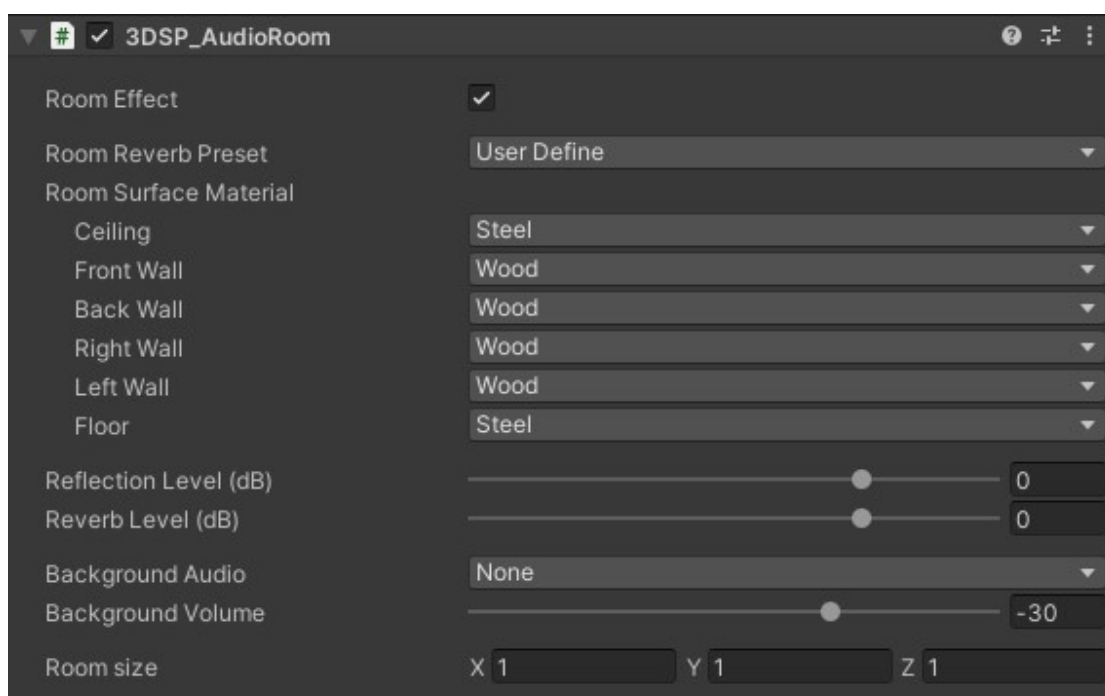
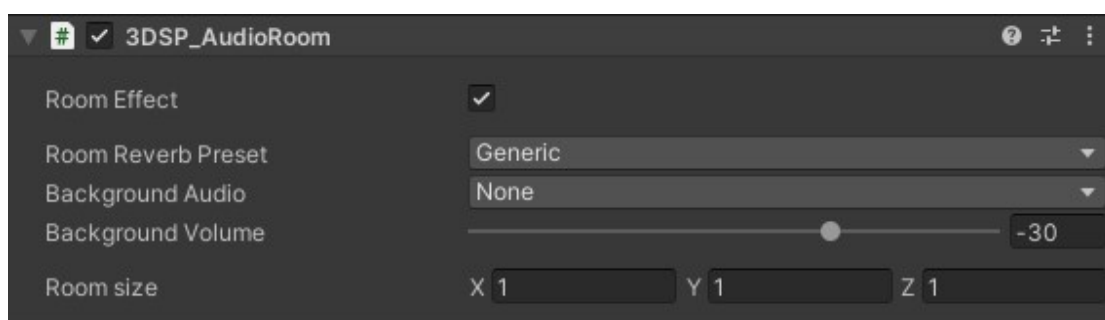
Notes:

- When **Overwrite Volume Roll-off** is selected, the default Unity distance curve will be set to a constant, 1.0.
- **Minimum Decay Volume (dB)** can be set in this application.

