

xCORE VocalFusion Speaker Evaluation Kit Quick Start Guide

IN THIS DOCUMENT

- ▶ Before you start
 - ▶ Load XVF3100 firmware
 - ▶ Setup
 - ▶ Evaluation
 - ▶ Voice Activity Detector
 - ▶ Keyword detection
 - ▶ Direction of Arrival indication
 - ▶ Useful information
-

Welcome to the **xCORE VocalFusion Speaker Evaluation Kit** quick start guide.

The xCORE VocalFusion Speaker Evaluation Kit demonstrates the voice capture and processing capabilities of the XMOS XVF3100 device in a 'smart speaker' application.

There are two versions of the VocalFusion Speaker Evaluation Kit available:

- ▶ xCORE VocalFusion Speaker circular kit (**XK-VF3100-C43**)
 - ▶ Providing 360° capture, for 'centre of the room' applications
- ▶ xCORE VocalFusion Speaker linear kit (**XK-VF3100-L33**)
 - ▶ Providing up to 180° capture, for 'edge of the room' applications

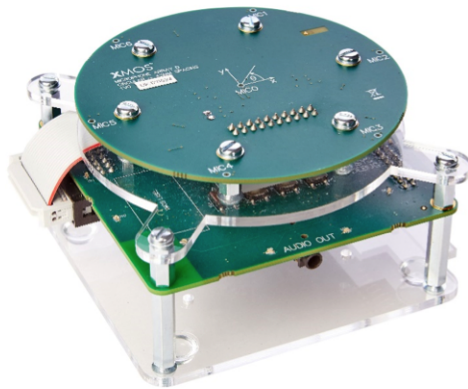
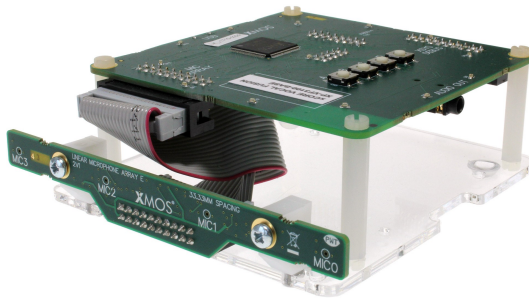


Figure 1:
xCORE
VocalFusion
Speaker
circular kit
(XK-VF3100-
C43)

Figure 2:

xCORE
VocalFusion
Speaker
linear kit (XK-
VF3100-L33)



This guide will explain how to setup and use the evaluation kit:

- ▶ We will use a USB connection from a host PC to playback audio content and to record both the raw microphone data and the extracted voice signal.
- ▶ We will demonstrate the XVF3100's class-leading voice isolation capabilities by comparing the extracted voice signal against the raw microphone data.
- ▶ Finally, we will demonstrate the voice activity detector, the always-listening keyword detection and direction-of-arrival functionality of the XVF3100.

1 Before you start

To complete this guide, you will also need:

- ▶ A mono powered speaker, with analogue audio input via a 3.5mm plug. (See notes at the end of this document for speaker recommendations.)
- ▶ A host laptop/PC (Windows, Mac or Linux) with internet access, some music content and a couple of free USB ports.
- ▶ XMOS xTIMEcomposer tools installed on the host PC. These are freely available from:
 - ▶ <http://www.xmos.com/tools>Use your existing MyXMOS login or create a new account to complete this step.
- ▶ Audio recording and editing software Audacity® installed on the host PC. This is freely available from:
 - ▶ <http://www.audacityteam.org>

2 Load XVF3100 firmware

The first step is to load firmware on to the evaluation kit. In this quick start guide we will use the simple command-line utility `xf1ash` which is provided as part of the xTIMEcomposer tools.

1. Download the latest XVF3100 evaluation firmware binaries from:

- ▶ <http://www.xmos.com/vfspeaker>

A MyX MOS account is required to complete this step.

2. Connect the evaluation kit to the host PC, using the supplied USB A to microB cable.
3. Connect the xTAG debug adaptor to the host PC.

Note: the evaluation kit may be supplied with either an xTAG2 or an xTAG3 debug adaptor.

- ▶ xTAG2 uses a USB A extension cable.
- ▶ xTAG3 uses a USB A to microB cable.

