

VocalFusion Stereo Dev Kit Quick Start Guide

IN THIS DOCUMENT

- ▶ Before you start
 - ▶ Setting up the VocalFusion Stereo Dev Kit
 - ▶ Evaluation
 - ▶ Direction of Arrival indication
 - ▶ Next steps
 - ▶ Useful Information
 - ▶ Troubleshooting
-

Welcome to the **VocalFusion Stereo Dev Kit** quick start guide.

The VocalFusion Stereo Dev Kit demonstrates the voice capture and processing capabilities of the XMOS XVF3500.

The kit includes a small form-factor linear microphone array which provides up to 180° capture for 'edge of the room' applications such as smart TV's, sound bars, set-top boxes and other near-to-wall products.

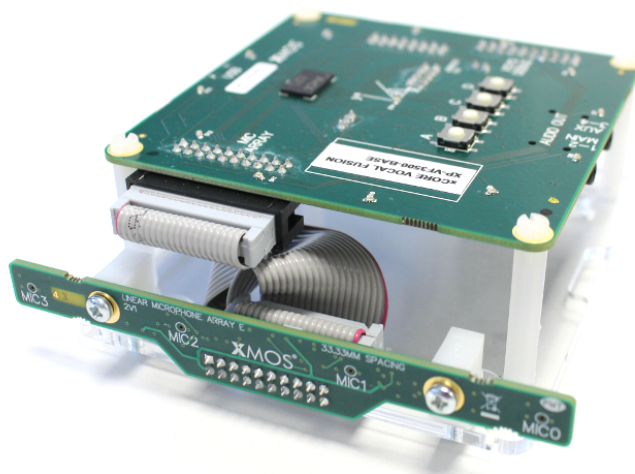


Figure 1:
VocalFusion
Stereo Dev
Kit

This guide will explain how to setup and use the **VocalFusion Stereo Dev Kit**:

- ▶ We will use a USB connection from a host PC to playback audio content and record both the raw microphone data and the extracted voice signal.
- ▶ We will demonstrate **VocalFusion Stereo's** class-leading voice isolation capabilities by examining the extracted voice signal in *Audacity* and comparing it against the raw microphone data.
- ▶ Finally, we will demonstrate the voice activity detector and direction-of-arrival functionality of the XVF3500.

1 Before you start

To complete this guide, you will need:

- ▶ **VocalFusion Stereo Dev Kit (XK-VF3500-L33).**

Note that base board revision 1V1 or higher, as indicated on the silk screen next to the 'JGB' code, is required.

- ▶ A host laptop/PC (Windows, Mac or Linux) with some stereo music content and two free USB ports.
- ▶ **Powered stereo speakers** with audio input via a 3.5mm analogue plug. (See notes at the end of this document for speaker recommendations.)
- ▶ Audio recording and editing software Audacity® installed on the host PC. This is freely available from:
 - ▶ <http://www.audacityteam.org>

2 Setting up the VocalFusion Stereo Dev Kit

For performance that will be representative of real-world conditions the following setup is recommended:



1. Place the VocalFusion Stereo Dev Kit on a horizontal surface. For example, on a table towards the edge of the room.
2. Place the powered speakers in the near-field next to (within 30cm) and behind the VocalFusion Stereo Dev Kit. Avoid pointing the speakers directly at the microphone array.
3. Connect the powered speakers to the 3.5mm **1-2 MAIN** socket on the VocalFusion Stereo Dev Kit.
4. Connect the VocalFusion Stereo Dev Kit to the host PC, using the supplied USB A to Micro B cable.
5. All done! The VocalFusion Stereo Dev Kit is now ready to use.

3 Evaluation

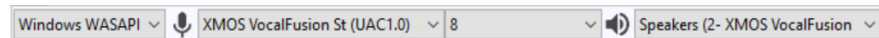
The VocalFusion Stereo Dev Kit and the speakers should remain static during your evaluation. If they are moved to new positions, then the adaptive algorithms will adjust to the new audio environment after approximately 10 seconds.

1. Open a music player on the host PC, select **XMOS VocalFusion St (UAC1.0)** as the playback device and play a stereo music file.

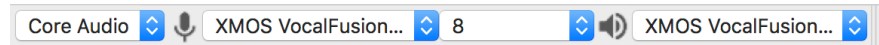
You should now hear this through your powered speakers. You can adjust the volume using either the music player, the **XMOS VocalFusion St (UAC1.0)** playback device, or the speakers.

2. Open *Audacity* on the host PC and configure the application to communicate with the VocalFusion Stereo Dev Kit.

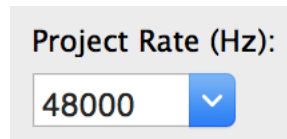
Windows users



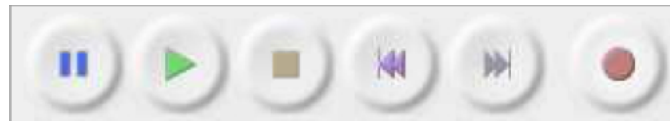
Mac users



3. Ensure that the number of recording channels is set to 8.
4. Set the project sample rate to 48kHz



5. In *Audacity*, click on the **Record** button (or press **r**) to start capturing the audio channels streamed from the VocalFusion Stereo Dev Kit.