Audio Room Settings

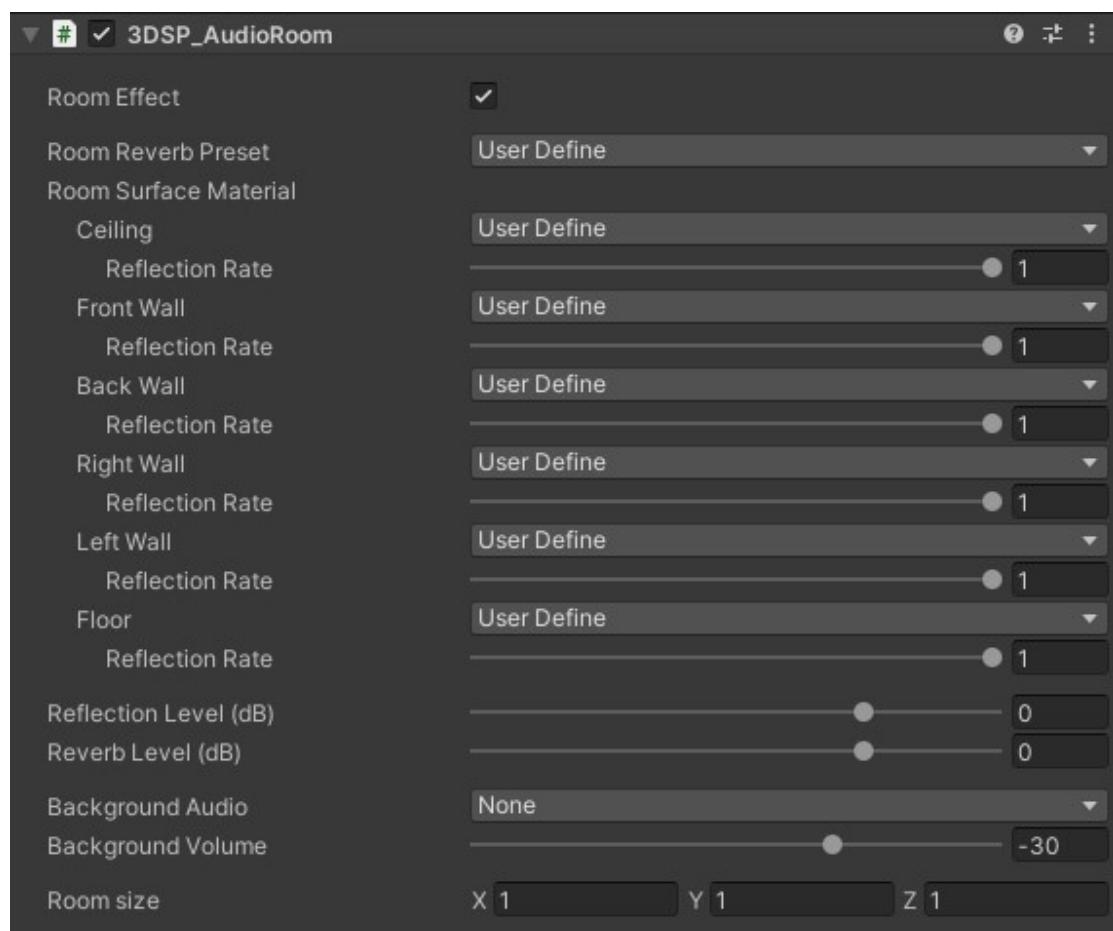
In VIVE 3DSP Audio Room, the reverberations can be represented via specific room sound effects. All audio sources inside a VIVE 3DSP Audio Room object are affected by the room properties such as size, materials, etc.

Specifically, several often-used room presets are provided in VIVE 3DSP Audio Room. Moreover, if **User Define** is selected, a list of materials and sound properties can be customized for the walls. This component can be selected via **Add Component > VIVE > 3DSP_AudioRoom**.

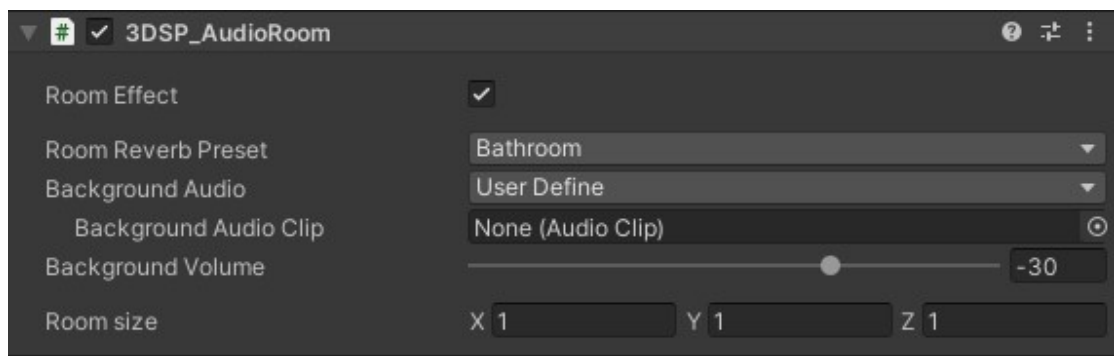


- **Room Effect:** This option can be turned on to apply the room reverberation effect.

