EXHIBIT 30

Audio Hardware Configurations

Overview

Alexa Voice Service (AVS) is Amazon's intelligent cloud service that allows you to voice-enable connected products that have a microphone and speaker. By integrating AVS, your customers immediately gain access to Alexa's core capabilities and a growing library of third-party skills.

Alexa customers expect a familiar experience. Use the design and implementation guidance below to ensure that your AVS integration meets their expectations:

- Application Examples – Identify your product's form factor and choose how customers will interact with Alexa.

- ASR Profile Selection – Determine the best automatic speech recognition (ASR) profile for your implementation.

- Hardware and Audio Algorithms – Review techniques used to optimize speech recognition.

- Third-party Resources – Learn about providers that offer hardware and software solutions.

As you integrate AVS, please make sure to reference our Amazon Developer Services Agreement (https://developer.amazon.com/support/legal/da) and Functional Requirements for AVS (/docs/alexa-voice-service/functional-requirements.html).

Typical Application Examples

Alexa allows customers to interact with products in the most natural way possible – with their voice. Whether you're building a hand-held device like a TV remote, a wearable like a smart watch, or you want a hands-free experience for your connected speaker or home intercom, AVS provides a way for your customers to speak to Alexa.

There are two ways to initiate an interaction with Alexa: touch and voice. Touch-initiated interactions rely on a physical control like the Amazon Fire TV remote or Amazon Tap. Voice-initiated interactions leverage the wake word "Alexa", like Amazon Echo.

This table highlights typical application examples for push-to-talk, tap-to-talk and voice-initiated products:

(http

(https://images-
na.ssl-images-
amazon.com/images
/G/01/mobile-
apps/dex/alexa/alex
a-voice-
service/docs/fire-tv-
web.png)

(https://images-
na.ssl-images-
amazon.com/images
/G/01/mobile-
apps/dex/alexa/alex
a-voice-
service/docs/tap-
web.png)

(https://images-
na.ssl-images-
amazon.com/images
/G/01/mobile-
apps/dex/alexa/alex
a-voice-
service/docs/echo-
web.png)

Application	Push-to-talk	Tap-to-talk	Voice-initiated (Wake Word)
Remotes	●		
Wearables	●	●	
Mobile Apps	●	●	●
Portable Speakers		●	●
Home Audio		●	●
Intercoms		●	●
Smart Home / Appliances	●	●	●
Automotive		●	●
Personal Computers		●	●
Smart TV / Set Top Boxes			●

Automatic Speech Recognition Profiles

Alexa uses a combination of automatic speech recognition (ASR) and natural language understanding (NLU) to understand customer speech and respond with precision. ASR converts customer speech into text, NLU converts that text into intents for Alexa to act on. At the end of this process Alexa sends directives to your product instructing it to perform an action, like playing music.

AVS provides a choice of three ASR Profiles tuned for different products, form factors, acoustic environments and use cases. The profile parameter is sent to Alexa in the payload of each Recognize event (/docs/alexa-voice-service/speechrecognizer.html#recognize) and is also used to indicate if the end of customer speech will be determined by your client or in the cloud (how the end of speech is determined is called speech endpointing).

(http

The following table highlights which ASR Profiles are commonly associated with different customer interactions.

	Push-to-talk	Tap-to-talk	Voice-initiated (Wake Word)	
Listening Range	Up to 2.5 ft.	Up to 5 ft.	Up to 5 ft.	Up to 20 ft.
ASR Profile	"CLOSE_TALK" (https://developer.amazon.com/public/solutions/alexa/alexa-voice-service/reference/speechrecognizer#recognize)	"NEAR_FIELD" (https://developer.amazon.com/public/solutions/alexa/alexa-voice-service/reference/speechrecognizer#recognize)	"NEAR_FIELD" (https://developer.amazon.com/public/solutions/alexa/alexa-voice-service/reference/speechrecognizer#recognize)	"FAR_FIELD" (https://developer.amazon.com/public/solutions/alexa/alexa-voice-service/reference/speechrecognizer#recognize)
Speech Endpointing	Client	Cloud	Cloud	Cloud

Hardware and Audio Algorithms

The correct hardware configuration and audio processing algorithms can improve your product's listening sensitivity for the wake word Alexa. This is especially true if your product is designed for music playback or intended for use in noisy environments. Any type of nonlinear processing on the audio input, such as traditional noise reduction algorithms or automatic gain controls, should **not** be used.

The following table highlights typical configurations:

	Push-to-talk	Tap-to-talk	Voice-initiated (Wake Word)	
Listening Range	Up to 2.5 ft.	Up to 5 ft.	Up to 5 ft.	Up to 20 ft.
Wake Word			●	●
# of Microphones	1	1	1+	2+
AEC			●	●
Beamforming			>2 Microphones	>2 Microphones

Acoustic Echo Cancellation (AEC)
In speech recognition systems, the term "acoustic echo" refers to the signal that is played out of a loudspeaker and captured by a microphone in the vicinity of the loudspeaker. The acoustic echo is a source of interference for the ASR engine since it is simultaneously captured along with the customer's voice at the microphone. The goal of AEC is to remove the acoustic echo component from the microphone signal, so that the customer's voice can be clearly understood by the ASR engine. The AEC algorithm functions by adaptively estimating the acoustic echo path (and thereby the acoustic echo) between the loudspeaker and microphone components. The estimated acoustic echo is then subtracted from the microphone signal to obtain a near echo-free microphone signal. An AEC-processed microphone signal is ideally free of acoustic echo.

(htt

However, because of system non-linearities and room acoustics, not all echo is typically removed. Only linear AEC should be applied for ASR. Nonlinear processing to further clean up the echo should not be used. In addition, any type of nonlinear processing on the output

path, such as compression or limiting, should be part of the reference audio sent to the AEC.

Beamforming
Beamforming is a signal processing technique for multi-microphone arrays that emphasizes the customer's speech from a desired direction while suppressing audio interference from other directions. These algorithms result in an increase in SNR and a reduction in reverberation in the audio signal from the desired direction that improves the accuracy of speech recognition systems, especially for far-field. Only linear processing based versions of beamforming should be used for ASR.

Third-Party Resources *

Visit Development Kits for AVS (https://developer.amazon.com/alexa-voice-service/dev-kits) for a complete list of third-party providers with wake word and audio processing solutions.

Disclaimer

Amazon makes no warranty or representation regarding, does not endorse, and is not in any way responsible for any third party solutions or any content or materials provided by such third parties. If you decide to visit any linked website, you do so at your own risk and it is your responsibility to review the terms of use, privacy policy and any other relevant legal notices on such site.

Resources

- Functional Design Guide (/docs/alexa-voice-service/functional-requirements.html)

- AVS API Overview (../alexa-voice-service/api-documentation-overview.html)

- Forum and Knowledge Base (https://forums.developer.amazon.com/spaces/38/index.html)

(http