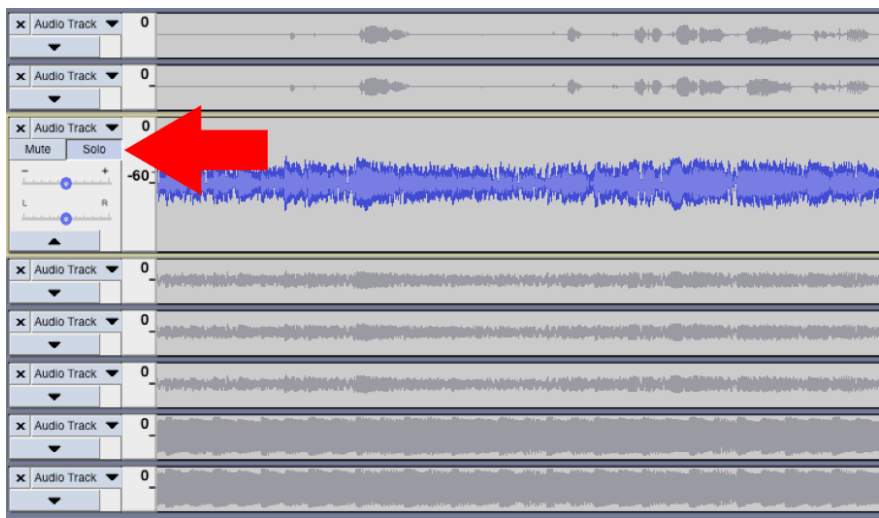
Audacity will record eight audio channels streamed from the VocalFusion Stereo Dev Kit:

- ▶ Audio Track 1: extracted voice signal, optimised for the human ear.
- ▶ Audio Track 2: extracted voice signal, optimised for Automated Speech Recognition engines.
- ▶ Audio Track 3: raw microphone 0.
- ▶ Audio Track 4: raw microphone 3.
- ▶ Audio Track 5: raw microphone 1.
- ▶ Audio Track 6: raw microphone 2.
- ▶ Audio Track 7: playback reference signal (left).

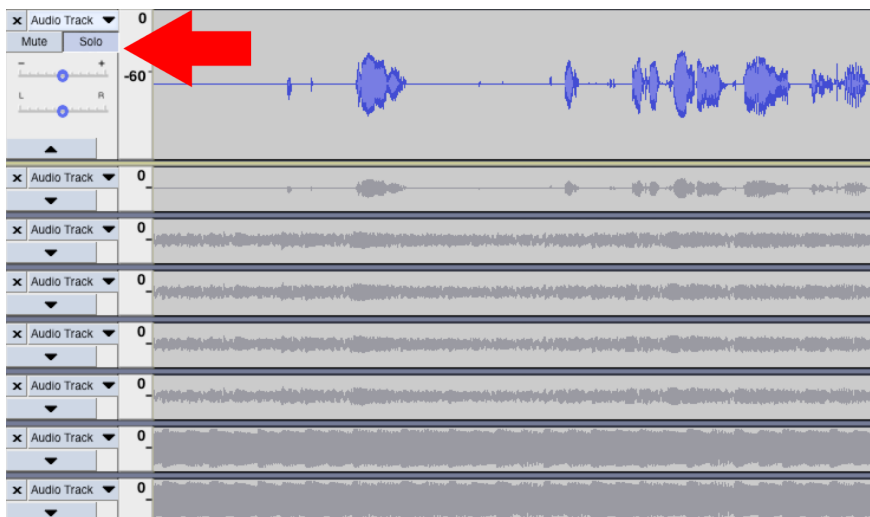
- ▶ Audio Track 8: playback reference signal (right).
6. Talk over the music content. Move around the room and continue talking.
 7. Stop playback from the music player.
 8. In *Audacity*, click on the **Stop** button (or press **space**) to stop recording.



9. In *Audacity*, highlight Audio Track 3 and click the **Solo** button (or press **SHIFT + s**) to only playback Audio Track 3.



10. In *Audacity*, click on the **Play** button (or press **space**) to playback Audio Track 3 only.
You will now hear the raw signal as captured by one of the microphones – a combination of the playback signal, your voice and any other background noise.
11. In *Audacity*, highlight Audio Track 1 and click the **Solo** button (or press **SHIFT + s**) to only playback Audio Track 1.



You will hear only your voice.

- ▶ **Stereo Acoustic Echo Cancellation** has removed the playback music.
- ▶ **Adaptive voice beam-forming** has isolated your voice.
- ▶ **Noise suppression** and **de-reverberation** has removed background noise.