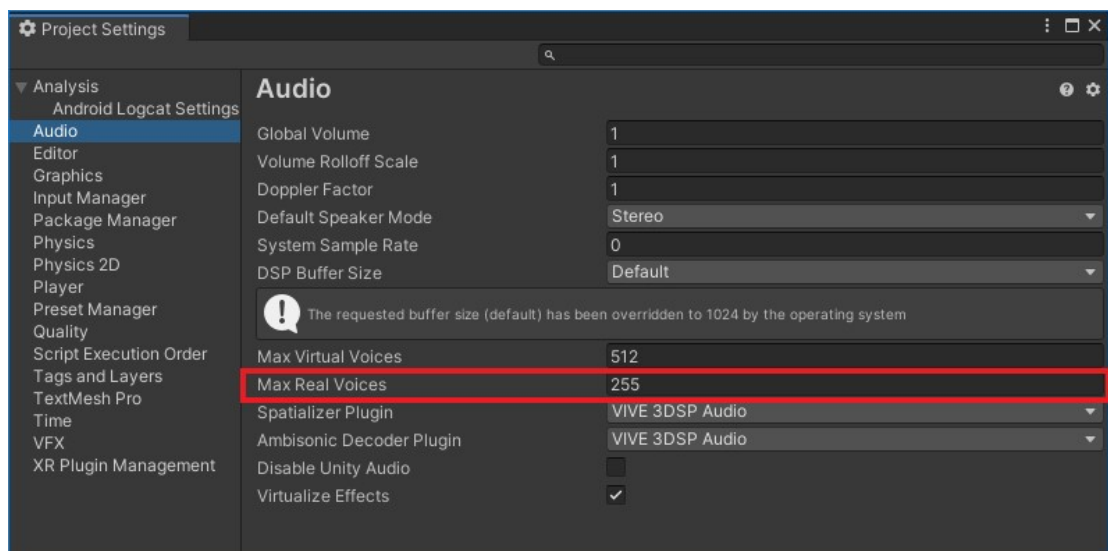
- **Room Surface Material** : This option shows the material of the walls. It will change the ingredients of the reflected sound. There are pre-defined [material presets](#) for each wall. A customized material can be chosen via **User Define** , and the parameter of the sound reflection ability of the wall can be set via **Reflection Rate** .



- **Reflection Level** : This option sets the strength of the early reflections, it ranges from -30 dB to 10 dB.
- **Reverb Level** : This option sets the strength of the late reverberations, it ranges from -30 dB to 10 dB.
- **Background Audio** : The background audio presents the sound effect of the ambient sounds in the environment. There are several default types in the background audio option. The preferred audio file can be chosen in **User Define** .



Note: The background audio is counted as one audio source in Unity. The limitation of the quantity of audio sources in Unity is adjustable in AudioManager settings. The maximum is 255 (setting in **Max Real Voices**).

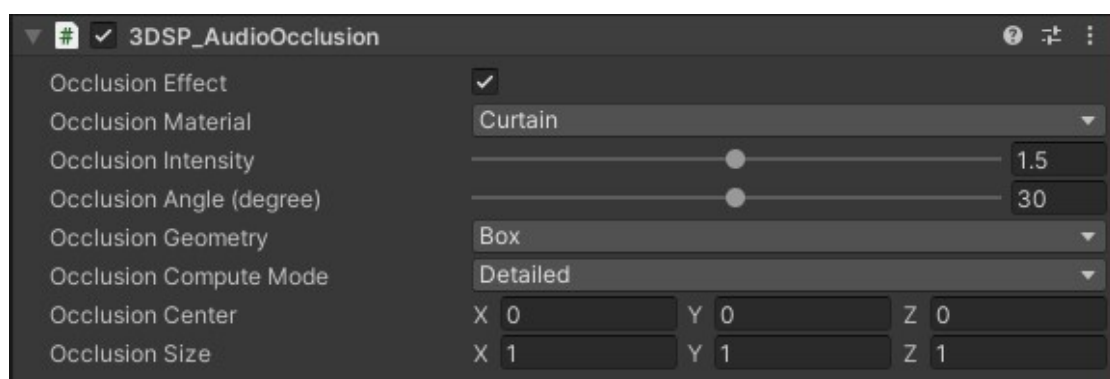


- **Background Volume** : This option sets the strength of the Background Audio, it ranges from -40 dB to 96 dB.
- **Room Size**: This option means the dimension of the room in meters. The factor, **Scale** in "Transform" of the attached object, is considered here.