4. Connect the xTAG debug adaptor to the evaluation kit; plug the xTAG in to the **XSYS DEBUG** socket.



5. Flash the XVF3100 firmware binary on to the evaluation kit.

- ▶ In this evaluation we will use the **VocalFusion Speaker Circular Evaluation Kit** and the firmware: `app_vf_spk_base_1i6o2_cir43_keyword.xe`

- For the **VocalFusion Speaker Linear Evaluation Kit** use the firmware:
app_vf_spk_base_1i602_133_keyword.xe

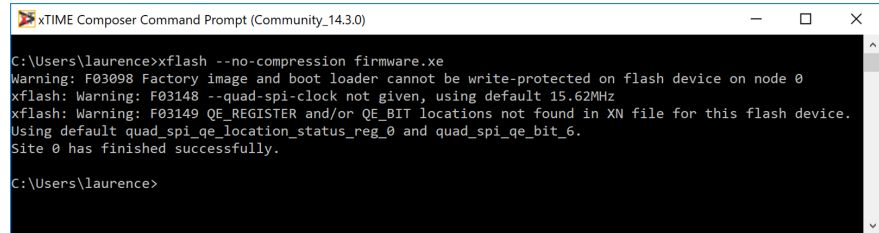
Windows users:

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
```



```
xTIME Composer Command Prompt (Community_14.3.0)
C:\Users\laurence>xflash --no-compression firmware.xe
Warning: F03098 Factory image and boot loader cannot be write-protected on flash device on node 0
xflash: Warning: F03148 --quad-spi-clock not given, using default 15.62MHz
xflash: Warning: F03149 QE_REGISTER and/or QE_BIT locations not found in XN file for this flash device.
Using default quad_spi_qe_location_status_reg_0 and quad_spi_qe_bit_6.
Site 0 has finished successfully.

C:\Users\laurence>
```

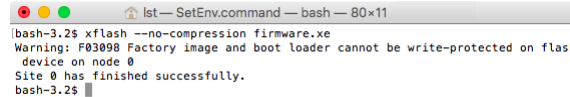
Mac users:

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
```

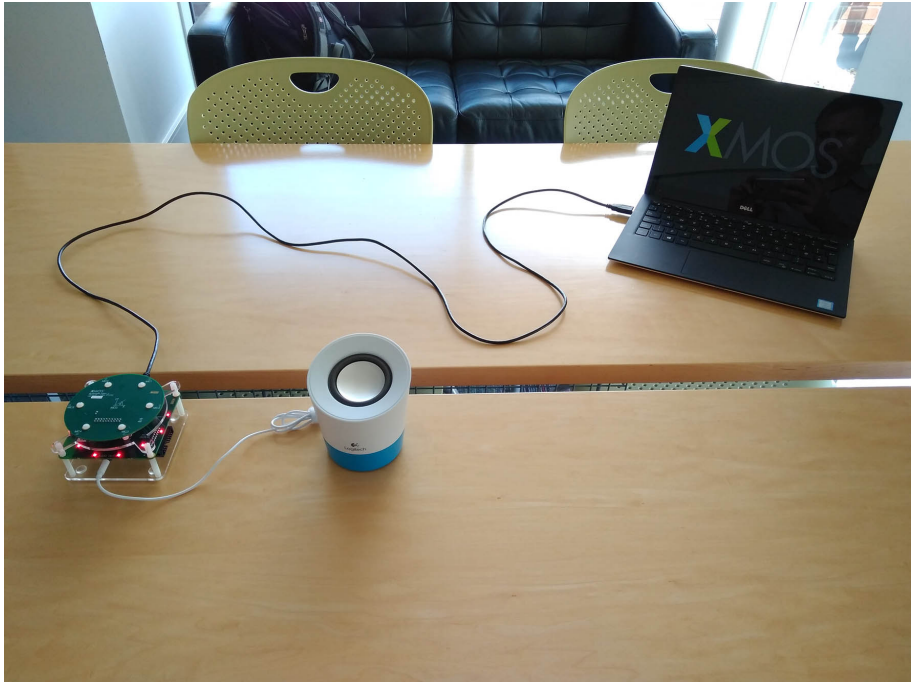


```
SetEnv.command — bash — 80x11
bash-3.2$ xflash --no-compression firmware.xe
Warning: F03098 Factory image and boot loader cannot be write-protected on flash
device on node 0
Site 0 has finished successfully.
bash-3.2$
```

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

3 Setup

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



1. Place the evaluation kit on a horizontal surface. For example:
 - ▶ If using the **xCORE VocalFusion Speaker Circular Kit**, place on a table in the middle of the room.
 - ▶ If using the **xCORE VocalFusion Speaker Linear Kit**, place on a table towards the edge of the room.
2. Place the powered speaker in the *near-field*:
 - ▶ For the **xCORE VocalFusion Speaker Circular Kit**, place the powered speaker next to (within 30cm), or under, the evaluation kit. Avoid pointing the speaker directly at the microphone array.
 - ▶ For the **xCORE VocalFusion Speaker Linear Kit**, place the powered speaker next to (within 30cm) and behind the evaluation kit. Avoid pointing the speaker directly at the microphone array.
3. Connect the powered speaker to the 3.5mm **AUDIO OUT** socket.