4 Direction of Arrival indication

The XVF3500 includes a **Direction of Arrival (DOA)** indicator. When the VocalFusion Stereo Dev Kit detects sound above a certain threshold, it will illuminate red LEDs at both the centre and along the edge of the board to indicate the DOA.

Try talking to the VocalFusion Stereo Dev Kit, both with and without music playing. Move across the front of the device as you do so. The central LED and an edge LED will illuminate when you talk, indicating the DOA of your voice.

5 Next steps

If you wish to take further steps in evaluating the VocalFusion Stereo Dev Kit, please refer to the **xCORE VocalFusion Control Users Guide**:

<https://www.xmos.com/published/vocalfusion-control-guide>

The **xCORE VocalFusion Control Users Guide** describes how to read and update the user configurable parameters on the VocalFusion Stereo Dev Kit.

6 Useful Information

6.1 Speaker Selection

The choice of stereo speakers used in your evaluation is important and can greatly affect overall system performance.

- ▶ The amplifier in the speakers should have linear gain. Non-linear gain (e.g. soft clipping) should be disabled or avoided.
- ▶ Any audio processing available on the speakers should be disabled.
- ▶ For low quality speakers it is best to use low volume settings to avoid non-linear distortions of the reference signal.

The Logitech Z130 stereo speakers, for example, work well for evaluating the VocalFusion Stereo Dev Kit.

6.2 Windows audio drivers

Windows includes a USB Audio Class 1 (UAC1) driver and so natively supports the VocalFusion Stereo Dev Kit.

When a configured dev kit is connected to a Windows PC, the native UAC1 driver is installed and the device details are stored in Windows' USB cache. If subsequently a different version of the dev kit is connected (for example, a dev kit with a firmware using the same VID and PID but exposing a different configuration of audio channels) then the Windows' USB cache is not updated and the USB audio connection to the dev kit will not operate correctly.

If this occurs, disconnect the dev kit and use the *USBDeview* utility (freeware available from http://www.nirsoft.net/utils/usb_devices_view.html) to uninstall all instances of XMOS USB audio devices. Then re-connect the dev kit; fresh instances of the UAC1 driver will then be correctly installed.

7 Troubleshooting

This section will guide you through reloading the XVF3500 firmware.

1. Install the XMOS xTIMEcomposer tools on the host PC. These are freely available from:
 - ▶ <https://www.xmos.com/tools>
2. Download the latest XVF3500 firmware binaries from:
 - ▶ <https://www.xmos.com/support/software/vocalfusion>A MyXMOS account is required to complete this step.
3. Connect the dev kit to the host PC, using the supplied USB A to Micro B cable.
4. Connect the xTAG debug adaptor to the host PC.

5. Connect the xTAG debug adaptor to the dev kit using the XSYS DEBUG socket.
6. Flash the XVF3500 firmware binary on to the dev kit.

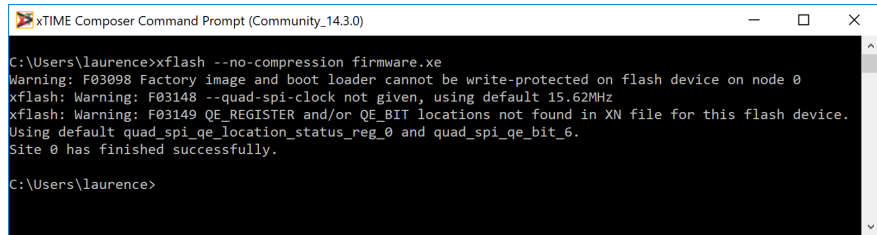
Windows users:

Open an xTIMEcomposer command prompt:

Start ► XMOS ► Command Prompt

Navigate to your folder containing the downloaded firmware binary and enter:

```
xflash --no-compression <filename>.xe
```

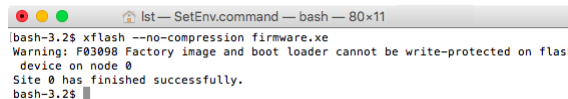
**Mac users:**

Open an xTIMEcomposer terminal session and run:

Finder ► Applications ► XMOS_xTIMEcomposer ► SetEnv.command

Navigate to your folder containing the downloaded firmware binary and enter:

```
xflash --no-compression <filename>.xe
```



7. Wait for the flashing process to complete. When complete, the dev kit will re-enumerate as a composite USB Audio Class 1 device.
8. All done! You may now close the xTIMEcomposer command prompt and unplug the xTAG debug adaptor from both the PC and the dev kit.



Copyright © 2018, All Rights Reserved.

Xmos Ltd. is the owner or licensee of this design, code, or Information (collectively, the "Information") and is providing it to you "AS IS" with no warranty of any kind, express or implied and shall have no liability in relation to its use. Xmos Ltd. makes no representation that the Information, or any particular implementation thereof, is or will be free from any claims of infringement and again, shall have no liability in relation to any such claims.