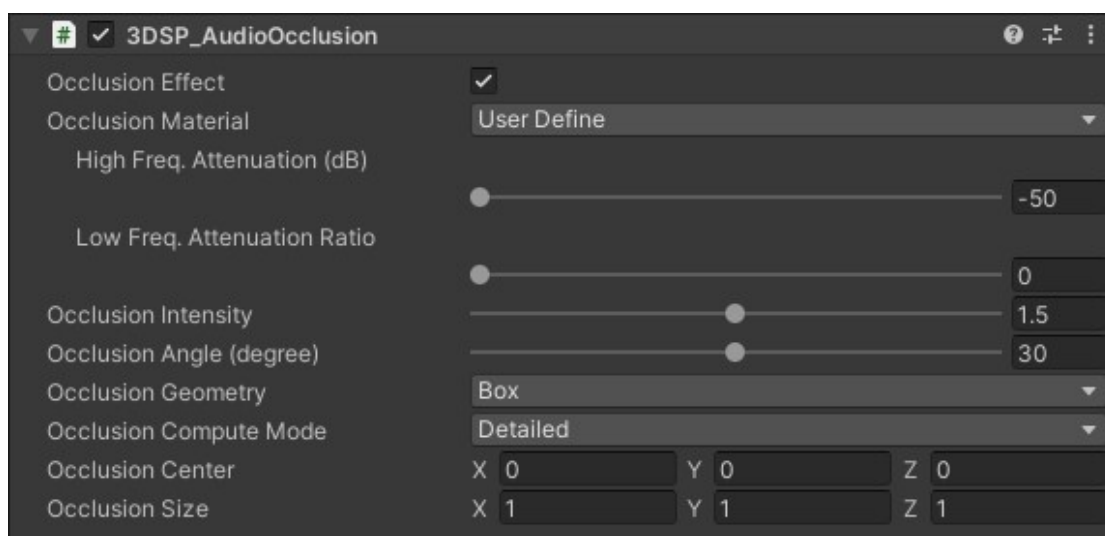
Audio Occlusion Settings

To have the occlusion effect, it is necessary to set the environment in both the VIVE 3DSP Audio Source and VIVE 3DSP Audio Listener. In audio source objects, the VIVE 3DSP Audio Occlusion component needs to be added to establish the sound occlusion effects.

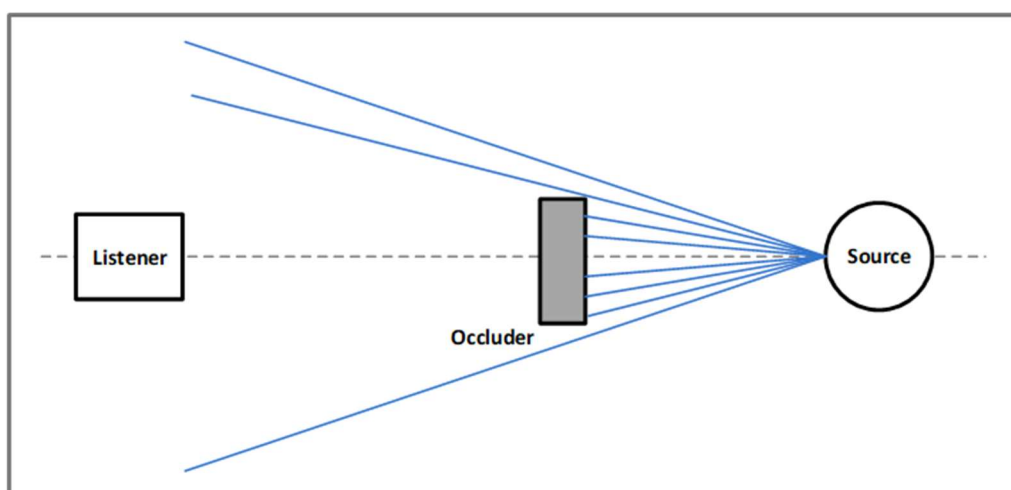
The VIVE 3DSP Occlusion component should be attached along with the object to enable the occlusion effect. This component can be selected via **Add Component > VIVE > 3DSP_AudioOcclusion**.



- **Occlusion Effect** : This option turns on or off the occlusion effect on the attached object.
- **Occlusion Material** : The material of an occluder can be changed here. There are some [material presets](#) in VIVE 3DSP. A customized material can be created via **User Define** by setting the **High Freq. Attenuation (dB)** and **Low Freq. Attenuation Ratio** .



- **Occlusion Intensity** : This option presents the strength of the occlusion effect. The higher the value, the more occlusion effect is applied.
- **Occlusion Angle** : This option presents the cover range of the occlusion effect. The higher the value, the more occlusion cover range is applied.
- **Occlusion Geometry** : The different shapes of the occlusion object can be set here. For calculating the occlusion effect, the geometrical information of the occluder is necessary. The values for Occlusion Center, Size, and Radius need to be entered.
- **Occlusion Compute Mode** : This is designed specifically for box-type geometric occlusion. It offers an option to select different computing modes to balance the coverage resolution and computation load.



- Very Light**: In this mode, it requires the least amount of calculation for geometric occlusion. There is only one ray emitted from the source to

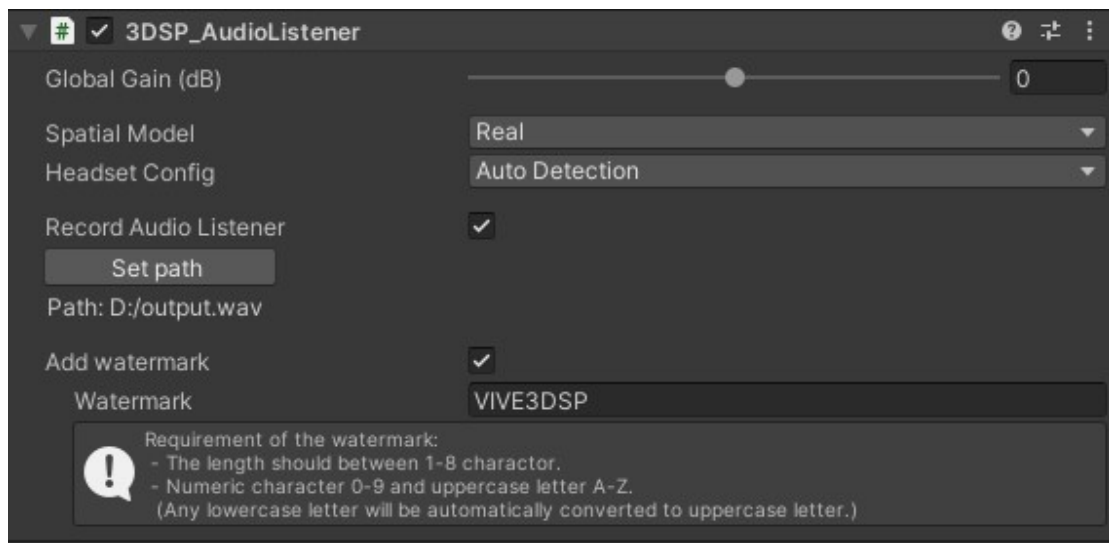
the listener, then evaluates whether the ray collided with the occlusion object or not.