Note that the **AUDIO OUT** socket provides a mono output, on the right channel only. This output is a mono mix any of stereo content streamed to the evaluation kit. The left channel output will always be muted.

4. Connect the evaluation kit to the host PC, using the supplied USB A to microB cable.

We recommend using the supplied USB cable. This allows the host PC to be located in the *far-field*, i.e. at least 1 m from the microphone array.

5. All done! The xCORE VocalFusion Speaker Evaluation Kit is now ready to use.

4 Evaluation

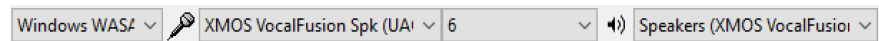
Both the evaluation kit and speaker should remain static during your evaluation. If they are moved to new positions, then the adaptive algorithms will adjust to the new audio environment after approximately 10 seconds.

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

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

2. Open *Audacity* on the host PC and configure the application to communicate with the evaluation kit.

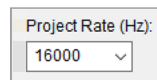
Windows users



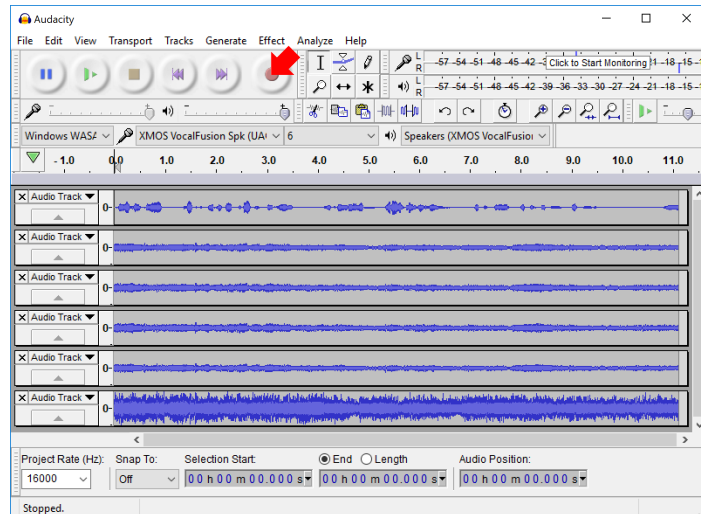
Mac users



3. Set the project sample rate to 16kHz



4. In *Audacity*, click on the **Record** button (or press **r**) to start capturing the audio channels streamed from the evaluation kit.



Audacity records the six audio channels streamed from the XMOS VocalFusion Speaker Evaluation Kit.

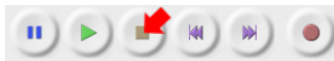
For the xCORE VocalFusion Speaker Circular Kit:

- ▶ Audio Track 1: voice, the extracted voice signal, optimised for the human ear
- ▶ Audio Track 2: raw microphone1 audio stream
- ▶ Audio Track 3: raw microphone3 audio stream
- ▶ Audio Track 4: raw microphone4 audio stream
- ▶ Audio Track 5: raw microphone6 audio stream
- ▶ Audio Track 6: mono playback audio stream

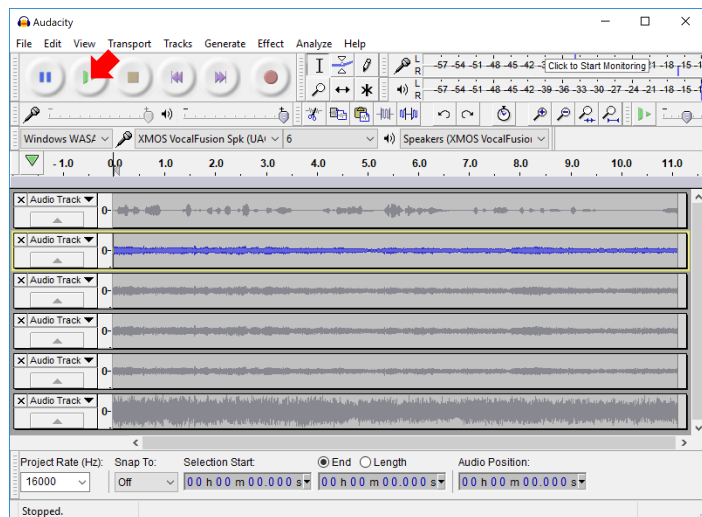
For the xCORE VocalFusion Speaker Linear Kit:

- ▶ Audio Track 1: voice, the extracted voice signal, optimised for the human ear
- ▶ Audio Track 2: raw microphone0 audio stream
- ▶ Audio Track 3: raw microphone1 audio stream
- ▶ Audio Track 4: raw microphone2 audio stream
- ▶ Audio Track 5: raw microphone3 audio stream
- ▶ Audio Track 6: mono playback audio stream

5. Talk over the music content. Move around the room and continue talking. Talk in a loud and quiet voice.
6. Stop playback from the music player.
7. In *Audacity*, click on the **Stop** button (or press **space**) to stop recording.