- b. **Simple**: There are 6 different rays emitted from the source to the listener in this mode. The occlusion ratio is evaluated by the number of rays that collide with the occlusion object.
 - c. **Normal**: There are 11 different rays emitted from the source to the listener in this mode. The occlusion ratio is evaluated by the number of rays that collide with the occlusion object.
 - d. **Detailed**: This is the most accurate mode for occlusion coverage calculation. It also takes subtle movements of the occlusion object into consideration when determining the extent of occlusion.
- **Occlusion Center** : This option means the center coordinates relative to the attached object.
 - **Occlusion size** : This option means the dimension of the occluded region in meters. The factor, **Scale** in “Transform” of the attached object, is considered here.

Note: The Occlusion component has its own collider calculator. The Unity collider component does not need to be added.

Audio Listener Settings

The VIVE 3DSP Audio Listener component should be attached along with the Unity Audio Listener to enable some 3D sound effects. This component can be selected via **Add Component > VIVE > 3DSP_AudioListener**.



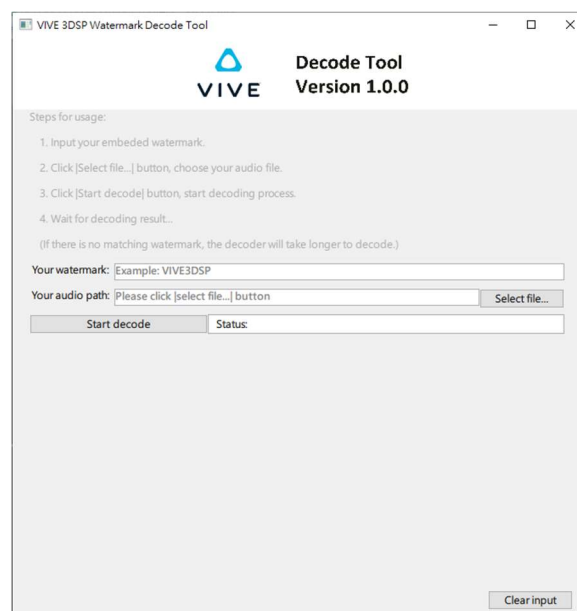
- **Global Gain**: This option sets the overall volume changes of the audio output from VIVE 3DSP Audio Listener. It ranges from **-24 dB** to **24 dB**.
- **Spatial Model**: This option controls the spatializer model.
 - a. **Real**: This options performs the sound response that is measured at the different direction from the real world.
 - b. **Lossless**: This options performs the sound response that represents the characters of the spatializer without any distortion.
- **Headset Config**: The option controls the targeted headset model. For the best audio experience, sound optimization has been done for most headset models.
 - a. **Generic**: Select this option if using non-VIVE VR devices and headphones.
 - b. **Auto Detection**: Audio response is optimized for VIVE VR devices. When enabled, the system will automatically detect the type of VIVE device and provide the optimized audio experience for that device. The feature supports the following VIVE devices: VIVE Pro, VIVE Pro 2, VIVE Focus 3, and VIVE XR Elite.
- **Record Audio Listener**: This option is selected to record what the listener is hearing in the scene and exports it as an audio file. **Set path** is applied to set the file location. If the **Record Audio Listener** option is turned off during Play Mode or the Play Mode is stopped, an audio file will be exported in the WAV audio format.
- **Add watermark**: This option lets you add a watermark in the audio system of the application for intellectual property rights and for identification

purposes. You can find the watermark decoder in the VIVE 3DSP SDK package. For more details, refer to [Decoding watermarks](#).

Watermark: The watermark must be **eight** characters or less, and can only contain **numbers** (0–9) and **UPPERCASE** letters (A–Z). (Any lowercase letter will be automatically converted to uppercase letter.)

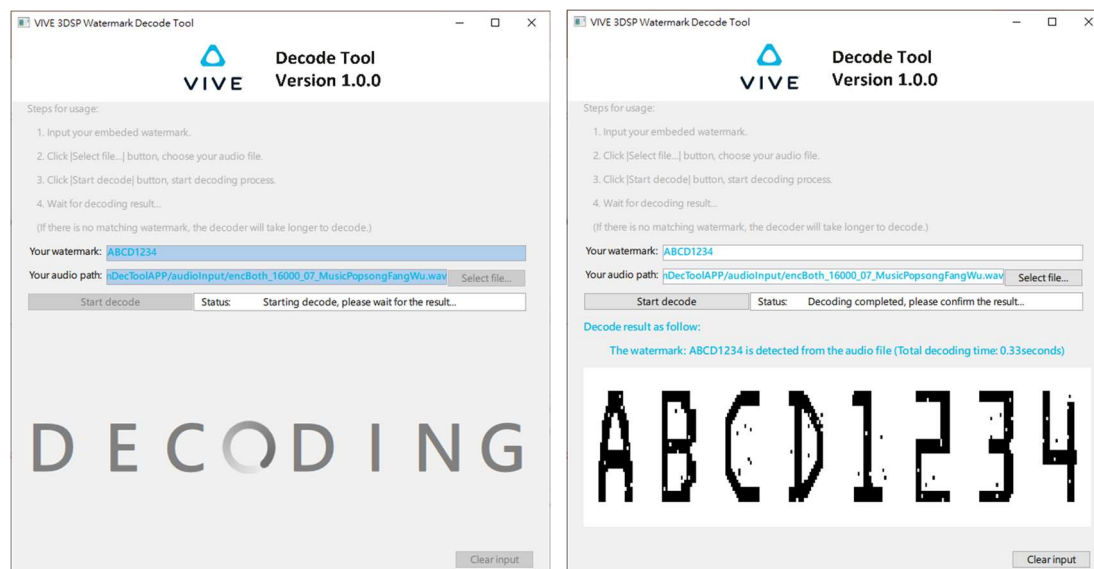
Decoding watermarks

The **VIVE 3DSP Watermark Decode Tool (appWmDecToolAPP.exe)** lets you check if there is a watermark in the audio file of the application. You'll need to prepare the audio file of the application first to be able to decode the watermark.



Decode the watermark

1. In **Your watermark** , enter the watermark that you used for the audio of the application.
2. In **Your audio path** , select the audio file of application.
3. Click **Start decode** . The status of the decoding process will then be shown on the window. The decoding process may take several minutes.

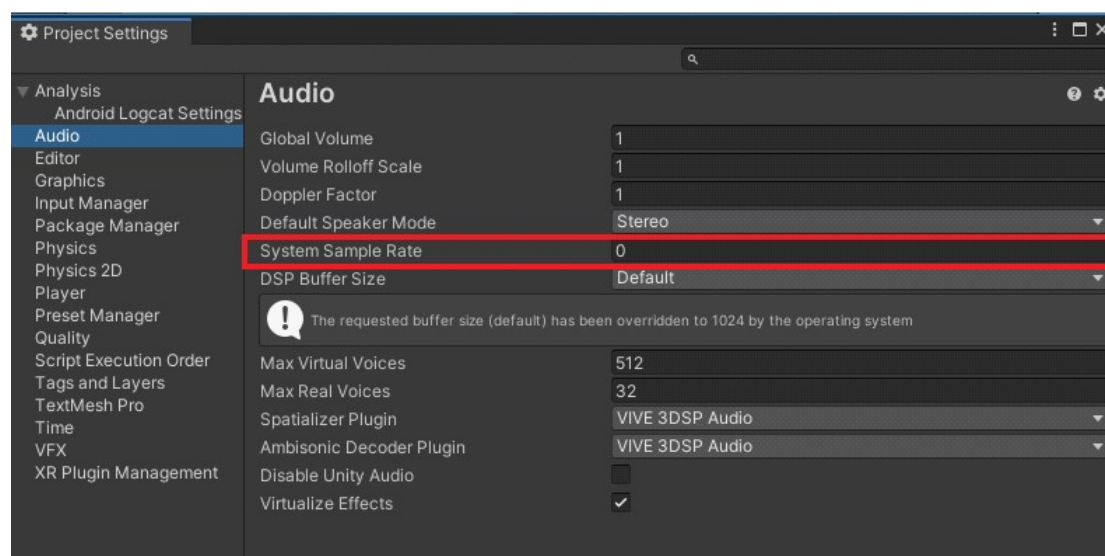


If the watermark entered is found in the audio file, it will be displayed on the window.