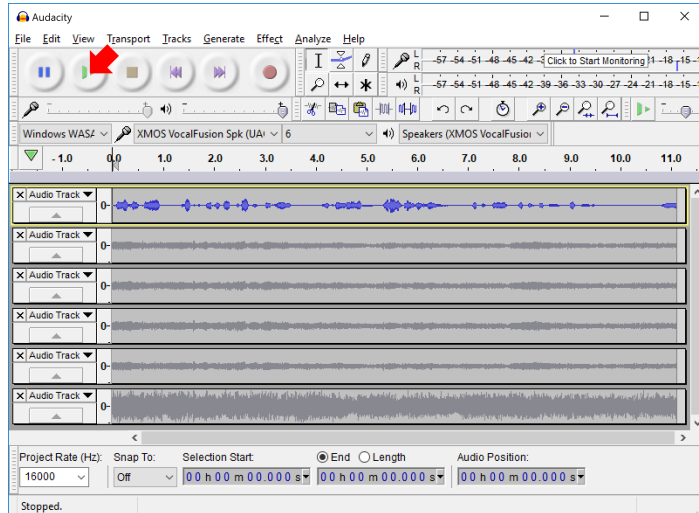
8. In *Audacity*, highlight Audio Track 2 and click the **Solo** button (or press **SHIFT + s**) to only playback Audio Track 2.



9. In *Audacity*, click on the **Play** button (or press **space**) to playback Audio Track 2 only.

You will now hear the raw signal as captured by one of the microphones – so a combination of the playback signal, your voice and any other background noise.

10. In *Audacity*, highlight Audio Track 1 and click the **Solo** button (or press **SHIFT + s**) to only playback Audio Track 1.



You will now hear your voice only.

- ▶ Acoustic Echo Cancellation (AEC) will have removed the playback music.
- ▶ Adaptive voice beam-forming will have isolated your voice.
- ▶ Noise suppression and de-reverberation will have removed background noise.
- ▶ Automatic gain control will always present your voice at a useful volume level.

5 Voice Activity Detector

The XVF3100 includes a Voice Activity Detector (VAD). In this demonstration, when voice activity is detected, the centre red LED will be momentarily illuminated.

- ▶ Try talking to the evaluation kit, both with and without music playing.

6 Keyword detection

The XVF3100 includes TrulyHandsfree™ voice control technology from Sensory. This capability continually monitors the extracted voice signal *listening* for a trigger key-word/phrase. This demonstration uses the trigger phrase **Hello Blue Genie**, optimized for US English pronunciation.

- ▶ Say “Hello Blue Genie”. Try this both with and without music playing.

When this phrase is *heard*, the demonstration illuminates the ring of red LEDs around the edge of the evaluation kit.

In an end product, the key-word/phrase may be customised (contact Sensory directly) and could be used to trigger other activities; such as muting the playback audio, or sending voice content to a third party voice service.

7 Direction of Arrival indication

The XVF3100 also includes a Direction of Arrival (DoA) indicator. In this demonstration, once “Hello Blue Genie” has been *heard* and you continue talking, a red edge LED will be illuminated to indicate the direction your voice is coming from. The Voice Activity Detector will then keep the same LED illuminated for as long as you continue talking.

8 Useful information

8.1 Speaker selection

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

- ▶ The speaker should be a mono speaker, so providing a single point audio source.
- ▶ The amplifier in the powered speaker should have linear gain. Non-linear gain (e.g. soft clipping) should be disabled or avoided.
- ▶ Any audio processing available on the speaker should be disabled.

An example powered speaker to use with the xCORE VocalFusion Speaker Evaluation Kit is the Logitech Z50. This onmi-directional (upwards facing) mono powered speaker has no sound processing.

8.2 Windows (audio) drivers

Windows includes a USB Audio Class 1 (UAC1) driver and so natively supports the xCORE VocalFusion Speaker Evaluation Kit.

When a configured evaluation 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 evaluation kit is connected (for example, an evaluation 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 evaluation kit will not operate correctly.

If this occurs, disconnect the evaluation 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 evaluation kit; fresh instances of the UAC1 driver will then be correctly installed.



Copyright © 2017, 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.