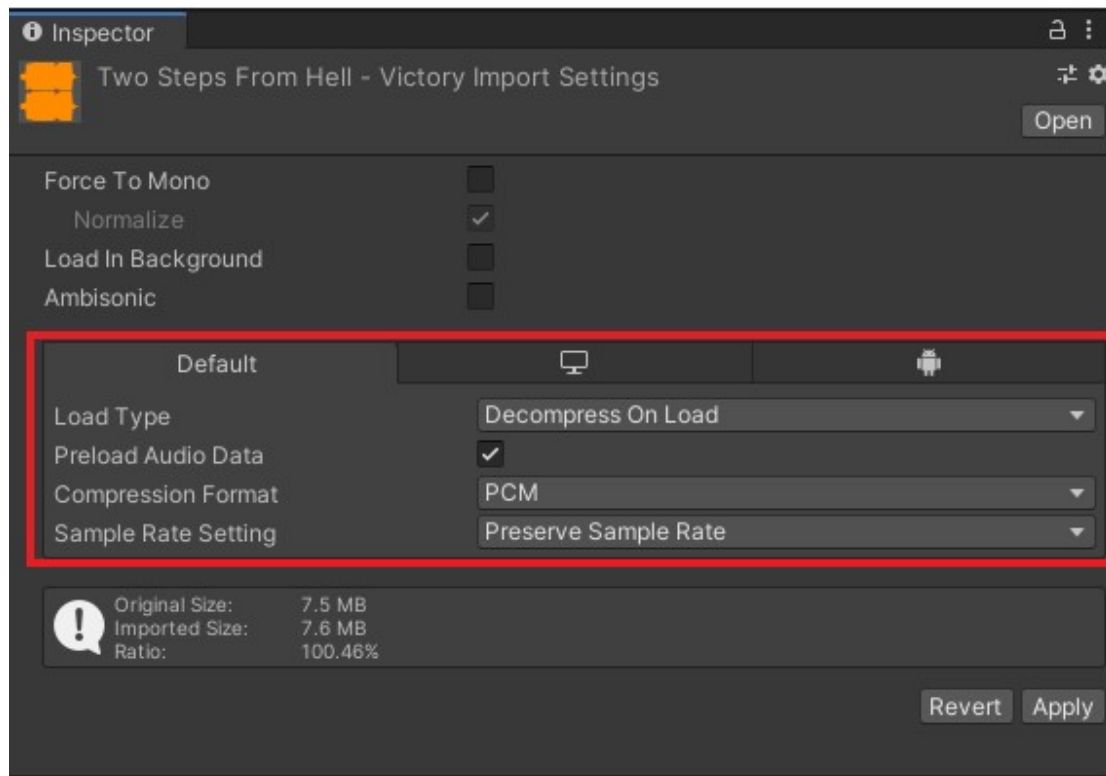
Hi-Res Audio Settings

VIVE 3DSP supports Hi-Res audio file for the audio source. This feature can be set by:

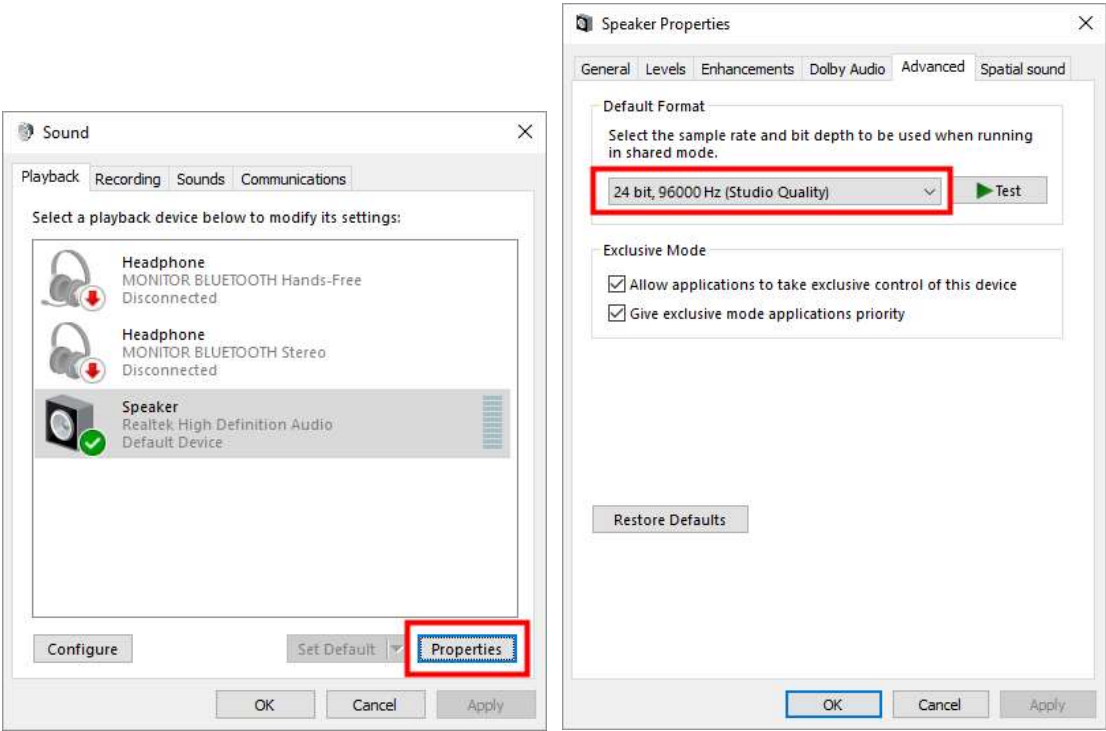
1. Set **System Sample Rate** to **96000** (maximum sample rate of Unity) by clicking **Edit > Project Settings > Audio** .



2. Modify the audio clip format by changing **Compression Format** to PCM in the **Inspector** window .



3. Modify the operating system's audio settings. Change the sample rate to **96000 Hz** in the system sound advanced settings.



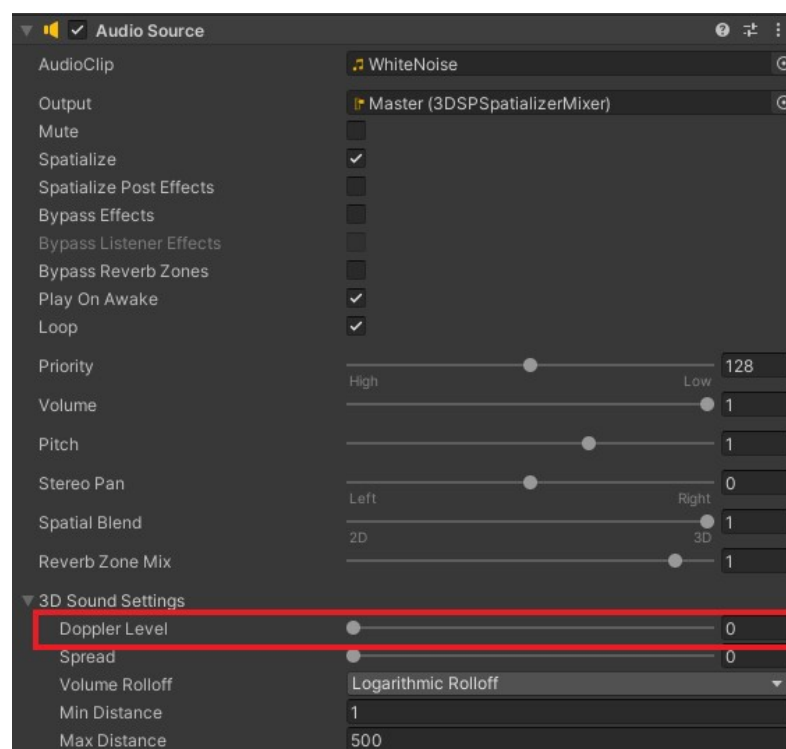
Troubleshooting

- **Memory consumption keeps increasing in Unity Editor**

There is an issue in Unity editor in the Unity version 2018 series. If the game object with a Unity **Audio Source** in a scene has been chosen, and the scene has been shown in the Unity Editor, the memory usage of Unity will sometimes keep increasing. This happens because if the sphere object is shown in the Unity scene, it keeps generating dummy meshes. This issue can be fixed by restarting Unity to delete the dummy meshes. Unity will crash or audio clippings will likely occur if the memory usage is too high.

- **Unity Built-in Doppler Effect might cause audio distortion**

Unity provides Doppler Effect and this can be found in the audio source object. However, the audio stream might be distorted by this built-in effect if the audio source object is moving very fast. For example, a quarter of the speed of sound, the audio stream could occasionally be rugged, therefore audio clipping occurs.



Appendix

Material preset list

Number	Material
1	Concrete
2	Porous Concrete
3	Marble
4	Brick
5	Glass
6	Wood
7	Gypsum
8	Curtain
9	Plywood
10	Steel
11	Stone Wall
12	Plaster Absorber_Concrete
13	Plaster Absorber_Brick
14	Plaster Absorber_Wood
15	Plaster Absorber_Gypsum
16	Timberboard_Concrete
17	Timberboard_Brick
18	Timberboard_Wood
19	Timberboard_Gypsum
20	Ceramic Tile_Concrete
21	Ceramic Tile_Brick
22	Ceramic Tile_Wood
23	Ceramic Tile_Gypsum
24	Corkboard_Concrete
25	Corkboard_Brick
26	Corkboard_Wood
27	Corkboard_Gypsum
28	Soft Curtain_Concrete
29	Soft Curtain_Brick

30	Soft Curtain_Wood
31	Soft Curtain_Gypsum
32	Curtain_Concrete
33	Curtain_Brick
34	Curtain_Wood
35	Curtain_Gypsum
36	Cotton_Concrete
37	Cotton_Brick
38	Cotton_Wood
39	Cotton_Gypsum
40	Plaster_Concrete
41	Plaster_Brick
42	Plaster_Wood
43	Plaster_Gypsum
44	Plywood_Concrete
45	Plywood_Brick
46	Plywood_Wood
47	Plywood_Gypsum
48	Glazed Tile_Concrete
49	Glazed Tile_Brick
50	Glazed Tile_Wood
51	Glazed Tile_Gypsum

VIVE 3DSP Audio SDK Forum Shortcut

If you have suggestions or questions, you can go to the [VIVE Audio SDK forum](#). It can be accessed by clicking **Window > VIVE 3DSP Audio SDK > Visit VIVE 3DSP Forum**.

